行政院國家科學委員會專題研究計畫　成果報告

適應性調變小波子頻道影像壓縮編碼研究

計畫類別：個別型計畫
計畫編號：NSC93-2213-E-216-007-
執行期間：93 年 08 月 01 日至 94 年 07 月 31 日
執行單位：中華大學電機工程學系

計畫主持人：辛錫進

計畫參與人員：蘇子傑 張皓清 梁章桓

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 94 年 9 月 12 日

行政院國家科學委員會專題研究計畫成果報告

計畫名稱：適應性調變小波子頻道影像壓縮編碼研究
計畫編號：NSC 93-2213-E-216-007
執行期限： 93 年 8 月 1 日至 94 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
計畫參與人員： 蘇子傑 張皓清 梁章桓

中文摘要

小波轉換只針對低頻訊號作連續分解；而我們所提出的調變小波轉換可以對任何頻帶的影像成份作連續分解，配合適應性調變頻率參數得以提高影像描述效率。本研究計畫是根據訊號之主瞬間頻率，應用調變小波封包轉換將影像分解成數個大小不同的子頻帶成份，再以調變小波轉換分析每個子頻帶影像，提高影像壓縮的品質。

關鍵詞：影像壓縮編碼；調變小波(封包)轉換；SPIHT 延伸；適應性調變頻率參數

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. An adaptive subband decomposition using the modulated wavelet packet transform has been developed, where the modulating frequencies adaptation is based on the energy spectral density in a scale recursive manner. In addition, the extended SPIHT coder has also been developed to code each complex-valued subband. Experimental results demonstrate that the modulated wavelet subband image coding is much preferable to both the wavelet coding and the JPEG standard.

Keywords: Image compression; Modulated wavelet (packet) transform; SPIHT extension; Adaptive modulating frequencies

I. 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.

II. Modulated Wavelet (Packet) Transform

Wavelet theory provides an efficient multiresolution analysis. The 1-D discrete wavelet transform (DWT) of signal $S_\ell(n)$ at resolution 2^ℓ is given by

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

where $S_{\ell+1}(n)$ is its approximation at the next coarser resolution $2^{\ell+1}$, $D_{\ell+1}(n)$ is the detail information between resolutions: 2^ℓ and $2^{\ell+1}$, ℓ is the resolution index with larger meaning coarser, $h(n) = <\phi, \phi_{-1,-n}>$, $g(n) = <\psi, \phi_{-1,-n}>$, $<\cdot,\cdot>$ is an inner product operator, ψ is a (mother) wavelet, ϕ is the corresponding scaling function, and $\phi_{-1,-n}(t) = 2^{-1/2}\phi(2^{-1}t - n)$. By using the inverse DWT (IDWT), which is defined below, $S_\ell(n)$ can be reconstructed form $S_{\ell+1}(n)$ and $D_{\ell+1}(n)$.

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

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

As a large class of signals with quasi-periodic patterns can be well represented by using the non-linear amplitude modulating -frequency modulating (AM-FM) functions of the form $S_\ell(n)\, e^{j2^\ell U n}$ [17], the following discrete modulated wavelet transform (DMWT) had been proposed in [15].

$$\begin{aligned} S_{\ell+1}(n) e^{j2^{\ell+1}Un} &= \sum_k S_\ell(k) e^{j2^\ell Uk}\, h_\ell(2n-k) \\ D_{\ell+1}(n) e^{j2^{\ell+1}Un} &= \sum_k S_\ell(k) e^{j2^\ell Uk} g_\ell(2n-k) \end{aligned} \tag{3}$$

where $h_\ell(n) = h(n) e^{j2^\ell Un}$, $g_\ell(n) = g(n) e^{j2^\ell Un}$ and U is the modulating frequency. $S_\ell(n)\, e^{j2^\ell Un}$ can be exactly reconstructed from its approximation $S_{\ell+1}(n)\, e^{j2^{\ell+1}Un}$ at the next coarser resolution $2^{\ell+1}$ and the detail $D_{\ell+1}(n)\, e^{j2^{\ell+1}Un}$ between resolutions: 2^ℓ and $2^{\ell+1}$ by using the inverse DMWT (IDMWT):

