行政院國家科學委員會專題研究計畫成果報告

計畫名稱： 以調變小波轉換為基礎的子頻道影像編碼系統─研發與實現
計畫編號： NSC 92－2213－E－216－013
執行期限： 92 年 8 月 1 日至 93 年 7 月 31 日
執行機構及單位名稱：中華大學電機系
主持人： 辛錫進
協同研究人員：藍兆禾
計畫參與人員： 曾國華　吳宜龍　李秋樺

一、中文摘要

根據資料特性將影像分解成數個大小不等的子頻帶成份後，利用 AM-FM 模型的特點可以更有效地描述子頻帶成份並提高壓縮品質，達到高品質可變位元率的影像壓縮結果。若配合調變小波轉換描述子頻道影像可以額外獲得的優點有：多解析度分析模式、空間域與頻域的區域性資料分析、調變小波係數間的低相關性與能量集中性、容易實現以及資料重建的準確性等。相對於傳統的小波轉換只能對低頻的資料成份做連續性分解，調變小波轉換卻可以對不同頻帶的影像成份做細部分解，且被分解的頻帶是由調變頻率參數所決定，選擇適當的調變頻率參數可以提高影像描述效率。我們提出一個根據影像資料之主瞬間頻率之對比，將影像分解成數個子頻帶成份，再依據每個子頻帶的主瞬間頻率之機率分佈選取適當的調變頻率參數各別進行調變小波轉換以提高影像壓縮品質。

關鍵詞：調變小波轉換，最佳調變頻率， AM-FM 模型

Abstract

High quality, embedded, variable rate image compression can be achieved by decomposing an image into subbands of different sizes, modeling with one significant AM-FM component as well as encoding with a distinct procedure for each subband. In this paper, an adaptive subband decomposition using the modulated wavelet packet transform is proposed, where the adaptation of modulating frequencies is based on the energy spectral density in a resolution-recursive framework; and, each subband can be well represented by the modulated wavelets. An extension of SPIHT algorithm that is needed for encoding the complex-valued modulated wavelet coefficients is also presented. Experimental results show that the modulated wavelet subband image coding is preferable to both the wavelet coding and the JPEG standard

Keywords: Modulated wavelet transforms; Optimal modulating frequency; AM-FM modeling

