

1

一個適用於雲端運算環境的高效能與
高效率影像壓縮技術

蔡正發教授、朱浚煌 分機：7906 / 屏科大資訊管理系 / cftsai@mail.npust.edu.tw

一、源起

近年來由於資訊科技的快速進步，許多網路應用服務已變得更加豐富，多媒體之應用在網際網路上受到極大的歡迎，導致使用者需求也更加多元化。人們在日常生活中大量地利用電腦來處理資訊作業，而在傳輸資料時數位內容特別是數位影像將造成儲存設備及網路頻寬的大量消耗，因此數位影像壓縮在現今的網路環境中即成為一重要且常見的技术。在過去數十年中，已有眾多類型的影像壓縮技術被提出來解決佔用傳輸成本的議題 [1]-[4]。一般而言，影像壓縮技術可分為「頻率領域(Frequency domain)」及「空間領域(Spatial domain)」，頻率領域利用一連串的訊號頻率轉換來消除較不重要的訊號以達到縮小影像容量之目的；空間領域主要利用向量量化 (Vector Quantization, VQ) 技術以降低影像中重複訊號所浪費的容量[5]-[13]。然而在此兩種不同技術中又可再分為「失真 (Lossy)」及「非失真(Lossless)」式壓縮。顧名思義，失真技術壓縮的圖片品質會較原始圖片稍差，但也因為失真技術利用犧牲些微圖片品質而可達到相當高的壓縮比例。而非失真技術之壓縮圖片可以百分之百毫無差別的還原成原始圖片，也因為如此，非失真壓縮圖片就無法達到如失真壓縮般的高壓縮比例。在諸多類型的壓縮技術當中[5]，我們針對擁有高壓縮率及高壓縮品質的向量量化 (Vector Quantization, VQ) [6]-[7]技術進行研究。下圖1表示一個在空間領域失真式影像壓縮技術所需基本步驟：首先，原始圖片經由向量量化將較具有代表性且容量較小的代表向量代替原始圖片保存於編碼簿(Codebook)中；而後，當要傳輸圖片時便傳輸被保存之編碼簿，並透過向量還原步驟還原成稍微失真之壓縮圖片。因此，如何改善向量量化技術，以更有效率的方式壓縮影像便是本發明的主要目標。

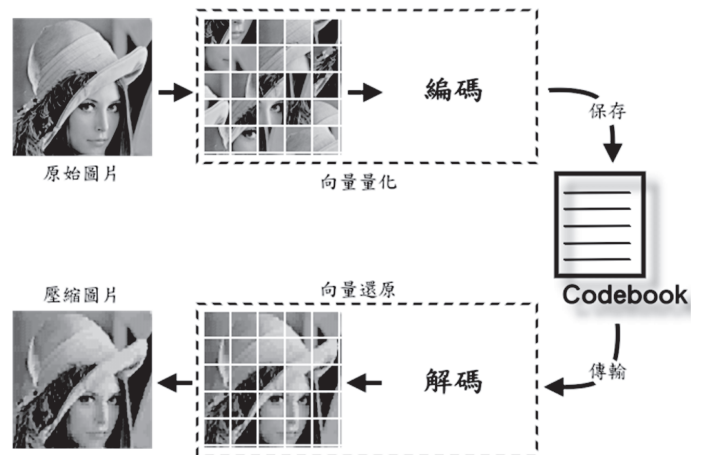


圖1 影像壓縮 / 解壓縮概念說明圖

在向量量化技術中最廣為人知的即是LBG演算法[8]，其優點為速度快及容易實作，但由於其分群過程依靠簡單的距離判斷規則及隨機決定初始編碼簿等特性，使其分群品質不甚良好。而SOM(Self-Organizing Map)[9]-[11]及HSOM (Hierarchical SOM)[12]此種類神經演算法雖然其以相當精細的訓練步驟已獲得到十分細膩的壓縮結果，但其最大之缺點即為速度過慢。本發明提出一個新的影像壓縮技術稱為ELSA，此方法主要採用兩階段步驟進行影像壓縮，第一階段係改良LBG隨機初始化編碼簿之方式，以更有效率的初始化方式快速獲得概略的初始編碼簿；而第二階段再套用一正規的LBG針對在第一階段得到的編碼簿進行處理以達成最後的優化效果。實驗結果顯示，本發明提出的影像壓縮技術在壓縮時間(Codebook Time)及壓縮品質(PSNR)上皆大幅優於其它已發表在國際上之著名之技術，如LBG、SOM、HSOM及INTSOM [2]等，同時也優於商業圖片編輯軟體Adobe Photoshop 9。

