ITU - Telecommunications Standardization Sector Document Q15-F-22
STUDY GROUP 16 Filename: q15f22.doc
Video Coding Experts Group (Question 15) Generated: 28 Nov. '98

Sixth Meeting: Seoul, Korea, 3-6 November, 1998

Question : Q.15/SG16

Source: LG Electronics
MPC Gr. Media Comm. Lab,
Woo-Myeon-Dong, Seocho-Gu, 137-724, Seoul, Korea

Title: A DCT-based Embedded Video Coder Using Rearrangement of DCT Coefficients

Author : Yeong-An Jeong, Chang-MO Yon, and Young-Man Park
Contact : Yeong-An Jeong
Tel: +82 2 526 7331
Fax: +82 2 571 8781
Email: jya@mail.lgcit.com

Purpose: Proposal

ABSTRACT

In this document, we present an efficient DCT-based embedded video coder. Especially, it is designed for I (intra)-frame and motion compensated residual(e.g. DFD) coding. Motivated by the wavelet structure of DCT, the proposed coder rearranges the DCT transformed image into 2-level wavelet pyramid structure and makes use of embedded zerotree coding. The proposed rearranged structure is well coupled with an embedded zerotree coding which has been mainly used in wavelet-based image coding. Resulting DCT-based embedded video coder shows the promising rate-distortion performance than conventional H.263+. Embedded feature of the proposed coder provides a precise rate control and scalable functionalities. Also, when restricted to I-frame coding, the performance of the proposed new coder is comparable to that of state-of-art still image compression methods such as JPEG and EZW-style coders.

1. INTRODUCTION

Recently, very low bit rate video coding has received considerable attention for the demands of visual communication applications through PSTN (Public Switched Telephone Network) in terms of both coding algorithms and standards activities. The ITU-T SG16/Q15 'Advanced Video Coding' group has been very successful in the development of the H.263 and H.263+ standard. However, Q.15/16 recognizes that there is still a potential need for a new video coding algorithm, currently named H.26L, capable of delivering much improved video quality and covering functionalities not supported by existing standards(e.g. H.261, H.262, H.263, H.263+, MPEG-4)[3][6][17]. The several enhanced functionality areas which should be supported by H.26L are currently identified as follows: 1) enhanced error robustness, including the needs of 3rd generation mobile networks 2) adaptable rate control mechanisms 3) applicability to various kind of source materials, such as fast motion, multiple scene changes, pans and tilts etc[18].

H.263 adopts the motion-compensated hybrid DPCM/DCT coding with considerable improvements to be suitable for very low bit rate transmission. The hybrid motion-compensated DCT structure is a part of all the existing video coding standards and still image coding standard (JPEG). This popularity can be attributed to the fact that the DCT performs well in a wide variety of coding situations and a substantial investment has been made in developing chipsets to implement DCT-based coders.

Generally, in very low bit rate video coding, the number of bits used for I-frame coding usually consumes a large bits in comparison with P(predictive)-frames and B(bidirectionally predicted)-frames. Furthermore, the coding efficiency of the next and subsequent frames such as P-frames and B-frames is strongly dependent on that of I-frame. Therefore, an efficient I-frame coding method is required for the improvement of overall coding performance.

Another major issue with video coding is rate control problem. In current H.263, it is difficult to estimate the number of coded bits generated for a given QP (Quantization Parameter). Hence, in order to adjust coded bits at the chosen target bit rates, iterative method is required and precise rate control is very difficult.

On the other hand, since a concept known as EZW (Embedded Zerotree Wavelet) was lately introduced by Shapiro[6], many wavelet-based image coders based on EZW have been proposed[1][2][6][10]. These coders commonly encode the position and magnitude information in the order of significance by using the spatial relationship of wavelet coefficients, called a zerotree. The merits of these techniques are that precise rate control by simply truncating the bit stream, scalability, and error robustness, and the better rate-distortion performance can be obtained than that of DCT-based coders.

Wavelet-based embedded image coders show an excellent rate-distortion performance in the area of still image coding using high resolution images (e.g. 512x512). But, in very low bit rate video coding using low resolution images (e.g. 176x144), the embedded zerotree coding performance is reduced due to the insufficient wavelet decomposition. Moreover, those methods are not compatible with most of the very low bit rate video compression standards based on DCT.

In traditional video compression methods, the algorithm for compressing DFD is the same as that for compressing the still image frames. They do not take into account that the characteristics of the two types of images are different. That is, DFD often contains more textured regions having significant middle and high frequencies than a still image containing the smooth region. As a result, it does not provided a significant coding gain anymore. Coding techniques different from traditional coding may provide higher coding gains.

In the point of both the compatibility with conventional video coding standards, the improved coding performance, and enhanced functionalities, we propose an efficient DCT-based embedded video coder designed for I-frame and DFD coding. The traditional still image coding method (Intra or DFD) is replaced with a new coding method. The proposed coder rearranges the DCT transformed image into 2-level wavelet pyramid structure and makes use of improved embedded zerotree coding. The proposed coder reflects the more mid- and high frequency. I expect it to be better able to represent the DFD using fewer coefficients as well as I-frame. These assumptions allow the proposed coder to avoid the artifacts most often produced by low bit rate DCT coding method.

This document is organized as follows. In Section 2, we introduce the concept of embedded zerotree coding. Section 3 explains the wavelet structure of DCT. The proposed DCT-based embedded video coder is described in Section 4 and 5. In Section 6, the simulation results are presented to evaluate the proposed coder. Finally, the conclusion of our proposal is given in Section 7.

