

## 差分演化法在最佳化向量量化式資訊隱藏之應用

杜淑芬<sup>1</sup>

黃冠雄<sup>2</sup>

<sup>1</sup> 中國文化大學資訊管理所 副教授 dsf3@ulive.pccu.edu.tw

<sup>2</sup> 中國文化大學資訊管理所 碩士班研究生 andymeo@gmail.com

### 摘要

在資訊隱藏的領域中，隱藏量(Hiding Capacity)是一項重要的議題，因此有一些學者提出結合影像壓縮技術的資訊隱藏方法，這類的方法會將秘密影像先行壓縮，再將壓縮後的影像藏入偽裝媒體中，在這些方法中最常被使用的影像壓縮技術為向量量化(Vector Quantization, VQ)。最近 Shie 等學者提出將秘密影像利用向量量化壓縮，並將壓縮後產生的索引表藏入編碼簿中，因此僅需將編碼簿傳給接收端，即可還原秘密影像。Shie 等學者的方法可提高藏密量與降低傳輸量，但其方法會破壞編碼簿，使得還原之秘密影像品質降低。本研究的目的便在於改善 Shie 等學者的方法，在同時保有該方法的優勢下，盡可能地提昇還原的秘密影像品質。本研究利用差分演化法改變編碼字的排列，可使索引表藏入編碼簿時，對編碼簿的改變量最小，使得還原的秘密影像品質比 Shie 等學者的方法更好。

**關鍵詞：**資訊隱藏、向量量化、最低位元置換法、差分演化法

## 差分演化法在最佳化向量量化式資訊隱藏之應用

### 1. 前言

在資訊隱藏(Information Hiding)的領域中，偽裝學(Steganography) (Petitcolas et al., 1999；陳文淵、卓江南，2003)即是將資訊秘密隱藏且使人無法察覺的隱藏技術，其中技術偽裝學(Technical Steganography)，主要是借助額外的媒體來隱藏秘密資訊。而良好的資訊隱藏技術必須滿足三點要求(呂慈純、陸哲明、張真誠，2007)：安全性(Security)、不可察覺性(Imperceptibility)與高資訊負載量(Payload)。在數位影像的偽裝學中，秘密影像會藏入至一張看似不重要的影像中，以確保攻擊者不易察覺，這張影像稱為掩蓋影像(Cover image)，而掩蓋影像藏入秘密影像後則稱為偽裝影像(Stego Image)。

掩蓋影像必須有足夠的空間來隱藏秘密影像，通常數位影像的資料量不會太小，為了能完全藏入秘密，必須選擇一張更大的影像來做為掩蓋影像，如此便造成網路傳輸的負擔，為了減輕網路負擔與加速傳輸時間，有許多學者致力於影像偽裝學的藏密量的提昇。由於人眼在辨識影像時，無法察覺細節上的差異，因此有學者提出利用失真壓縮的方法，將秘密影像先行壓縮，降低秘密影像的資料量，進而可降低掩蓋影像的資料量。在眾多的失真壓縮技術中，向量量化(Vector Quantization, VQ)由於具有極高的壓縮率與不錯的還原品質，因此近年來有一些學者提出結合向量量化的資訊隱藏法(Chen et al., 1998；Hu, 2003；Lin and Shie, 2004；Shie et al., 2010)。

其中 Shie et al. (2010)提出將秘密影像利用向量量化壓縮，並將壓縮後所產生的索引表(Index Table)藏入編碼簿(Codebook)中，因此傳送端僅須將編碼簿傳送出去，接收端便可還原與解壓縮秘密影像，不僅提高了藏密量同時大幅減少了資訊傳輸量，他們的研究也指出比起其他類似的方法，其所還原的秘密影像品質亦較佳，只要編碼簿的品質不錯，便可還原出非常近似的秘密影像。不過他們利用最低位元置換法將索引表藏入或替換掉編碼簿的最低位元，很明顯地，編碼簿會受到破壞，因此接收端用來還原影像的編碼簿已經與壓縮影像時所用的編碼簿不同，即使傳送端所使用的編碼簿品質良好，接收端所還原的影像品質勢必會被影響。

