Journal of Science and Engineering Technology, Vol. 1, No. 2, pp. 57-67 (2005)

# 可變動區塊之小波轉換影像編碼

吳家琪<sup>1</sup> 施銘鎧<sup>2</sup>
<sup>1</sup>國立台灣海洋大學通訊與導航工程系 基隆市中正區北寧路2號
<sup>2</sup>大葉大學電機工程學系
彰化縣大村鄉山腳路112號

### 摘要

小波轉換(wavelet transform)可提供具有對影像多重解析(multi-resolution)的功用,所 以被廣泛地使用在影像壓縮技術上。而爲了增加影像的壓縮比率,一般將小波轉換後的小波係 數經由向量量化(vector quantization)來進行壓縮,而使用向量量化的主要原因是由於它具有 很高的壓縮比率,而且在解壓縮時也相當容易。但在實用上,影像壓縮法除了壓縮比率外,編 碼速度也是考量的重點。所以在本研究中,我們提出用四分樹分割法作爲影像的前處理 (preprocessing),將影像的背景和細節部份分割成尺寸大小不同的區塊,然後經小波轉換後, 再將每一個子頻帶區塊的小波係數能量值(variance)做爲位元分配的依據。而依照分配到的位 元選擇最適當的編碼簿(codebook),並由編碼後所得到的失真值來調整位元率。這樣的壓縮方 式下,在低位元率時依然保有良好的視覺品質,而且由於編碼簿變得較小,所以運算速度也相 對的加快許多。本文所提方法的成果與效能及處理時間將於模擬結果處展示。

# Variable Blocksize Wavelet-Transform Image Coding

JIA-CHYI WU<sup>1</sup> and MING-KAI SHIH<sup>2</sup>

<sup>1</sup>Department of Communications and Guidance Engineering, National Taiwan Ocean University 2 Pei-Ning Rd., Keelung, Taiwan <sup>2</sup>Department of Electrical Engineering, Da-Yeh University 112 Shan-Jiau Rd., Da-Tsuan, Changhua, Taiwan

#### ABSTRACT

The wavelet-transform procedure has received considerable attention because it provides a multiresolution representation of images. The wavelet transform is now a widely used technique in image compression. By way of a vector quantization technique, the wavelet transformed coefficients can be compressed further. For real-time image compression, achieving a high-data compression ratio and coding efficiency simultaneously are both important. In this study, a quadtree segmentation method for image preprocessing is proposed. A quadtree segmentation algorithm is used to divide a given image, in which regions with image detail will be segmented into blocks with smaller sizes, and the background of the image will be assigned larger block sizes. After the wavelet

transform, to select an appropriate size for the codebook, a bit assignment is applied in association with the variance of each subband image block. A rate-distortion concept is also adopted to adjust the compressed bit rate. For this proposed compression scheme, the simulation results indicate that an acceptable visual quality and a high compression ratio can be achieved simultaneously. Furthermore, due to the small size of the codebook, the computational time can be reduced. A system performance analysis is also demonstrated in this study.

Key Words: ouadtree segmentation, wavelet transform, bit allocation, vector quantization

# 一、引言

存在現有靜態影像壓縮標準 JPEG (joint photographic experts group)中的缺點有:(1)當位元率低於 0.25 bpp 時, JPEG 產生的失真太大,即會有區塊效應(block artifacts) 產生,讓人無法接受;(2) JPEG 用數位餘弦轉換(discrete cosine transformation, DCT)做失真壓縮,用 Huffman 編碼 及算術編碼做無失真壓縮,使得壓縮演算法不是單一的;(3) 無法針對影像中感興趣的部分做高解析度的處理。為了解決 上述問題,新一代的靜態影像壓縮標準 JPEG2000 因而產 生,而 JPEG2000 不使用 DCT 做為核心技術,而是以小波 轉換為其核心。JPEG2000 的特性為:(1)有更高的壓縮比 並保有較佳的影像品質;(2)可依影像品質或不同解析度做 漸進式影像傳輸(progressive image transmission);(3)可 對某一特定區域的影像品質做局部強化;(4)都使用小波轉 換來做失真與無失真壓縮;(5)可將影像與一般文件混合壓 縮。

近年來應用小波轉換於影像壓縮的研究非常多,小波轉換的多重解析應用也成為影像壓縮的相關研究之一。而在現有的小波轉換應用於影像壓縮技術中,有利用小波轉換結合向量量化的壓縮法,另外還有如 EZW (embedded zerotree wavelet)、SPIHT (set partitioning in hierarchical trees)、SLCCA (significance-linked connected component analysis)等等針對小波係數做處理的壓縮法,本文即針對前者的壓縮法做改善及效能的提升。