二、Introduction

The main obstacle to image applications such as storage and transmission over a band-limited channel is the huge amount of data required to represent an image directly. There has been increasing demand for image compression with the rapid growth in modern communications and computer technologies; this is a natural trend. State of the art techniques can compress typical images by factors of 10 to 50 without significantly degrading the image quality, depending on the specific application and encoder/decoder complexity [1]-[4]. The Joint Photographic Experts Group (JPEG) [5] shows good performance at moderate to high bit rates of compression measured in bits per pixel (bpp). Multiresolution representation is well suited to the properties of Human Visual System (HVS); thus, the wavelet subband coding has shown promising results at low to moderate bit rates [6]. Wavelet transform provides many advantages such as multiresolution analysis, joint spatial-spectral localization, fast de-correlation with compact energy distribution in the wavelet domain, and exact reconstruction, which are beneficial to the image compression task [7]. Many competitive wavelet coders including embedded zero-tree wavelets (EZW) of Shapiro [8], set partitioning in hierarchical trees (SPIHT) of Said and Pearlman [9], morphological representation of wavelet data (MRWD) of Servetto et al. [10], and group testing for wavelets (GTW) of Hong and Ladner [11] have been developed. In addition, wavelet based coding techniques have been adopted as the underlying methods to implement the JPEG 2000 standard [12]. Wavelet transform analyzes an image with a decomposition procedure, which is recursively performed on the low frequency component only; on the other hand, wavelet packet transform applies the decomposition procedure to both the low and high frequency components to generate a larger family of subband components [13]-[14]. Recently, the modulated wavelet transform acting as an extension of wavelet transform had been proposed [15]. Images can be represented by using the modulated wavelet transform to gain an adaptable zooming in the frequency regions of importance depending on the energy spectral density (ESD), instead of being constrained in the fixed low frequency region centered at zero frequencies. As a result, the modulate wavelet transform provides a flexible as well as adaptable representation framework that can open a broad range of image applications. This paper extends the previous work in [16], where the modulated wavelet subband coding with a fixed Gabor filterbank was introduced. Specifically, an adaptable rather than fixed subband decomposition is proposed to improve the coding performance; and, a simplified version of the extended SPHIT algorithm is developed to encode the complex valued modulated wavelet coefficients.Image compression is crucial to many applications dealing with storage and transmission of visual data. To design effective coders, it is essential to recognize the importance of data representation. The wavelet representation has many properties that are beneficial to the image compression task; namely multi-resolution analysis, spatial-spectral localization, fast de-correlation with energy compaction of wavelet coefficients, easy implementation and exact reconstruction. Many competitive wavelet-based image coders had been proposed, e.g. embedded zero-tree wavelets (EZW) of Shapiro, set partition in hierarchical trees (SPIHT) of Said, and morphological representation of wavelet data (MRWD) of Servetto. In this project report, a new, modulated wavelet transform (MWT) based approach to image representation for compression is presented. In contrast to wavelet transform (WT) that is focused on the low frequency decomposition in a multi-resolution manner, MWT successively decomposes images with zooming in the frequency regions centered at selectable modulating frequencies; thus, MWT is an extension of WT and may be used to improve the representation performance by selecting suitable modulating frequencies. Motivated by the amplitude-modulation frequency-modulation (AM-FM) modeling of Bovik et al. the analysis of dominant instantaneous frequency (DIF) seems essential to the selection of suitable modulating frequencies for applying MWT to image representation. For computation simplicity, we had taken the average of local DIFs as the modulating frequency of MWT and the results demonstrated the potential of MWT for image compression in terms of the computed PSNR-entropy values shown as the rate-distortion curves.

三、Modulated Wavelet Transform (MWT)

Let $\psi(x)$ be a valid wavelet and $\phi(x)$ the scaling function. Finite energy signal $f(x)$ can be written by

$$f(x)=\sum_n S_J(n)\phi_{Jn}(x)+\sum_{\ell\le J}\sum_n D_\ell(n)\psi_{\ell n}(x) \quad (1)$$

where $S_J(n)$ and $D_\ell(n)$ are the scaling and wavelet coefficient representing the approximation at the coarsest resolution 2^J and the detail at resolution 2^ℓ respectively. $S_\ell(n)$ and $D_\ell(n)$ can be obtained from $S_{\ell-1}(n)$ by wavelet transform given by

$$\begin{aligned}S_\ell(n)&=\sum_k S_{\ell-1}(k)h(2n-k)\\ D_\ell(n)&=\sum_k S_{\ell-1}(k)g(2n-k)\end{aligned} \quad (2)$$

where $h(n)=<\phi,\phi_{-1,-n}>$, $g(n)=<\psi,\phi_{-1,-n}>$, $<\cdot,\cdot>$ is the inner product; moreover, $S_{\ell-1}(n)$ can be exactly reconstructed from $S_\ell(n)$ and $D_\ell(n)$ by the inverse wavelet transform given by

$$S_{\ell-1}(n)=\sum_k S_\ell(k)\tilde{h}(n-2k)+\sum_k D_\ell(k)\tilde{g}(n-2k) \quad (3)$$

where $\tilde{h}(n)=h(-n)$ and $\tilde{g}(n)=g(-n)$.