在向量量化編碼中，影像會被切割成固定大小的區塊，每個區塊會與編碼簿中的編碼字一一比對，找到相近的編碼字後，該區塊便會用近似編碼字的索引編號取代，最後整張影像便成為一個索引表。很顯然地，就同一本編碼簿而言，若將編碼字的順序重排，則每個編碼字的編號也會改變，影像編碼後形成的索引表也會有所不同。在 Shie 等學者的方法中，他們將秘密影像編碼後產生的索引表直接替換掉編碼簿的最後幾個位元平面，如果索引表的位元串與編碼簿最後幾個位元平面的位元串剛好相符，那麼編碼簿可以說完全不會受到破壞。既然不同的編碼字順序會產生不同的索引表，如果我們能找到其中一種編碼字的順序，其所產生的索引表的位元串，與編碼簿的最後幾個位元平面的位元串相近，那麼便可改善 Shie 等學者方法中，接收端影像還原品質降低的缺點。

本研究的目的是，便是希望索引表位元串最接近編碼簿的編碼字順序。若編碼簿有  $N_c$  個編碼字，則編碼字的可能排列方式就有  $(N_c)!$  種，一本向量量化的編碼字不會太少，有可能有 128 個或 256 個，甚至 1024 個，以 128 個編碼字而言，便有 128! 種排列方式，

已是一個非常大的天文數字，若採用動態規劃法(Dynamic Programming) (Chang et al., 2003)或貪婪演算法(Greedy Algorithm) (Chang et al., 2002)這類的決定性演算法(Deterministic Algorithm)，在這麼大的解答空間去搜尋最佳解，求解效率往往不佳，而非決定性演算法(Non-Deterministic Algorithm)由於具備平行運算的特性，比起決定性演算法更適合求解解答空間龐大的問題。本研究採用差分演化法(Differential Evolution, DE)做為求解最佳編碼字順序的方法。最佳化方法應該要能找到全域最佳解，收斂速度要夠快，同時控制參數要少以便容易使用(Onwubolu and Davendra, 2009)，而差分演化法在許多測試問題上，都能符合上述所說的特性，這也是本研究採用差分演化法的原因。

