

利用分頻轉換編碼與類神經網路之影像壓縮*

劉仁俊

國立屏東商業技術學院電腦與通訊系

renjean@npic.edu.tw

摘要

本文提出利用影像轉換編碼之離散餘弦轉換分頻處理，搭配 Kohonen 自組織特徵映射的方法，來進行靜態影像的資料壓縮。使用影像轉換編碼的目的為擷取出輸入影像在頻域中的特徵；而分頻的過程是將此轉換結果的直流與交流係數區隔開來。直流之係數可直接傳輸或儲存；交流之係數可利用低通濾波器來完成初步的壓縮，並用來做神經網路之訓練樣本，如此可大幅減輕訓練之負擔，並提高信號雜訊比。向量量化被用在需求高壓縮比的環境中，自組織特徵映射可用來產生向量量化所需要的編碼簿，此方法的優點在於能夠產生順序之編碼簿以保留其位置排列之特性，並能顯著地降低編碼簿的空間維度使壓縮能夠更有效率。高壓縮比以及優異的重整品質是本文的最終目的，產出之結果無論在通訊、資料壓縮、多媒體和網路等應用方面都會有重要貢獻。

關鍵詞：自組織特徵映射、離散餘弦轉換、影像壓縮、類神經網路、向量量化

一、緒論

隨著通訊系統搭配多媒體時代的來臨，數位系統被大量的運用，所使用的資料量特別是影像也急速增加，在資料儲存與傳輸的處理上常形成瓶頸，為降低資料的處理量，影像壓縮成為一個重要的課題[1][2]，在許多重要的場合有著關鍵作用，包括文件、醫學影像、衛星影像、氣候探測、傳真、通訊系統、以及在軍事與太空等應用。此外，影像壓縮是目前增加影像感測器空間解析度及制

定廣播電視標準的當然過程，無論是在印刷、出版、視訊會議和傳播等的應用上，均為多媒體處理技術的關鍵，其有效的管理、控制、儲存與傳輸等的相關應用都在與日俱增。

資料壓縮又稱為源編碼(Source coding)，目的在於加速傳輸速度，降低儲存所需的空間，使其能提昇影像處理、傳輸和儲存的效能。方式大致可分成兩類：非失真(Lossless)編碼和失真(Lossy)編碼。非失真壓縮編碼能保留影像之原始資訊，倘若忽略小數點的誤差，在還原後能得到與原始完全一樣的影像。失真壓縮編碼則是由於壓縮還原後的影像與原始影像有所差異，但此方法之壓縮效率極高，故衡量不同壓縮方式優劣的重點是在於如何訂定一個標準，以及判斷被壓縮的程度。一般若以人類的視覺為基準則為主觀標準，此種方式乃嘗試各種不同的壓縮比，歸納其在人眼中可感受的壓縮後影像品質為準而定出壓縮程度。理論研究上，則是計算出原影像和壓縮後影像間誤差，以此誤差的能量大小做為一參考標準，此方式則較為客觀。本計劃將利用一套客觀之標準來衡量執行之成效[3]。主要是以信號雜訊比和壓縮比來做為衡量的依據。

資料壓縮的過程中，影像轉換編碼(Transform coding)通常是第一步驟。其方法是將原始訊號轉換成另一種表示法，這個表示法可以經由逆轉換(Inverse transform)恢復成原始的訊號。而轉換後的能量較原始訊號更為集中，且是集中在相當少的轉換係數上，若是除去原始資料的重覆多餘性(Redundancy)，便能利用此集中的能量對影像資料做壓縮與編碼。此過程也常在影像處理

*本研究由國科會計畫 NSC95-2221-E-251-002 補助

中用來作特徵擷取。此外影像轉換編碼可作為向量量化(Vector quantization, VQ)的前處理，目的是擷取原始輸入影像的特徵、降低向量量化的維度(Dimension)以及減少運算上的複雜度。

現今最常用的影像轉換編碼方法有離散傅立葉轉換(DFT)、離散餘弦轉換(DCT)[4]-[9]、和離散小波轉換(DWT)等[10][11]。本文選擇以 DCT 作為主要的影像轉換編碼，原因在於其計算上的快速和簡易性。