本文的研究目的是改善小波轉換結合向量量化壓縮法 的缺點,其中包括有:(1)在做小波轉換前需先讀入一整幅 來源影像到記憶體,如此一來會增加記憶體的成本;(2)大 的影像區塊需要較大的編碼簿,所以會使編碼簿的訓練時間 增加。因此我們希望在不太增加系統的複雜性與演算法之前 提下,藉由四分樹分割法來做爲影像輸入的前處理,這樣一 來不但能解決上述的問題,而且能使小區塊的影像特性一 致,更有利於能量集中。由模擬結果可知,此方法還能提高 編碼增益。本文研究相關的系統架構將分別於下一節中敘 述,系統模擬結果之分析與比較將於第三節描述,最後第四 節提供整個系統的結論。

# 二、系統架構

在影像壓縮的過程中,是將各個來源影像做處理,使得 在接收端的重建影像品質,其失真度為人眼所能接受的前提 之下,用最少的位元來傳送。本文中是以改進小波轉換 (wavelet transform)搭配向量量化(vector quantization)的 影像壓縮法,在做小波轉換與向量量化壓縮前,先做一個前 處理,即利用四分樹分割法來做分割,再經由小波轉換得到 每個小區塊的子頻帶,依每個子頻帶的能量值(variance) 來決定其重要性,作為分配位元數的依據,最後再以向量量 化做壓縮,圖1為整個系統的原理架構圖。

### (一)四分樹影像分割法

一般生活中的影像圖片,大多包含有屬於較平滑 (smooth)的背景影像區域,同時也包含些許較高頻、變化 較大的細節(detail)與邊緣(edge)部分;所以,在進行 影像壓縮前,將屬於平滑的背景部分,分割成較大的影像區 塊,而屬於高頻的細節部分,則分割成較小的影像區塊。Teng 和 Neuhoff [17] 藉由分割法把影像分割成不同大小的區 塊,使每一個區塊能較趨近於數學上的統計穩定(stationary) 狀態。即使分割後的小區塊內,其像素間的灰階值較爲相 近,以利於我們能較準確的用簡單的數學模式來描述 [22]。

而加入四分樹分割法當前處理有三個優點,一是使每個 影像區塊的統計特性一致,以利於轉換後的能量更集中;二 是可縮短處理時間,雖然多了前處理的時間,但是小區塊做



#### 圖 1. 本文系統架構

小波轉換會比大區塊做小波轉換的時間來的短,尤其最重要 的是在小波轉換之後,利用向量量化做壓縮時,其編碼簿的 尺寸將可以大大縮小,如此可縮短編碼簿的訓練時間;三是 以小區塊的影像資料去做小波轉換,會比一次讀入一整幅影 像資料去做小波轉換,更節省記憶體的需求量。

本文的四分樹影像分割法是利用區域區分法來做分 割,但不再對相鄰子區域間進行合併或融合的處理。四分樹 的分割方式是對於方形的影像連續分割成越來越小的四個 象限,最後整個影像將分割出各種不同大小尺寸的方塊。為 了不使分割的處理過程太繁複,本文中的四分樹分割法是採 用由上而下(top-down)[20]的分割方式,而在四分樹分 割法中是以最大值最小值法(maxmin),來作爲分割與否的 準則。首先我們定出臨界值 θ 的大小,然後以一開始的影 像區塊中,找出最大的像素灰階值為 xmax,和最小的像素灰 階值為 xmin,當(xmax-xmin)> θ 時,表示目前影像區塊內 的活動狀況較多,即此區塊內包含有邊界的高頻訊號,故需 要將此影像區塊細分成四個影像區塊,再對四個影像區塊做 上述的檢測,然後重複此步驟直到我們所設定的區塊大小或 已達臨界值的終止條件時才停止。在本文中的最小像素區塊 是設定為1個像素的大小區塊。

## (二)小波轉換

影像轉換編碼(transform coding)是將原始訊號經過轉換,成為另一種表示法,這個表示法可以經由逆轉換(inverse transform)恢復成原始的訊號。而且,轉換後的能量較原始訊號更為集中,因此利用此特性可對影像資料做壓縮編碼。而轉換編碼本身並不能達到影像資料壓縮的目的,轉換的目標是將原訊號取樣間的相關性打散,以去除原始資料的重覆多餘性(redundancy),並將訊號的能量重新分佈而且集中在少數的轉換係數上。

而小波轉換是透過一個原型函數(prototype function)  $\Psi(t)$ 的放大、收縮和平移所形成的基底函數來描述訊號, 而這些基底函數的放大收縮可以看成是頻率的高低變化,因 此小波轉換在訊號的分析上非常具有彈性。小波轉換的觀念 是由傅立葉轉換的大小變化特性(scaling property)而來。 由於傅立葉轉換將訊號從時域轉換至頻域中時,我們就無法 知道時間的資訊,因此對於暫態的訊號,我們將無法分析。 而爲了彌補此一缺點,便引入視窗的觀念進去,一般稱爲短 時傅立葉轉換(short-time FT, STFT),即利用一視窗移動函 數(window function),來取得訊號,但視窗固定之後,便 無法改變,但是對於高頻訊號需用小視窗,而低頻訊號需用 大視窗而言,若欲分析的訊號包含高頻及低頻訊號,則此方 法將不可行。而小波轉換的出現,剛好可彌補此缺點,小波 轉換是用一尺度(scale)參數及平移(shift)參數所形成的 基底函數來描述訊號,尺度參數的大小可看做頻率的高低變 化,而平移參數包含時間的訊息,因此可得到時域和頻域的 資訊。圖2為小波與各種轉換間的解析度比較。 原型函數Ψ(t) 可表示為:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \tag{1}$$