$$S_\ell(n) e^{j2^\ell Un} = \sum_k S_{\ell+1}(k) e^{j2^{\ell+1}Uk} \tilde{h}_{\ell+1}(n-2k) + \sum_k D_{\ell+1}(k) e^{j2^{\ell+1}Uk} \tilde{g}_{\ell+1}(n-2k) \tag{4}$$

where $\tilde{h}_{\ell+1}(n) = h_{\ell+1}(-n)$ and $\tilde{g}_{\ell+1}(n) = g_{\ell+1}(-n)$.

Both DMWT and DWT provide the perfect reconstruction property, which is a requirement for image synthesis. DWT analyzes signals with a successive zooming in the low frequency region only. However, DMWT provides a successive zooming in the frequency region centered at selectable modulating frequencies. Thus, DMWT acting as an extension of DWT may improve the analysis and synthesis tasks by selecting suitable modulating frequencies.

As the discrete wavelet packet transform (DWPT) obtained by applying DWT to both the low and high frequency components, the discrete modulated wavelet packet transform (DMWPT) can be obtained in a similar way. It is noted that the widely used, efficient, pyramid- (tree-) structured algorithms for implementing DWT (DWPT) are also applicable to DMWT (DMWPT) provided that the associated filters are modulated accordingly. The 2-D DMWT (DMWPT) can be obtained by using the tensor product of tow 1-D DMWT (DMWPT) with a pair of modulating frequencies; one is in the horizontal direction and the other is in the vertical direction.

III. Modulated Wavelet Subband Coding

Wavelet transform is focused on the low frequency decomposition only, which may not be suitable for natural images with a large portion of textures consisting of middle/high frequency components. Modulated wavelet transform, which is derived from a combination of AM-FM modeling and wavelet transform together with all the advantages of both techniques is suitable for the aforesaid images. Motivated by the AM-FM

representation [17], in which a filterbank with frequency and orientation selectivity had been utilized, an adaptive modulated wavelet subband image coder is proposed.

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.

The above decomposition scheme leads to an adaptive DMWPT with a quadtree structure. The side information describing the resulting decomposition structure can be encoded by scanning the tree-node symbols from left to right, top to bottom. Two symbols: 1 and 0 are used, where symbol 1 means the corresponding subband is to be decomposed further and symbol 0 means no need for further decomposition. As an example shown in Fig. 1, where the analytic image is decomposed into subbands: a ~ m, the corresponding side information by raster scanning the node symbols is '101010000010000 00'. It is noted that the nodes on the bottom level called leaves are not decomposed further, thus the symbols of leaves that are all zeros can be omitted.

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. Fig. 2 shows a 3-level DMWT with indexing for the transform coefficients, based on which the transform coefficients are encoded in a specific order. 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.

After an image is decomposed into

subbands of different sizes by using the adaptive DMWPT, each subband is encoded uniquely by using the extended SPIHT algorithm with a global initial threshold value T_0 in the hierarchical modulated wavelet trees. The resulting bit stream could be formulated as shown in Fig. 3, where the header portion is composed of the side information about the adaptive DMWPT based subband decomposition structure, the number of DMWT decomposition levels, the associated modulating frequencies, and the initial threshold: T_0.

IV. Experimental Results

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. 4 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. 4(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. 4(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. 5 shows the comparison on a fingerprints image. By comparing the decoded images shown in Fig. 5(a)-(c), and the rate distortion curves shown in Fig. 5(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. 6(a)-(d). The rate distortion curves shown in Fig. 6(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. 4(d), 5(d) and 6(e).

V. 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.

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.

