論文 96-1-1-1

AN IMAGE SEQUENCE CODING USING MOTION-COMPENSATED TRANSFORM TECHNIQUE BASED ON THE SUB-BAND DECOMPOSITION

Hoon Paek*, Rin Chul Kim*, Sang Uk Lee* Regular Members

움직임 보상 기법과 분할 대역 기법을 사용한 동영상 부호화 기법

正會員 白 勳*, 金 鱗 澈**, 李 商 郁*

ABSTRACT

In this paper, by combining the motion compensated transform coding with the sub-band decomposition technique, we present a motion compensated sub-band coding technique(MCSBC) for image sequence coding. Several problems related to the MCSBC, such as a scheme for motion compensation in each sub-band and the efficient VWL coding of the DCT coefficients in each sub-band are discussed. For an efficient coding, the motion estimation and compensation is performed only on the LL sub-band, but the discrete cosine transform(DCT) is employed to encode all sub-bands in our approach. Then, the transform coefficients in each sub-band are scanned in a different manner depending on the energy distributions in the DCT domain, and coded by using separate 2-D Huffman code tables, which are optimized to the probability distributions of each sub-band. The performance of the proposed MCSBC technique is intensively examined by computer simulations on the HDTV image sequences. The simulation results reveal that the proposed MCSBC technique outperforms other coding techniques, especially the well-known motion compensated transform coding technique by about 1.5 dB, in terms of the average peak signal to noise ratio.

要 約

본 논문에서는 동영상 부호화를 위하여 움직임 보상 기법과 분할 대역 기법을 사용한 MCSBC 부호화 기법을 제안하였다. 또한 MCSBC에 관한 여러가지 문제들, 즉 각 분할 대역에의 움직임 보상 기법에 관한 문제, 각 분할 대역의 DCT 계수에 대한 효율적인 비트 할당등을 다루었다. MCSBC의 효율적인 부호화를 위하여 먼저 원신호에 대역 분할을 수행한 후, 움직임 보상 기법은 저대역의 영상 신호에만 적용하였고, 모든 대역에 이산 여현 변환(DCT)를 적용하였다. DCT가 적용된 블럭들은 각 대역 신호의 특징에 따라 최적화된 주사 방법 및 비트 할당을 사용하여 부호화한다. 이러한 MCSBC 기법은 고화질 TV 용 동영상에 적용하여 모의 실험을 수행하였다. 모의 실험 결과, 제안한 MCSBC 기법은 일반적인 움직임 보상 동영상 부호화 기법에 비하여 약 1.5 dB의 성능향상을 확인할 수 있었다.

*서울대학교 제어계측공학과 신호처리연구실
**한성대학교 정보공학과
論文番號：95070-0218
接受日字：1995年 2月 18日

I. Introduction

The motion compensated transform coding(MCTC)[1]-[6] is one of the most effective and popular techniques currently used for encoding the moving image sequence, since it can exploit both the temporal and spatial redundancies. The MCTC technique has been chosen for commercial application, such as a full digital high-definition television(HDTV)[3] and several moving image coding standards[5,6]. However, one of problems with the MCTC is that the visually annoying blocking effect is observed in the reconstructed image, especially when the bit rate is relatively low[2,3].

In order to alleviate the blocking effect, various efforts have been made so far. These efforts include the post-processing of the reconstructed image[7], the use of lapped orthogonal transform[8], instead of the commonly used discrete cosine transform(DCT)[2], and incorporation of the sub-band coding(SBC)[9,10,11] with the motion compensation[12,13,14], *etc*. Among them, the SBC incorporating the motion compensation, *i.e.*, the motion compensated sub-band coding(MCSBC), seems to be very effective technique since it takes advantage of both the SBC and motion compensation techniques.

The MCSBC consists of the motion compensation, the sub-band decomposition, and the quantization module for encoding the sub-bands. In designing the MCSBC, however, there are several problems to be considered, such as a method to combine the motion compensation with the sub-band decomposition, an efficient coding technique to exploit the characteristics of each sub-band, and the bit rate allocation scheme to each sub-band, *etc*. Irie and Kishimoto[12] compared various MCSBC techniques and reported that an adaptive DCT coding technique for the lowest sub-band and the DPCM or PCM with dead zone for other sub-bands could provide an adequate performance. Gharavi[13] considered two different schemes on how to combine the motion compensation with the sub-band decomposition for the MCSBC technique, and proposed an efficient coding technique for the higher sub-bands. But, Irie and Kishimoto[12], and Gharavi[13] did not take account of an optimal bit rate allocation for each sub-band. In other words, the main disadvantage of both coding techniques is that it would be difficult to allocate the bit rate for each sub-band optimally, since the coding scheme for each sub-band is quite different. Hang *et. al*[14] described the MCSBC technique, emphasizing a real-time implementation of the coder based on parallel processing. In [14], the bit rate allocation for each sub-band is provided by the dynamic channel allocation algorithm. However, the dynamic channel allocation algorithm may not provide the optimal bit rates to each sub-band[14]. Moreover, the coding technique proposed by Hang *et. al*[14] is believed to be inappropriate and complex, since the motion estimation and compensation is performed on all sub-bands.