For signals of the form $f(x)e^{jUx}$, from equation (1) we have

$$\begin{aligned}f(x)e^{jUx}&=\sum_n(S_J(n)e^{j2^JUn})\phi_{Jn}(x)e^{j2^JU(2^{-J}x-n)}\\ &\quad+\sum_{\ell\le J}\sum_n(D_\ell(n)e^{j2^\ell Un})\psi_{\ell n}(x)e^{j2^\ell U(2^{-\ell}x-n)}\end{aligned} \quad (4)$$

where coefficients $S_\ell(n)e^{j2^\ell Un}$ and $D_\ell(n)e^{j2^\ell Un}$ can be obtained by

$$\begin{aligned}S_\ell(n)e^{j2^\ell Un}&=\sum_k S_{\ell-1}(k)e^{j2^{\ell-1}Uk}h(2n-k)e^{j2^{\ell-1}U(2n-k)}\\ D_\ell(n)e^{j2^\ell Un}&=\sum_k S_{\ell-1}(k)e^{j2^{\ell-1}Uk}g(2n-k)e^{j2^{\ell-1}U(2n-k)}\end{aligned} \quad (5)$$

which is called modulated wavelet transform with modulating frequency U . Its inverse transform can be obtained from equation (3) as follows,

$$\begin{aligned}S_{\ell-1}(n)e^{j2^{\ell-1}Un}&=\sum_k S_\ell(k)e^{j2^\ell Uk}\tilde{h}(n-2k)e^{j2^{\ell-1}U(n-2k)}\\ &\quad+\sum_k D_\ell(k)e^{j2^\ell Uk}\tilde{g}(n-2k)e^{j2^{\ell-1}U(n-2k)}\end{aligned} \quad (6)$$

The 2D extension can be obtained by a tensor product of two 1D MWT.

四、Adaptive MWT-Based Subband Decomposition

The analytic image obtained by adding an imaginary part via the 2-D Hilbert transform [18] is utilized for the compression task. Specifically, the analytic image $t(x,y)$ and the corresponding real valued image $s(x,y)$ are uniquely related by $t(x,y)=s(x,y)+jH[s(x,y)]$, where $H[\cdot]$ denotes the 2-D Hilbert transform acting in the, say, $\vec{e}=[1\,0]^T$ direction; the Fourier transforms of $s(x,y)$ and $H[s(x,y)]$ are related by $F\{H[s(x,y)]\}=-j\,\mathrm{sgn}(\Omega^T\ \vec{e})F\{s(x,y)\}$; the spectrum of $t(x,y)$ is supported only in quadrants I and IV of the frequency plane: $\Omega = [u\ v]^T$; and the spectral redundancy of $s(x,y)$ can be removed. Multicomponent AM-FM modeling represents the analytic image as sums of nonlinear functions, each of the form $f(x,y)\exp[jU(x,y)]$, where $f(x,y)$ and $\nabla U(x,y)$ (i.e. the gradient of $U(x,y)$) are the amplitude and frequency modulating functions, respectively. As a flexible decomposition without excessive side information to describe the resulting structure, the DMWPT-based subband decomposition algorithm is presented below.

Step 1: For a real valued image, remove the DC component, perform the Hilbert transform, and form the analytic image. Specifically, if the Hilbert transform is performed horizontally, subbands with negative horizontal frequencies are all zeros, which can be ignored by down sampling the analytic image by 2 in the horizontal direction for the rest of the compression task; and, the data size of the (critically down sampled) analytic image is equal to that of the original real valued image.

Step 2: Compute energy spectral density (ESD) of the analytic image, find the position of the maximum ESD (in the frequency plane), take it as the modulating frequency, and perform DMWT to generate four subbands.

Step 3: For each subband, compute the respective ESD and decide whether to decompose it further or not. Specifically, let M be the global maximum of the ESD, if there is a local maximum that is greater than αM ($\alpha < 1$) with a distance (from the global maximum) greater than a given threshold β , then the subband needs to be further decomposed into four smaller subbands by DMWT with adapted modulating frequency according to the position of the global maximum M .

Step 4: Repeat Step 3 until there is no subband with more than one significant local maximum in the ESD, or the subband size reaches to the minimum size given a priori.