## 2. 文獻探討

### 2.1 向量量化編碼法

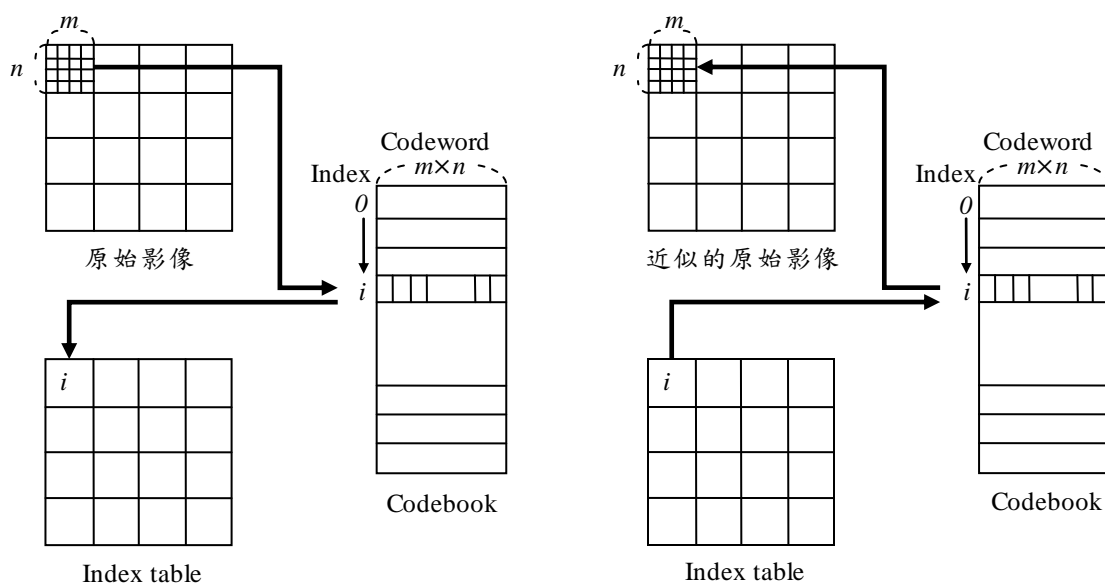


圖 1 向量量化編碼法之壓縮影像流程圖

圖 2 向量量化編碼法之還原影像流程圖

向量量化編碼法(Vector Quantization, VQ)最早是由 Gray 於 1984 年提出(Gray, 1984 ; 張真誠、黃國峰、陳同孝, 2000)，做法就是先藉由 LBG 演算法(Linde et al., 1980 ; Chang and Hu, 1998)產生一本由許多具有代表性向量(編碼字, Codeword)所組成的編碼簿(Codebook)，主要就是從一堆編碼字(向量)樣本中，用分群的方法找出能代表每個分群的編碼字，來形成一本編碼簿。接下來，如圖 1 將原始的數位影像分割成許多大小相同的  $m \times n$  區塊，每一區塊都是由  $m \times n$  個像素所組成且依序由左至右、由上至下地組成一個向量，其向量的維度即  $m \times n$ 。在進行向量量化編碼時，只要從這本具有代表性的編碼簿中，以歐幾里得距離(Euclidean Distance)找出影像中每個區塊與編碼簿最相近的編碼字，再分別用最相近編碼字的索引  $i$  來分別代表每個影像區塊的向量，如此就完成了整張影像向量量化編碼的壓縮，其壓縮碼也稱為索引表(Index Table)。解壓縮時，如圖 2，只要用索引表中的各索引值  $i$  去找出編碼簿中各索引所對應的編碼字來進行還原，即可得到近似的原始影像。

由於向量量化編碼法是用一個量化值來取代多維度的向量值，如此就具有明顯的資料壓縮效果且也因此造成部分原始資料失真，但也不失為一個簡單且有效率的壓縮技術。所以，一本大小適中且具有代表性的編碼簿才可以確保向量量化壓縮後的影像品質。

## 2.2 強化型差分演化法

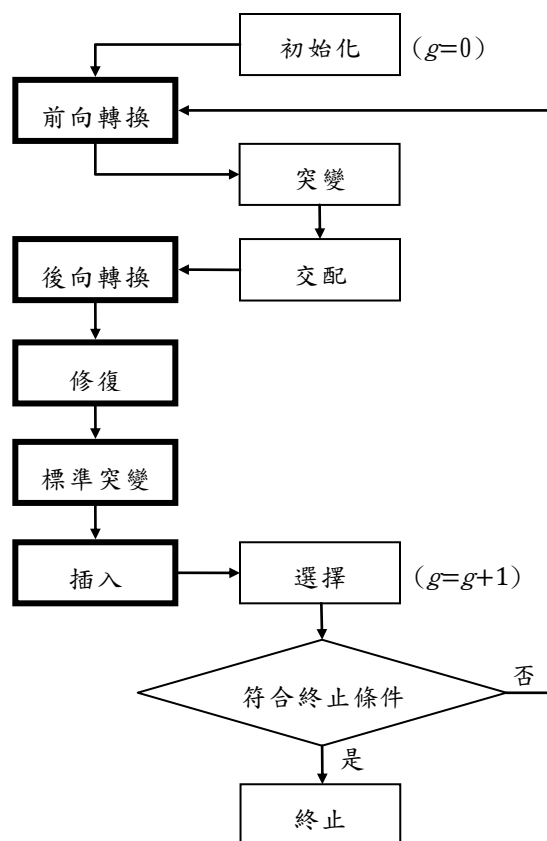


圖 3 強化型差分演化法流程圖

差分演化法(Differential Evolution, DE)是由 Price and Stom 於 1995 年提出(Price et al., 2005)，其為一族群式的演化式計算方法，它將具有  $n$  個決策變數的問題解答視為分佈在  $n$  維空間上的點，並以  $n$  維向量來表示它們。一開始會先隨機選  $N_p$  個向量組成初始的族群，而族群中的每個個體(稱為目標向量)會與各目的試驗向量競爭，勝出者可成為下一世代族群的成員。試驗向量則是經由下列兩種運算產生：

(1) 突變(Mutation)：利用方程式(1)產生突變向量  $v_i$ ：

$$v_i = x_{r_0} + F \cdot (x_{r_1} - x_{r_2}) \quad (1)$$

其中  $i$  為目標向量在族群中的編號，且  $i \neq r_0 \neq r_1 \neq r_2$ ， $F \in (0,1+)$  為一縮放因子(Scale Factor)。

(2) 交配(Crossover)：以目標向量  $x_i$  與突變向量  $v_i$  進行重組，產生試驗向量  $u_i$ 。重組的方式有兩種，一種稱為指數交配(Exponential Crossover)，另一稱為二項交配(Binomial Crossover)。