In this paper, we present the MCSBC technique for image sequence coding. Also, several issues relating to the MCSBC, such as a scheme for motion compensation in each sub-band and the efficient variable word length(VWL) coding of the DCT coefficients in each sub-band are discussed. The coding scheme considered in this paper to encode each sub-band is based upon the DCT. Although it is known that the sub-band decomposition of the input image partially removes the statistical redundancy existing in the input image, we have found that there still exists a significant spatial redundancy in each sub-band. Thus, it is argued that the DCT is still capable of providing the necessary energy compaction in each sub-band. In addition, we shall show that the VWL coding in the DCT domain is more efficient than in the spatial domain, since the energy distribution in the DCT domain differs from sub-band to sub-band. In our approach, depending on the energy distributions in the DCT domain, the DCT coefficients in each sub-band are scanned in a different manner and coded by separate 2-D Huffman code tables, which are optimized to the statistics of each sub-band. The performance of the proposed MCSBC technique is examined on the several HDTV image sequences and compare with the MCTC technique[4,5]. The specific MCTC technique considered in this paper is described in section Ⅱ for the sake of completeness. In section Ⅲ, two different MCSBC techniques are compared and discussed. Also, the bit rate allocation strategy suitable for the MCSBC is described. Section Ⅳ discusses the problems related to the improvement of the efficiency of the motion compensation and VWL coding of the DCT coefficients in each sub-band. Section Ⅴ provides the computer simulation results on the several image sequences and discussions. Finally, section Ⅵ gives our conclusions.

Ⅱ. Motion compensated transform coding

There are many different MCTC techniques in the literature[1,3,4,5,6]. However, the MCTC technique chosen in this paper is similar to the ITU/TS H.261 standard[4,5]. As shown in Fig. 1, the basic block diagram for the MCTC technique consists of the motion compensation, the DCT, and the quantizer including the VWL coding. In the MCTC coder, the chrominance is subsampled by horizontally 4:1 and vertically 2:1, resulting in 1/8 of the luminance in size. Then, the input frame is formatted into the macro blocks(MB), which consist of 32×16 luminance and their chrominance counterpart, namely two 8×8 C_r and C_b. Then the MB is further divided into ten 8×8 blocks.

The MCTC technique performs first the motion estimation on the luminance MB by using the full search block matching algorithm with the search area of horizontally ±

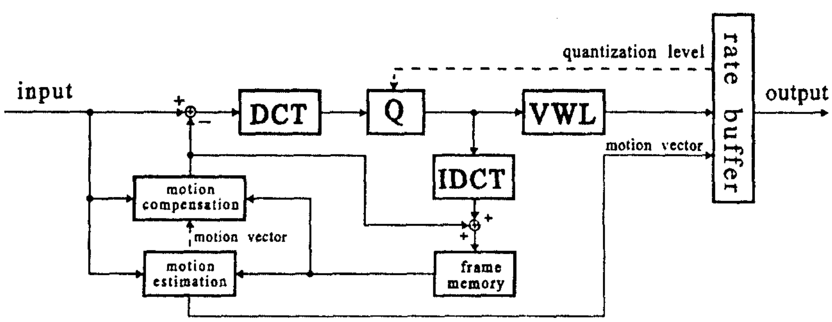


Fig. 1. A block diagram for the MCTC.

15 pels and vertically ±7 lines. Then, the motion compensated prediction error, which is called the displaced frame difference(DFD), is transformed by 8×8 DCT. The DCT coefficients are quantized using the uniform quantizer with dead zone and scanned in a zig-zag manner to make the quantized coefficients serialized. In our approach, we use the same quantizer proposed in the ITU/TS H.261 standard[4,5]. Then, the non-zero quantized coefficients together with zero runs before non-zero values are coded with the 2-D Huffman code table[4,5], which is also provided in the ITU/TS H.261 standards[4,5]. The VWL coded transform coefficients are transmitted to the decoder, together with the motion vector which is represented by 9 bits/MB. With the conditional replenishment[3], only the MB's that change significantly from frame to frame are sent to the decoder.

