以人類視覺之小波轉換為基礎的影像壓縮

程世州 張哲瑋 陳世融 蔡明學 郭盈君 黃金本
銘傳大學　資訊傳播工程學系
e-mail: hcptw@mcu.edu.tw

摘要

近年來隨著電腦網路、多媒體的普遍化，影像資料需求大量的增加，如何有效的儲存、傳送影像成為重要的課題。雖目前存在許多的影像壓縮方法，但這些方法皆以去除影像中的統計相關冗餘為主，甚少考量人類的視覺冗餘，也因此無法達到更高的壓縮效果。小波轉換的多重解析表示法為一公認符合人類視覺的模型，它可以由粗略到精細的描述物件，且可有很好的壓縮效果，而整數小波轉換，除了具有小波轉換的特性，更具備了方便計算以及快速處理的優點。由 Jayant 提出 JND(Just-Noticeable-Distortion) 的視覺編碼概念，由此概念延伸出一對影像信號視覺上的門檻值，如低於此門檻值則為人類視覺無法查覺的信號。本文提出以整數小波轉換為基礎結合 JND 的影像壓縮方法。

關鍵詞：影像壓縮、整數小波轉換、視覺模型

1.前言

現在的電腦已經擁有了大量的記憶空間、快速的處理能力、及傳輸速度，因此有能力來處理體積龐大的影像資料。而且，如今網際網路非常的盛行，影像資料已被廣泛的應用在網路上面，配合著文字的描述，以方便使用者在網路上進行各種應用。可是影像資料的體積實在是太龐大了，以一張 640x480 點的影像資料為例，它需要約 300k 個位元組的空間來儲存；若要存放一百張這樣的影像，那麼就需要花費約 30M 個位元組，至於傳輸方面，假設利用每秒 33.6k bps 速度的數據機來傳送先前的一張影像，需要約 1.2 分鐘，傳送一百張就要二個小時的時間。

整體而言，這些影像資料在空間及時間方面的需求，對於使用者是一項沈重的負擔，如果可以將影像資料事先加以壓縮後，再傳輸或儲存，這樣不但可以節省記憶空間，並且可以減少傳輸時間。

2.文獻探討

近年的研究已指出，以小波轉換為基礎的壓縮方法，會比以離散餘弦轉換為基礎的壓縮方法，有更好的壓縮效果。然而，這些小波編碼方法並沒有善用人類視覺的特性，去有效的移除視覺上的冗餘，以得到更好的壓縮效能。本研究即利用這個人類視覺系統(HVS)的特點，配合影像壓縮演算法，來達成視覺上較佳的影像壓縮效果;採用整數小波轉換方法，除保留傳統小波轉換的好處之外，且減少耗時的浮點數運算，由於都是進行整數的計算，可提升壓縮系統的處理速度，再配合人類視覺系統模型(JND model)，且針對各個次頻帶作水平、垂直、斜線方向的考慮，

使它對量化感知冗餘的去除和改進編碼效率非常有助益，藉由控制量化的級距(step size)值，能夠使影像即使在低位元率下也能有好的影像品質。

2.1 視覺冗餘

視覺冗餘[1,5,9,10]在影像中基本上是由於在空間上的對比與灰階度變化的刺激，以及人類視覺系統的敏感度不一致所引起的。而去移除視覺冗餘的方法有:利用人類視覺系統(HVS)的敏感度去調整量化器的步距(step size);及試著有效率的利用空間遮罩以達到隱藏失真。儘管如此，由於缺乏一個有效的量化預估模型去計算影像品質，所以沒有任何一個影像編碼方法可以提供一個簡單和有效的方法去移除視覺冗餘。

理想的恰可查知失真(JND)提供每一訊號被編碼依錯誤能見度的門檻，在此門檻以下重見錯誤是察覺不出的。恰可查知失真(JND)輪廓在影像是一個部分訊號特性的函數，像背景強度流明改變的活動力和優勢的空間頻率。這些特性在恰可查知失真(JND)輪廓需要一個有效的視覺模型。一旦一個影像的恰可查知失真(JND)輪廓獲得，視覺的失真量可以被估計，而且每個訊號的視覺重要性也可以被估計，因此，一個有效的視覺模型從輸入的影像去產生真正的恰可查知失真(JND)輪廓是不可或缺的。

次頻帶編碼是一個接近人類視覺模型的影像壓縮的方法。次頻帶基本的想法是分解空間的影像訊號成為窄的頻率次頻帶，每個頻帶會被分別的編碼，就像是有目的的結合人類視覺系統(HVS)和影像編碼演算法一樣。

2.2 整數小波轉換

小波轉換[2,3,4]的架構分成兩個部分，一個是小波分解（wavelet analysis），另一個是小波合成（wavelet synthesis）。小波分解的部分，小波分解首先要有一個原始數列的資料，這個原始數列是一個無窮數列。以影像壓縮為例，數列裡的每個元素的值，可以是影像中每個圖點的灰階值。有了這樣的一個原始數列之後，要將此數列分別進行低頻濾波以及高頻濾波分析這兩個步驟。如果將影像看成由具有平順的資料以及具細微的資料組合而成，那麼我們將原始數列經過低頻分析濾波的步驟之後，所產生的低頻數列，則將保留原始數列的一致性的資料，而將原始數列經過高頻分析濾波的步驟之後，所產生的高頻數列，則將保留原始數列的高度變化性的資料。因此小波分解的步驟其實就是將原始資料中的一致性資料與高度變化性資料分解出來。而在小波合成的部分，將低頻數列經過低頻合成濾波之後的數列，與高頻數列經過高頻合成濾波之後的數列，相加起來，就可以還原成原始的數列。

根據統計結果發現，一般的影像，大部分相鄰的區域會有相似的灰階度值，這意味著影像中一致性的資料佔大多數，而高度變化性的資料則佔少數，以量化的觀點來看，將原始的影像資料，經過小波分解得到低頻數列與高頻數列，高頻數列裡大部分的值都將非常接近 0，因此高頻數列具有相當大的壓縮空間。於是將小波分解與小波合成之間，再加上壓縮和解壓縮的步驟，這樣子便構成了小波壓縮的架構。