2. EMBEDDED ZEROTREE CODING

In this Section, we give a brief description of the EZW. Reference [6] describes the algorithm in further detail.

EZW coding is a proven technique for coding wavelet transform coefficients. Besides superior compression performance, the advantages of EZW coding include simplicity, an embedded bitstream, scalability, error robustness, and precise rate control. These features enable EZW coding to address the H.26L or H.263++ functionalities.

EZW coding is based on three key ideas: 1) exploiting the self-similarity inherent in the wavelet transform to predict the location of significant information across scales; 2) successive approximation quantization of the wavelet coefficients; and 3) universal lossless data compression using adaptive arithmetic coding.

EZW coding is applied to coefficients resulting from a DWT (Discrete Wavelet Transform) having a hierarchical subband decomposition. The DWT decomposes the input image into a set of subbands of varying resolutions. The coarsest subband is a lowpass approximation of the original image, and the other subbands are finer-scale refinements. In the hierarchical subband system such as that of the wavelet transform, with the exception of the highest frequency subbands, every coefficient at a given scale can be related to a set of coefficients of similar orientation at the next finer scale. The coefficient at the coarse scale is called the parent, and all coefficients at the same spatial location and of similar orientation at the next finer scale are called children. For the lowest frequency subband, in the example, the parent-children relationship is defined such that each parent node has three children, one in each subband at the same scale and spatial location but different orientation.

EZW introduced a data structure called a zerotree, built on the parent-children relationship. The zerotree structure takes advantage of the principle that if a wavelet coefficient at a coarse scale is insignificant (quantized to zero) with respect to a given threshold *T*, then all wavelet coefficients of the same orientation at the same spatial location at finer wavelet scales are also likely to be insignificant with respect to that *T*.

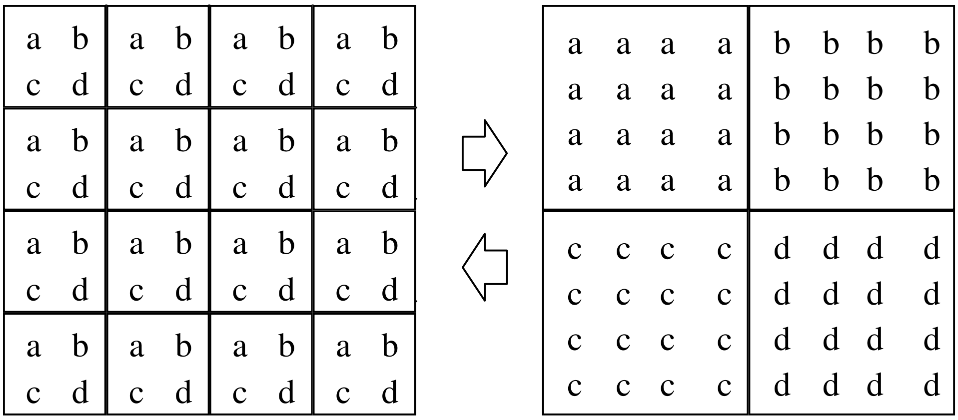
The zerotree structure is similar to the zigzag scanning and end-of- block (EOB) symbol in JPEG (or “LAST’ in H.263) commonly used in coding DCT coefficients. EZW scans wavelet coefficients subband by subband. Parents are scanned before any of their children, but only after all neighboring parents have been scanned. Each coefficient is compared against the current threshold *T*.

A coefficient is significant if its amplitude is greater than *T*; such a coefficient is then encoded using one of the symbols negative significant or positive significant. The zerotree root symbol is used to signify a coefficient below *T*, with all its children in the zerotree data structure also below *T*. The isolated zero symbol signifies a coefficient below *T*, but with at least one child not below *T*.

For significant coefficients, EZW further encodes coefficient values using a successive approximation quantization (SAQ) scheme. Coding is done bit-plane by bit-plane. The successive approximation approach to quantization of the wavelet coefficients leads to the embedded nature of an EZW coded bitstream.

3. WAVELET STRUCTURE OF DCT

In EZW, an image is decomposed into a set of band pass filtered subimages by using DWT. Intuitively, this process seems to be related to the basis function decomposition that occurs in DCT. Both approaches provide a localized frequency representation of the original image. The difference between the two approaches lies primarily in the manner where the output data are organized. Figure 1 illustrates the relationship between wavelet and DCT.



(a) DCT Transformed Image (b)Wavelet Transformed Image

Figure 1. The relationship between wavelet and DCT

Consider a DCT that is applied to 2x2 non-overlapping blocks of an image. Referring to Figure 1(a), each DCT transformed blocks contain four DCT coefficients, labeled a (the lowest frequency) to d (the highest frequency). Also, now consider decomposing the same image by using 2x2 analysis filters and decimating by a factor of two in each direction. It can be shown that the DCT representation is entirely equivalent to a wavelet representation[3][13][14].

As shown in Figure 1(b), the highest subband image merely corresponds to a grouping of the lowest frequency coefficients in each DCT transformed block. The other DCT coefficients are similarly grouped to form the remaining subband image. Generally, although the subband localization is slightly poor, a DCT on NxN blocks of an image can be viewed as a wavelet decomposition with an uniform NxN-subband decomposition.