The overall transmission rate is controlled by employing the rate buffer and smoothed by adjusting the quantizer step size, according to the buffer state by the unit of the MB, also as in the ITU/TS H.261 standards[4,5].

Ⅲ. Motion compensation in the sub-bands

In this section, we shall describe two different MCSBC techniques considered in this paper. The two techniques are differentiated by configuration of the motion compensation and the sub-band decomposition.

The first technique, which is called the MCSBC-I, decomposes the input image into sub-bands before the motion compensation. On the other hand, the second technique called the MCSBC-II, performs first the motion compensation and then sub-band decomposition next. In our approach, the input image is decomposed into 4 sub-bands, *i.e.*, LL, LH, HL, HH sub-band) employing the well-known Johnston 32-order filter[10]. In fact, two different filters, *i.e.*, Johnston 32-order filter[10] and the perfect reconstruction 64-order filter[11], are tested in our approach. However, we found that two filters yield the similar performance, if the quantizer is included in the analysis/synthesis filtering process. Let us describe the two techniques in more detail.

A. Configurations

In the MCSBC-I, as shown in Fig. 2, the input image is decomposed into 4 sub-bands first, then the motion compensation is applied on the sub-bands. However, the motion compensation scheme, which shall be discussed in more detail later, could be implemented differently in the MCSBC-I. In the first approach, the motion estimation is performed only on the LL sub-band of block size 16×8 with the search area of horizontally ±15 pels and vertically ±7 lines, but the motion compensation is performed on all sub-bands with the motion vector obtained from the LL sub-band in the same spatial location. Subsequently, the 8×8 DCT is applied on the DFD's of each sub-band. Note that, in our approach, the same quantization and transmission codeword allocation schemes proposed by the ITU/TS H.261 standard[4,5] are used for each sub-band of the MCSBC-I for the sake of convenience.

On the other hand, the MCSBC-Ⅱ, as shown in Fig. 3, performs first the motion estimation and compensation on the 32×16 luminance blocks with the search area of horizontally ±31 pels and vertically ±15 lines. Note that the search area in the MCSBC-Ⅱ is

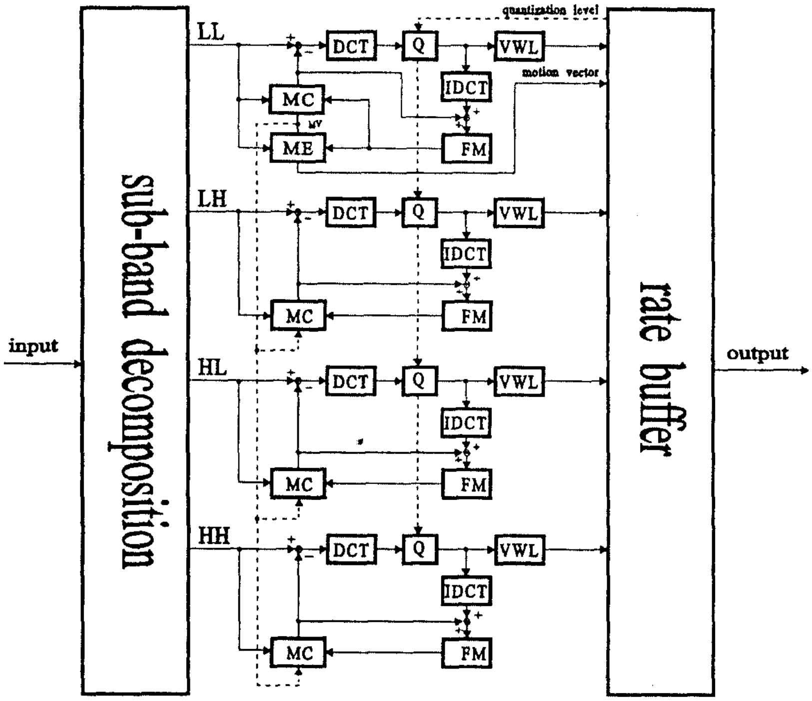


Fig. 2. A block diagram for the MCSBC-I.