小波轉換主要架在三個主要的基礎理論之上，分別是階層式編碼架構（pyramid structme）、濾波器組理論（filter bank theory）以及次頻帶編碼（subband coding），可以說小波轉換統合了此三項技

術，因此小波轉換能將各種交織在一起的不同頻率組成的信號，分解成不相同頻率的信號，能有效的應用於編碼、解碼、檢測邊緣、資料壓縮，及將非線性問題線性化。

另外，整數小波轉換在壓縮應用上有幾點優點：

一、多重解析度分解提供了不同解析度下影像的資訊，並且轉換後的能量大部分集中在低頻部分，方便了對不同解析度下的小波係數分別設計量化編碼方案，以在提高圖像壓縮比的情況下保持好的視覺效果。小波轉換的多重解析度特性正適用於網路的應用，因為可以瀏覽一張影像根據解析度的不同由粗糙到細膩，不會因為網路的壅塞而無法看到整張影像。

二、小波分解和還原演算法是遞迴的，易於硬體實現。在一般的小波轉換過程中，需要經過浮點數相乘的運算，而使壓縮系統經常需要透過較長的時間去處理，整數小波轉換有別於一般傳統小波轉換，它可以在轉換步驟中都是以整數的方式表現出來，達到精確的還原影像，而不像傳統小波轉換整數係數的輸入，浮點數的係數輸出，因此更有利於壓縮及傳輸。

整數小波[6]分解的過程中，*x(n)*代表原始影像被分解的數列，*s(n)*對應到透過低頻分析濾波器輸出的平滑部分參數，以及 d(n)代表透過高頻分析濾波器輸出的細節部份的參數。接著設定濾波器係數，對於高頻分析濾波器，設定係數為{-1,2,-1}；對於低頻分析濾波器，設定係數為 {-1,2,6,2,-1}，在設定分析濾波器系數時，高頻分析濾波器的係數比低頻分析濾波器的係數短，所以會先獲得細節訊號 d(n)的輸出，然後再獲得平滑訊號 s(n)的輸出。它的公式如下：

$$d(n) = x(2n) - \left\lfloor \frac{x(2n-1) + x(2n+1)}{2} \right\rfloor \qquad (1)$$

$$s(n) = x(2n+1) + \left\lfloor \frac{d(n) + d(n+1)}{4} \right\rfloor \qquad (2)$$

使用整數小波轉換的優點是對於冗餘的資訊可以被偵測，且可以簡單的排除它們，相對提高壓縮系統效率。

2.3 區域特性的視覺編碼

在知覺視覺模型裡，主要使用了兩種方法:一個是神經生理學，另一個是精神物理學。神經生理學透過研究神經的細胞回應：精神物理學是研究人類視覺系統的巨觀特性的學問。人類的視覺系統被認為是一整個黑盒子特性，就像敏感性和遮罩的性質一樣，透過精神物理學主觀的實驗去鑑定。

區域性調適的感知影像編碼[11]是調整對比敏感性模型作為知覺的圖像編碼器，編碼器使用相同的圖像分解和相同的強度。多頻的分解，是使用普遍化的二次的鏡影濾波器組(GQMF)來分解。在多頻的分解中在被計算和被移除之後，輸入影像會被 GQMF 分解成數個子頻段。雖然 GQMF 只是一個人類視覺的多頻道的天然的近似模型，它也為頻段提供不同的空間頻率和方向(orientation)。區域性調適的感知編碼器的量化，是需要利用影像的區域變化，遮罩門檻進行計算就可以了。因此對於每個頻段都需要量化器的步距(step size)在區域中可提供屏蔽而不用傳送大量側邊訊息(side information)。

2.4 視覺模型

對整張影像建立視覺模型[5]JND 的公式如下:

$$\mathrm{JND}_{fb}(x,y) = \max\{f_1(bg(x,y), mg(x,y)), f_2(bg(x,y))\} \quad (3)$$

$$f_1(bg(x,y), mg(x,y)) = mg(x,y)\alpha(bg(x,y)) + \beta(bg(x,y)) \quad (4)$$

$$f_2(bg(x,y)) = \begin{cases} T_0 * (1 - (bg(x,y)/127)1/2) + 3 \\ \quad for \quad bg(x,y) \le 127 \\ \gamma * (bg(x,y) - 127) + 3 \\ \quad for \quad bg(x,y) > 127 \end{cases} \quad (5)$$

$$\alpha(bg(x,y)) = bg(x,y) * 0.0001 + 0.115 \quad (6)$$

$$\beta(bg(x,y)) = \lambda - bg(x,y) * 0.01 \quad \text{for} \quad \begin{matrix} 0 \le x < H \\ 0 \le y < W \end{matrix} \quad (7)$$

其中 $T_0 = 17$、$\gamma = 3/128$、$\lambda = 0.5$ 且 $bg(x,y)$ 和 $mg(x,y)$ 分別代表平均背景亮度(luminance)和亮度在像素 (x,y) 周圍的最大平均加權。H 和 W 表示影像的高和寬。函數 $f_1(x,y)$ 是模仿空間遮罩效應。變數 $\alpha(x,y)$ 和 $\beta(x,y)$ 是和背景亮度相關的函數。背景亮度值由 $f_2(x,y)$ 來決定。$mg(x.y)$ 橫跨在 (x,y) 的像素決定，是由計算在像素周圍的亮度改變的四個方向平均加權。如圖 2-4.1，四個運算子，$G_k(i,j)$，k=1,…,4，i,j=1,…,5，將用來執行這個計算，當離中心像素距離增加時，加權係數減少。

$$mg(x,y) = \max_{k=1,2,3,4}\{|\mathrm{grad}_k(x,y)|\} \quad (8)$$

$$grad_k(x,y) = \frac{1}{16}\sum_{i=1}^{5}\sum_{j=1}^{5} p(x-3+i, y-3+j) * G_k(i,j)$$
$$for \ 0 \le x < H, \ 0 \le y < W,$$

(9)