向量量化[12]-[15]是一種在編碼端將 n 維的輸入資料壓縮成 m 維($m \leq n$)的編碼輸出，解碼端部份則又將 m 維的編碼資料恢復成 n 維原始資料的壓縮技術。因此，向量量化在影像處理及訊號處理的應用上，可以得到較佳的壓縮比率，並維持不錯的重建品質。向量量化的輸入資料為多維度的向量，目的是在編碼簿(Codebook)中搜尋較為相近似的預設向量，即碼字(Code-Words)或碼向量(Code Vector)，替代原始的輸入向量，而量化後的輸出則為碼字在編碼簿中的索引(Index)；因此，原本多維度的輸入向量將被簡化為一維或二維的量化輸出索引值，達到簡化編碼和壓縮資料量的目的，這亦是不同於純量編碼之處。向量編碼的關鍵在於編碼簿的好壞，這牽涉到編碼簿本身的大小、碼向量的維度以及編碼簿中碼向量的通用性，而讓編碼簿具有較佳分類碼向量一直是最終的目標。因此，其研究的重點和可能改進之處不外乎能夠快速搜尋的向量量化、碼向量索引值的壓縮編碼以及混合型式的向量量化這幾個方向。

類神經網路(Artificial Neural Networks, ANN)[16]-[18]的發展提供了向量量化一個非常有效的工具，並能增進編碼簿的能力與強健性(Robustness)。不同於以上之方法，類神經網路是以模擬人腦思維的模式所發展的理論，具有強大的學習(Learning)、容錯(Fault Tolerance)、平行處理(Parallel Processing)和聯想記憶(Associative Memory)等能力，可廣泛

地應用在不同的領域當中，例如影像處理、信號處理、型態識別、系統鑑別、及非線性系統最佳化等，經多年的驗證都有不錯的成果產生。

一般的類神經網路是由多個神經元(Neurons)組成，輸入資料在乘上連接權值後經過一非線性之轉換函數而得到單層之輸出，並以層層相連的方式，如輸入層、隱藏層及輸出層等串接組成一網路架構。在神經生理學研究發現，人腦由巨量的神經細胞組成，但並非所有的神經細胞都起同樣的作用；不同空間區域的神經細胞分工有所不同，各區域對應不同的外部刺激；相似的外部刺激在大腦的相鄰近位置上會引起最大的響應，因此輸出神經元之間發生空間耦合，產生了交互作用。據此，Kohonen 於 1980 年[19]提出了另一種類神經網路的架構，稱之為自組織特徵映射(Self-Organization Feature Map；簡稱為 SOFM 或 SOM)，來模擬前述人腦的相鄰神經細胞間的耦合功能。

SOM 具有「物以類聚」的特性，各神經元間會相互影響，以競爭性學習(Competitive Learning)的方法和優勝者激發(Winner-Take-All)的模式來決定網路中各神經元的狀態。訓練的過程中，輸出層具有網路拓樸(Topology)，鄰近區域(Neighborhood)的輸出神經元間會相互影響；在學習完畢後，其輸出神經元相鄰近者會具有相似的響應，也具有相似的連結權值。鄰近區域間交互作用的關係，以網路拓樸決定鄰近區域的神經元方向位置，而優勝神經元鄰近的其他神經元，則依「物以類聚」的特性與優勝神經元同時做移動，以達到分類的目的。SOM 是採用無監督式學習(Unsupervised Neural Network)的網路模式，從原始輸入資料中取得訓練範例，而從中學習範例的內在聚類(Clustering)規則，以應用於其他的問題。因此在訓練完成之後，能對全新的輸入變數推論或分類它與那些訓練範例隸屬於同一聚類，極為適合在聚類分析和形體識別的應用

上。因而利用 SOM 類神經網路作分類的最大優點之一是不需事先知道輸入原始資料中特徵歸屬(Membership)的情況。在向量量化問題的應用上此點是相當重要的，不僅對不同的輸入資料能有正規化(Generalization)的效果，且能對量化後的空間維度作彈性的分配與調整。

雖然在應用 DCT 和類神經網路等方法於影像壓縮已有許多之探討[20]-[23]，本文主要之貢獻與新創之部分在於結合各方法之優點，形成一套整合型之處理過程；此外加入分頻處理的觀念，能更細緻的分析影像的內容。針對轉換後特徵顯著的部分保留較完整的訊息，以提昇處理之品質；而其他部份則做較高幅度之壓縮，以提高壓縮比，並改善系統之效率。因而本研究在品質不變的前提下，能有效提高資料傳輸之速度並降低資料儲存所須之空間。

二、系統架構

本研究所提出之整合式分頻 DCT 和以 SOM 做向量量化架構下的失真壓縮流程如圖 1 所示。

主要分成幾個部份。第一部份是前處理，包含影像向量或區塊化、DCT 與低通濾波器。影像向量化是把影像分割為固定大小的區塊，並針對每一區塊做個別處理，以降低運算之複雜度，這個過程一般稱之為向量