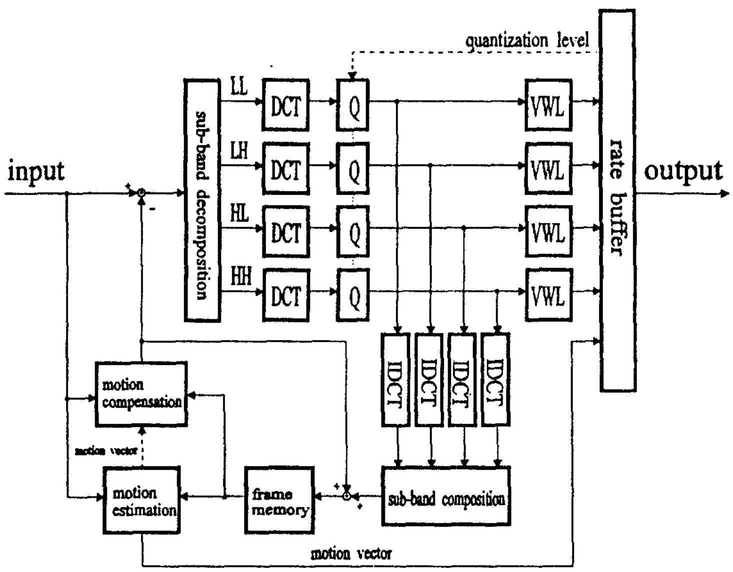


Fig. 3. A block diagram for the MCSBC-Ⅱ.

determined so as to make the ratios of search area to image size in both cases of MCSBC-I and MCSBC-II equal, because the MCSBC-II performs the motion estimation on the full-band image, while the MCSBC-I performs the motion estimation on the LL sub-band, which is, in fact, sub-sampled by 4 : 1, compared to the full-band image. Then, the DFD's are decomposed into 4 sub-bands. The DCT is applied on each sub-band and the transform coefficients are quantized and coded using the same method employed in the MCSBC-I.

In both MCSBC techniques, the chrominance is coded by the similar techniques used for the luminance. More specifically, without the sub-band decomposition, the chrominance is motion-compensated by using the motion vector obtained from the luminance in the same spatial location. Then, the DCT is performed on the DFD's, and the transform coefficients are quantized and coded by using the same techniques for the luminance.

B. Issues in the rate allocation

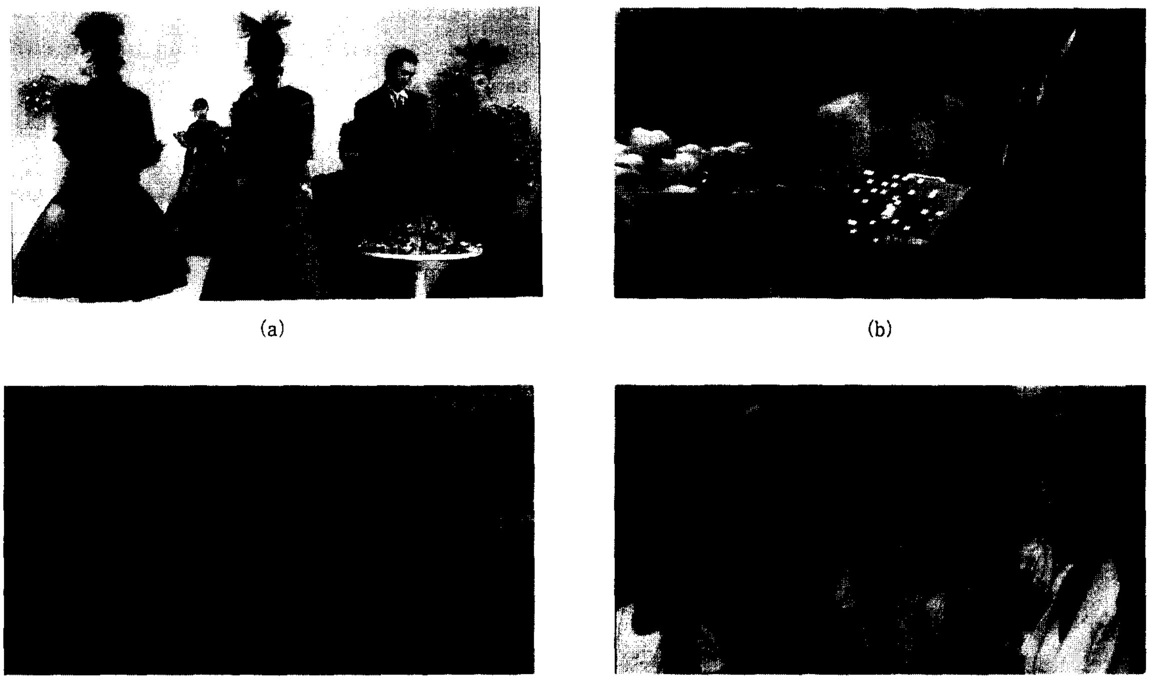
In designing the MCSBC's, an optimal bit rate allocation to each sub-band is one of the most important problems. Conventionally, the bit rate is allocated by the optimal bit rate allocation schemes[9,17]. However, note that the conventional bit rate allocation scheme is only applicable to the fixed word-length(FWL) coding, but not to the VWL coding, which is known to provide better performance than the FWL coding[15]. Thus, we need to pay a careful attention to the bit rate allocation to each sub-band for the VWL coding. For VWL coding, Hang *et. al*[14] attempted to allocate the bit rates for each sub-band by using the dynamic channel allocation algorithm. In the dynamic channel allocation algorithm, the bit rate for each sub-band is allocated based on the statistics obtained from the previous frame. However, the ITU/TS H.261 standard coder[4,5], which is used separately for encoding each sub-band[14], may not operate adequately at extremely low or high rate. Hence, the dynamic channel allocation algorithm may not provide a satisfactory performance, especially when the bit rates deviate greatly from sub-band to sub-band[14]. Thus, we shall briefly discuss about the different bit rate allocation scheme, which is suitable for our purpose.