其中 a 為尺度參數, b 為平移參數, 而小波轉換可表示為:

$$W(a,b) = \int_{-\infty}^{\infty} f(t)\psi_{a,b}^{*}(t)dt$$
(2)

反小波轉換(inverse wavelet transform) 為:

$$f(t) = \frac{1}{c} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\left|a\right|^2} W(a,b) \psi_{a,b}(t) dadb$$
(3)

c 為一常數。

由於現行的訊號大都是數位的,所以我們將小波轉換轉 變為離散形式,把尺度及平移參數 $a \cdot b$ 取樣,令  $a = a_0^m, b = nb_0 a_0^m$ ,則小波的原型函數為:

$$\psi_{a,b}(t) = \psi_{a_0^m, nb_0 a_0^m} = a_0^{-\frac{m}{2}} \psi(a_0^{-m}t - nb_0)$$
(4)

m,n 為任意整數,所以離散小波轉換可表示為:

$$f(t) = \sum_{m \in n} \sum_{n \in N} d_{m,n} \psi_{m,n}(t)$$
(5)



圖 2. 小波與各種轉換間的解析度比較

其中 d<sub>mn</sub> 稱爲小波係數 (wavelet coefficient), 而

$$d_{m,n} = \left\langle \psi_{m,n}(t), f(t) \right\rangle = \frac{1}{a_0^{m/2}} \int f(t) \psi(a_0^{-m}t - nb_0) dt \qquad (6)$$

但影像多是以二維的形式存在,所以要把一維小波轉換 推廣到二維小波轉換,一般把二維訊號用一維小波轉換做分 解,使用 Mallat 運算法,可將訊號一層層做分解。當輸入 一數位影像時,先對列的資料經由低通和高通濾波器做分 頻,再分別做降取樣(downsample),之後再對行的資料做 低通和高通濾波,再做降取樣,則會得到四個子頻帶。若要 再往下一層做分解,可把上一層得到的低頻訊號再依上述方 法做分解;而重建過程則依相反程序進行,其中每一層在重 建後,要做升取樣(upsample),即在每個序列之間內插零。 訊號的分解是將訊號經由與濾波器做摺積(convolution), 經由低通濾波器(lowpass filters)的運算可得到低頻訊號, 經由高通濾波器(highpass filters)的運算則可得到高頻訊

#### (三) 位元分配

位元分配的目的有二:一是有效分配位元,以增加壓縮 效率;二是減少編碼簿的搜尋時間,即先找出編碼簿可能之 大小,如此可減少整個編碼簿的搜尋時間。

把影像經四分樹分割後,接著把每一個區塊做小波轉換,並分別把轉換後每個子頻帶的能量值計算出來,然後依 據能量值來決定各個頻帶的重要性,計算出各個頻帶所需要



圖 3. 二維訊號之一階小波分解重建架構

分配的位元率  $b_k$  (bit per pixel)。方程式如下:

$$b_{k} = \frac{R \times \log(Var_{k}) \times bs_{k} \times bs_{k}}{\log(\sum_{i=1}^{n} Var_{i}) \times BS \times BS}$$
(7)

R表示原始影像區塊中每一個像素所需要的位元數, Var<sub>k</sub>表示某一個子頻帶影像區塊的能量值, n 為所有子頻帶影像區塊的個數, bs<sub>k</sub>為第 k 個變異數的子頻帶影像區塊大小, BS 則是原始影像的大小。在得到每一個頻帶的位元率之後, 利 用此位元率找出最接近的編碼簿大小, 其方程式如下:

$$b_k = \frac{\log_2 C \times bs_k \times bs_k}{d \times BS \times BS} \tag{8}$$

C表示碼向量(code vector)的個數,d則是編碼簿的維度。 決定編碼簿的大小之後,即用 LBG(Linde-Buzo-Gray)演 算法找出各個子頻帶的編碼簿,並計算其平均失真值,當平 均失真值大於預設值時,就把位元率調高,即碼向量增多或 維度減小。而預設的平均失真值,由下面的方程式求出:

$$dis_k = W_k \times \log \frac{1}{b_k} \tag{9}$$

其中  $dis_k$  表示第 k 個區塊的預設失真值,  $W_k$  為每個區塊的 權重值。而權重值可根據表 1 中的結果, 選擇粗體字的權重 值, 另外在表 1 中所求出來的值, 是依據 LENNA 影像經四 分樹分割後, 得到的最大區塊, 再經三階小波轉換後, 依照 變異數的大小, 由上面所提之方程式求出。往後將依此權重 值來預設失真值。表 1 為不同權重值下的 PSNR (peak signal-to-noise ratio) 比較表。