化(Vectorization)，而 DCT 則是將向量化後的影像方塊進行轉換，以得到影像方塊的頻域資訊，這些頻域資訊在經過分頻與低通濾波的處理才完成初步的資料簡化；上述步驟我們統稱為前處理(Preprocessing)。第二個部份則是依據 SOM 的訓練結果來完成向量量化，編碼簿和碼向量索引即是在這個階段產生。第一個部份與第二個部份是屬於編碼端的運作範圍，也是整個壓縮過程的核心。後面的步驟則屬於解碼端的運作範圍。第三個部份則是在解碼端搜尋接收之索引所對應的碼向量；第四個部份則將替代的碼向量作 DCT 之逆轉換，並加以重組來恢復原始影像。

在使用 DWT 分頻的方法時，整套系統流程與圖一相似，只要將圖中 DCT 與 IDCT 改成 DWT 與 IDWT 即可。另外在分頻處理上不若 DCT 僅分為 DC 與 AC，DWT 在頻率上的解析度極高，可將細節(Detail)和近似(Approximation)的部份分別採用不同的編碼簿來處理，以供傳輸與儲存用途；重建時再將其結合即可。

以下將就上述步驟做詳細說明。

三、離散餘弦轉換

假設原始輸入影像之大小為 $M \times M$ ，為降低影像在處理上的複雜度，並提高普及化之特性，可先將原始影像分割成等大小且不重疊的影像方塊，假設每個影像方塊有 $N \times N$ 個像素(Pixel)，則每一個影像方塊可視為向量維度為 $N \times N$ 的影像向量，這個過程我們稱之為向量化(Vectorization)。之後，每個向量都

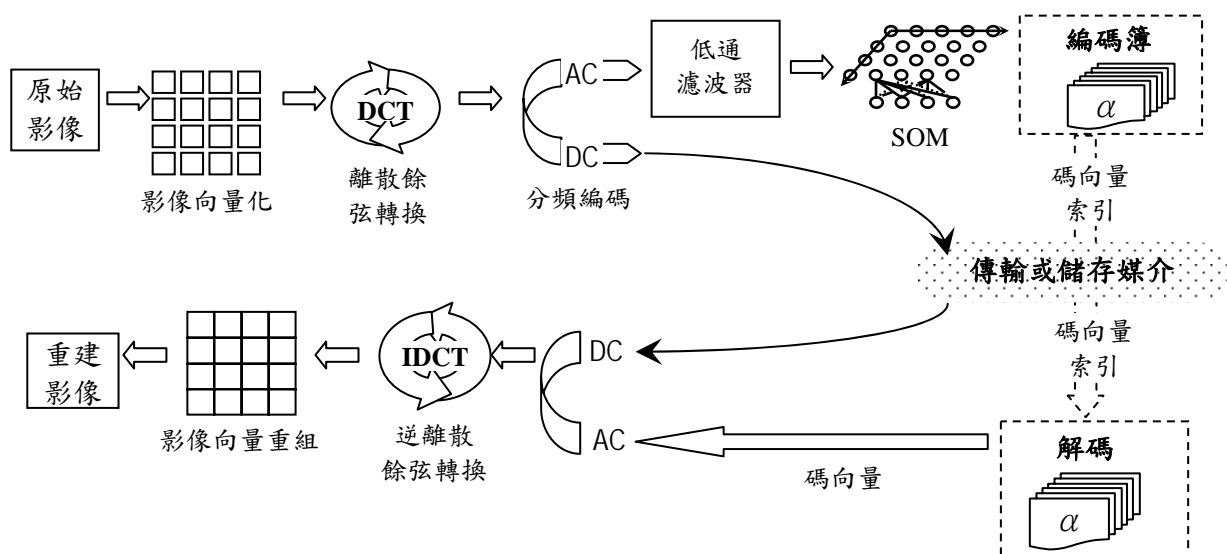


圖 1 利用 Sub-band DCT 與 SOM 影像壓縮與重建之架構流程圖。

要經過離散餘弦轉換(DCT),轉換的目的是將原訊號取樣間之相關性打散,使訊號的能量重新分佈而且只集中在少數幾個轉換後的係數當中,如此許多係數在量化後將為 0 或趨近於 0,故可予以忽略。

DCT 在影像處理的應用上,一個大小為 $N \times N$ 之影像方塊 $f(x, y)$ 其二維的 DCT 轉換之定義為:

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cdot \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (1)$$

式中 α 之定義為

$$\alpha(u), \alpha(v) = \begin{cases} \sqrt{\frac{1}{N}} & u, v = 0 \\ \frac{2}{\sqrt{N}} & u, v = 1, 2, \dots, N-1 \end{cases} \quad (2)$$

DCT 轉換為一可逆之線性轉換,而其二維的反轉換 IDCT 為

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u, v) \cdot \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (3)$$