It is noted that the bit rate allocation procedure indicates that the quantizers used for encoding of the DCT coefficients in all sub-bands have the same step size for optimal quantization[15,16,17], yielding an equal average distortion in each sub-band. Based on above notion, first, each sub-band is divided into 16×8 blocks. Then, we construct the MB by grouping four 16×8 blocks each of which belongs to the different sub-bands and two different 8×8 chrominance blocks, in the same spatial location. Thus, the MB consists of ten 8×8 blocks, which is identical to the one for the MCTC described in section Ⅱ. The quantizers with the same step size are used to encode all the DCT coefficients in the MB, which consists of eight Y blocks, one Cr block, and one Cb block, so that the bit rate for each sub-band is allocated optimally in the sense described previously. In both the MCSBC techniques, the VWL coded transform coefficients and the motion vector encoded with 9 bits/MB are sent to the decoder, and the buffer control strategies are the same as the MCTC described in section Ⅱ.

C. Comparisons

Now, let us compare two MCSBC techniques with the viewpoint of the position of sub-band decomposition. In order to evaluate the performance of both techniques, computer simulations on the real image sequences are conducted. In the simulation, three test image sequences, namely "MODEL", "FRUITS" and "CATHEDRAL", are used as shown in Fig. 4. In Fig. 4, it can be seen that "CATHEDRAL" sequence has a scene change at 10-th frame. The size of all test images is 1280×720. In the rest of the paper, in the simulation, the overall bit rate is chosen to be about 0.27 bits/pel, yielding a total bit rate of 9.45 Mbps. As a measure of the reconstructed image quality, the conventional peak signal to noise ratio(PSNR) in dB is used.

In order to examine the energy packing efficiency of two different sub-band decomposition techniques, we compare the energy of each sub-band. Notice that in this paper energy means mean square value. In terms of coding gain, the energy packing efficiency is of importance consideration, since more energy compaction in the LL sub-band improves the coding efficiency, as in the DCT coding[2]. Fig. 5 shows the average energies of all sub-bands and the energies of the LL sub-band image by two MCSBC techniques on the 40 frames of the "MODEL" sequence. From Fig. 5, it is seen that although the average energy of all sub-band of the MCSBC-I is similar to that of the MCSBC-Ⅱ, the MCSBC-I is better than the MCSBC-Ⅱ, in terms of the energy packing in the LL sub-band. In addition, since the sub-band decomposition is performed on the DFD image, the MCSBC-Ⅱ



(a) (b) (c) (d)

Fig. 4. Original images: (a) MODEL; (b) FRUITS; (c) 10-th frame of CATHEDRAL; (d) 11-th frame of CATHEDRAL.

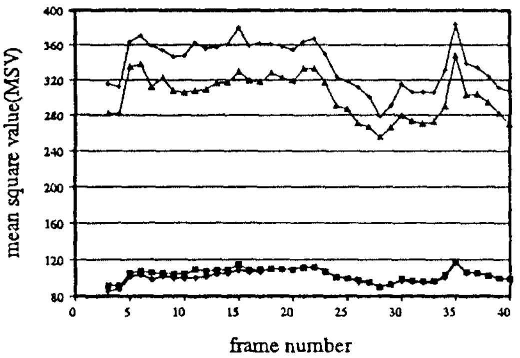


Fig. 5. Comparison of the average of all sub-bands and the energies of the LL sub-band on the "MODEL" sequence by the MCSBC's.