#### (四)向量量化

一般將連續的信號在時間上數位化的過程,我們稱為取 樣,而在信號振幅上數位化的過程,則稱為量化。在純量量 化中,它是對各個取樣點逐點量化,但大多數的信號,其各 取樣値間存在著相關性,即若知道某個取樣値的參數,可由 它來推測鄰近取樣値的參數,此種量化方式稱為向量量化, 也就是說把一維的純量量化延伸到多維度的量化方式。

值得注意的是,把數位影像經過小波轉換後,並沒有降 低資料量,主要是將訊號的大部分能量集中在少數的基底 上,即我們轉換後所得到的小波係數,再藉由量化編碼的方 式來達到資料壓縮的目的。而向量量化是一種簡單且廣泛使

權重	5	10	15	100	500
LL3 PSNR	37.20	32.25	32.25	21.48	21.48
С	SQ=6	SQ=5	SQ=5	32	32
d				4	4
LH3 PSNR	38.24	30.70	28.41	28.41	28.41
С	SQ=4	32	16	16	16
d		4	4	4	4
HL3 PSNR	33.10	32.74	28.90	28.90	28.90
С	SQ=4	32	16	16	16
d		4	4	4	4
HH3 PSNR	36.82	30.10	30.10	30.10	30.10
С	32	8	8	8	8
d	4	4	4	4	4
LH2 PSNR	34.13	31.46	29.90	29.90	29.90
С	32	16	8	8	8
d	4	4	4	4	4
HL2 PSNR	36.10	33.58	29.98	29.98	29.98
С	32	16	8	8	8
d	4	4	4	4	4
HH2 PSNR	34.10	32.26	32.26	32.26	32.26
С	16	8	8	8	8
d	4	4	4	4	4
LH1 PSNR	35.60	34.14	34.14	34.14	34.14
С	8	4	4	4	4
d	4	4	4	4	4
HL1 PSNR	34.60	34.60	34.60	34.60	34.60
С	4	4	4	4	4
d	4	4	4	4	4
HH1 PSNR	43.45	43.45	43.45	43.45	43.45
С	128	128	128	128	128
đ	16	16	16	16	16

表 1. 不同權重值下的 PSNR 比較表

用的失真壓縮技術,因為它具有高的壓縮率,且在解碼端做 解碼時也相當容易,所以我們選用向量量化來做為壓縮的工 具。

向量量化的編碼方式,是輸入一多維度的向量,在編碼 簿(codebook)中搜尋較相近的預設向量,即碼字 (code-words)或碼向量(code vector),來取代原始的輸入 向量,而量化後的輸出則爲碼向量在編碼簿中的索引值,因 此可降低原來所需的位元數,而達到簡化編碼與壓縮資料量 的目的,這是與純量編碼最大不同之處。在影像壓縮的應用 上,輸入的影像向量,以失真度準則在編碼簿中搜尋出最適 當的碼向量,一般失真度準則是求兩向量間的歐基里得距離 (Euclidean distance)爲最小,在找到適當的碼向量後,以 其索引值來取代原始輸入的影像向量做為編碼與傳送,而在 接收端或解碼端則將接收到的索引值,再從編碼簿中搜尋其 對應的碼向量,然後以搜尋得到的碼向量來重建回原始影 像。向量量化編碼演算法的基本原理和流程為:

- 步驟一:將小波轉換後所得到的各個子頻帶影像區塊,分割 成 N×N 大小的區塊,其向量維度(dimension) n= N×N;如 4×4 的區塊大小,可視為 1×16 維度的向 量或 16×16 大小的區塊,可視為 1×256 維度的向 量。令每一個影像向量為 X<sub>i</sub>, i = 1,2,...,N。其中 N 是影像分割後的區塊總數。
- 步驟二:令找到的碼向量為 $\hat{X}_j$ , j = I,2,...,C,其中 C 是碼 向量的總數,把形成的影像向量  $X_i$ 與預設編碼簿 中的碼向量 $\hat{X}_j$ 做比對,以失真度準則來選擇最接 近影像向量的碼向量,其中碼向量的維度與影像向 量相同。而失真度準則是計算兩向量間的歐基里得 距離的平方,公式為:

$$d(X_i, \hat{X}_j) = \frac{1}{n} \sum_{k=1}^n (x_k - \hat{x}_k)^2$$
(10)