As an intriguing experiment, Xiong et al.[8] showed that a DCT can also be coupled with an embedded zerotree quantizer. For that purpose, they treated each 8x8 DCT block as a 3-level wavelet decomposition and applied the EZW to each DCT block. But that method requires heavy overhead information for each block and cannot take advantage of the spatial correlation of DC. Therefore, rate-distortion performance is not so much good as EZW[6].

For solving those problems, this paper describes more improved zerotree structure which can sufficiently take advantage of the spatial correlation of DC between neighboring DCT blocks and more reflect the middle and high frequency for low bit rate coding.

4. GENERAL OVERVIEW OF THE PROPOSED CODER

This section gives an overview of the proposed encoder. The proposed DCT-based embedded video coder is shown in Figure 2.

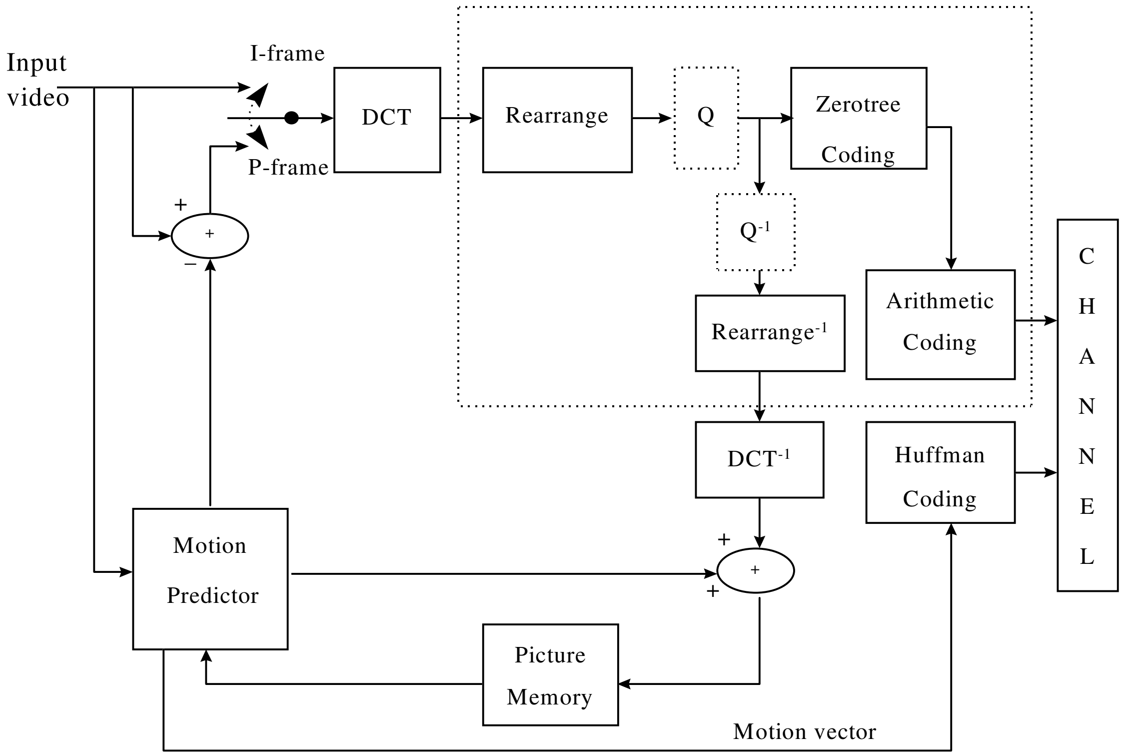


Figure 2. The proposed DCT-based embedded video coder

The distinctive feature of the proposed coder is quantization method as seen in dashed line upper Figure 2. we adopt an embedded zerotree coding after rearrangement process to quantize the DCT coefficients. Frames of video are encoded either as I-frame or P-frame, where the former is used for the first frame and the latter is used for the remaining frames of the video sequences.

First input image (I-frame) is decomposed into non-overlapping 8x8 blocks, which are then transformed with DCT. The DCT coefficients of each block are rearranged into 2-level wavelet pyramid structure. The rearranged DCT coefficients are quantized by using embedded zerotree coding and the results are coded by using arithmetic coding.

All succeeding frames of the video sequence are encoded from the preceding frame as forward-predicted P-frames as used in H.263+. Therefore, the proposed coder can use the block motion estimation, ANNEX D (Unrestricted Motion Vector mode) or ANNEX F (Advanced Prediction Mode), optionally. Next, each block is predicted using the block motion compensation scheme of H.263+. Motion compensated residuals are transformed by DCT and rearranged for subsequent processing as was done to I-frame.

In a H.263+ video coder, for those blocks where prediction fails (such as fast motion, multiple scene changes, pans and tilts etc), the INTRA mode is selected and the original image block is coded. That information is sent as overhead. Since we wish to avoid heavy overhead, we follow a different approach for coding INTRA blocks. After DCT of INTRA blocks, the average DC value is calculated over each frame. DC of INTRA blocks is subtracted by the average DC after rearranging process. This technique reduces the unnecessary scanning for zerotree coding from high threshold to low threshold considering differenence natural image between DFD image.

After all, INTRA blocks and INTER blocks(e.g. motion compensated residual) is processed simultaneously for embedded zerotree coding.

We have not yet implemented a rate control scheme for the proposed video coder. However, since my coder is embedded coder, I can easily adjust coding bit rates to be the same as that of H.263+ coder for each frame.