二、設計概念

本發明提出之ELSA技術主要採用以下步驟進行影像壓縮：(1)利用動態分裂樹產生初始編碼簿、(2)提出一個更有效率的失真度評估方法、(3)編碼簿優化收斂。以下為各步驟之說明以及演算流程。

1. 利用動態分裂樹產生初始編碼簿

如圖2(a)所示，起初只有一個代表區塊A代表所有的像素區塊。在圖2(b)中，將失真度最大的代表區塊(假設是A)以LBG演算法重新分群出兩個新的代表區塊A'及B'。在圖2(c)中將失真度最大的代表區塊(假設是B')同樣以LBG重新分出兩個新的代表區塊B''與C''。

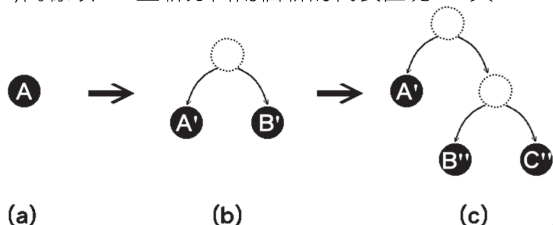


圖2 動態分裂樹

在此步驟當中ELSA在每次的分裂過程針對失真度最高的代表區塊套用LBG(收斂門檻值： ϵ ，中心點數量：2)將原本的代表區塊重新計算出兩個新的代表區塊，並且重複此準則逐一產生出所需的總代表區塊數量為止。

2. 一個更有效率的失真度評估方法

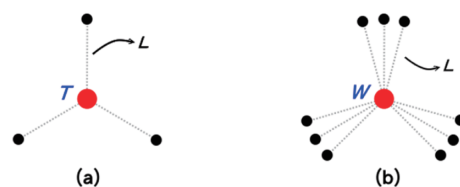
以逐步分裂失真度最大的代表區塊來進行編碼簿初始化，本發明提出一個較MSE更有效率的方法做為評斷失真度之用。首先，MSE的定義為：先計算圖片中每一個原始像素與壓縮後的像素之差距，並取平方值之加總，最後再除以總像素數量。因此，針對某一代表區塊計算之失真度MSE(n)如公式(1)所示，其中 $X_i(n)$ 為屬於第n個代表區塊的第i個像素區塊， $C(n)$ 為第n個代表區塊， $B(n)$ 為屬於第n個代表區塊的像素區塊數量。

$$MSE(n) = \frac{\sum_{i=1}^{B(n)} [x_i(n) - c(n)]^2}{B(n)} \quad (1)$$

而本發明提出的新失真度評估方法為：TSE(Total Square Error)如公式(2)所示，與MSE不同的地方在於TSE最後並不除以總像素數量。

$$\text{TSE}(n) = \sum_{i=1}^{B(n)} \left[x_i(n) - c(n) \right]^2 \quad (2)$$

圖3演示一個MSE與TSE差異的比較範例，以MSE求得代表區塊 T 之失真度為 $\frac{3L^2}{3} = L^2$ 、代表區塊 W 之失真度為 $\frac{9L^2}{9} = L^2$ 。亦即若以MSE公式來評估代表區塊 T 及 W 會得到相同的失真度；以TSE求得代表區塊 T 之失真度為 $3L^2$ 、代表區塊 W 之失真度為 $9L^2$ 。若以TSE來評估，則失真度較大的代表區塊 W 將較代表區塊 T 更值得進行分裂。在評估代表區塊 T 及 W 之失真度上，我們認為理應同時考慮像素區塊數量所帶來的影響。因此，求取總誤差的TSE即較求取平均誤差的MSE來得恰當。