其中 $p(x,y)$ 代表在 (x,y) 的像素。平均背景亮度，$bg(x,y)$，經由一個加權的低通運算子計算，$B(i,j)$，i,j=1, …, 5，如圖 2-4.2。

$$bg(x,y) = \frac{1}{32}\sum_{i=1}^{5}\sum_{j=1}^{5} p(x-3+i, y-3+j) * B(i,j) \quad (10)$$

0	0	0	0	0
1	3	8	3	1
0	0	0	0	0
-1	-3	-8	-3	-1
0	0	0	0	0

G1

0	0	1	0	0
0	8	3	0	0
1	3	0	-3	-1
0	0	-3	-8	0
0	0	-1	0	0

G2

0	0	1	0	0
0	0	3	8	0
-1	-3	0	3	1
0	-8	-3	0	0
0	0	-1	0	0

G3

0	1	0	-1	0
0	3	0	-3	0
0	8	0	-8	0
0	3	0	-3	0
0	1	0	-1	0

G4

圖 2-4.1 計算四個方向流明改變加權平均的運算元(G1) 垂直 90 度 (G2) 45 度 (G3) 135 度 (G4) 水平 0 度

1	1	1	1	1
1	2	2	2	1
1	2	0	2	1
1	2	2	2	1
1	1	1	1	1

圖 2-4.2 計算平均背景流明的運算元

2.6 訊息編碼

算術編碼法與哈夫曼編碼法都是屬於無失真資料壓縮[12]，亦即壓縮過的資料可以完整重建成原來的模樣。這兩種方法所依據的理論基礎相當類似，都是利用出現機率較高的符號，使用較少位元的字碼來表示，出現機率較低的符號，則使用

較多位元的字碼來表示。

哈夫曼編碼法是利用二元編碼樹的方式來編碼，它會將整個編碼空間以最佳整數個位元的方式來分配。是一種不定長度的編碼，主要是根據狀態出現機率來編碼，出現越頻繁的狀態用較短的碼，出現機率越少的狀態用較長的碼，這樣整體的平均碼會大大的減少。

算術編碼則是用一個實數來表示要壓縮的資料，換句話說，它將整個編碼空間以精確至小數的方式來分配，以達到更佳的編碼效果。算術編碼法最大的特點在於不需要每個符號都編碼並且送出結果(霍夫曼法每個輸入符號都要編碼成整數個位元)，算術編碼是用一個介於 0 到 1 之間的實數來代表編碼到目前的情形，當輸入符號持續進入編碼時，這個實數會跟著改變，由於存在 0 到 1 之間的實數共有無限多個，因此任何待壓縮檔案都能用一個獨一無二的實數來代表它。

3. 提議的方法

本專研主要的目的是利用人類視覺系統的特性，以有效的進行整數小波轉換後的次頻帶編碼，使壓縮後的影像具高感知品質，且可以被維持在非常低的位元率下。

為了達到在理想的給定觀看距離條件下，編碼失真是不顯著的，利用恰可察覺的失真(JND)輪廓的方法將從兩方面著手：一為開發一個有效的人類視覺模型，經由被量化的感知冗餘，獲得一個準確的恰可查知失真(JND)輪廓。另一個是設計次頻帶編碼演算法，以移除感知的冗餘。功能方塊如圖 3

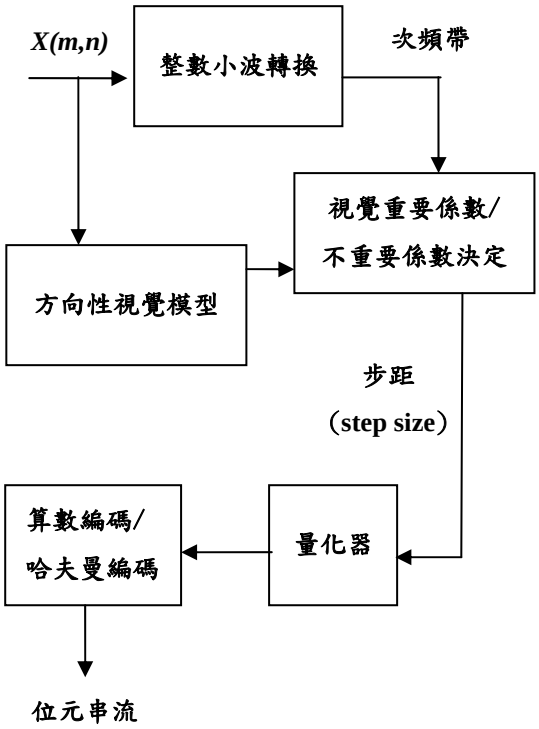


圖 3 人類視覺影像編碼方塊圖

3.1 參考文獻之視覺模型

為了可以跟本文所提出的方向性視覺模型做比較，所以模擬參考文獻中的視覺模型[5, 7, 8]，實作流程為：將整張原始圖經過整數小波轉換方法放入陣列 A，再將陣列 A 經過 JND 視覺模型，得到一個陣列 B，再將陣列 A 與陣列 B 做視覺冗餘的移除。

3.2 方向性視覺模型

方向性視覺模型參考[5]的作法，應用整數小波轉換的多重解析的特性，將整張原始圖經過整數小波轉換方法放入陣列 A，接著再將原始圖放入上述的 JND 後得到陣列 B，將陣列 B 分別進行水平運算、垂直運算，以及雙斜交叉運算並依序放入新的陣列 C 的右上角、左下角及右下角處，將會得到一個對應於整數小波轉換後的分頻階層圖，再將陣列 A 與陣列 C 做視

覺冗餘的移除。

水平運算演算法

1.將原圖經過 JND 視覺模型→2.每兩水平的點兩兩相乘且開根號→3.依照小波轉換的階層排列，一階是兩兩相加〈如 2×1矩陣〉除以 2，二階是八個相加〈如 4×2 矩陣〉除以 8，三階則有 32 個相加〈如8×4矩陣〉除以 32，四階則是 128 個相加〈如16×8矩陣〉再除以 128→4.依序放回每一階小波轉換所對應的階層區塊處。

垂直運算演算法

與水平運算相似，只是在水平運算演算法步驟二中改為上下兩個垂直的點相乘且開根號，以及在步驟三中，相加矩陣改為1×2、2×4、4×8、8×16，其他作法同水平運算，最後再依序放回每一階小波轉換所對應的階層區塊處。