5. TECHNICAL DETAILS OF THE PROPOSED CODER

5.1 Rearrangement of the DCT coefficients into 2-level wavelet pyramid structure

This technique makes it possible for DCT to be well coupled with a embedded zerotree coding as was mainly used in wavelet image coding. As illustrated in Figure 3, DCT coefficients of each 8x8 DCT block are rearranged into 2-level wavelet pyramid structure.

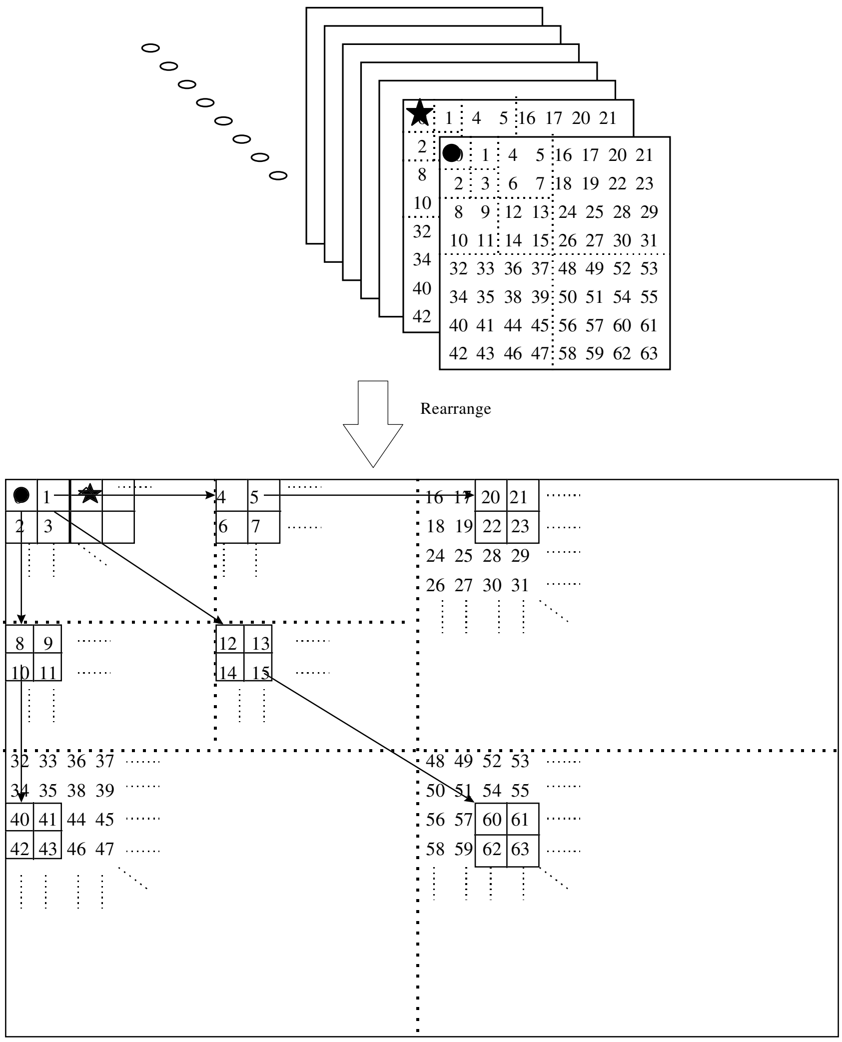


Figure 3. The rearrangement of DCT coefficients into 2-level wavelet pyramid structure

In the proposed structure, the parent-children relationship (zerotree) is defined as follows. A parent's direct children correspond to the pixels of the same spatial orientation in the next finer level of the pyramid. The tree is defined in such a way that each node has either no children or four children which always form a group of 2x2 adjacent pixels. In Figure 3, the arrows are oriented from the parent node to its four children. The special pixels (indicated by the circle and star etc.) in the highest level of the pyramid are the tree roots and have no children.

The difference between the proposed 2-level wavelet pyramid structure by rearrangement of DCT coefficients and that by wavelet transform is frequency spectrum feature by uniform and hierarchical decomposition method, respectively. By iterating on the lowest frequencies, the hierarchical wavelet decomposition does not capture the middle and high frequency range well[13].

While, by rearranging the coefficients of the DCT which splits the frequency spectrum into uniformly spaced bands, the proposed structure captures the middle and high frequencies better than hierarchical wavelet decomposition instead of slight loss of low frequencies. This feature is more profitable than hierarchical wavelet decomposition in very low bit rate video coding. Generally, high frequency coefficients are not coded because of limited channel bandwidth. Hence, the proposed structure can improve the coding efficiency by reflecting the middle and high frequencies.

In case of I-frame coding, the average of DC is subtracted from each DC components in the coarsest(highest) subband and transmitted as an overhead per frame. Because DC coefficients of neighboring blocks are highly correlated, this technique reduces the unnecessary scanning for zerotree coding from high threshold to low threshold.

5.2 Embedded zerotree coding

An embedded zerotree coding is applied to quantize the rearranged DCT coefficients as was done to the wavelet coefficients in [2]. SPIHT[2] slightly modifies and enhances some of the ideas underlying EZW. The difference is that SPIHT exploits intraband dependancy by encoding the activity masks while the EZW exploits interband dependancy by encoding the zerotrees.

This schemes adopt a successive quantization procedure which is achieved not in one step but in several successive steps. The DCT coefficient is only quantized up to the precision of the initial threshold T. Then the quantization with T is refined with a smaller threshold T/2, T/4, T/8, and so on, successively. Further detail is described as follows.