▲ 圖3 MSE與TSE之差異比較圖

3. 編碼簿優化收斂

在以動態分裂樹快速產生初始編碼簿之後，我們套用LBG演算法對於初始編碼簿進行訓練。在此步驟中的LBG與正規的LBG不同之處在於正規的LBG之初始編碼簿是經由隨機產生的，而此步驟的LBG之初始編碼簿已由動態分裂樹策略快速以及概略性的產生。由於正規LBG隨機初始編碼簿的性質，使得之後的訓練過程將花費較久的時間才能達到收斂。而此步驟的LBG之初始編碼簿已大抵得到適度的調整，因此在之後的訓練過程中將花費較少的訓練時間即可達到收斂條件。而在初始編碼簿的品質上，由動態分裂樹產生之編碼簿也較隨機產生之編碼簿良好。

三、技術開發

本發明揭示一種影像壓縮技術，以智慧型動態擴張樹的概念將壓縮圖片所需的調色盤顏色一一訓練出來，進而成為容量較小但足以用來代表原始圖片的像素編碼簿。在非頻率領域之有損式影像壓縮技術中最廣為人知的即是LBG(Linde, Buzo and Gray)，其優點為速度快及容易實作，但由於其分群過程依靠簡單的距離判斷規則及隨機決定初始編碼簿等特性，使其分群品質不甚良好。而SOM(Self-Organizing Map)及HSOM(Hierarchical SOM)此種類神經演算法雖然其以相當精細的訓練步驟來得到十分細膩的壓縮結果，但最大的缺點就是速度過慢。然而，實驗結果顯示，本發明提出的影像壓縮技術在壓縮時間(Codebook Time)及壓縮品質(PSNR)上皆大幅優於其他已發表在國際上之著名演算法，如 LBG、SOM、HSOM等，同時也優於商業圖片編輯軟體Adobe Photoshop 9。

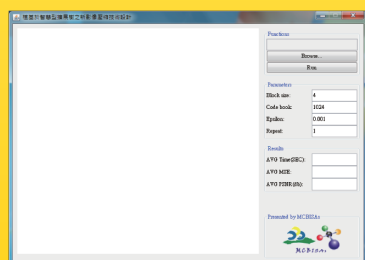
本發明運作流程係包含以下步驟：

1. 將一原始輸入影像分割成數個原始區塊，再將各該原始區塊轉換為原始向量；
2. 設定一失真度參數及一代表區塊參數；
3. 將該數個原始向量設定為同一個群組；
4. 一前置分群步驟，係以分割式分群演算法，將目前具有最高失真度之群組內的原始向量分為兩群，重複進行，直至總群組數到達設定之代表區塊參數；
5. 以該數個群組所對應之形心點作為初始形心點並以該些初始形心點共同對所有原始向量進行分群，以獲得數個最終形心點，並將該些最終形心點所分別對應之向量共同儲存於一編碼簿

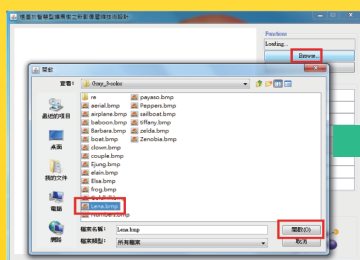
綜上所述，本發明提出一個較前述所有演算法進步且有效的灰階影像壓縮方法。其優點如下所示：(1)低時間複雜度、(2)高壓縮品質、(3)架構簡單、(4)低監督複雜度。此外，據實際測試結果，本發明之壓縮品質也優於商業圖片編輯軟體Adobe Photoshop 9。總之，(1)ELSA所需的輸入參數與LBG一樣少、(2)ELSA與LBG最大的差異在於編碼簿的初始策略、(3)相較於其它類神經網路技術，ELSA與LBG之訓練架構皆簡單許多。然而，因為ELSA可快速得到訓練得宜的初始編碼簿，使得之後的訓練過程可較LBG更快達到較佳的收斂狀態。由實驗結果指出，ELSA之壓縮品質與訓練編碼簿時間都較其它比較演算法大幅領先許多。