其中  $X_i$  為第 i 個影像向量,  $\hat{X}_j$  為第 j 個碼向量,  $x_k$  為第 i 個影像向量中的第 k 個分量,  $\hat{x}_k$  為第 j 個 碼向量中的第 k 個分量。

步驟三:將碼向量的索引值當成量化結果送出。

重複步驟二及步驟三,直到所有的輸入影像向量都被處 理完,才換下一個影像區塊。圖4 為向量量化的編解碼過程。



#### 圖 4. 向量量化的編解碼過程

編碼簿的產生一般是將要壓縮的影像集中選出幾張具 有代表性的影像做為訓練集(training set),然後以此訓練集 訓練出編碼簿。這樣的優點是我們沒有必要知道關於這些影 像的統計資料或相關資訊。所以,要編碼一張特定的影像 時,最佳化的編碼簿就是使用這張影像做爲訓練集所產生的 碼向量。因爲這樣產生的編碼簿,具有該張影像所特有的影 像特性,所以足以用此編碼簿來取代。這樣的編碼簿一般稱 爲區域編碼簿(local codebook),而且使用區域編碼簿通常 可以有相當好的壓縮效率。

本研究中我們使用 LBG 演算法 [11] 來產生編碼簿, 它是 Lloyd-Max 演算法的推廣,其主要精神是從訓練向量中 以分群法找出代具代表性的碼向量組成一編碼簿表每個群 的碼向量,再由這些,以下就 LBG 演算法在影像資料壓縮 時的運算步驟做一說明:

- 步驟一:從一個向量訓練集開始,首先由預設的編碼簿大 小,對向量訓練集均勻或隨意的取出所需的向量個 數,此即為初始的編碼簿,然後選擇一失真度準則 及設定訓練停止條件。
- 步驟二:計算所有的訓練向量與編碼簿中碼向量的距離,並 將訓練向量歸類到距離最短的碼向量,如此可將訓 練集分爲幾個群。計算分群過程中所產生的平均失 真值,然後由前一次與此次的平均失真值比例作爲 演算法的停止條件,其式子如下:

$$\frac{D^{(l-1)} - D^l}{D^{(l-1)}} \le \varepsilon \tag{11}$$

其中 $D = \frac{1}{N} \sum_{i=1}^{N} d(X_i, \hat{X}_i)$ ,  $X_i$  為第i 個影像向量,

 $\hat{X}_i$ 為第i個重建影像向量, $D^{(l-1)}$ 代表前一次的平均失真值, $D^l$ 代表此次的平均失真值, $\varepsilon$ 為預設臨界值。

步驟三:將歸類到碼向量的所有訓練向量做平均,得到新的 向量,以此來更新編碼簿的碼向量。重複上述步驟 直到演算法達到終止條件為止。

## 三、模擬結果分析與比較

失真壓縮的編碼法允許重建訊號或影像有一些失真,藉 此來換取比無失真壓縮編碼法更高的壓縮比。假設原始訊號 為 x(t),而重建訊號為 y(t),則重建訊號的失真度為原始訊號與重建訊號的差值,即為:

$$e(t) = x(t) - y(t)$$
 (12)

$$\sigma_e^2 = E\left\{ \left[ e(t) - \mu_e \right]^2 \right\}$$
(13)

通常誤差平均值  $\mu_e = 0$  ,  $\sigma_e^2$  稱為均方誤差 (mean squared error, MSE),而編碼訊號重建的品質評估方式是以 原始訊號的方差  $\sigma_x^2$  對重建誤差方差  $\sigma_e^2$  的比值取對數,稱 爲訊號雜訊比 (signal-to-noise ratio, SNR)。即為:

$$SNR(db) = 10 \log\left(\frac{\sigma_x^2}{\sigma_e^2}\right)$$
(14)

影像振幅的平均值通常為非零的正數,因此,上式可用 最大的振幅值  $x_{max}$ 來代替  $\sigma_x$ ,以得到峰值雜訊比為:

$$PSNR(db) = 10\log\left(\frac{x_{\max}^2}{\sigma_e^2}\right)$$
(15)

例如影像像素以 8-bits 表示,則 x<sub>max</sub> = 255,因此

$$PSNR(db) = 10\log\left(\frac{255^2}{\sigma_e^2}\right)$$
(16)

本文是用 MSE 與 PSNR 來衡量重建訊號的品質和失真 程度。而 MSE 及 PSNR 可對誤差提供一個較客觀公平的依 據,再藉由人眼視覺感官來做一個主觀性的判斷。

首先我們測試影像分割的處理方式,一是可分為在小波 轉換前面或後面做分割;二是可使用固定區塊或可變動區塊 的方式來進行。本文所使用之模擬配備為 P3-1GHz CPU, 512M SDRAM, windows2000, MATLAB 模擬軟體。表 2 為固定區塊於小波轉換前後做分割的結果。圖 5 為表 2 所列 結果之影像圖。表 3 為可變動區塊於小波轉換前後做分割的 結果。圖 6 為表 3 所列結果之影像圖。表 4 為可變動區塊與 固定區塊於小波轉換前做分割的結果。