由於原始的 DE 是設計來求解連續型數值的問題，後續便有一些學者提出改良的方

法，使 DE 可求解離散型數值的問題(Onwubolu and Davendra, 2009；Lichtblau, 2009；Tasgetiren et al., 2009；Liang et al., 2009；Zelinka, 2009)。Onwubolu and Davendra 於 2009 年提出強化型差分演化法(Enhanced Differential Evolution, EDE) (Onwubolu and Davendra, 2009)，採用一種非常直覺的改良方式，使 DE 可求解離散型數值的問題，其流程如圖 3 所示，粗框的部分即為 Onwubolu and Davendra 在原有的 DE 中加入的運算。他們將離散型數值的向量經由方程式(2)的前向轉換(Forward Transformation)公式，轉換成連續型數值向量：

$$x'_{j,i} = -1 + \frac{x_{j,i} \cdot f \cdot 5}{10^3 - 1} \quad (2)$$

其中  $j$  代表向量中元素的編號， $i$  為向量在族群中的編號， $f$  為縮放因子(Scaling Factor)。如此一來，便可使用原始的 DE 來進行演化，但在計算解答品質時，必須以原來的離散型數值，因此在選擇階段前，會將連續型數值向量經由方程式(3)的後向轉換(Backward Transformation)公式，轉換成離散型數值向量：

$$x_{j,i} = \text{int} \left[ \frac{(1+x'_{j,i}) \cdot (10^3 - 1)}{f \cdot 5} \right] \quad (3)$$

其中  $\text{int}[\cdot]$  代表四捨五入的函數。為了修正轉換過程中產生的不合法解答，以及提升解答的品質，Onwubolu and Davendra 另外提出了修復、標準突變，與插入三個運算。

### 2.3 Shie 等學者的方法

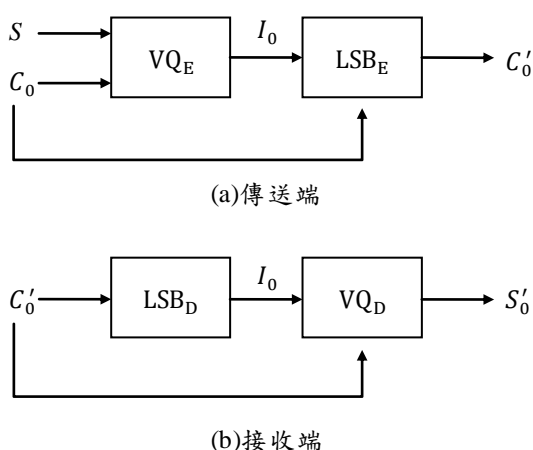


圖 4 Lin and Shie 的方法之流程簡圖：(a)傳送端 (b)接收端

本研究主要針對 Shie et al. (2010)的方法進行改善，而此篇則是架構在 Lin and Shie (2004)的方法之上，因此我們先介紹 Lin and Shie (2004)的方法。圖 4 為 Lin and Shie 提出的方法的流程簡圖，為了說明起見，我們假設只有單一張秘密影像，並且利用 LBG 演算法產生了編碼簿  $C_0$ 。根據  $C_0$  將秘密影像  $S$  編碼成索引表  $I_0$ ，最後  $I_0$  藏入編碼簿  $C_0$  的最低位元，得到藏有索引表的編碼簿  $C'_0$ ，在接收端取得藏有索引表的編碼簿  $C'_0$  後，直接取出原本的索引表  $I_0$ ， $I_0$  再根據編碼簿  $C'_0$  解碼，取得還原之秘密影像  $S'_0$ 。Shie et al. (2010) 進一步利用最低位元置換法(Least-significant-bit(LSB) substitution)，將  $C'_0$  藏入一張有意義的影像中。

Lin and Shie (2004)的方法可提昇藏密量與降低通訊量，然而接收端使用  $C'_0$  解碼，而非  $C_0$  解碼，造成解碼影像的品質下降，這點在 Shie et al. (2010)的文章中亦有提到，並建