系統簡介：

本發明為實作一新的適用於雲端運算環境的高效能與高效率影像壓縮技術，系統介面由資料呈現區及使用者操作區所構成。本系統主畫面如圖4所示，操作時首先載入欲執行影像壓縮之圖片，如圖5所示，然後再輸入參數，如圖6所示，而圖7乃為執行影像壓縮之結果。



▲ 圖4 本系統之起始主畫面



▲ 圖5 載入圖片



▲ 圖6 參數設定



▲ 圖7 執行結果

四、技術競爭力

現今資訊之應用不斷蓬勃發展，其中影像壓縮於網際網路多媒體傳輸中扮演重要之角色。本發明係關於一種編碼簿產生方法，特別是一種用於影像壓縮技術之編碼簿產生方法。本影像壓縮技術揭示一種影像壓縮技術，以智慧型動態擴張樹的概念將壓縮圖片所需的調色盤顏色一一訓練出來，進而成為容量較小但足以用來代表原始圖片的像素編碼簿，成功降低訓練所需執行時間及迭代次數，其於影像壓縮之執行時間與影像壓縮品質於各種不同之測試影像皆優於LBG、SOM、HSOM等國際著名之影像壓縮方法，同時壓縮品質也優於商業圖片編輯軟體Adobe Photoshop 9。就本參選作品之實施成本而言，其具有(1)低時間複雜度、(2)高壓縮品質、(3)架構簡單、(4)低監督複雜度等特性，故其實施成本低。

五、研發成果

為驗證本發明所提出之ELSA技術是否優於其它比較方法，故使用壓縮技術評估準則來測量壓縮品質及壓縮時間，並以九張影像進行測試。在各個演算法、各個圖片、各個編碼簿大小各執行30次的獨立運算來取得平均測試結果。演算法程式皆由Sun JAVA 程式語言實作，測試環境為Intel Pentium4 3.2 GHz CPU、1GB RAM以及Microsoft Windows XP Pro作業系統。而所有測試圖片之影像尺寸為512X512像素、以切割單位4X4的區塊大小來進行前置切塊。此外，由於HSOM無法處理不可開平方得整數之編碼簿大小，因此有些測試表格中的欄位並無測試資料(N/A, Not Available)。其中，Time為影像壓縮時間，越低代表系統運算速度越快；PSNR(Peak Signal-to-Noise Ratio)為一常用來評比壓縮品質之公式，PSNR越高代表壓縮品質越好。測試結果如表1所示。圖8顯示執行ELSA在六種不同之圖片之影下壓縮結果畫面(a)Lena、(b)Airplane、(c)Peppers、(d)Boat、(e)Sailboat、(f)Ejunga。圖9顯示ELSA(亦稱ET-LBG)與LBG、SOM、HSOM在不同大小編碼簿時使用(a)Lena、(b)Airplane、(c)Peppers、(d)Boat、(e)Sailboat、(f)Ejunga圖片之影像壓縮品質PSNR (dB) 與執行時間(秒)之比較，而圖10則為ELSA(亦稱ET-LBG)在不同大小編碼簿時之影像壓縮品質PSNR(dB)比較。由圖8-10可知，ELSA在不同大小編碼簿時之影像壓縮品質PSNR(dB)與執行時間(秒)之皆較比較LBG、SOM、HSOM為佳。