DCT 和離散傅利葉轉換(DFT)一樣,都有快速演算法則可供使用,但是 DCT 只需作實數運算,DFT 則包含了實數與虛數的運算。且 DCT 在資訊集中的能力上,提供較佳的表達方式,故大多數的轉換編碼系統是以 DCT 為主,被認為是數位語音影像與信號處理的最佳工具之一。

在分頻編碼法(Sub-Band Coding, SBC) [24]-[26]裡,輸入訊號經由轉換被分離出幾個不同的訊號,每個訊號各自包含於有限範圍的頻率、頻譜。我們就稱這些訊號為次頻帶(Sub-Band)訊號。每一個次頻帶可以被分開個別做編碼的工作。以 DCT 編碼的觀念也與分頻編碼類似,只不過它的每一個次頻帶只含單一頻率,即是該頻率之基底函數以相對應之轉換係數加權後的結果。採用 SBC 的優點在於依照每個次頻帶而使用不同數量的位元數,故每個次頻帶的量化係數及重建誤差就可分別控制,因此重建誤差的頻譜形狀可以控制成頻率的函數。依此推論,我們可以依據每個次頻帶的特性使用不同的編碼方法。在比較低頻的次頻帶中,因為它含有原訊號

的大部份能量,因此我們可以使用較高的位元率來編碼;一般在最低頻的次頻帶,使用無失真編碼法即可完整重建該次頻帶。至於高頻的次頻帶,它所含的能量很少,因此可以使用低位元率做編碼,為較粗糙的壓縮編碼方式。

基於二維 DCT 轉換後具有分離高低頻能量的特性,以 8×8 的方塊為例,其頻域由低頻至高頻的排列方式如圖 2 表示。左上角 0 的位置,一般稱為直流係數(DC);1 至 63 稱為交流係數(AC);由 0 至 63 依序排列方式,呈鋸齒狀(Zig-Zag)或"之"型排列。

以 DCT 轉換後的結果來觀察,大部分影像方塊約有 80% 以上的能量是集中在 DC 部分,為避免較高的失真,可把 DC 直流係數直接編碼後傳遞出去或儲存起來,而 AC 交流係數則被用來作進一步的處理。因直接傳送 DC 的數值在過程上會造成嚴重負擔,故 Cham[27]提出一種方法來已 AC 值估算 DC 值,使 DC 值完全不用送出,大幅降低了 BPP(Bit per pixel),故本方法亦將於本計劃中做驗證。

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

圖 2 空間頻率係數排列順序。

此外再就 AC 之係數觀察,影像大部份的資訊、特徵和能量集中在係數矩陣的左上角即低頻部分,而較少的影像特徵是分散在高頻的部分;故就影像壓縮而言,當我們著重於影像的重建品質時,可保留較多的高頻係數以獲得較多的影像特徵和資訊;若著重的是壓縮比率,則需保留最少的低頻係數做為重建端的影像訊息來源。保留係數的方法只需以一低通濾波器即可完成。利用低通濾波器不僅能控制被保留係數的數量,又能減低影像向量的維度,以利於降低向量量化搜尋空間的複雜度,並增加編碼簿的分類能力。

四、SOM 與向量量化編碼

SOM 類神經網路的基本架構不同於傳統的類神經網路架構，它只需要輸入層與輸出層的結構，如圖 3 所示。輸入層即為訓練樣本的輸入向量，其神經元數目依各問題輸入特徵向量而不同；輸出層則代表聚類 (Clustering)，其神經元數目則可依所期望之輸出種類來決定。對於影像的向量量化而言，輸入向量為影像向量或特徵向量，而經訓練之後輸出層聚類的結果即形成編碼簿，每一個輸出神經元所代表的就是編碼簿中的碼向量索引，而各神經元所對應之碼向量即為連結權值。