由上述結果中可看出固定區塊比可變動區塊,在差不多的位元率與較小的固定區塊下,PSNR 值稍低,且需花費較

			D.1703 HJ.	-9-1H>14
固定區塊	DEND (db)	Pata (hpp)	MSE	Time
(8×8)	r SIVK (ub)	Kate (opp)	MSE	(minute)
小波轉換前				
做分割	32.89	0.300	33.39	21.48
小波轉換後				
做分割	32.10	0.300	40.51	4.75



小波轉換前做分割



小波轉換後做分割

圖 5. 表 2 所列結果之影像圖

衣・局り変動画塊於小波轉換則夜敞刀割結為	<b>公割結果</b>
----------------------	-------------

可變動區塊	PSNR (db)	Rate (bpp)	MSE	Time (minute)
小波轉換前 做分割	33.07	0.300	32.08	6.28
小波轉換後 做分割	33.07	0.300	32.10	2.40



小波轉換前做分割



小波轉換後做分割

圖 6. 表 3 所列結果之影像圖

## 表 4. 可變動 / 固定區塊大小於小波轉換前做分割的結果

小波轉換前 做分割	PSNR (db)	Rate (bpp)	MSE	Time (minute)
可變動區塊 大小	33.07	0.300	32.08	6.28
固定區塊 (128×128)	28.62	0.300	89.30	1.30
固定區塊 (64×64)	29.31	0.300	76.27	1.80
固定區塊 (32×32)	31.30	0.300	47.75	3.87
固定區塊 (16×16)	32.40	0.300	37.43	7.28
固定區塊 (8×8)	32.89	0.300	33.39	21.48

多時間,而在差不多的位元率與較大的固定區塊下,雖然可 節省較多時間,但相對的 PSNR 值較低。故我們選擇可變動 區塊來做分割。而在小波轉換前後做可變動區塊分割下,發 現效果差不多,但在小波轉換後做可變動區塊分割的時間花 費較短,原因為經小波轉換後的子頻帶區塊中,小波係數值 較為相近,所以可節省訓練編碼簿的時間。但是由圖 7 所 示,為在較低位元率下所重建之影像,我們發現在小波轉換 後做可變動區塊分割的效果較差,即有較明顯的區塊效應產 生。

圖 7 中,在小波轉換前做分割所得到的結果為:平均位 元率 *Rs* = 0.1612 Bits/Pixel,重組影像的 *MSE* = 47.399830, *PSNR* = 31.373036;而在小波轉換後做分割所得到的結果 為:平均位元率 *Rs* = 0.1787 Bits/Pixel,重組影像的 *MSE* =41.092031,*PSNR* = 31.993228,所以若要求即時處理且不 需要太高的影像解析度,而且不考慮記憶體的成本時,則可



小波轉換前做分割



小波轉換後做分割

圖 7. 可變動區塊於小波轉換前後在較低位元率下所重建 之影像

選擇在小波轉換後做分割處理;若記憶體為考量重點,且需 使用較低位元來傳送,則可選用在小波轉換前做分割處理。 本文中為考慮輸入的資料量為不可預知,所以在有限的記憶 體之下,我們在小波轉換前做可變動影像區塊分割之處理。 而實際作法為將原始影像經由四分樹分割法把影像分割成 不重疊且大小不等的影像區塊,在此我們設定分割臨界值為 140,把得到的所有區塊中,128×128、64×64、32×32 的所 有影像區塊做三階小波轉換,16×16 的影像區塊做二階小波 轉換,8×8 的影像區塊做一階小波轉換,4×4 的影像區塊直 接用向量量化做壓縮,2×2 的影像區塊則用純量量化壓縮, 而1×1 像素則不做壓縮,直接傳送。

經由計算所有小波轉換後各個子頻帶的變異數來做位 元分配,再依分配到的位元數計算預設失真值來做位元的調 整,然後用所分配到的位元數與預設失真值訓練出所屬影像 區塊之編碼簿,並用此編碼簿做向量量化之壓縮處理;在接 收端或解碼端的部分,則依所收到的編碼簿索引值,跟預先 傳送之編碼簿做一比對,然後依所對應之碼向量做影像重 建,則可得到近似的原始影像。表5為此上述方法在不同位 元率下的 PSNR 值。表6為本文跟其他研究論文之比較。

在表 5、6 中,我們並沒有考慮傳送編碼簿所需之額外 位元,因爲若測試足夠多的影像資料後,能找一通用編碼簿 來取代區域編碼簿。假如考慮傳送編碼簿所需之額外位元, 則需 3.38 Bits/Pixel。在本文的向量量化是以區域編碼簿來 實現,我們雖然增加了四分樹分割法的步驟,但卻可改善小 波轉換結合向量量化的 *PSNR* 達 2db 左右,且是在節省 0.07bpp 的位元率之下,所以增加一前處理是可行的。另外,