雙斜交叉演算法

1.將原圖經過 JND 視覺模型→2.每次取四個點，以左上和右下相乘再開根號加上右上和左下相乘開根號為一組。→3.一階就是以一組除以 2，二階為四組除以 8，三階十六組除以 32，四階則是六十四組除以 128。→4.依序放回每一階小波轉換所對應的階層區塊處。

3.3 量化

量化的演算法，為了將經過整數小波轉換和 JND 後的圖片，移除不必要的冗餘值，量化的規則是將每一階的值依照各個區塊的平均去取的，平均數越大的將移去較少的值，而平均比較小的值移去較多的值，這是因為平均數較大的區塊含有比較大量的值，反之，平均比較小的區塊裡面的值比較少，因此移掉數量較多。演算法歸納如下:

1.將 JND 過後圖片的每一階的每一區塊算出各個平均值並放入 1 個 3×4 的陣列 *X* 裡。→2.陣列 *X* 的每一行的最大值求出放入 1×4 陣列 MAX 裡，每一行的最小值放入 1×4 陣列 MIN 裡。→3 將陣列 MAX 減去陣列 MIN 可得到一個新的陣列 S，因為每一階分為三個區塊，因此將陣列 S 去除以 3，可得到每一階的三個範圍，最小範圍為 MIN 至 MIN+S，中間範圍為 MIN+S 至 MIN+2×S，最大範圍為 MIN+2×S 至 MAX，如圖 3-3.1，接著使用陣列 X 去判斷範圍，因為最高階層的值比最低階層的值重要，因此採用了遞減規則，即是第一階範圍為 4 至 6，第二階範圍為 3 至 5，第三階為 2 至 4，依此類推。以 LINA 影像為例可得到如圖 3-3.2。以第一階為例：陣列 X 裡的第一行值落在範圍(MIN 至 MIN+S)的值被設定為 6，值落在範圍二(MIN+S 至 MIN+2S)的值設定為 5，依此類推。→4.依照陣列 SS 裡面的值來決定量化的範圍，假設某一區塊的值為 4，即是指這個區塊裡面的所有值變為 4 的倍數，值為 4、5、6、7 都將它改為 4，依照這樣可以刪去不少的值，而得到量化的結果。

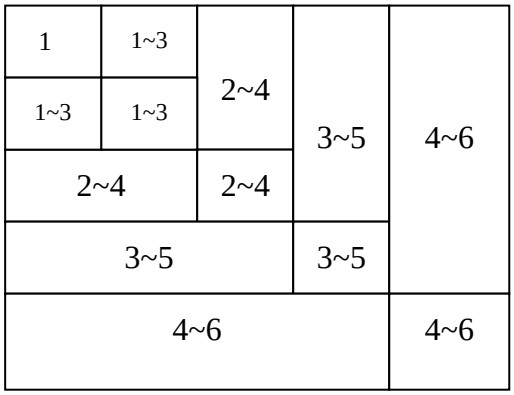


圖 3-3.1 區塊量化值 每一區塊範圍圖

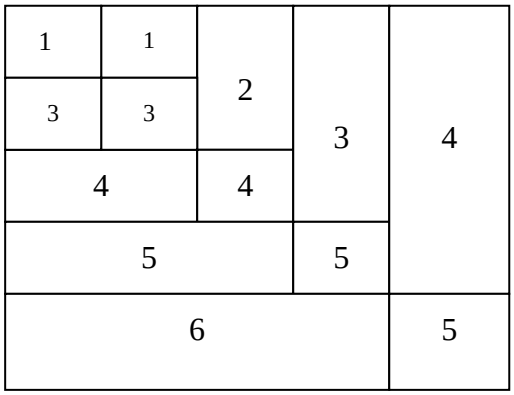


圖 3-3.2 區塊量化值 LENA 範例

4. 實驗結果與討論

本節提供的實驗結果，實驗環境為 IBM P4 3.0G 512M/2.5 筆記型電腦，專業版 Windows XP 系統軟體，Matlab 7.0 應用軟體。實驗結果分為四個部分，一為對影像進行整數小波轉換的分析及對各個次頻帶量化並且重建，二則是為了解量化的效果，使用了二張不同複雜度的影像進行實驗，先將影像進行多階的整數小波轉換，再對各個次頻帶進行各種不同步距的設定的量化器，以進行編碼及解碼，再重建影像以了解在各個不同步距時對影像重建品質的影響。三是影像進行整數小波轉換，再經由兩個不同的 JND 視覺模型方法，然後經過編碼再還原回來，藉以比較兩方法之間的優劣。

本實驗結果對第三部分影像重建品質採客觀的 PSNR 及 PSPNR，再加上主觀視覺評估方法。PSNR 是一個常見用來判斷經過影像處理後影像品質的好壞標準，其定義如下：

$$PSNR = 10 \times \log_{10} \frac{255^2}{MSE} \tag{10}$$

$$MSE = (\frac{1}{m \times n}) \sum_{x=1}^{m} \sum_{y=1}^{n} (p(x,y) - \hat{p}(x,y))^2 \tag{11}$$

基本上，PSNR 的值越高表示影像品質越好，但這並不是絕對的，因為假如將原圖中幾個黑的像素轉白，PSNR 的值仍然很高。所以，本實驗還加入了一個多考慮 JND 因素的判定標準 PSPNR，方法如下[5]：

$$PSPNR = 10 \times \log_{10} \frac{255^2}{\sqrt{\{[\,| p(x,y) - \hat{p}(x,y) | - JND_{fb}(x,y)]^2 \times \delta(x,y)\}}} \tag{12}$$

$$\delta(x,y) = \begin{cases} 1, & if \mid p(x,y) - \hat{p}(x,y) \mid > JND_{fb}(x,y) \\ 0, & if \mid p(x,y) - \hat{p}(x,y) \mid \le JND_{fb}(x,y) \end{cases} \quad for \quad 0 \le x < H, \quad 0 \le y < W \tag{13}$$