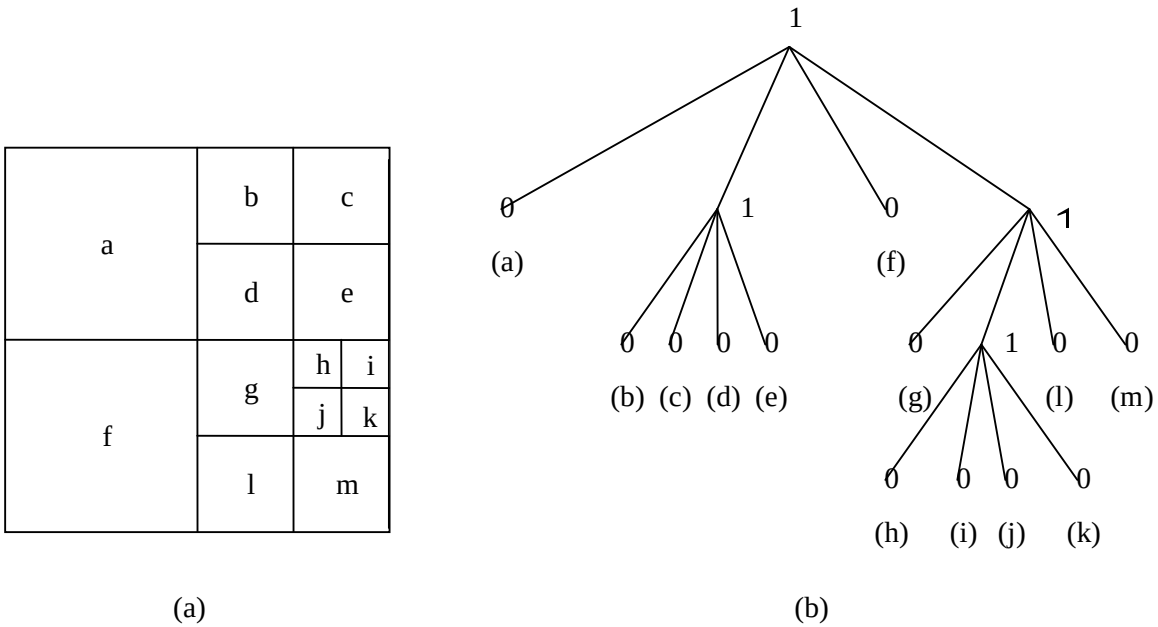
[14] D. Engle, A. Uhl, "Adaptive Object Based Image Compression Using Wavelet Packets," VIPromCom-2002, 4th EURASIP, IEEE Region 8 International Symposium on Video/Image Processing and Multimedia Communications, pp. 183-187, 2002.

[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.

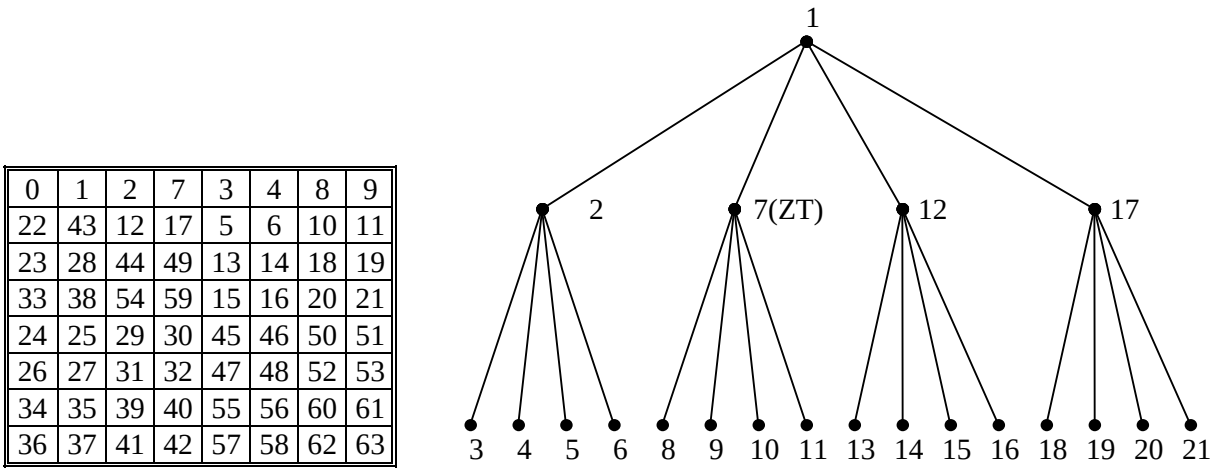
[17] J.P. Havlicek, D. Harding and A.C. Bovik, "The Multi-component AM-FM Image Representation," IEEE Trans. Image Processing, vol. 5, 1094-1100,1996.