五、SPIHT Extension

Since the subbands of images obtained by using the adaptive DMWPT are complex valued, the original SPIHT algorithm developed by Said et al. needs to be extended to encode these subbands in the hierarchical modulated wavelet domain. Four symbols are used to identify the status of transform coefficients: IP, NP, SP and ZT, which stand for insignificant pixel, newly significant pixel, significant pixel and zero tree, respectively. Initially, all the modulated scaling and wavelet coefficients at the coarsest resolution are set IP and ZT, respectively. The complex value is represented by the magnitude and angle (i.e. in the polar form). The magnitude of complex valued transform coefficients is used for the comparison with a given sequence of successively smaller threshold values to sort out the significant coefficients in the status check pass. The sequence of threshold values can be obtained by using the recursive equation: $T_k = T_{k-1}/2$, where the initial value T_0 must be greater than or equal to half the maximum magnitude of the transform coefficients.

六、實驗結果

The proposed, modulated wavelet subband coding (MWSC) system with extended SPIHT algorithm has been evaluated on several 512 x 512 gray scale images, including Barbara, fingerprints, and a SAR image. The performance is compared with the wavelet-based SPIHT and the JPEG standard. In MWSC, the critically down sampled analytic image, which is constructed via the use of the 2-D Hilbert transform in the horizontal direction, is decomposed into subbands using the adaptive DMWPT; the adaptability strategy is based on the respective ESD with parameters: $\alpha = 0.5$, $\beta = 1$ radian, and the minimum size of subbands is: 64 x 32; the coding sequence of the decomposed subbands is in a zigzag order: from low-to-high frequency subbands; the angle quantity of the complex valued transform coefficients represented in the modulated wavelet trees is uniformly quantized with 6-bit resolution; the number of bits representing the initial angle information of the newly significant pixels is 3. The number of decomposition levels in both wavelet and modulated wavelet transforms is 4. Daubechies orthogonal wavelet D2 is used. The sequence of successively smaller threshold values in the original SPIHT as well as its extension is obtained by $T_k = 0.5T_{k-1}$; $k = 1,2,\ldots$, with the initial T_0 equal to half the maximum amplitude of the transform coefficients. The compression distortion is measured by the peak signal to noise ratio (PSNR) in dB. The compression rates measured in bits per pixel (bpp) and PSNR values are plotted as the rate distortion curve for performance comparison.

Fig.1 shows the comparison on natural Barbara image at different bit rates. Part of the decoded images from the JPEG standard, wavelet-based SPIHT, and MWSC at 0.2 bpp are shown in Fig. 1(a)-(c), respectively. By comparing visually, MWSC improves the reconstruction result on the textured regions with significant stripes in specific directions. Their respective rate distortion curves shown in Fig. 1(d) demonstrate that MWSC is preferable for images with large portions of textures.

The compression of fingerprints image is one of the most important issues, which demands the best solution. Without any compression, the storage of digitized fingerprints of a person may be in the order of mega bytes. Fig. 2 shows the comparison on a fingerprints image. By comparing the decoded images shown in Fig. 2(a)-(c), and the rate distortion curves shown in Fig. 2(d), wavelet-based SPIHT outperforms JPEG at low to moderate bit rates (< 1 bpp); and MWSC is the superior for this kind of images.