其中，PSPNR 值越高也是代表影像品質越好。有了以上兩個客觀的評斷標準後，再加入了主觀的視覺評估方法，此方法為經由十個人的主觀視覺來判斷影像的好壞，將他分為三類，若全數的人都認為還原圖與原圖一樣的話，則列為優，半

數以上不到全數者列為尚可，半數以下則列為劣。

4.1 整數小波轉換

對一張大小為 M×N 個圖點的影像進行壓縮，並不是把整張影像的灰階值放進一個大小為 M×N 的一維陣列裡，進行一維的小波轉換，因為這樣子只能發揮影像單一方向的灰階度關聯性。為了要充分利用二維影像中，橫向及縱向這兩個方向灰階度的關聯性，提高壓縮的效果。例如一張大小為 512×512 的 LENA 影像，如圖 4-1 (a)，經過橫向小波轉換以及縱向小波轉換的一至三階後，得到影像各別如圖 4-1 (b)，(c)，(d)所示，可以發現到各個對應影像的各次頻帶(subband)的內容。經過了一次的二維小波轉換後，可以發現到，其實橫向低頻縱向低頻區的資料其實還保留了部分與原始影像的一致的資料，可以說低頻區是原始影像的縮小版，因此這部分的資料還有可以壓縮的空間，可以再進一步地進行二維小波轉換。

一般的情況下，一張 512×512 個圖點大小的影像，大約進行四階二維小波轉換就可以了，不需要進行到資料區的寬度或高度變成 1 個圖點才停止。到了第四階解析度低頻區的資料，因為解析度蠻低的，因此那個區域裡面並不像原始影像中有大量的一致性資料，可以說相鄰的點的數值資料都非常的雜亂無章的，所以如果再進行小波轉換也不能在壓縮上得到好處，反而會浪費時間在小波轉換上。因此往後所討論的多樣解析度二維小波轉換，指的就是經過四次小波轉換的多樣解析度二維小波轉換。

(a) (b)

(c) (d)

圖 4-1 LENA 小波轉換圖 (a) 原圖 (b), (c), (d)小波轉換圖

4.2 量化步距影響

在本實驗針對經小波轉換後各個次頻帶給不同步距(step size)的量化後，評估對人類視覺的影響。

實驗一

對一自然的影像 LENA 進行實驗，對不同次頻帶採取遞增性的量化步距，對各種不同條件下所得主觀的視覺品質評估如表 4-1.1 所示，表中數值代表著各級次頻帶的量化位準值，在圖 4-2.1(b)與(c)中，經相關研究人員與原圖(a)比較後，無法察覺任何失真，在(d)中，需仔細觀察後，才會發現影像有些微失真，在(e)中，已經可以很明顯察覺出失真了

表 4-1. 1　LENA 品質評估表

重建 影像	第 四 級	第 三 級	第 二 級	第 一 級	察覺 失真
(b) 量化值	2	3	4	5	否
(c) 量化值	3	4	5	6	否
(d) 量化值	4	5	6	7	是
(e) 量化值	5	6	7	8	是

(a)

(b)

(c)

(d)

(e)

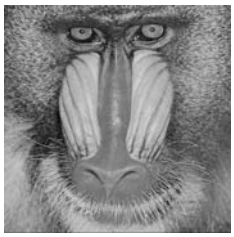
圖 4-2.1 LENA 量化重建圖 (a) 原圖 (b),(c),(d),(e)重建影像

實驗二

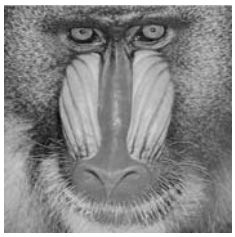
對一含較多紋路(texture)的 BABOON 影像進行實驗，本實驗作法與實驗一相同，但使用另一含較多紋路的 BABOON 影像，在各種不同的量化條件下進行許多實驗，對各種不同條件下所得主觀的視覺品質評估，如表 4-2.2 所示，在圖 4-2.2(b)、(c)、(d)、(e)中，經過相關研究人員與原圖(a)觀察後，不易察覺有任何失真，而在(f)中，仔細觀察影像中的鼻樑兩邊，可以發現已有些微的失真，在(g)中，失真較明顯，但在毛髮部份，還是察覺不出有所改變，證明了人類視覺對於多紋路的影像敏感度較低。

表 4-2. 2　BABOON 次頻帶品質評估表

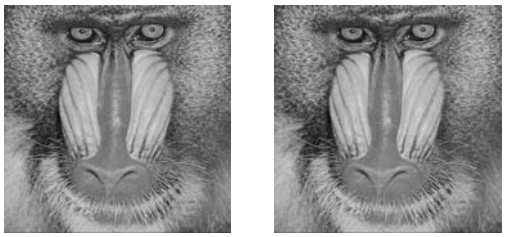
重建 影像	第 四 級	第 三 級	第 二 級	第 一 級	察覺 失真
(b) 量化值	2	3	4	5	否
(c) 量化值	3	4	5	6	否
(d) 量化值	4	5	6	7	否
(e) 量化值	5	6	7	8	否
(f) 量化值	6	7	8	9	是
(g) 量化值	7	8	9	10	是



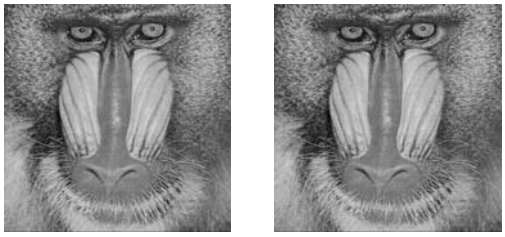
(a)



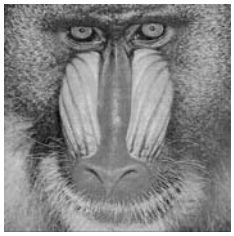
(b)



(c) (d)



(e) (f)



(g)