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



(a) (b)

Fig. 1. Example of the adaptive DMWPT subband decomposition; (a) subbands: a ~ m; (b) the representative quad-tree with node symbols.



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

(a) (b)

Fig. 2. (a): 3-level DMWT with indexing for the transform coefficients; (b): one of the hierarchical, modulated wavelet trees with the corresponding indices.

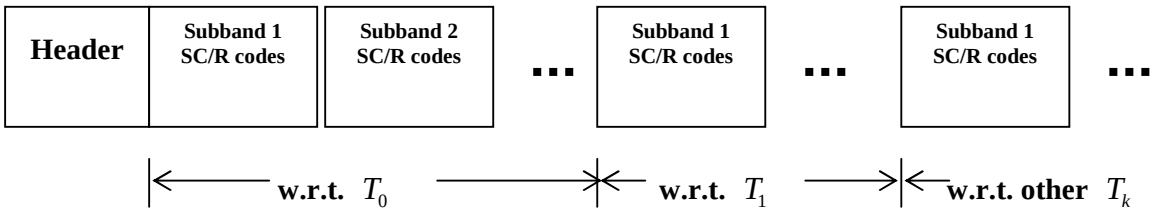
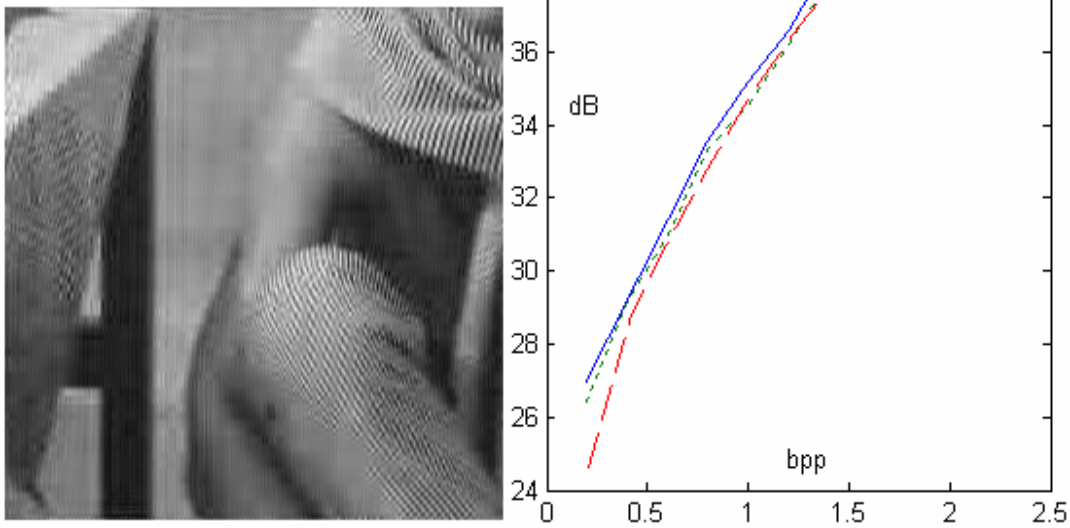


Fig. 3. Bit stream structure (SC/R: status check/refinement).

(a) (b)



(c) (d)

Fig. 4. (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).