議 $I_0$ 藏入 $C_0$ 時，最多僅能使用編碼字的每個元素的末3個位元，否則接收端的解碼影像會與秘密影像有較大的差異。

### 3. 研究方法

#### 3.1 問題定義

在 Shie 等學者的方法中，索引表藏入編碼簿後，通常會對編碼簿造成破壞，使得接收端解碼的影像品質，會不如使用未破壞的編碼簿所解碼的影像。然而，若調整編碼字的順序，我們有可能降低索引表藏入編碼簿時所造成的破壞，茲舉一例說明。考慮圖 5 的例子，假設秘密影像  $S$  是一張  $4 \times 4$  大小的影像，並切割成 4 個不重疊的  $2 \times 2$  區塊，圖中有兩本編碼簿  $C_1$  與  $C_2$ ，其編碼字完全相同，只是順序不同。秘密影像  $S$  以  $C_1$  編碼得到索引表  $I_1$ ，若以  $C_2$  編碼則得到的索引表  $I_2$ 。而圖 6 是編碼簿  $C_1$ 、 $C_2$  與索引表  $I_1$ 、 $I_2$  的二進位值，觀察圖 6(a) 發現，若將索引表  $I_1$  藏入  $C_1$  的第一個編碼字的最末位元，編碼簿  $C_1$  完全不會被破壞；反觀圖 6(b)，索引表  $I_2$  與  $C_2$  第一個編碼字的最末位元串完全不同，在藏入  $I_2$  後， $C_2$  便會有所改變。因此本研究的主要目的，即是提出調整編碼字順序的方法，使得索引表藏入編碼簿時，編碼簿被修改的幅度最小。

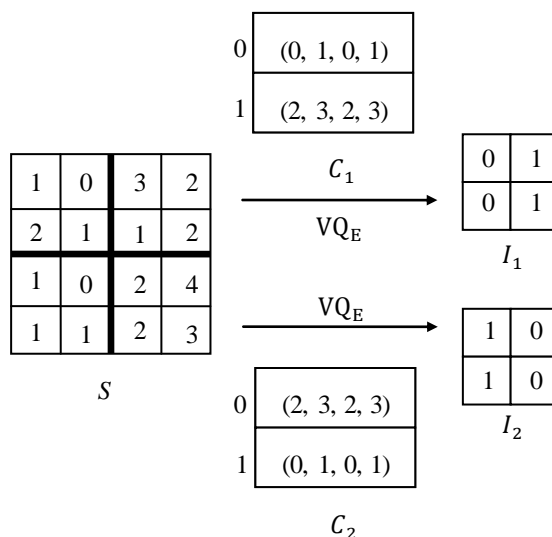


圖 5 秘密影像進行向量量化編碼

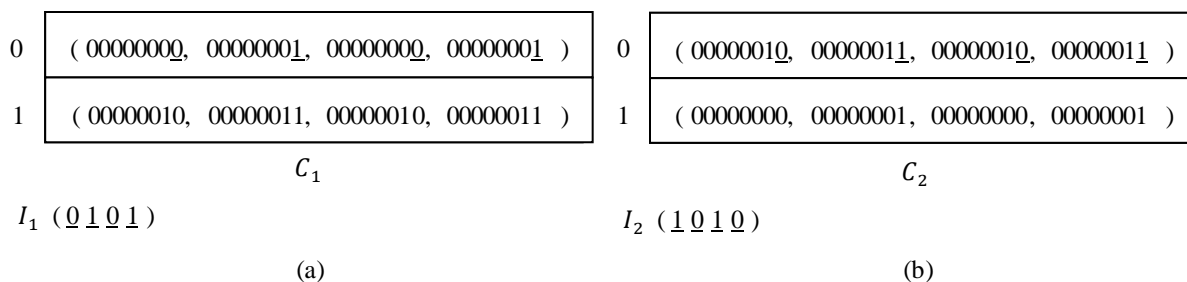


圖 6 編碼簿與索引表之二進位表示

為了說明本研究方法，我們先定義本研究所使用到的符號，令  $S$  代表一張大小為  $w \times h$

的灰階影像， $LBG(\cdot)$ 代表產生編碼簿的演算法，根據方程式(4)，我們可產生代表  $S$  的編碼簿  $C_0$ ：

$$C_0 \leftarrow LBG(S, N_c) \quad (4)$$

其中  $N_c$  代表編碼簿大小， $C_0 = \{V_i \mid i=1, 2, \dots, N_c\}$ ，編碼字  $V_i = (x_1, x_2, \dots, x_{m \times n})$  為一向量， $m \times n$  為影像區塊的大小，亦代表編碼字的維度。令  $VQ_E(\cdot)$  代表向量量化編碼的函式， $VQ_D(\cdot)$  代表向量量化解碼的函式，方程式(5)與方程式(6)即一般向量量化編碼與解碼的過程：

$$I_0 \leftarrow VQ_E(S, C_0) \quad (5)$$

$$S_0 \leftarrow VQ_D(I_0, C_0) \quad (6)$$

令  $Permute(\cdot)$  代表產生一集合元素各種排列的函數，利用方程式(7)產生  $C_0$  的所有排列：

$$C \leftarrow Permute(C_0) \quad (7)$$

其中  $C = \{C_j \mid j=1, 2, \dots, (N_c)!\}$ ，且  $C_j = (V_1^j, V_2^j, \dots, V_{N_c}^j)$ ，很明顯地， $C$  中的每個元素所產生的索引表皆不同，而每個索引表解碼的結果皆相同，如方程式(8)與方程式(9)所示：