SPIHT algorithm maintains lists of insignificant sets, insignificant pixels, and significant pixels, and is initialized with the list of insignificant pixels being the nodes in the highest layer and the list of insignificant sets being the sets of the subtree descendantss of each such node.

In a “sorting” pass, the algorithm works its way down the list of insignificant pixels first, testing their magnitude against the current threshold, outputting their significance, and, when one is significant, outputting its sign and moving it to the list of significant pixels. Then, it moves through the list of insignificant sets, performing the magnitude test to all the coefficients in the current set (which initially are full subtrees of descendants. When all the coefficients in a set are insignificant, that fact can be indicated by outputting only one bit. When a given set contains a significant pixel, it is partitioned into subsets (subtrees rooted at the descendants of the current root) and those in turn are tested for significance. As before, when significant pixels are identified, their sign is transmitted and they are moved to the list of significant pixels.

After a pass has been completed through the lists of insignificant set and pixels, a “refinement” pass is carried out. Those coefficients are refined by one bit. After an entire pass is made, the threshold is decreased by a factor of two and another sorting pass is initiated. The output bitstream thus consists of the result of significance tests and signs and refinement bits of currently significant coefficients. The location of the coefficients being refined or classified as significant is never explicitly transmitted; it is known implicitly since both the encoder and decoder share the same algorithm, and since all branching decisions are outputed as the encoder searches the coefficients.

By proceeding in this manner, the algorithm transmits approximations of the most important coefficients (those with largest magnitude) first, and refines the values of all significant coefficients one bit at a time. This results in a progressive transmission which always selects the most important information which yields the largest distortion reduction to be transmitted first. Since the DCT is unitary and the Euclidean norm is preserved, it can be shown that such a transmission scheme is the optimal way to decrease the MSE (Mean Square Error) in the reconstructed image as explained in [2].

Also, since the output bitstream is organized in an order of decreasing importance, target bits can be easily achieved by simply truncating the bit stream according to the desired coding budget.

5.3 Entropy coding

Adaptive arithmetic coding can be applied to the output bitstream to further compress the result of the significant tests results, signs, and refinement bits. This method provides a very efficient way to encode these bits, since it can exploit effectively the statistical dependency between the significance of adjacent pixels. In this work, adaptive arithmetic coding algorithm of Witten et al.[4] is used.

6. EXPERIMENTAL RESULTS

A. Intra-frame and Still Image Coding

In this subsection, the performance of the proposed coder is evaluated in the point of I-frame coding of video sequences and still image coding. The first Y frame of QCIF (176x144) MPEG4 test sequences ("Hall Monitor", "Container", "News", "Foreman", "Table", "Weather") and Barbara image (512x512) are used respectively.

The first experiment presents the results that compare the proposed coder to H.263+ coder and EZW-style coders. For comparison of coding performance at a wide range of bit rate, H.263+ coder coded the I-frame by changing the QP from 6 to 31. The H.263 results were produced with the publicly available TMN encoder software[17]. For fair comparison, the proposed coder and other representative wavelet-based embedded image coders[2][6] are run with 2-level wavelet pyramid structure (marked as "L2") at a pre-specified same bit rates. Table 1 shows the average PSNR improvement of the proposed coder to H.263+ coder and representative EZW-style coders[2][6]. It can be seen that the proposed coder has the improvement by the average PSNR gain listed in Table 2. Figure 4 shows the rate-distortion performance in the point of objective image quality and subjective image quality. In all the test image having different image characteristics, the proposed coder shows the better rate-distortion performance in a wide range of bit rate.

We noticed that the EZW's performance is worst due to insufficient wavelet decomposition, its higher overhead for encoding the zerotrees in very low bit rate, and disusage of intraband correlation in 2-level wavelet decomposition.

This is a consequence of the reflection of middle and high frequency. Although EZW[6] and SPIHT[2] method based on wavelet transform have no blocking effect, ringing effect come to visible prominently in very low bit rate. At high compression ratios, ringing is a major problem of the wavelet transform.

For reducing the ringing effect at very low bit rate, Martucci et al.[1] used the different filter lengths at each stage of the decomposition, that is, Daubechies' 9-3 tap filter bank was used at the first level of the decomposition of luminance and followed by the 2-tap Haar filter for the remaining three levels of the decomposition of luminance.

Table 2 shows the performance comparison of proposed coder with the quoted results in [1]. Here, VM60 represents the MPEG-4 Verification Model Version 6.0[15]. The VM is a slightly modified DCT-based H.263 coder used by MPEG-4 as a reference for evaluation of new tools and algorithms. As you seen Table 2, the proposed coder shows the improved rate-distortion performance than EZW method with 4-level pyramid structure using two different filter. This result proves that the proposed rearrangement structure of DCT is quite well coupled with the embedded zerotree coding.

The next experiment shows the coding results to still image. Coding results for Barbara at different rates are listed in Table 3. This table shows the results quoted from baseline JPEG[9][11], improved JPEG[7] which optimizes JPEG's quantization table, EZDCT[8], EZW[6] and SPIHT[2] (with a 2-level wavelet decomposition). The performance of the proposed new coder is comparable to that of state-of-art still image compression methods.