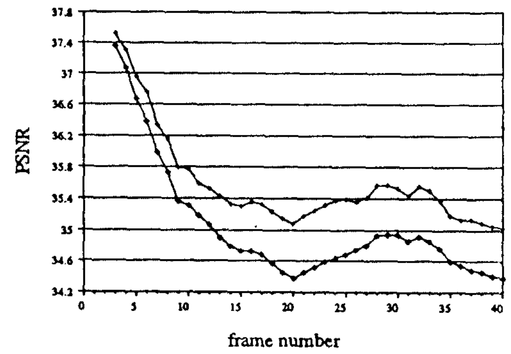


Fig. 6. Performance comparion between the MSSBC-I and MCSBC-Ⅱ on the "MODEL" sequence.

yields more high frequency components than the MCSBC-I. Therefore, a better performance of the MCSBC-I over the MCSBC-Ⅱ is expected, since the MCSBC-I fully utilizes the advantage of sub-band decomposition on the input image for more energy compaction in the LL sub-band.

The performance comparison of the MCSBC-I and MCSBC-Ⅱ on the "MODEL" sequence is shown in Fig. 6. In Fig. 6, it is observed that the MCSBC-I shows better performance than the MCSBC-Ⅱ by about 0.6 dB, in terms of average PSNR. Thus, since the MCSBC-I yields better performance than the MCSBC-Ⅱ, in terms of the energy packing efficiency and the PSNR evaluation, we restrict our attention to the MCSBC-I in the rest of the paper. In the next section, we shall describe the schemes to improve the performance of the MCSBC-I further.

Ⅳ. Motion compensated sub-band coding

In the MCSBC-I, since each sub-band exhibits quite different statistics, the motion compensation and the DCT may not work adequately for encoding each sub-band. Thus, in this section, we shall compare the various motion compensation schemes in the sub-bands first. Then, we shall propose adaptive scanning methods for VWL coding of the DCT coefficients of each sub-band efficiently.

A. Comparison of the motion compensation schemes

In the MCSBC-I, we can consider three different motion compensation schemes. The first scheme(TYPE-Ⅰ), similar to Hang, *et. al*[14], is to perform the motion estimation and compensation separately on each sub-band. In the second scheme(TYPE-Ⅱ), the motion information is extracted from the LL sub-band only, but the motion compensation is performed on all the sub-bands using the motion vector obtained from the LL sub-band in the same spatial location, which is similar to the technique of Gharavi[13]. While the motion compensation is performed on all sub-bands in the previous two schemes, the third scheme(TYPE-Ⅲ) employs the motion estimation and compensation only on the LL sub-band, which is similar to Irie and Kishimoto[12]. However, notice that, in the TYPE-Ⅲ, the higher sub-bands, i.e., LH, HL and HH sub-band, are intra-frame encoded.

To evaluate the efficiency of the use of the LL motion vector on higher sub-bands in the TYPE-Ⅱ, we compare, for example, the energies of the HL bands according to three coding schemes in Fig. 7. In Fig. 7, it can be seen that the energies of higher sub-bands by the TYPE-Ⅱ increase as the frame progresses, while the TYPE-I and TYPE-Ⅲ show almost similar energy levels for all frames. In fact, the performance of the TYPE-Ⅱ is similar to the TYPE-Ⅲ in the beginning. However, as the frame progresses, each sub-band image is degraded due to the quantization error. Thus, it seems that since the motion vector extracted from the LL sub-band would not be adequate for the compensation of the higher sub-bands, the efficiency of the motion compensation on the higher sub-bands is significantly degraded in the TYPE-Ⅱ. Therefore, we conclude that the TYPE-Ⅱ, which is similar to the technique of Gharavi[13], is inadequate for image sequence coding.

In comparison of the TYPE-I and TYPE-Ⅲ, as shown in Fig. 7, the TYPE-I looks better choice than the TYPE-Ⅲ as far as the energy is concerned. However, the situation is reversed if we consider the overheads for encoding the motion vectors in the TYPE-I. Table I shows the average entropies of the higher sub-bands for both techniques, which are, in fact, obtained from three different moving image sequences, namely, "MODEL", "FRUITS" and "CATHEDRAL". Advantage of employing the motion compensation on the higher sub-bands in the TYPE-I, especially for LH and HL sub-bands, is clearly observed. Notice that the motion compensation is performed on the higher sub-bands in the TYPE-I, while the higher sub-bands are intra-coded in the TYPE-Ⅲ. But, let us define a gain, in terms of average bits per 16 ×8 macro block, which is obtained from the motion compensation on the higher sub-bands in the TYPE-I. The gain can be calculated from multiplying the number of pixels in the 16×8 macro block by difference of the entropies of two techniques, which are also presented in the Table I. However, it is noted that three additional motion vectors should be transmitted in the TYPE-I. In our approach, it is assumed that a motion vector in each sub-band is encoded with 9 bits if the motion vector is not zero. The result shown in the Table I indicate that the overhead to encode three motion vectors in the TYPE-I completely wipes out the gain obtained from the motion compensation on the higher sub-bands, indicating the TYPE-Ⅲ is a better choice than the TYPE-I.