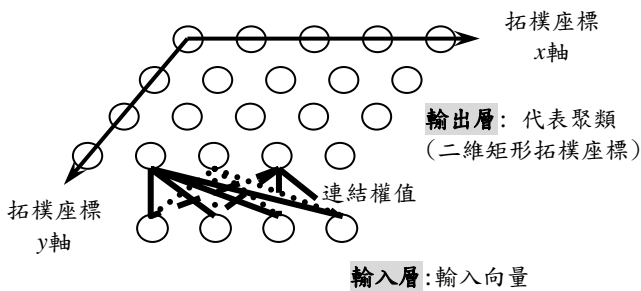


圖 3 SOM 類神經網路的網路架構

SOM 是採取非監督式的學習演算法，將高維度的輸入向量映射至一維或二維的輸出神經元陣列，針對各輸入向量不需事先確定所求的分類輸出結果，由神經網路自身對輸入資料作出綜合、歸納和統計，調整各神經元間的連結權值，使網路自適應地發展成為對不同輸入信號模式作出特殊的響應。網路拓樸是指輸出神經元的排列方式，一般大多採二維的矩形或六角形輸出模式。SOM 的輸出層神經元間相對位置具有相互影響的意義存在，對於不同的拓樸型式會有不同的分類結果。我們可以將輸入向量視為 n 維的影像向量，映射至一至二維的輸出神經元陣列，藉由這個映射的動作，將原本存在於訓練樣本輸入向量之間的相對距離關係，能保留在輸出神經元之間。簡單地說，彼此相鄰近的輸入向量，透過權值的學習，使得輸出神經元之間保有拓樸特性，映射在 SOM 的網路拓樸上。依此精神，其演算學習的法則首先是初始連結權值，即初始編碼簿，以隨機的方式建構我們所需要的 VQ 編碼簿。而決定優勝神經元的方法是計算輸入與權值間最近的歐基里德距離 (Euclidean Distance)，即

$$\| \mathbf{x} - \mathbf{w}_c \| = \min_i \{ \| \mathbf{x} - \mathbf{w}_i \| \} \quad (4)$$

式中 \mathbf{x} 為輸入向量， \mathbf{w}_i 為輸出神經元 i 的權值向量， \mathbf{w}_c 為勝利神經元 c 的權值向量，符號 $\| \bullet \|$ 則代表歐基里德距離。權值的調整是根據勝利神經元所在的位置來決定，若將輸入向量與權值向量都當成時間的函數，則

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_{ci}(t) \cdot [\mathbf{x}(t) - \mathbf{w}_i(t)] \quad (5)$$

式中， $h_{ci}(t)$ 為勝利神經元鄰近區域的函數，是控制鄰近區域大小與鄰近距離關係式的函數，我們選擇高斯函數為基礎，使不同的鄰近距離會有不同的調整比重，其表示式為

$$h_{ci} = \alpha(t) \cdot \exp\left(-\frac{\| r_c - r_i \|^2}{2\sigma^2(t)}\right) \quad (6)$$

式中， $\alpha(t)$ 為學習速率 (Learning Rate)， $\alpha(t)$ 與 $\sigma(t)$ 均是隨時間遞減的函數，所以鄰近距離將愈縮愈小以使各個類別更為聚集。

網路拓樸使得 SOM 之輸入特徵向量若彼此間近似程度較高，那麼其映射在輸出層之間的距離也較為接近。因此 SOM 在保持拓樸關係的前提下實現輸入空間的維數壓縮，完成類似於向量量化的功能。事實上，利用 Kohonen 的 SOM 可以構成一個向量量化器，其輸出便是向量量化，也就是聚類的結果，而輸出神經元的權值，則構成了 VQ 的編碼簿。

對於 SOM 類神經網路所構成的 VQ 編碼簿，與 LBG 向量量化演算法相較，SOM 類神經網路的編碼效率較佳。另外，SOM 類神經網路還具有收斂特性，不易受初始編碼簿的影響，生成的 VQ 編碼簿結構性強而且相對於訓練樣本集有較好的通用性之優點。

五、影像重建與效能評估

在 SOM 訓練完成之後所產生的編碼簿需預先儲存在資料庫中以便解碼時能夠使用。在影像重建時需先將所接收之碼索引還原成 AC 碼向量，再將 DC 值對應回原先的 AC 碼向量，之後經過逆 DCT 轉換就能產生原始影像的近似重建。

失真壓縮的編碼法允許重建信號或影像有些許差異，藉以換取比無失真壓縮編碼法更低的位元率與資料率。假設原始信號為 $f(x, y)$ 而重建信號為 $g(x, y)$ ，則重建信號的失真度或誤差可以用 $e(x, y)$ 表示，即

$$e(x, y) = f(x, y) - g(x, y) \quad (7)$$

標準差(Standard Deviation)的定義是為

$$\sigma_e = \sqrt{E\{[e(t) - \mu_e]^2\}} \quad (8)$$

式中 μ_e 為誤差平均值，一般是為0。 σ_e 又稱為均方根誤差(RMSE)。對編碼信號重建品質的客觀評量常是用信號雜訊比(Signal-to-Noise Ratio, SNR)，以原始信號的方差 σ_x^2 對重建誤差方差 σ_e^2 的比值取對數值，即

$$\text{SNR (dB)} = 10 \log \left(\frac{\sigma_x^2}{\sigma_e^2} \right) \quad (9)$$