圖 4-2. 2 BABOON 量化重建圖
(a) BABOON 原圖 (b),(c),(d),(e),(f),(g)重建影像

在這兩個實驗中，發現 LENA 影像在第一級次頻帶的量化步距由 2 跟 3 開始遞增時，重建品質非常好，無法由人類視覺判斷出差別，而當 BABOON 影像在第一級次頻帶的量化步距由 2 開始遞增，一直到由第一級次頻帶的量化步距由 5 開始遞增時，其重建影像與原影像觀察不出有所差異。在表 4-2.1 與表 4-2.2 比較之下，可以看出 BABOON 影像比 LENA 影像有較大的量化步距範圍，證明了人類視覺對較多紋路的影像的敏感度比較低，以及證明了一張影像中的確有許多的視覺冗餘資訊，當移除這些冗餘後，無法由肉眼分辨出與原影像的差別，已達到初步的壓縮效果，在往後研究中，由於量化的最佳值會因影像內容的不同有所差異，未來將對視覺模型運用到編碼方法上，利用視覺模型編碼，可以依照影像的特性，計算它的恰可查察失真值以控制量化時步距的值，以達最佳視覺品質效果。

4-3 提議方法之效能比較

在本實驗結果中，藉由客觀的評估方式來評斷文獻視覺模型方法與方向性視覺模型方法效能的比較。

實驗一

以 LENA 為主的視覺模型方法與方向性視覺模型方法之比較，由表 4-3.1 與表 4-3.2 可以明顯的比較出兩者之間的差異，在相同的 PSPNR 之下，雖然改善方法編碼出來的 bit rate 沒有比文獻方法的小，但方向性視覺模型方法的算數編碼的 bit rate 也都和原始方法效能近似。

表 4-3. 1 LENA 文獻視覺模型方法客觀值

Huffman	Arithmetic	PSNR	PSPNR
0.49 bpp	0.34 bpp	42.06dB	42.28dB

表 4-3. 2 LENA 方向性視覺模型方法客觀值

Huffman	Arithmetic	PSNR	PSPNR
0.50 bpp	0.41 bpp	39.72dB	42.28dB

(a)


(b) (c)

圖 4-3. 1 LENA 還原圖 (a)原圖 (b)視覺模型方法 (c)方向性視覺模型方法


(b) (c)

圖 4-3. 2 PEPPERS 還原圖 (a)原圖 (b)視覺模型方法 (c)方向性視覺模型方法

實驗二

以 PEPPERS 為主的視覺模型方法與方向性視覺模型方法之比較由表 4-3.3 與表 4-3.4 可以看到，在相同的 PSPNR 之下，方向性視覺模型方法的算數編碼之 bit rate 都比視覺模型方法來的小，由此可知，方向性視覺模型方法的效能在此張圖是比視覺模型方法來的好。

表 4-3. 3 PEPPERS 文獻視覺模型方法客觀值

Huffman	Arithmetic	PSNR	PSPNR
0.91 bpp	0.38 bpp	43.64dB	44.55dB

表 4-3. 4 PEPPERS 方向性視覺模型方法客觀值

Huffman	Arithmetic	PSNR	PSPNR
0.61 bpp	0.22 bpp	38.39dB	44.55dB

(a)

經由以上兩個實驗可以發現，雖然在第一個實驗中 bit rate 並沒有比原始方法的好，但差距不大，可以說是非常的接近，但在實驗二中的提議的方向性視覺模型方法一則是遠勝於原始方法許多，由此可見提議的方法是有效的，且使用整數小波轉換也比傳統小波轉換之速度快許多。

4.4 文獻視覺模型方法與方向性視覺模型方法之高低門檻值比較

實驗一 : 以 LENA 為主的門檻值比較

經過本實驗後，在主觀方面，可以看出在圖 4-4.1(b)(c)中在邊界部分有些模糊，但基本的顏色、輪廓還是非常的清晰。而在圖 4-4.2 的部分，幾乎沒辦法分辨出原圖與還原之後有什麼差別。至於客觀方面，在表 4-4.1、表 4-4.2、表 4-4.3、表 4-4.4 相互比較之後，可以分別看到在相同的 PSPNR 值，不論是高低門檻值，兩者的 bit rate 值都是非常接近。

表 4-4. 1 LENA 高門檻值還原圖之客觀評估表(文獻視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
0.25 bpp	0.19 bpp	32.25dB	37.48dB

表 4-4. 2 LENA 高門檻值還原圖之客觀評估表(方向性視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
0.24 bpp	0.17 bpp	35.56dB	37.48dB

(a)

(b) (c)

圖 4-4. 1 LENA 高門檻值還原圖 (a) 原圖 (b)文獻視覺模型方法 (c)方向性視覺模型方法

表 4-4. 3 LENA 低門檻值還原圖之客觀評估表(文獻視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
1.07 bpp	0.99 bpp	43.52dB	58.59dB

表 4-4. 4 LENA 低門檻值還原圖之客觀評估表(方向性視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
1.13 bpp	0.97 bpp	44.01dB	58.59dB

(a) (b)

圖 4-4. 2 LENA 低門檻值還原圖 (a) 文獻視覺模型方法 (b)方向性視覺模型方法

實驗二：以 BABOON 為主的門檻值比較

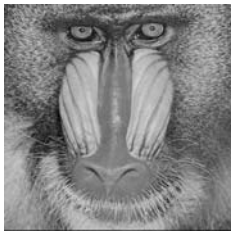
在圖 4-4.3(b)(c)中，可以看到只有在鬍子的兩側以及鼻子兩側有些許的模糊，其他部分可以說是與圖 4-4.3(a)完全一樣，而在低門檻值圖 4-4.4(a)(b)，可以說是與原圖完全一樣。而在表 4-4.5 與表 4-4.6 則是可以看出視覺模型方法無論是在哈夫曼編碼或是算數編碼都是方向性視覺模型較好的；至於表 4-4.7 與表 4-4.8，在哈夫曼編碼方面，方向性視覺模型比視覺模型方法小了一點，但在算數編碼方面，方向性視覺模型則是比視覺模型方法大了許多。

表 4-4. 5 BABOON 高門檻值還原圖之客觀評估表(文獻視覺模型方法)

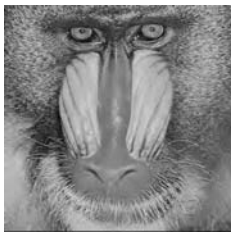
Huffman	Arithmetic	PSNR	PSPNR
1.63 bpp	1.23 bpp	32.38dB	32.20dB