$$I_j \leftarrow VQ_E(S, C_j) \quad (8)$$

$$S_0 \leftarrow VQ_D(I_j, C_j) \quad (9)$$

若將索引表藏入編碼簿，則接收端解碼出來的影像便可能與  $S_0$  不同。令  $LSB_E(\cdot)$  與  $LSB_D(\cdot)$  分別代表最低位元置換法的隱藏與取出的函式，方程式(8)與方程式(10) ~ (12)代表以  $C$  中的某個編碼簿將  $S$  編碼，並將索引表藏入編碼字最低  $d$  個位元，之後在接收端解碼的過程(即圖 4(b)的流程)：

$$C'_j \leftarrow LSB_E(I_j, C_j, d) \quad (10)$$

$$I_j \leftarrow LSB_D(C'_j, d) \quad (11)$$

$$S'_j \leftarrow VQ_D(I_j, C'_j) \quad (12)$$

方程式(12)所得到的解碼影像  $S'_j$  與方程式(9)的  $S_0$  通常不會相同，在向量量化中，如果編碼簿的品質夠好，則  $S_0$  會與  $S$  很接近，但方程式(12)使用被破壞的編碼簿解碼，所得到的解碼影像可能不像  $S$  那麼好，而每個可能的  $S'_j$  與  $S$  的差異不盡相同，本研究的目的便是希望找出  $C^*$ ，其中  $C^*$  定義如方程式(13)：

$$C^* \leftarrow \arg \min_{C_j \in C} MSE(S, S'_j) \quad (13)$$

$MSE$ (Mean-Square Error)函數定義如下：

$$MSE = \left(\frac{1}{w \times h}\right) \sum_{x=1}^w \sum_{y=1}^h (p_{xy} - \hat{p}_{xy})^2 \quad (14)$$

其中  $p_{xy}$  代表原始秘密影像  $S$  的像素， $\hat{p}_{xy}$  代表解碼影像  $S'_j$  的像素。

### 3.2 秘密影像藏入流程

圖 7 為本研究的秘密影像藏入流程圖，茲說明如下：

#### 1. 產生初始編碼簿 $C_0$

首先本研究利用 LBG 演算法，訓練出一本秘密影像  $S$  的編碼簿  $C_0$ ，並以  $C_0$  所含的編碼字集合做為產生最佳化編碼簿  $C^*$  的基礎。

#### 2. 應用強化型差分演化法產生最佳化編碼簿 $C^*$

如上一小節所述，本研究希望藉由調整 $C_0$ 的編碼字順序，提昇接收端解碼影像的品質， $C_0$ 會有 $N_c$ 個編碼字，因此共有 $(N_c)!$ 種排列方式，假設 $N_c = 128$ ， $128!$ 已是一個天文數字，更何況更大的 $N_c$ 值。因此本研究採用強化型差分演化法求解，比起決定性演算法(Deterministic Algorithm)而言，可較有效率地找到近似最佳解。強化型差分演化法將問題的解答編碼成向量，假設一向量 $x(x_1, x_2, \dots, x_{N_c})$ 為演化族群中的一個個體，其代表的編碼字排列為 $(v_{x_1}, v_{x_2}, \dots, v_{x_{N_c}})$ 。另外，本研究採用 MSE(方程式(14))做為評估解答品質好壞的適應值函數。

### 3. 將秘密影像 $S$ 進行向量量化編碼

秘密影像  $S$  與編碼簿  $C^*$  經由向量量化編碼階段，產生索引表  $I^*$ 。

### 4. 將索引表 $I^*$ 藏入編碼簿 $C^*$ 中

索引表  $I^*$  藏入編碼字的每個元素的末  $d$  個位元，其中

$$d = \frac{\lceil \log_2 N_c \rceil \times w \times h}{(m \times n)^2 \times N_c} \quad (15)$$