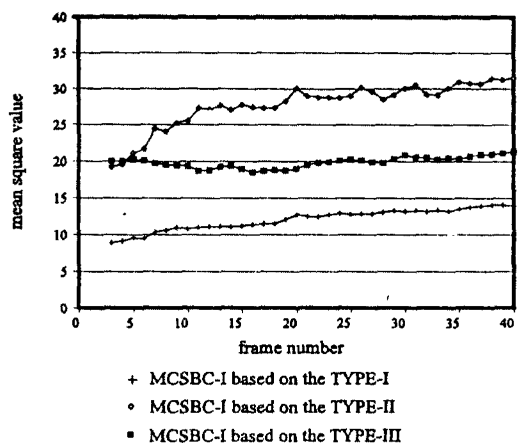
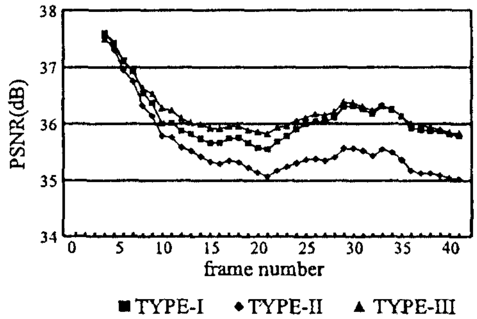
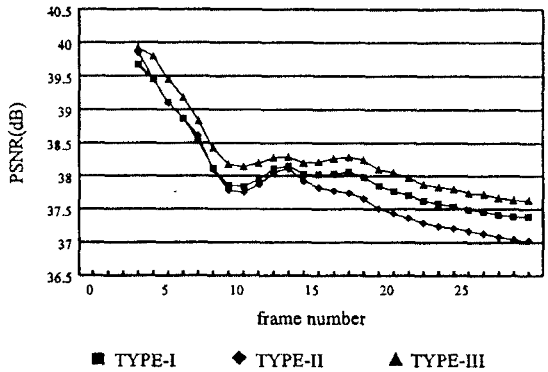


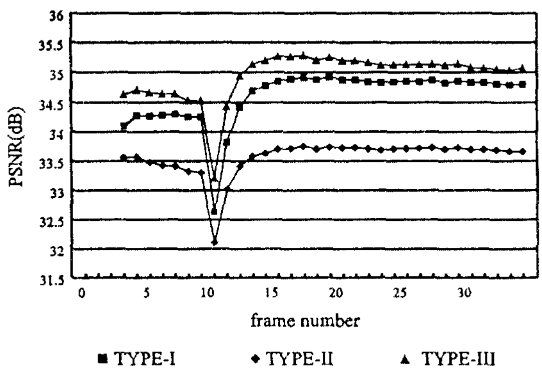
Fig. 7. Comparison of the energies of the HL sub-band according to motion estimation and compensation schemes on the "MODEL" sequence.



(a) "MODEL" sequence



(b) "FRUITS" sequence



(c) "CATHEDRAL" sequence

Fig. 8. Performance comparison among three different motion estimation and compensation scheme for the MCSBC-I on the (a) "MODEL" sequence, (b) "FRUITS" sequence and (c) "CATHEDRAL" sequence.

In Fig 8, we compare the PSNR performance of three different motion compensation schemes discussed so far on the three image sequences. Note that the average bit rate's are equal for three cases. As is expected from Fig. 7, the TYPE-Ⅱ yields the worst performance. We also abserve that the performance of the TYPE-Ⅲ is always better than that of the TYPE-I for all three image sequences. Also notice that the TYPE-Ⅲ is considerably simpler to implement than the TYPE-I. Thus, we restrict our attention to the TYPE-Ⅲ in the rest of the paper, which we shall call the MCSBC-Ⅲ for convenience.

B. Adaptive scanning method

In the MCTC techniques described in section Ⅱ, the 8×8 quantized DCT coefficients are scanned by a zig-zag fashion, and coded into the VWL codes using single 2-D Huffman code table. The notion behind the serialization of the DCT coefficients by the zig-zag scan is to improve the efficiency of the VWL coding. More specifically, in the case of full-band image, since more energy is concentrated to the lower sequency components in the DCT domain, the zig-zag scan from low sequency to high sequency makes the probability of occurring the consecutive zeros high. As a result, the VWL coding, in which the non-zero coefficients together with zero runs preceding non-zero values are coded by 2-D Huffman code table, becomes very efficient. However, in the case of the sub-band, the VWL coding by the zig-zag scanning of the DCT coefficients may not be an efficient one, since each sub-band exhibits a different energy distribution. Hence, in our approach, we employ an adaptive scheme that can efficiently encode the DCT coefficients in each sub-band. The adaptivity is accomplished by employing different scanning methods and separate 2-D Huffman code tables for each sub-band. Let us describe the adaptive scheme in more detail.