Seq.	H.263	EZW(L2)	SPIHT(L2)
Hall Monitor	1.62	3.66	0.61
Container	1.55	4.75	1.13
News	1.35	3.14	0.80
Foreman	1.27	2.28	0.46
Table	0.86	2.63	0.47
Weather	2.49	2.94	0.29

Table 1. The average PSNR gains of proposed coder over H.263+ and EZW-style coders at a wide range of bit rate

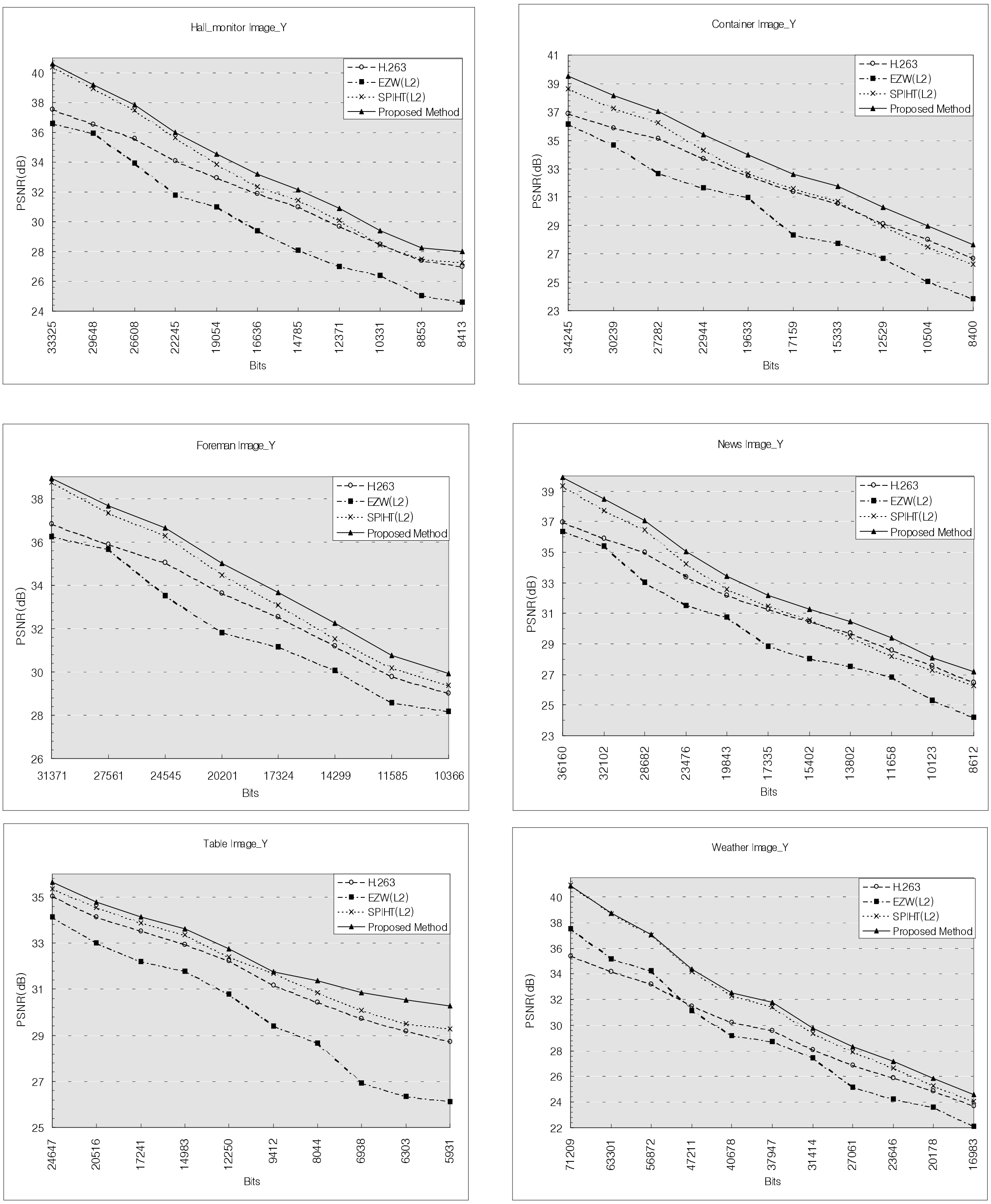


Figure 4. The rate-distortion performance

Seq	Bits	VM60	EZW[1]	Proposed Coder
News	11658	28.6	29.38	29.4
	23476	33.38	34.68	35.06
Foreman	12130	30.11	30.85	31.06
	24545	35.05	35.49	36.65

Table 2. Performance comparisons of the proposed coder with MPEG4 VM and EZW coder of [1]

Rate	JPEG	I_JPEG	EZDCT[8]	EZW(L2)	SPIHT(L2)	Proposed Coder
1	33.1	35.2	36.1	33.58	36.12	36.79
0.75	31	33	33.7	31.62	33.42	34.39
0.5	28.3	30.1	30.82	28.02	30.19	31.18
0.25	25.2	26	26.83	24.09	25.32	27.08

Table 3. Performance comparisons of the proposed coder with JPEG, 2-level EZW coder[2][6], and EZDCT[8] in Barbara

B. Video Coding

This section presents results compared the proposed coder to H.263+ coder. The H.263+ results were produced with the publicly available TMN encoder software [17] as used in subsection A. As described earlier, the motion model used by the proposed coder is identical to the model used in the H.263+. We have not yet implemented a rate control scheme for the proposed video coder. However, since our coder is embedded coder, we can easily adjust the coding bit rates to be the same as that of H263+ for each frame. First the TMN was run, then the proposed coder was run with rate control set to match the bit rate achieved by the TMM. The H.263+ coders were run with QP_{intra} =16 and QP_{inter} having 4 to 25 for all frames. Frame rates is 10Hz.