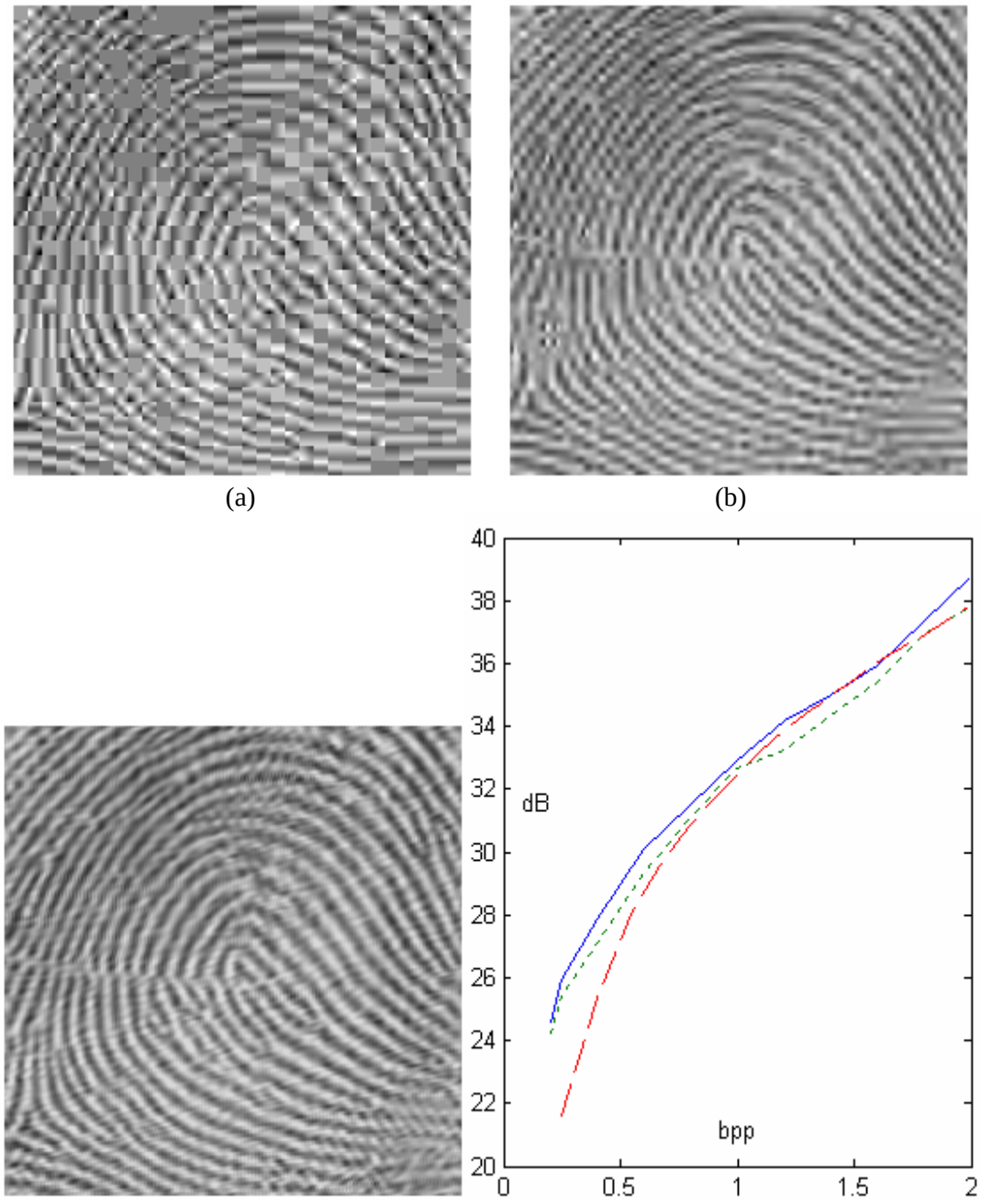
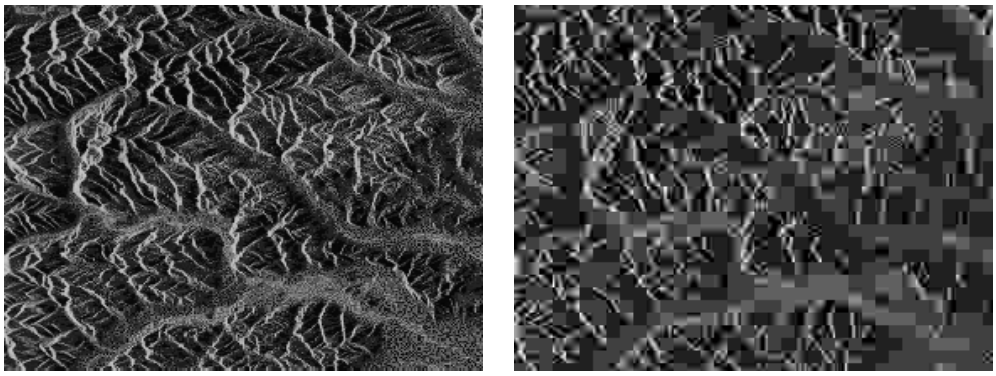
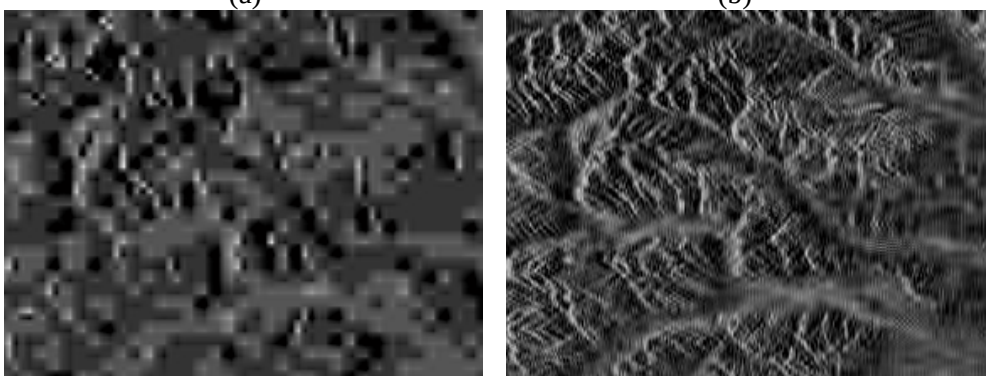