表 4-4. 6 BABOON 高門檻值還原圖之客觀評估表(方向性視覺模型方法)

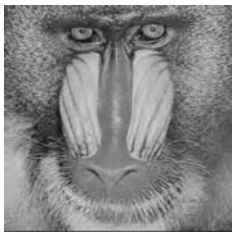
Huffman	Arithmetic	PSNR	PSPNR
2.15 bpp	1.44 bpp	32.20dB	32.20dB



(a)



(b)



(c)

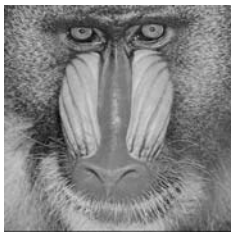
圖 4-4. 3 BABOON 高門檻值還原圖 (a) 原圖 (b)文獻視覺模型方法 (c)方向性視覺模型方法

表 4-4. 7 BABOON 低門檻值還原圖之客觀評估表(文獻視覺模型方法)

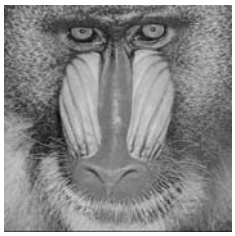
Huffman	Arithmetic	PSNR	PSPNR
3.72 bpp	2.85 bpp	42.92dB	58.27dB

表 4-4. 8 BABOON 低門檻值還原圖之客觀評估表(方向性視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
3.57 bpp	3.15 bpp	43.61dB	58.27dB



(a)



(b)

圖 4-4. 4 BABOON 低門檻值還原圖 (a) 文獻視覺模型方法 (b)方向性視覺模型方法

實驗三：以 AIRPLANE 為主的門檻值比較

主觀方面，可以發現到圖 4.4.5 的(b)與(c)在雲層的部分，變成些許的模糊，較看不出雲朵的層次，但其他主體，如飛機、山脈部分，則還是清晰可見，甚至於飛機上的字還是一樣的清楚。而在低門檻值圖 4-4.6 部分，可以說是與原圖不分軒輊，難以分辨。至於客觀方面，不論是高門檻值的表 4-4.9 與表 4-4.10，或是低門檻值的表 4-4.11 與 4-4.12，在一樣的 PSPNR 之下，方向性視覺模型與文獻視覺模型的值都是非常的接近，不相上下。

表 4-4. 9 AIRPLANE 高門檻值還原圖之客觀評估表(文獻視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
0.29 bpp	0.22 bpp	36.22dB	37.37dB

表 4-4. 10 AIRPLANE 高門檻值還原圖之客觀評估表(方向性視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
0.31 bpp	0.22 bpp	35.73dB	32.20dB

(a)

(b)

(c)

圖 4-4. 5 AIRPLANE 高門檻值還原圖 (a)原圖 (b)文獻視覺模型方法 (c)方向性視覺模型方法

表 4-4. 11 AIRPLANE 低門檻值還原圖之客觀評估表(文獻視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
1.04 bpp	0.92 bpp	45.15dB	60.83dB

表 4-4. 12 AIRPLANE 低門檻值還原圖之客觀評估表(方向性視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
1.05 bpp	0.92 bpp	45.01dB	60.83dB

(a)

(b)

圖 4-4. 6 AIRPLANE 低門檻值還原圖 (a) 文獻視覺模型方法 (b)方向性視覺模型方法

實驗四：以 SAILBOAT 為主的門檻值比較

在這個實驗，則是可以看到在圖 4-4.7(c)中的雲和水面光線倒影的部分都有較多的模糊，而在圖 4-4.7(b)中，則是較接近圖 4-4.7(a)。而在圖 4-4.8 中，則是兩張圖都非常的清晰,看不出與圖 4-4.7(a)有什麼差別。而在客觀方面，高門檻值時是視覺模型方法較好，而低門檻值時，則是方向性視覺模型方法較好，如表 4-4.13，表 4-4.14，表 4-4.15，表 4-4.16 所示。

表 4-4. 13 SAILBOAT 高門檻值還原圖之客觀評估表(文獻視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
0.52bpp	0.35 bpp	33.70dB	36.30dB

表 4-4. 14 SAILBOAT 高門檻值還原圖之客觀評估表(方向性視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
0.70 bpp	0.48 bpp	33.40dB	36.30dB

(a)

(b)

(c)

圖 4-4. 7 SAILBOAT 高門檻值還原圖 (a)原圖 (b)文獻視覺模型方法 (c)方向性視覺模型方法

表 4-4. 15 SAILBOAT 低門檻值還原圖之客觀評估表(文獻視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
2.55 bpp	2.26 bpp	43.54dB	64.22dB

表 4-4. 16 SAILBOAT 低門檻值還原圖之客觀評估表(方向性視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
2.20 bpp	2.08 bpp	43.54dB	64.22dB

(a) (b)

圖 4-4. 8 SAILBOAT 低門檻值還原圖 (a) 文獻視覺模型方法 (b)方向性視覺模型方法

實驗五：以 PEPPERS 為主的門檻值比較

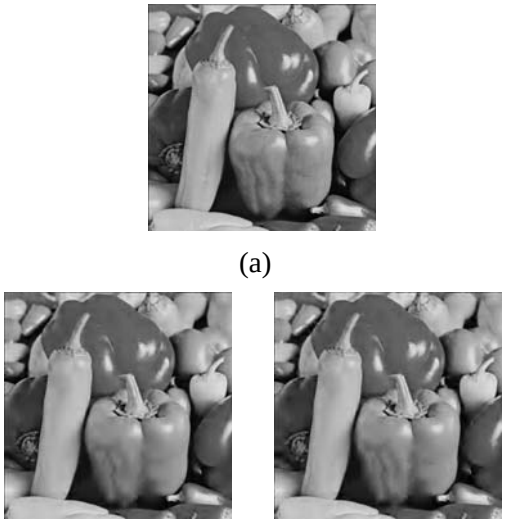
主觀方面，圖 4-14 的(b)與(c)中的主體的細紋和線條都有些許的模糊，而在圖 4-15 低門檻值的部分，依然如前面實驗的結果，與原圖非常相近，無法分辨何者為原圖，何者為壓縮還原圖。客觀方面，經由表 4-15 以及表 4-16 可以得知在相同的 PSPNR 之下，無論是高門檻值或低門檻值，除了在賀夫曼編碼方面有些許差距外，在算數編碼則是改善方法一的 bit rate 都與改善方法二非常接近，差異值都在 0.03 以下。