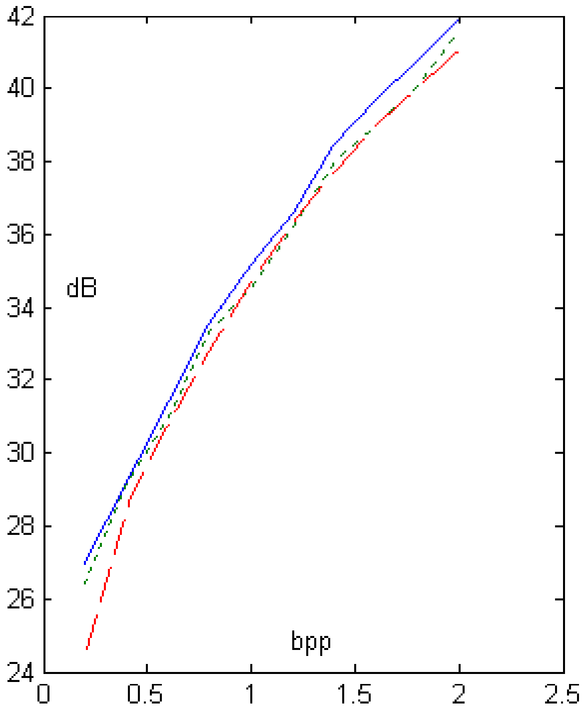
Finally, the comparison on a SAR image containing large portions of irregular textures is presented in Fig. 6. Part of the original and decoded images is shown in Fig. 3(a)-(d). The rate distortion curves shown in Fig. 3(e) demonstrate that wavelet-based SPIHT is inferior to JPEG at low to moderate bit rates, and vice versa at moderate to high bit rates; however, MWSC is still the superior in terms of the rate distortion curves as well as visual comparison.

The residual correlation can be exploited further by using arithmetic coding. In our experiments, the use of arithmetic coding improves the performance with a gain of about 0.2–0.4 dB over the non-arithmetic coded versions of both the wavelet-based SPIHT and MWSC; nevertheless, MWSC still outperforms the wavelet-based SPIHT with similar rate distortion curve improvements like Fig. 1(d), 2(d) and 3(e).

(a)

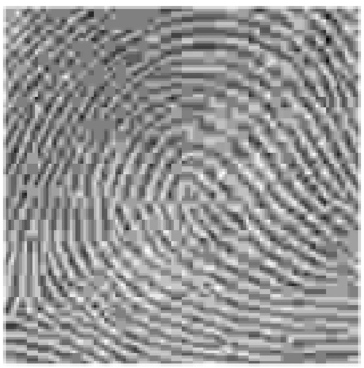
(b)

(c)

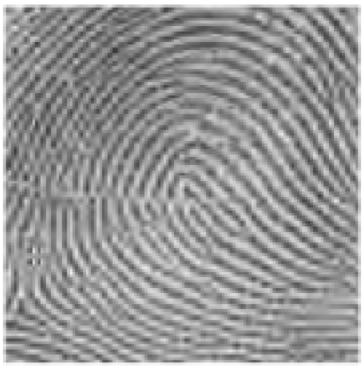


(d)

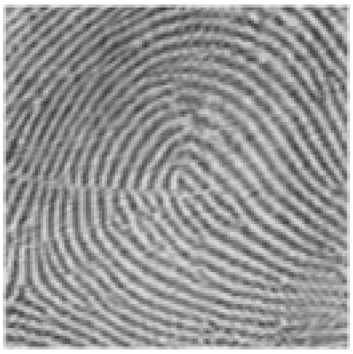
Fig. 1: (a) ~ (c): Part of the decoded Barbara image at 0.2 bpp using the JPEG standard (PSNR=24.6 dB), the wavelet based SPIHT without arithmetic coding (26.4 dB), and the proposed MWSC (26.9 dB), respectively; (d): the corresponding rate distortion curves (dashed: JPEG, dotted: wavelet based SPIHT, solid: the proposed method).



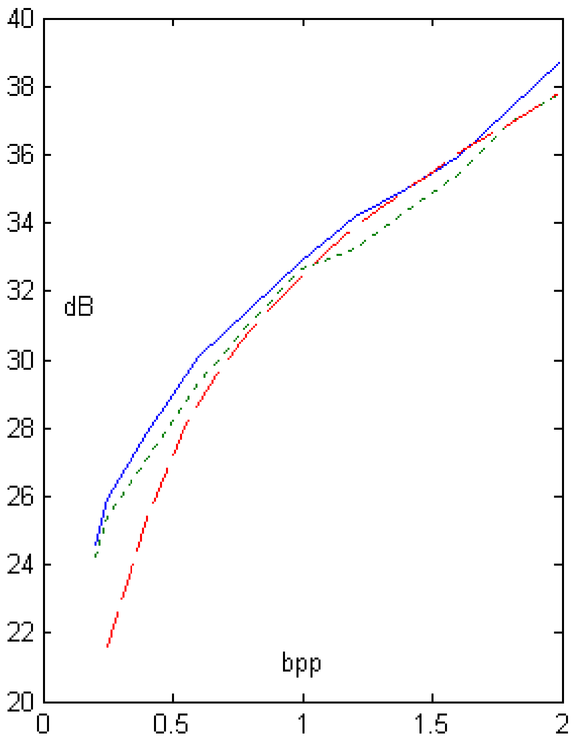
(a)