最後，便產生一本藏有索引表  $I^*$  的編碼簿  $C^{**}$ 。

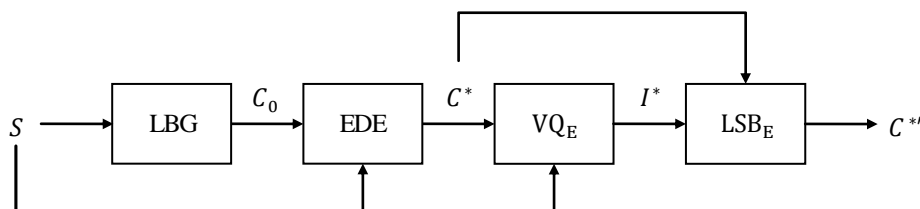


圖 7 秘密影像藏入流程圖

### 3.3 秘密影像取出流程

圖 8 為接收端取出秘密影像的流程，接收端取得  $C^{**}$  後，便從  $C^{**}$  中每個編碼字的每個元素的末  $d$  個位元，讀出秘密影像的索引表  $I^*$ ， $I^*$  與  $C^{**}$  輸出向量量化解碼函數 ( $VQ_D$ )，即可得出解碼影像  $S^{**}$ 。

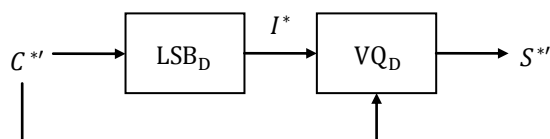


圖 8 編碼簿取出秘密影像流程圖

## 4. 實驗結果

圖 9 為本實驗之秘密影像，影像大小均為  $512 \times 512$ ，影像區塊大小為  $8 \times 8$ ，而差分演化法的相關參數值設定如表一所示。本研究使用兩種不同大小的編碼簿，分別為 256 與 512，並使用 PSNR(Peak-Signal-to-Noise Ratio) 做為衡量解碼影像品質的指標，PSNR 的方程式如下：

$$PSNR = 10 \log_{10} \frac{255^2}{\left(\frac{1}{w \times h}\right) \sum_{x=1}^w \sum_{y=1}^h (p_{xy} - \hat{p}_{xy})^2} \quad (16)$$



其中 $p_{xy}$ 代表原始秘密影像 $S$ 的像素， $\hat{p}_{xy}$ 代表解碼影像的像素。

表二與表三列出兩種編碼簿大小下，本研究方法與 Shie 等學者方法的解碼影像 PSNR 值，由表二與表三可看出，本研究的解碼影像品質皆優於 Shie 等學者方法的解碼影像品質，足見本研究提出改變編碼字的排列順序，的確可提昇解碼影像的品質。



圖 9 秘密影像(影像大小：512×512)

表一：差分演化法參數

參數	參數值
族群大小( $N_p$ )	50
DE 策略	DE/rand/1/bin
縮放因子( $F$ )	0.7
交配率( $C_r$ )	0.8
停止條件	最大世代數 100

表二：實驗結果一( $N_c = 256$ )

影像	Shie 等學者	本研究方法
Lena	28.932	28.948
Pepper	28.735	28.747
Airplane	27.598	27.615
Gold Hill	27.512	27.521
Boat	26.519	26.533

表三：實驗結果二( $N_c = 512$ )

影像	Shie 等學者	本研究方法
Lena	30.087	30.104
Pepper	29.876	29.893
Airplane	28.680	28.706
Gold Hill	28.511	28.523
Boat	27.540	27.564

## 5. 結論

Shie et al.(2010)所提出的方法，將秘密影像利用向量量化壓縮，再將壓縮後產生的索引表藏入編碼簿中，所以接收端在取出編碼簿時便可從中取出索引表，進而解碼秘密影像，雖然可提高藏密度與降低傳輸量，但缺點就是索引表在藏入編碼簿時會破壞編碼簿，使得還原的秘密影像品質降低。本研究目的便在於降低藏入編碼簿時的破壞程度，在同時保有該方法的優勢下，盡可能地提昇還原的秘密影像品質。本研究方法應用差分