Codebook Size		128		256		512		1024	
Images	Method	PSNR	Timecost	PSNR	Timecost	PSNR	Timecost	PSNR	Timecost
Lena	LBG	29.549	5.18	30.447	10.32	31.262	18.26	32.109	29.96
	SOM	29.677	227.98	30.581	434.13	31.471	832.02	32.436	1624.91
	HSOM	N/A	N/A	30.642	67.58	N/A	N/A	33.022	132.73
	ET-LBG	29.697	3.21	30.758	5.60	31.929	8.87	33.386	14.87
Airplane	LBG	30.527	7.64	31.416	11.60	32.313	17.74	33.313	26.94
	SOM	31.115	229.30	32.124	424.82	33.156	831.29	34.319	1621.73
	HSOM	N/A	N/A	32.150	68.44	N/A	N/A	34.684	134.65
	ET-LBG	31.173	3.15	32.306	5.46	33.542	8.65	35.029	14.86
Peppers	LBG	29.650	5.31	30.441	9.83	31.197	16.66	31.980	25.89
	SOM	29.777	227.66	30.600	436.20	31.390	828.78	32.305	1620.84
	HSOM	N/A	N/A	30.613	68.95	N/A	N/A	32.594	142.48
	ET-LBG	29.798	3.09	30.712	5.29	31.660	8.76	32.698	14.04
Boat	LBG	29.068	7.19	29.896	11.50	30.733	17.96	31.670	26.60
	SOM	29.329	223.80	30.238	429.68	31.214	824.28	32.281	1610.44
	HSOM	N/A	N/A	30.175	67.88	N/A	N/A	32.475	134.40
	ET-LBG	29.311	3.21	30.326	5.37	31.425	9.07	32.688	14.76
Sailboat	LBG	27.484	6.35	28.188	9.91	28.931	15.41	29.776	23.19
	SOM	27.713	231.38	28.500	434.88	29.334	837.35	30.298	1653.13
	HSOM	N/A	N/A	28.502	68.90	N/A	N/A	30.421	135.15
	ET-LBG	27.689	2.97	28.564	5.18	29.484	8.71	30.543	14.21
Ejunga	LBG	31.542	5.48	32.587	10.60	33.576	18.41	34.623	28.88
	SOM	31.887	231.73	32.905	440.23	33.992	833.84	35.130	1678.81
	HSOM	N/A	N/A	32.95	67.83	N/A	N/A	35.461	135.76
	ET-LBG	31.932	3.14	33.096	5.46	34.311	9.70	35.687	15.67

本發明專利可應用於如防洪、土石流災情連續影像攝影壓縮後需傳至雲端等應用，產值佳。本發明同樣可適用於如市場區隔(Market Segmentation)、鏡頭自動對焦(Auto-Focusing)、影像的壓縮、資料分群(Data clustering)、資料採礦(Data mining)、圖樣辨識(Pattern recognition)、音訊辨識(audio recognition)、虹膜辨識(IRIS recognition)、人臉辨識(Face recognition)等應用。本專利在產品上應用之廣度極佳。本發明專利及以此為基礎衍生之相關影像壓縮技術已於2012-2014三年中共參加五次如阿基米德、波蘭華沙、義大利、台北等著名國際發明展，已獲得多面金牌與銀牌；其次，目前已有數家影像處理相關產業公司與發明者合作，欲努力將本發明技術應用於該等公司之產品中。此外，本發明已獲得中華民國與美國發明專利共兩件。

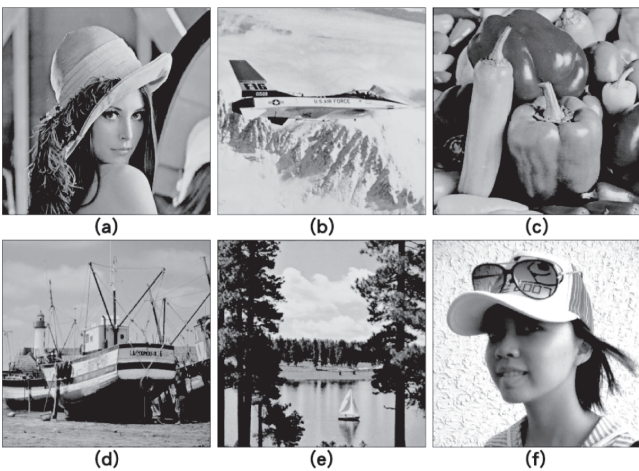


圖8 執行ELSA在六種不同之圖片之影下壓縮結果畫面(a)Lena、(b)Airplane、(c)Peppers、(d)Boat、(e)Sailboat、(f)Ejunga