Rat (bpp)	PSNR	MSE
0.159	31.26	48.67
0.170	31.56	45.45
0.193	31.88	42.20
0.231	32.32	38.10
0.250	32.52	36.40
0.289	32.90	33.22
0.296	33.02	32.42
0.351	33.40	29.70
0.383	33.50	29.10
0.460	34.00	26.50
0.600	34.10	25.35
0.670	34.40	23.35
0.070	34.40	23.44

表 5. 上述方法在不同位元率下的 PSNR 值

表 6. 本文跟其他研究論文之比較

編碼方法	位元率	PSNR
	(bpp)	(db)
本文	0.300	33.02
(不考慮傳送編碼簿所需之額外位元)	0.460	34.00
小波轉換/向量量化 [3]	0.370	30.85
邊界於子頻帶/向量量化 [12]	0.500	34.14

若向量量化的所有編碼簿無法滿足預設平均失真値時,則退 化為純量量化,所有的純量量化其最小位元數設為1。我們 除了以 LENNA 的人物圖像做測試外,還測試了一些自然圖 像。表7與圖8為測試 bridge 影像之結果。表8與圖9為測 試 airplane 影像之結果。

表 7. 測試 bridge 影像之結果

可變動區塊	PSNR (db)	Rate (bpp)	MSE	Time (min.)
小波轉換前做分割	27.13	0.303	125.9	11.40
小波轉換後做分割	22.40	0.301	373.1	3.74



小波轉換前做分割



小波轉換後做分割

圖 8. 測試 bridge 影像之結果

表 8. 測試 airplane 影像之結果

可變動區塊	PSNR (db)	Rate (bpp)	MSE	Time (min.)
小波轉換前做分割	31.14	0.308	50.1	8.40
小波轉換後做分割	23.03	0.301	324.0	3.90



小波轉換前做分割



小波轉換後做分割

圖 9. 測試 airplane 影像之結果

在此,我們討論一下額外位元的問題,因為用四分樹做 影像前處理的分割,所以需要額外紀錄位置並傳送到解碼 端,以LENNA、bridge、airplane 為例,分割後的結果如表 9、10、11 所示。

由表 9 中可知總共分割了 1627 個區塊,但由於四分樹 具有結構性,所以只要根據節點的層數即可算出相對座標的 真正位置。因此以每個區塊最少需要 1 個位元來計算的話, 在整個影像中需多出 1627/262144,即多出 0.00621bpp,所 以對於整個影像所需的位元數來說,這樣的壓縮方式還是有 價值的。

表 9. LENNA 影像做四分樹分割後各個區塊大小的個數

區塊大小	個數	區塊大小	個數
256 × 256	0	8 × 8	374
128 × 128	1	4 × 4	448
64 × 64	16	2 × 2	471
32 × 32	98	1 x 1	36
16 × 16	183		

表 10. bridge 影像做四分樹分割後各個區塊大小的個數

區塊大小	個數	區塊大小	個數
256 × 256	0	8 × 8	1030
128 × 128	0	4 × 4	1311
64 × 64	5	2 × 2	1199
32 × 32	32	1 x 1	212
16 × 16	377		

表 11. airplane 影像做四分樹分割後各個區塊大小的個數

區塊大小	個數	區塊大小	個數
256 × 256	0	8 × 8	712
128 × 128	2	4 × 4	1924
64 × 64	17	2 × 2	2504
32 × 32	27	1 × 1	1184
16 x 16	174		

由表 10 中可知總共分割了 4186 個區塊,因此以每個區 塊最少需要 1 個位元來計算的話,在整個影像中需多出 4186/262144,即多出 0.016bpp。

由表 11 中可知總共分割了 6544 個區塊,因此以每個區 塊最少需要 1 個位元來計算的話,在整個影像中需多出 6544/262144,即多出 0.025bpp。

# 四、結論

本文中,主要是在小波轉換結合向量量化的影像壓縮流 程中,做一效能的提昇與改善,而為了不增加系統複雜度與 演算法之前提下,我們提出以四分樹分割做為影像輸入的前 處理過程,藉由把影像分割成不重疊且大小尺寸不等的區 塊,如此可節省記憶體需求量,並使影像區塊特性一致,以 利後續處理,而更重要的是小區塊只需較小的編碼簿,這樣 一來可縮短編碼簿的訓練時間。由模擬結果可看出,在 0.3bpp時, PSNR 值還比小波轉換/向量量化多2個 db,雖 然位元率的提高,對 PSNR 值的增加有限,尤其在較高位元 率時,增加的幅度並不多,但是在較低位元率時,依然仍保 有一定的 PSNR 值,且能縮短運算時間。