演化法來改變編碼字的排列，使得索引表在藏入編碼簿時，對編碼簿的改變量最小，因此還原的秘密影像品質比 Shie 等學者的方法更好。

### 誌謝

本研究由國科會計畫 NSC 100-2221-E-034 -011 -補助支持，特此誌謝。

### 參考文獻

1. 呂慈純、陸哲明、張真誠，2007「多媒體安全技術」，全華圖書。
2. 張真誠、黃國峰、陳同孝，2000「電子影像技術」，松崗電腦圖書資料股份有限公司。
3. 陳文淵、卓江南，2003「資訊隱藏技術之研究」，勤益學報，第二十一卷，第一期：117-132 頁。
4. Chang, C.C. and Hsiao, J.Y. and Chan, C.S. "Finding Optimal Least-Significant-Bit Substitution in Image Hiding by Dynamic Programming Strategy" Pattern Recognition (36) 2003, pp:1583-1595.
5. Chang, C.C. and Hu, Y.C. "A Fast LBG Codebook Training Algorithm for Vector Quantization" IEEE Transactions on Consumer Electronics (44:4) 1998, pp:1201-1208.
6. Chang, C.C. and Lin, M.H. and Hu, Y.C. "A Fast and Secure Image Hiding Scheme Based on LSB Substitution" International Journal of Pattern Recognition and Artificial Intelligence (16:4) 2002, pp:399-416.
7. Chen, T.S. and Chang, C.C. and Hwang, M.S. "A Virtual Image Cryptosystem Based upon Vector Quantization" IEEE Transactions on Image Processing (7:10) 1998, pp:1485-1488.
8. Daley, W.M. "Data Encryption Standard (DES)" National Technical Information Service (46:3) 1999, Springfield, VA22161.
9. Denning, D.E.R. "Cryptography and Data Security" Addison-Wesley Publishing Company 1982.
10. Gray, R.M. "Vector Quantization" IEEE ASSP Magazine 1984, pp:4-29.
11. Hu, Y.C. "Grey-Level Image Hiding Scheme Based on Vector Quantization" IEEE Electronics Letters(39:2) 2003, pp:202-203.
12. Liang, Y.C. and Tasgetiren, F. and Pan, Q.K. and Suganthan, P. "Discrete/Binary Approach" Studies in Computational Intelligence (175) 2009, pp:139-162.
13. Lichtblau, D. "Relative Position Indexing Approach" Studies in Computational Intelligence(175) 2009, pp:81-120.
14. Lin, S.D. and Shie, S.C. "Secret Image Communication Scheme Based on Vector Quantization" IEEE Electronics Letters (40:14) 2004, pp:859-860.
15. Linde, Y. and Buzo, A. and Gray, R.M. "An Algorithm for Vector Quantizer Design" IEEE Transactions on Communications (28:1) 1980, pp:84-95.

16. Onwubolu, G.C. and Davendra, D. "Differential Evolution: A Handbook for Global Permutation-Based Combinatorial Optimization" *Studies in Computational Intelligence* (175) 2009.
17. Petitcolas, F.A.P. and Anderson, R.J. and Kuhn, M.G. "Information Hiding - A Survey" *Proceedings of the IEEE* (87:7) 1999, pp:1062-1078.
18. Price, K.V. and Storn, R.M. and Lampinen, J.A. "Differential Evolution A Practical Approach to Global Optimization" *Natural Computing Series* 2005.
19. Shie, S.C. and Lin, S.D. and Jiang, J.H. "Visually Imperceptible Image Hiding Scheme Based on Vector Quantization" *Information Processing and Management* (46) 2010, pp:495-501.
20. Tasgetiren, F. and Chen, A. and Gencyilmaz, G. and Gattoufi, S. "Smallest Position Value Approach" *Studies in Computational Intelligence* (175) 2009, pp:121-138.
21. Zelinka, I. "Discrete Set Handling" *Studies in Computational Intelligence*(175) 2009, pp:163-205.

# The Optimal Vector Quantization of the Data Hiding Scheme Using Differential Evolution

Shu-Fen Tu<sup>1</sup>

Kuan-Hsiung Huang<sup>2</sup>

<sup>1</sup>Graduate Institute of Information Management, Chinese Culture University  
Taipei, Taiwan, R.O.C.  
dsf3@ulive.pccu.edu.tw

<sup>2</sup>Graduate Institute of Information Management, Chinese Culture University  
Taipei, Taiwan, R.O.C.  
andymeo@gmail.com

## Abstract

Hiding capacity is an important issue for stegagnography; hence some researchers utilized image compression techniques to decrease the size of secret images and consequently increase the hiding capacity. Recently, Shie et al. utilized a common image compression technique, called Vector Quantization (VQ), to compress the secret image. The compressed result, i.e. an index table, is then embedded into the least significant bits of the codebook. The receiver can extract the index table and decode the secret image with the stego-codebook. Their method can increase the hiding capacity and decrease the transmission traffic. However, the stego-codebook degrades the decoded image quality. On the basis of Shie et al.'s method, we proposed permuting the order of codewords in the codebook, so that the bit stream of the index table can be close to the least significant bits of the codebook. There are thousands upon thousands of possible permutations; therefore, we adopt differential evolution to find out the near-optimal permutation. The experimental results show that our decoded image quality is better than that of Shie's et al.'s.

**Keywords:** Information Hiding, Vector Quantization, Least Significant Bit Substitution, Differential Evolution