(b)

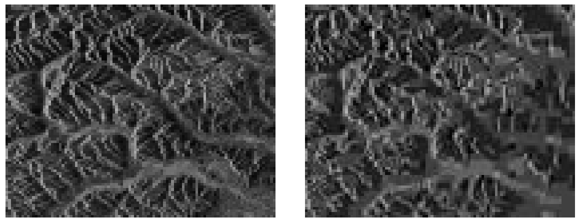


(c)

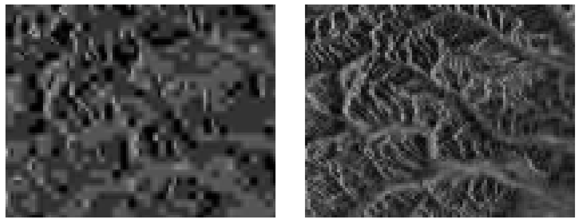


(d)

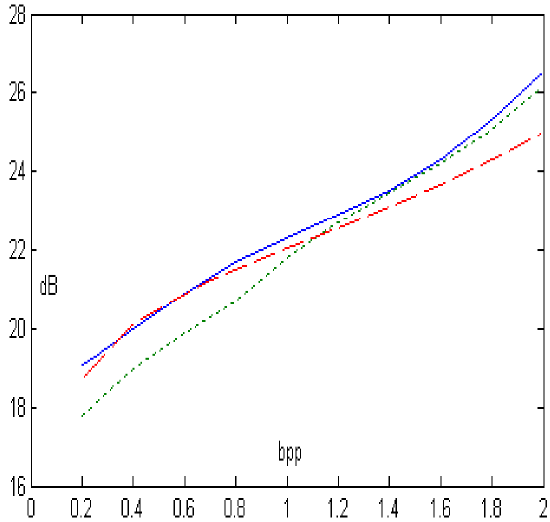
Fig. 2: (a) ~ (c): Part of the decoded fingerprints image at 0.25 bpp using the JPEG standard (PSNR=21.6 dB), the wavelet based SPIHT without arithmetic coding (25.4 dB), and the proposed MWSC (25.9 dB), respectively; (d): the corresponding rate distortion curves (dashed: JPEG, dotted: wavelet based SPIHT, solid: the proposed method).



(a) (b)



(c) (d)



(e)

Fig. 3: (a) ~ (d): Part of the original SAR image, decoded image at 0.2 bpp using the JPEG standard (PSNR=18.7 dB), the wavelet based SPIHT without arithmetic coding (17.8 dB), and the proposed MWSC (19.1 dB), respectively; (e): the corresponding rate distortion curves (dashed: JPEG, dotted: wavelet based SPIHT, solid: the proposed method).

六、計畫成果

An adaptive subband image coding system based on the modulated wavelet packet and modulated wavelet transforms is presented. It consists of three stages: adaptive subband decomposition, adaptive modulated wavelet transform, and embedded coding in the modulated wavelet trees. In the first stage, the analytic image is decomposed into subbands via the adaptive modulated wavelet transform in a resolution recursive manner, which leads to the adaptive modulated wavelet packet transform with a top-down quadtree structure; In the second stage, each of the decomposed subbands containing one significant AM-FM component is represented in the modulated wavelet tree; the associated modulating frequencies in the first two stages are adapted based on the respective energy spectral densities; In the last stage, a simplified version of the extended SPIHT algorithm is proposed to encode the transform coefficients. Experimental results demonstrate that the modulated wavelet based subband coding system is preferable to both the wavelet based and the JPEG standard for images with significant energies in the middle-high frequency regions, in terms of the rate distortion curves and visual comparisons. Moreover, it is a highly parallel processing, which is a substantial advantage for the hardware implementation; and there is no the so-called blocking effects that are usually to be found on the decoded JPEG images with compression at low bit rates.