We ran experiments using three color components (Y, Cb, Cr) of various test sequences at QCIF resolution. The test sequences are coded at a wide range of bit rate. The News, Foreman sequences are suitable test material for video telephone systems. The other sequences, Hall Monitor is more suited to remote monitoring applications. Weather is hybrid natural and synthetic sequence. Bowing sequences is for sign language and lip reading application of H.26L.

Table 4 shows some of the numerical coding results. Average PSNR calculated over the entire coded sequences shows that our new coder achieves the promising performance to the H.263+. It can be seen that the proposed coder has a comparable average PSNR than H.263+ in nearly all the cases. Interestingly, this results tells that our coding method works very efficiently on DFD although it was developed for still image coding.

Hall_Monitor_QCIF_Sequence(300), 10Hz, Intra_QP = 10, Average PSNR of Y_Frames except for INTRA

Bit Rate(kbps)	Inter_QP	H.263+ with baseline(dB)	Proposed Coder(dB)
70.73	4	38.79	39.68
57.61	5	37.52	37.66
37.43	7	35.19	36.07
22.83	10	33.04	33.80
14.75	15	31.84	32.10
9.08	25	30.17	30.05

Foreman_QCIF_Sequence(300), 10Hz, Intra_QP = 10, Average PSNR of Y_Frames except for INTRA

Bit Rate(kbps)	Inter_QP	H.263+ with baseline(dB)	Proposed Coder(dB)
222.77	4	37.52	38.08
178.47	5	36.44	36.89
120.68	7	34.28	34.69
76.97	10	32.02	32.08
48.57	15	29.92	29.68
28.56	25	27.54	27.47

News_QCIF_Sequence(300), 10Hz, Intra_QP = 10, Average PSNR of Y_Frames except for INTRA

Bit Rate(kbps)	Inter_QP	H.263+ with baseline(dB)	Proposed Coder(dB)
108.84	4	37.88	39.05
87.33	5	36.69	36.86
59.34	7	34.39	34.84
37.63	10	32.34	32.51
23.36	15	30.70	30.34
13.38	25	28.74	28.52

Weather_QCIF_Sequence(300), 10Hz, Intra_QP = 10, Average PSNR of Y_Frames except for INTRA

Bit Rate(kbps)	Inter_QP	H.263+ with baseline(dB)	Proposed Coder(dB)
127.42	4	36.00	38.24
104.00	5	34.69	35.88
72.66	7	31.86	32.83
47.76	10	29.45	30.29
29.66	15	27.46	27.92
16.74	25	26.02	26.25

Bowing_QCIF_Sequence(300), 10Hz, Intra_QP = 10, Average PSNR of Y_Frames except for INTRA

Bit Rate(kbps)	INTER_QP	H.263+ with baseline(dB)	Proposed Coder(dB)
33.98	7	35.78	36.27
23.41	10	34.01	34.54
16.80	15	32.77	33.40
12.15	25	31.09	31.79

Table 4. Numerical coding results of proposed coder

More detail is provided in Figure 5, which plot the luminance PSNR against frame number for foreman sequence. The bit rate of proposed coder is the same of H.263+ baseline(120.68kbps). In case of H.263+ with ANNEX I, the bit rate is 121.26kbps. In most of the plots, the proposed coder shows a consistent PSNR gain over H.263+ baseline and H.263+ with ANNEX I for the majority of coded frames. Specially, there are noticeable PSNR gain around frames 186-210 that scene change and fast motion occurs in these period. This result shows that our proposed coder is very efficient to areas where motion compensation fails, as are the complex and fast motion or scene change.