表 4-4. 17 PEPPERS 高門檻值還原圖之客觀評估表(文獻視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
0.24 bpp	0.12 bpp	35.46dB	36.97dB

表 4-4. 18 PEPPERS 高門檻值還原圖之客觀評估表(方向性視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
0.32 bpp	0.15 bpp	35.00dB	36.97dB



(a)

(b) (c)

圖 4-4. 9 PEPPERS 高門檻值還原圖 (a) 原圖 (b)文獻視覺模型方法 (c)方向性視覺模型方法

表 4-4. 19 PEPPERS 低門檻值還原圖之客觀評估表(文獻視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
1.71 bpp	0.72 bpp	44.64dB	61.76dB

表 4-4. 20 PEPPERS 低門檻值還原圖之客觀評估表(方向性視覺模型方法)

Huffman	Arithmetic	PSNR	PSPNR
1.58 bpp	0.71 bpp	44.17dB	64.22dB

(a) (b)

圖 4-4. 10 PEPPERS 低門檻值還原圖 (a) 文獻視覺模型方法 (b)方向性視覺模型方法

由以上的五個實驗，可以發現在主觀方面，無論是低門檻值或高門檻值，方向性視覺模型方法和文獻視覺方法的圖可以說都是非常的接近，高門檻值時，兩者失真的地方都約略相同，低門檻值時，兩者則都是一樣清晰，與原圖可以說是不分上下。而在客觀方面，大致上高門檻值時可以說是文獻視覺方法較好，低門檻值則是方向性視覺模型方法較好，但無論是高門檻值或是低門檻值，除了在細微訊號較多的 BABOON 圖裡差距較大，其他圖在兩方法二間的差距都不大。至於編碼方面，可以說是算數編碼大獲全勝，在所有的情況裡，算數編碼都要比哈夫曼編碼來的小。

5.結論

本研究已對整數小波轉換與視覺影像編碼相關文獻進行深入的研究，並提出整數小波轉換來加快演算的速度、JND 的視覺編碼模型及移除人類視覺所不易察覺高頻部份的量化，結合了這些方法後，將一張 512×512 大小的圖片，在人類視覺無法察覺失真的情況下，做出最好的壓縮。

經由參考文獻[3]，並設法改進，首先想到的辦法就是整數小波轉換，在文獻[3]裡面它使用的是一般的小波轉換，前面也提到在一般的小波轉換過程中，需要經過浮點數相乘的運算，利用整數小波轉換可以將整個影像在壓縮的過程中節省了相當多的時間，因為不用浮點數部份的運算，正是整數小波轉換快速的原因。

加快了演算的速度後，接著的關鍵就是如何去定義一個準確的 JND，然後利用 JND 去對影像中冗餘的部份進行量化或是設計新的演算法來移除視覺的冗餘，來獲得壓縮的效果。因為目前沒有一個標準的視覺模型，只有一些視覺特性的研究報告可以被用來把視覺模型具體化。但經過了 JND 的實驗後，發現了在改進編碼效率對於大多數的真實影像是可行的，當然一種方法並不能證明這就是移除視覺冗餘的最好方法，因此將圖片的每一階的三個區塊使用了不同方向去計算，接著在考量壓縮率、視覺影響的情況下互相比較。

本研究已經能將影像在兼顧壓縮率與人類視覺下，從中找出一個最完美的平衡點，並已客觀與主觀的兩種判斷，去辨別目前的實驗結果的好壞。到目前為止，研究結果是令人滿意的，期待未來能完整定義出一個標準的視覺模型，以增加實用性。

參考文獻

[1]N. Jayant and J. Johnston and R. Safranek, “Signal compression based on models of human perceptionn,” *Proceedings of the IEEE*, vol. 81(10), 1993

[2]M.D. Adams and R.K. Ward, “Symmetric-extension-compatible reversible integer-to-integer wavelet transforms”, *IEEE Trans. on Signal Processing*, vol. 51, no. 10, pp.

2624-2636, 2003.

[3]M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, “Image coding using wavelet transform,” *IEEE Trans. on Image Processing*, vol. 1, pp. 205-220, 1992.

[4]A.R. Calderbank, I. Daubechies, W. Sweldens, and B.L. Yeo, “Wavelet transforms that map integers to integers,” *Applied and Computational Harmonic Analysis*, vol. 5, pp. 332-369 ,1998.

[5]C. H. Chou and Y C Li, ”A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile,” *IEEE Trans. on Circuits Sand systems for Video Technology*, vol. 5(6) pp. 467-476, 1995

[6]H. Kim and C. C. Li, “Lossless and lossy image compression using biorthogonal wavelet transforms with multiplierless operations,” *IEEE Trans. On Circuits and Systems—II: Analog and Digital Signal Processing*, vol 45(8), pp. 1113-1118, 1998

[7]I. Höntsch and L. J. Karam, “Locally adaptive perceptual image coding,” *IEEE Trans. on Image Processing*, vol 9(9), pp. 1472-1483, 2000

[8]I. Höntsch and L. J. Karam, “Adaptive image coding with perceptual distortion control,” *IEEE Trans. on Image Processing*, vol 11(3), pp. 213-222, 2002

[9]Z. Yu, ”Perceptual coding of digital images,” *Proceeding ICSD 2004*, pp. 1167-1172, 2004

[10] Z Liu and L J Karam, “Jpeg encoding with perceptual distortion control,” *Human Information processing Research* , pp. 637-640, 2003

[11] Ingo Höntsch, *Member, IEEE,* and Lina J. Karam, *Member, IEEE* “Locally Adaptive Perceptual Image Coding” *IEEE Transactions on Image Processing, vol. 9, 2000* pp. 1472-1483

[12] 戴顯權編著,"資料壓縮(二版)", 紳藍出版社, 2002.