This project result had been submitted to and accepted by the **Pattern Recognition Letters.**

七、參考文獻

[1] K. R. Rao and J. J. Hwang, Techniques and Standards for Image, Video and Audio Coding, Prentice Hall, 1996.

[2] H.G. Musmann, P. Pirsch, and H. J. Grallert, " Advances in Picture Coding," Proc. IEEE, vol. 73, pp. 523-548, April, 1985.

[3] R. J. Clarke, Transform Coding of Images, New York: Academic, 1985.

[4] O. J. Kwon and Rama Chellappa, "Region Adaptive Subband Image Coding," IEEE Trans. Image Processing, vol. 7, no. 5, pp. 632-648, May, 1988.

[5] W. B. Pennebaker and J. L. Mitchell, JPEG Still Image Data Compression Standards, New York: Van Nostrand, 1993.

[6] Sonja Grgic, Mislav Grgic and Branka Zovko-Cihlar, "Performance Analysis of Image Compression Using Wavelets," IEEE Trans. Industrial Electronics, vol. 48, no. 3, pp. 682-695, 2001.

[7] S. G. Mallat, "A Theory for Multiresolution Signal Decompositions: The Wavelet Representation," IEEE Trans. PAMI, 11, pp. 674-693, 1989.

[8] J. M. Shapiro, "Embedded Image Coding Using Zero-Trees of Wavelet Coefficients," IEEE Trans. Signal Processing, vol. 40, pp. 3445-3462, 1993.

[9] A. Said and W. A. Pearlman, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," IEEE Trans. Circuits Syst. Video Tech. vol. 6, pp. 243-250, 1996.

[10] S. D. Servetto, K. Ramchandran, M. T. Orchard, "Image Coding Based on A Morphological Representation of Wavelet Data," IEEE Trans. Image Processing, vol. 8, pp. 1161-1174, 1999.

[11] E. S. Hong, R. E. Ladner, "Group Testing for Image Compression," IEEE Trans. Image Processing, vol. 11, pp. 901-911, 2002.

[12] Bryan E. Usevitch, "A Tutorial on Modern Lossy Wavelet Image Compression: Foundations of JPEG 2000," IEEE Signal Processing Magazine, pp. 22-35, Sep. 2001.

[13] F. G. Meyer, A. Z. Averbuch, and J.-O. Stromberg," Fast Adaptive Wavelet Packet Image Compression," IEEE Trans. Image Processing, vol. 9, pp. 792-800, 2000.

[14] D. Engle, A. Uhl, "Adaptive Object Based Image Compression Using Wavelet Packets," VIPromCom-2002, 4th EURASIP, IEEE Region 8 International Symposium on Video/Image Processing and Multimedia Communications, pp. 183-187, 2002.

[15] Hsi-Chin Hsin, "Texture Segmentation Using Modulated Wavelet Transform," IEEE Trans. Image Processing, vol. 9, pp. 1299-1302, 2000.

[16] Hsi-Chin Hsin and Ching-Chung Li, "Image Coding with Modulated Wavelets," Pattern Recognition Letters, vol. 24, pp. 2391-2396, 2003.

[17] J.P. Havlicek, D. Harding and A.C. Bovik, "The Multi-component AM-FM Image Representation," IEEE Trans. Image Processing, vol. 5, 1094-1100,1996.

[18] J. P. Havlicek, J. W. Havlicek and A. C. Bovik, "The Analytic Image," Proc. IEEE, pp. 446-449, 1997.