此外,本文所探討的方法與過程,也不一定只限於小波 轉換結合向量量化的影像壓縮法,其實可以把向量量化的壓 縮方式替換為其他的壓縮法,如 EZW、SPIHT、SLCCA, 這些都是對小波係數做量化的處理方式,我們一樣可結合四 分樹分割法來做前處理,雖然小波轉換與離散餘弦轉換的一 個重要改善為,小波轉換不像離散餘弦轉換是一種區塊編 碼,所以不會有區塊效應,但是若小波轉換的部分以硬體方 式來實現,那節省記憶體將是考量的一大重點,所以由測試 結果可知,只有在較低位元率下才會有區塊效應產生。整個 系統中可改進的方向為:

在本文中所使用的編碼簿是屬於區域編碼簿(local codebook),我們並沒有考慮它的額外位元,而在實用上必須是通用編碼簿(global codebook),所以必須要有更多的測試影像資料來找出資料的通用性,而這在未來也需做進一步的探討。

另外,向量量化的訓練法也有很多快速搜尋的方法被提 出,我們可以研究改善編碼簿的訓練方式,或朝其他的量化 方式做研究,以增加壓縮率。而在影像前處理的作法上,則 是朝著應用其他的影像擷取方法,以尋求更好的影像壓縮系 統之效能,使能廣泛應用於實際生活上的各種需求。

## 參考文獻

- 李朝欽(民90),應用小波轉換與四分樹切割之影像合成,大葉大學電機工程系碩士論文。
- 邱俊德(民88),應用自組織映射圖於影像壓縮,大葉 大學電機工程系碩士論文。
- Antonini, M., M. Barlaud, P. Mathieu and I. Daubechies (1992) Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1(2), 205-220.
- Chen, C. T. (1989) Adaptive transform coding via quadtree-based variable blocksize DCT. *IEEE International Conference on Acoustic, Speech and Signal Processing*, 3, 1854-1857.
- Gersho, A. (1982) On the structure of vector quantizers. IEEE Transactions on Information Theory, 28(2), 157-162.
- Gersho, A. and R. M. Gray (1991) Vector Quantization and Signal Processing, Kluwer Academic Publishers, Boston, MA.
- 7. Gonzalez, R. C. and P. Wintz (1997) Digital Image

Processing, Addison Wesley, Reading, MA.

- Gray, R. M. (1984) Vector quantization. *IEEE Acoustic,* Speech, and Signal Processing Magazine, 1, 4-29.
- Hu, Y. C. and C. C. Chang (1999) Variable rate vector quantization scheme based on quadtree segmentation. *IEEE Transactions on Consumer Electronics*, 45(2), 310-317.
- Jain, K. (1989) Fundamental of Digital Image Processing, Prestice-Hall, Englewood Cliffs, NJ.
- Linde, Y., A. Buzo and R. M. Gray (1980) An algorithm for vector quantizer design. *IEEE Transactions on Communications*, 28(1), 84-95.
- Mohsenian, N. and N. M. Nasrabadi (1994) Edge-based subband VQ Techniques for images and video. *IEEE Transactions on Circuits and Systems for Video Technology*, 4(1), 53-67.
- Nasrabadi, N. M. and R. A. King (1988) Image coding using vector quantization: a review. *IEEE Transactions on Communications*, 36(8), 957-971.
- 14. Ramchandran, K. and M. Vetterli (1993) Best wavelet packet bases in a rate-distortion sense. *IEEE Transactions on Image Processing*, 2(2), 160-175.
- Shoham, Y. and A. Gersho (1988) Efficient bit allocation for an arbitrary set of quantizers. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(9), 1445-1453.

- Sullivan, G. J. and R. L. Baker (1994) Efficient quadtree coding of images and video. *IEEE Transactions on Image Processing*, 3(3), 327-331.
- Teng, Y. and D. L. Neuhoff (1995) A new quadtree predictive image coder. *IEEE International Conference on Image Processing*, 2, 73-76.
- Torres, L. and J. Huguet (1994) An improvement on codebook search for vector quantization. *IEEE Transactions on Communications*, 42(2/3/4), 208-210.
- Vaisey, J. and A. Gersho (1987) Variable block-size image coding. *IEEE International Conference on Acoustics*, *Speech, and Signal Processing*, 12, 1051-1054.
- Vaisey, J. and A. Gersho (1992) Image compression with variable block size segmentation. *IEEE Transactions on Signal Processing*, 40(8), 2040-2060.
- Wang, C. Y., S. J. Liao and L. W. Chang (1997) Wavelet transform coding using variable blocksize vector quantization with optimal quadtree segmentation. *IEEE Region 10 Annual Conference, TENCON* '97. 2, 2-4.
- Wu, J. C. and D. G. Daut (1997) Adaptive non-stationary DPCM image coding with variable blocksize. SPIE 1997 symposium on Visual Communications and Image Processing, 3024, 447-458.

收件:93.07.14 修正:93.08.27 接受:93.11.16