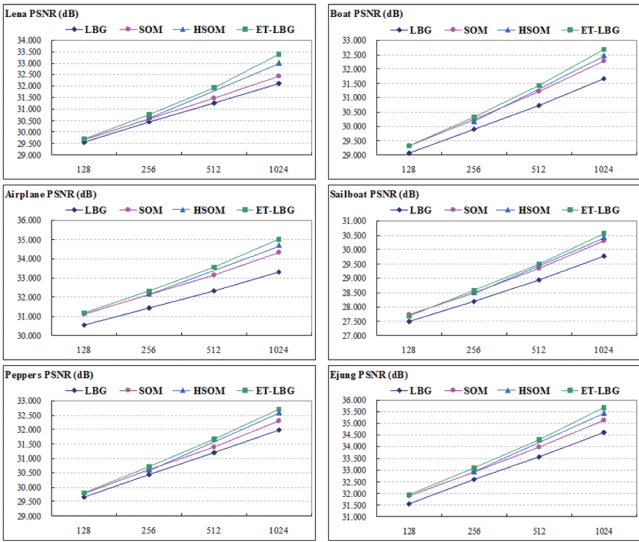


圖10 ELSA (亦稱ET-LBG) 在不同大小編碼簿時之影像壓縮品質PSNR (dB)比較

圖9 ELSA (亦稱ET-LBG)與LBG、SOM、HSOM在不同大小編碼簿時使用(a)Lena、(b)Airplane、(c)Peppers、(d)Boat、(e)Sailboat、(f)Ejunga 圖片之影像壓縮品質PSNR(dB) 與執行時間(秒)之比較

致謝

作者感謝屏東科技大學研發處之協助及中華民國科技部研究計畫之經費補助，計畫編號為NSC 102-2410-H-020-009。

參考文獻

[1]Tsai, C.F. and Ju, J.H., "ELSA: A New Image Compression Using An Expanding-Leaf Segmentation Algorithm," Lecture Notes in Computer Science, Vol. 5579, pp. 624-633, 2009.
[2]Tsai, C.F. and Ju, J.H., "INTSOM: Gray Image Compression Using An Intelligent Self-Organizing Map," Studies in Computational Intelligence, Vol. 199, pp. 31-40, 2009.
[3]Tsai, C.F. and Lin, Y.J., "LazySOM: Image Compression Using an Enhanced Self-Organizing Map," Lecture Notes in Computer Science, Vol. 5414, pp. 118-129, 2008.
[4]Tsai, C.F., Jhuang, C.A. and Liu, C.W., "Gray Image Compression Using New Hierarchical Self-Organizing Map Technique," In: Proceedings of the 3rd International Conference on Innovative Computing Information and Control, pp. 544-549, 2008.
[5]Sayood, K., "Introduction to Data Compression," 2nd Edition, Morgan Kaufmann Publishers, San Francisco, 2000.
[6]Gray, R.M., "Vector Quantization," IEEE Acoustics Speech and Signal Processing Magazine, Vol. 1, No. 2, pp. 4-29, 1984.
[7]Gersho, A. and Gray, R.M., "Vector Quantization and Signal Compression," Kluwer Academic Publishers, Boston, MA, 1992.
[8]Linde, Y., Buzo, A. and Gray, R., "An Algorithm for Vector Quantizer Design," IEEE Trans. on Communications, Vol. 28, pp. 84-95, 1980.
[9]Kohonen, T., "Self-organizing Maps," Springer-Verlag, Berlin, 1995.
[10]Kohonen, T., "The Self-Organizing Map," In: Proceedings of the IEEE, Vol. 78, No. 9, pp. 1464-1480, 1990.
[11]Madeiro, F., Vilar, R. and Neto, B., "A Self-Organizing Algorithm for Image Compression," In: Neural Networks, Vol. 28, pp. 146-150, 1998.
[12]Barbalho, M., Duarte, A., Neto, D., Costa, F. and Neto, A., "Hierarchical SOM Applied to Image Compression," In: Proceedings of International Joint Conference on Neural Networks, Vol. 1, pp. 442-447, 2001.
[13]Lehmann, E.L. and Casella, George, "Theory of Point Estimation," 2nd Edition, Springer-Verlag, New York, 1998.