而影像幅度的平均值通常為非零的正數，為避免計算上的負擔，如果輸入影像是以8個位元來表示，則最大的幅度值 $x_{\max}=255$ ，可用來代替式(9)中的 σ_x ，所得到之結果稱之為峰值雜訊比(PSNR)，即

$$\text{PSNR (dB)} = 10 \log \left(\frac{255^2}{\sigma_e^2} \right) \quad (10)$$

信號雜訊比之目的主要在觀察重建品質，具有客觀之標準；在壓縮比上也應採用一套標準來評估效能。最常使用的方式為計算每個像素所需要的位元，即所謂的bpp(Bits per pixel)。在結果檢討時，可以固定bpp的方法來比較重建品質PSNR，或固定PSNR來比較壓縮比bpp。

六、模擬結果

為驗證本研究之適用性，我們初步先以兩張影像Lena和Building作為模擬示範，採用DCT的分類方法。原始影像之大小為 512×512 ，若將其切割成每個有 4×4 大小的影像方塊，則共有 128×128 個方塊。若分別對這兩張影像的所有小方塊做DCT轉換，再對轉換後之係數予以低通濾波，藉以將影像的高低頻資訊分離出來，擷取影像的特徵，並減低影像向量的維度。為觀察此結果之特性，針對低通濾波後所保留之係數，我們不做SOM類神經網路分類而直接以IDCT逆轉換來還原影像，可比較PSNR之效果。表1和表2為分別對Lena和Building保留6和8個DCT係數值的比較，我們分別以DCT6和DCT8來表示。當保留6或8個DCT係數時，其還原之PSNR值仍可維持34dB以上的高重建品質影像。由此可推得DCT影像轉換的效率極高，原本影像向量維度是 $4 \times 4 = 16$ ，只要取6或8個維度即可，失真完全是在可以接

受的範圍之內。被保留的係數將做為SOM類神經網路的輸入向量。

表1 Lena取不同DCT係數後還原之比較。

Low-Pass Filter			
DCT6		DCT8	
RMSE	PSNR	RMSE	PSNR
4.2711	35.5201	2.8360	38.8381

單位：PSNR (dB)

表2 Building取不同DCT係數後還原之比較。

Low-Pass Filter			
DCT6		DCT8	
RMSE	PSNR	RMSE	PSNR
4.9887	34.1710	3.2334	37.9109

在SOM訓練樣本的選取上，主要採自類型不同的影像特徵，目的在於產生通用的編碼簿(Global Codebook)，如此所得的編碼簿才能保證各種影像都有平均水準的效率。在編碼簿大小的選取上，我們選擇128、256及512大小的編碼簿作為模擬結果分析的依據，意即分別需要7、8及9個位元來編所有的碼向量，因此其位元率(Bits Per Pixel)分別為0.4375bpp、0.5bpp及0.5625bpp。

DCT轉換後保留之係數是直接送至SOM做分類或向量量化的工作，DCT6和DCT8對SOM各有5和7個輸入，我們選取了上百個具代表性的影像方塊來訓練SOM，並建立編碼簿。由表1和表2的結果得知，DCT6與DCT8的PSNR差了3.5dB以上，但經SOM類神經網路壓縮測試的結果，DCT6與DCT8卻只剩下1dB左右的差距，對於這樣的結果是非常令人興奮且值得探討的。DCT8包含了比DCT6較多的影像資訊，所以其影像重建品質PSNR值應比DCT6來的高。由於在低通濾波器決定保留多少個DCT係數值取捨之下，DCT6比DCT8的向量維度低，對類神經網路於訓練生成編碼簿時，可減低計算上的複雜度，強化所生成編碼簿的結構性及提昇編碼簿的分類效率；故為提昇SOM類神經網路系統的學習成

效，輸入向量的向量維度不宜過高。我們發現若比較對神經網路的訓練與運算負擔，DCT6 確實比 DCT8 更具優勢，且能提供編碼簿較好的編碼效率。故在整個重建品質相似的前提下，DCT6 應為較佳的選擇。本計劃的執行將對更多相關的參數作探討，如不同大小的影像方塊，輸入向量，以及編碼簿等，以對其結果加以分析比較。