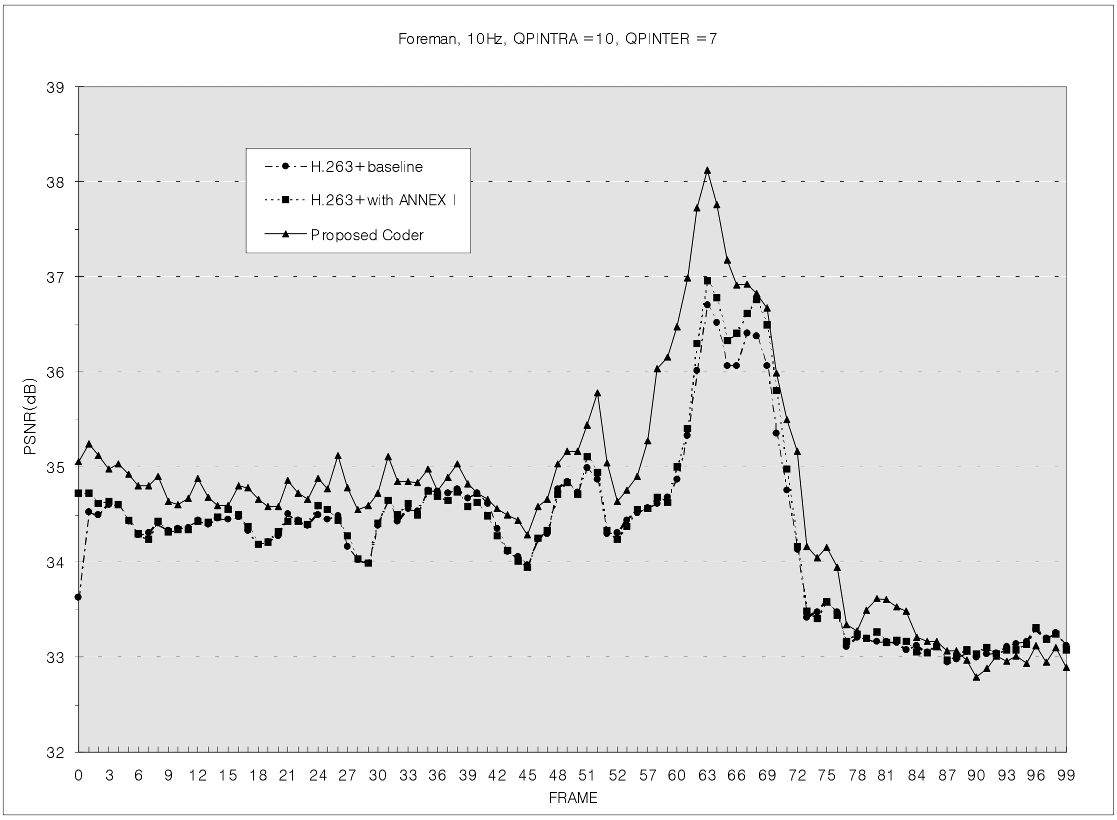


Figure 5. The PSNR-curve of luminance in foreman sequence

C. Subject Image Quality

The reconstructed sequences are not given in this document. It can be seen by D1 device in ITU-T Seoul Meeting. Our coding strategy allows residual bits to be concentrated in the areas where motion compensation fails, as are the complex and fast regions or scene change. The proposed coder provides the reconstructed image with the reduced blocking artifacts. We will demonstrate that the proposed coder outperforms the H.263+ standard in visual quality.

7. CONCLUSION AND FUTHER STUDIES

In this document, we presented an preliminary investigation of the proposed DCT-based embedded video coder. Especially, the rearranged 2-level wavelet pyramid structure of DCT is well coupled with an embedded zerotree coding which has been mainly used in wavelet-based image coding. Furthermore, embedded feature of the proposed coder provides an enhanced functionalities such as precise rate control, scalable functionalities, and enhanced error robustness as well as improved coding efficiency. Also, when restricted to I-frame coding, the performance of the proposed new coder is comparable to that of state-of-art still image compression methods such as JPEG and EZW-style coders.

In DFD coding, our proposed coder has a room for improvement by using band adaptive pre-processing of the rearranged 2-level pyramid structure of DCT. Simulation results will be shown in next meeting.

REFERENCES

[1] S. A. Martucci, I. Sodagar, T. Chiang, and Y. Q. Zhang, "A Zerotree Wavelet Video Coder," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 7, no. 1, pp. 109-118, Feb 1997.

[2] A. Said and W. A. Pearlman, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243-250, June 1996.

[3] D. M. Baylon and J. S. Lim, "Transform/subband analysis and synthesis of signals", MIT Research Laboratory of Electronics Technical Report, June, 1990.

[4] I. H. Witten, R. M. Neal, and J. G. Clearly, "Arithmetic Coding for Data Compression," *Commun. ACM*, vol. 30, pp. 520-540, June 1987.

[5] ITU-T, Draft Recommendation H.263--Video Coding For Low Bit Rate Communication.

[6] J. M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3445-3463, Dec. 1993.

[7] S. Wu and A. Gersho, "Rate-constraned picture adaptive quantization for JPEG baseline coders", in Proc. ICASSP'93, Vol. 5, pp. 389-392, April, 1993.

[8] Z. Xiong, O. Guleryuz, and M. T. Orchard, "A DCT-Based Embedded Image Coder," *IEEE Signal Processing Letters*, vol. 3, pp. 289-290, Nov. 1996.

[9] W. B. Pennebaker, JPEG Still Image Data Compression Standard, New York: Van Nostrand Reihold, 1993.

[10] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," IEEE Trans. on Signal Processing, Vol. 3, No. 3, pp. 572-588, 1994.

[11] ISO/IEC 10918-3, ITU-T Rec. T.84, *Information technology* - Digital compression and coding of continuous-tone still image: Extensions, 1996.

[12] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," IEEE Trans. Image Processing, Vol. 1, pp. 205-220, Apr. 1992.

[13] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[14] M. Rabbani and P. W. Jones, *Digital Image Compression Technique*. SPIE Tutorial Text Series Vol TT7 : SPIE Optical Engineering Press, 1991.

[15] MPEG-4 video verification model V.60 ISO/IEC JTC/SC29/WG11 Coding of Moving Pictures and Associated Audio MPEG97/Sevilla, February 1997.

[16] ITU-T Standardization Sector of ITU, "Video Coding for Low Bit rate Communication," ITU-T Recommendation H.263, March 1996.

[17] Image Processing Lab, University of British Columbia, "TMN (H.263+) encoder/decoder, version 3.0," TMN (H.263+) codec, September 1997.

[18] ITU-T Standardization Sector of ITU, "Call For Proposal For H.26L Video Coding" SG16/Q15, Geneva, 26 January - 6 February 1998.

[19] Y. A. Jeong and C. K. Cheong, "A DCT-Based Embedded Image Coder Using Wavelet Structure of DCT for Very Low Bit Rate Video CODEC" *Proc. of International Conference on Consumer Electronics*, vol. 1, pp. 40-41, Los Angeles, June 2-4 1998.

[20] Y. A. Jeong and C. K. Cheong, "A DCT-Based Embedded Image Coder Using Wavelet Structure of DCT for Very Low Bit Rate Video CODEC" *IEEE Trans. on Consumer Electronics*, vol. 44, no. 3, pp. 500-508, August 1998.