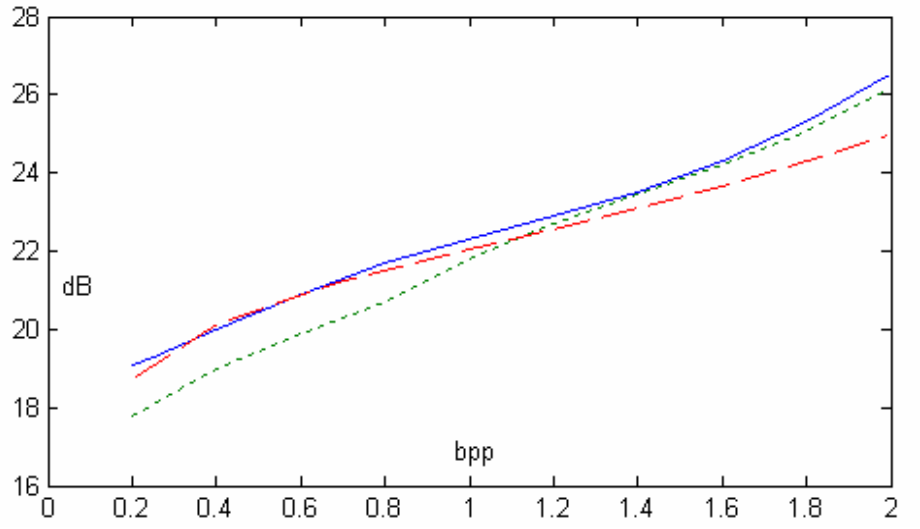
Fig. 5. (a) ~ (c): Part of the decoded fingerprints image at 0.25 bpp using the JPEG standard (PSNR=21.6 dB), the wavelet based SPIHT without arithmetic coding (25.4 dB), and the proposed MWSC (25.9 dB), respectively; (d): the corresponding rate distortion curves (dashed: JPEG, dotted: wavelet based SPIHT, solid: the proposed method).



(a) (b)



(c) (d)



(e)

Fig. 6. (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).