若加入分頻處理的過程則能得到更好的結果。對 Lena 和 Building 模擬測試的結果分別如表 3 和表 4 所示。利用分頻編碼的觀念，將 DCT 轉換後的高低頻係數分離，針對轉換後大部份影像特徵集中在 DC 係數值，所以 DC 係數值我們選擇不做任何處理，直接送至解碼端；對於 AC 係數值，做向量量化處理，同樣利用 SOM 類神經網路訓練生成 AC 編碼簿。若同樣依照低通濾波保留序號在前面的八個或六個轉換後之係數值，採取分頻編碼的壓縮架構其重建品質可高出 1~3dB，主要的原因在於我們選擇不對 DC 係數值做任何處理的緣故。再者，保留相同的轉換係數值，對採用分頻編碼的壓縮架構而言，其 SOM 類神經網路的輸入向量維度因取出 DC 係數值而少了一個維度，因此，SOM 類神經網路的輸入向量維度減少，權值空間較為平滑，學習與訓練過程的運算量與複雜度降低，所以比較容易尋得最佳化的編碼簿，這也是本計劃採用分頻編碼壓縮架構的目的。雖然位元率提高了約 0.5 個 bpp，也就是說以保留 7、8 及 9 個位元來編所有的碼向量，其位元率分別為 0.9375bpp、1bpp 及 1.0625bpp。但以此相似壓縮比的程度而言，單純使用 DCT 是無法達到如此高的 PSNR。

此外如表 5 和表 6 選取 64、128、256 大小的 AC 編碼簿，因權值空間變得比較平滑，向量維度減低，在 128 與 256 大小的 AC 編碼簿，其編碼簿以趨於收斂，相對的其影像重建之 PSNR 值差異不大，使得我們可以選取較小的 AC 編碼簿也能有較佳重建品質。因此，雖然對於 DC 係數值不做處理會使得壓縮比率增加，但我們卻可以較少的位元去編碼 AC 編碼簿，兩相平衡之下，其整體的壓縮比率並不會高出許多，而在解碼端因編碼簿碼向量的個數變少，也能達到快速搜尋的效果。

表 3 分頻重建 Lena 後之結果。

Figure Name	Codebook Size	Low-Pass Filter	
		6	8
		PSNR	PSNR
<i>SOM</i>	256	26.4964	29.2003
<i>SOM with Sub-band</i>	256	30.0530	31.3677

表 4 分頻重建 Building 後之結果

Figure Name	Codebook Size	Low-Pass Filter	
		6	8
		PSNR	PSNR
<i>SOM</i>	256	25.9524	27.0373
<i>SOM with Sub-band</i>	256	29.3075	33.3410

圖 4 至圖 5 分別為 Lena 和 Building 在不同狀況下之重建影像。圖 4 為使用 Kohonen 之 SOM 編碼，取 256 個 codeword 之結果，PSNR 為 32.41dB；圖 5 為使用 Kohonen 之 SOM 編碼，取 256 個 codeword 外加分頻處理的結果，PSNR 提昇至 36.53dB。無論就數字上或視覺上的觀察，我們都可發現加了分頻處理會使 PSNR 大幅提昇，同時因為 256 個 codeword 和 128 個 codeword 之 PSNR 並無太大差異，在 VQ 的複雜度上亦能降低許多。



圖 4 利用 SOM 與 256 個 codeword 重建 Lena 之影像。



圖 5 利用 SOM 與 256 個 codeword 加上分頻重建 Lena 之影像。



圖 6 利用 SOM 與 128 個 codeword 重建 Building 之影像。



圖 7 利用 SOM 與 128 個 codeword 加上分頻重建 Building 之影像。

七、結論

本文將以神經網路之向量量化與分頻 DCT 為基礎，對影像壓縮之法則開發及其效能加以研究。在資訊時代科技不斷日新月異的今天，數位影像的應用將更加廣泛與需要，但因有許多瓶頸的存在使得很多領域都尚待推廣與開發，例如影音電話，無線影音傳輸，和影像資料庫之儲存等，若影像壓縮技術在運算速度、壓縮比、信號雜訊比與功率消耗等功能上能夠更進一步的提昇，則許多應用上的問題都將能一一突破。本文所提出之方法整合了現今所使用各種方法之優點，並改進了其缺點與限制，大幅提高壓縮比與信號雜訊比，因而在資料之傳輸上能夠提昇速度，在資料之保存上亦能降低儲存所須之空間，因此本文對影像處理的基礎理論和實際應用都會有長足的貢獻。而神經網路的並行處理架構能以即時的方式處理問題，只需要簡單的硬體設備與法則就能快速的執行。

參考文獻

- [1] A. K. Jain, *Fundamental of Digital Image Processing*, Prestice-Hall, Englewood Cliffs, NJ, 1989.
- [2] W. K. Pratt, *Digital Image Processing*, 3rd ed., Wiley-Interscience, New York, 2007.
- [3] C. Liu; J. Lin; K. Wu and C. Wang, "Objective image quality measure for block-based DCT coding," *IEEE Transactions on Consumer Electronics*, vol. 43, no. 3, pp. 511-516, Aug. 1997.
- [4] N. Ahmed, T. Natarajan and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Comput.*, vol. C-23, pp. 90-93, Jan. 1974.
- [5] K. R. Rao and P. Yip, *Discrete Cosine Transform - Algorithms, Advantages, Applications*, Academic Press, 1990.
- [6] H. Yamaguchi, "Adaptive DCT coding of video signals," *IEEE Transactions on Communications*, vol. 41, no. 10, pp. 1534-1543, Oct. 1993.
- [7] P. Dalessandro and R. Lancini, "Video

- coding scheme using DCT-pyramid vector quantization,” *IEEE Transactions on Image Processing*, vol. 4,no.3, pp.309 –319, March 1995.
- [8] S. Arazaki, M. Saigusa, S. Hashiguchi, M. Ohki, M. Uchiyama and F. Itoh, “Image data compression by DCT with adaptive run-length coding,” *IEEE Transactions on Consumer Electronics*, vol. 37,no.4, pp. 860 –866, Nov. 1991.
- [9] Li W. S., Wang, Z. S.; He, Z. Y., “Neural network based real-time computation of DCTs and DSTs,” *Electronics Letters*, vol. 32, no.9, pp. 1795 –1796,12 Sept. 1996
- [10] Z. Xiong; K. Ramchandran, M. T. Orchard and Y. Zhang,“A comparative study of DCT- and wavelet-based image coding,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no.5, pp. 692-695, Aug. 1999.
- [11] Hong Wang; Ling Lu; Da-Shun Que; Xun Luo, “Image compression based on wavelet transform and vector quantization,” *Proceedings of 2002 International Conference Machine Learning and Cybernetics*, vol. 4, pp. 1778-1780, Nov. 2002.
- [12] N. M. Nasrabadi and R. A. King, “Image Coding using Vector Quantization: a review,” *IEEE Trans. on Communication*, vol. 36, no. 8, pp. 957-971, Aug. 1988.
- [13] Dong Sik Kim, and Sang Uk Lee, “Image Vector Quantizer Based on a Classification in the DCT Domain,” *IEEE, Trans. on Communication*, vol. 39, no. 4, Apr 1991.
- [14] R. M. Gray, “Vector Quantization,” *IEEE Acoust., Speech, Signal Processing Magazine*, pp. 9-31, Apr. 1984.
- [15] A. Gersho and Robert M. Gray, *Vector Quantization and Signal Compression*. London:Kluwer, 1992.
- [16] D. E. Rumelhart, G. E. Hinton and R. J. Williams, *Parallel Distributed Processing*, Cambridge, Massachusetts: The MIT Press, 1986.
- [17] R. P. Lippmann, “An Introduction to Computing with Neural Nets,” *IEEE ASSP Magazine*, pp. 4-22, Apr. 1987.
- [18] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed., Prentice Hall, 1998.
- [19] T. Kohonen, *Self-Organization Maps*, Springer-Verlag, New York, 1997
- [20] K. S. Ng and L. M. Cheng, “Artificial Neural network for Discrete Cosine Transform and Image Compression,” *Proceedings of the 4th IEEE ICDAR*, pp. 675-678, 1997.
- [21] Veleva, L.V.; Kunchev, R.K. ,“Adaptive speech coding with DCT and neural net vector quantisation,” *Electronics Letters* vol. 29, no. 8, pp.704–705, 15 April 1993.
- [22] Harandi, M.T.; Gharavi-Alkhansari, M., “Low bitrate image compression using self-organized Kohonen maps,” *Proceedings of 2003 International Conference on Image Processing*, vol. 2, pp. II-267-70, Sept. 2003.
- [23] Chang, P.R.; Hwang, K.S.; Gong, H.M., “A high-speed neural analog circuit for computing the bit-level transform image coding,” *IEEE Transactions on Consumer Electronics*, vol. 37, no. 3, pp. 337-342, Aug 1991.
- [24] Da Silva, E.A., Ghanbari, M., ”A DCT-based aliasing cancellation method in subband coding,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 3, no. 5, pp. 384-387, Oct. 1993.

- [25] Sundsbo, I.; Ramstad, T.A., "Synthesis filterbank with low hardware complexity for subband image coding," *IEEE Transactions on Image Processing*, vol. 7, no.12, pp. 1717-1724, Dec. 1998.
- [26] Mukherjee, J.; Mitra, S.K.; "Image resizing in the compressed domain using subband DCT," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 7, pp.620-627, July 2002.
- [27] Tse, F.-W.; Cham, W.-K.; Liu, J.Z.; "DC coefficient restoration technique and its application to image coding," *IEE Proceedings on Vision, Image and Signal Processing*, vol. 149, no. 5, pp. 272-282, Oct. 2002.