In order to gain an insight into the scanning method for the DCT coefficients suitable for each sub-band, we examine the energy distributions in the DCT domain for each sub-band on the "MODEL" sequence as shown in Fig. 9. From Fig. 9, we can see that the energy distributions in the DCT domain differ from sub-band to sub-band. The LL sub-band shows more energies in the low sequency components in the DCT domain, which is similar to the full-band image. On the other hand, the energies are concentrated to the bottom-left corner, top-right corner, and the right

791.4	467.0	312.0	150.4	110.1	82.5	57.1	41.9
354.0	245.0	160.3	106.7	71.0	58.5	41.1	32.9
229.8	170.0	107.1	80.9	61.1	44.5	35.5	25.9
159.6	110.6	81.8	53.4	44.8	28.9	26.3	19.4
132.8	87.7	51.1	36.5	29.1	21.7	17.3	13.4
88.2	59.8	40.8	25.8	19.3	15.1	11.0	9.4
60.0	36.8	27.4	18.2	13.5	9.3	7.6	6.1
45.9	27.7	18.0	12.9	8.5	7.0	5.5	4.8

(a)

8.2	7.0	6.5	6.2	6.0	6.1	5.8	6.0
7.8	7.6	6.8	6.6	6.3	6.3	6.7	6.2
8.8	8.0	7.4	6.8	6.4	6.3	6.7	6.2
10.2	9.0	7.8	6.8	6.3	6.3	6.0	5.9
13.7	10.8	8.6	8.3	7.1	7.0	6.5	5.8
18.3	13.3	11.3	8.7	8.4	7.3	6.3	6.2
23.0	16.2	13.1	10.8	8.6	7.6	7.3	5.8
33.5	20.5	17.9	13.2	10.4	8.9	7.6	6.4

(b)

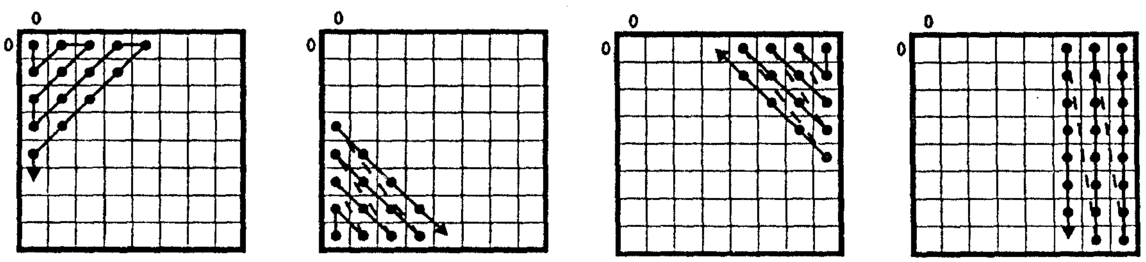
1.6	2.6	2.7	3.1	3.7	9.1	18.7	28.0
1.5	2.0	2.0	2.3	3.1	6.2	12.2	22.4
1.4	1.8	1.9	2.1	2.8	5.7	10.8	18.8
1.3	1.6	1.7	2.0	2.3	4.7	8.7	14.9
1.3	1.5	1.6	1.8	2.3	3.9	7.1	9.8
1.2	1.4	1.5	1.7	1.9	3.1	5.3	7.7
1.1	1.3	1.6	1.6	1.8	2.5	4.4	5.5
1.2	1.3	1.4	1.4	1.7	2.4	3.5	4.3

(c)

1.3	1.5	1.7	1.7	1.8	3.0	4.4	4.3
1.3	1.4	1.4	1.5	1.8	2.6	3.7	4.2
1.2	1.4	1.4	1.5	1.7	2.4	3.7	3.9
1.1	1.3	1.4	1.4	1.7	2.5	3.2	3.7
1.2	1.3	1.3	1.4	1.7	2.2	3.1	3.5
1.1	1.3	1.4	1.4	1.6	2.2	2.8	3.5
1.1	1.3	1.4	1.5	1.6	2.3	3.2	3.7
1.1	1.3	1.3	1.4	1.6	2.1	3.1	3.8

(d)

Fig. 9. An example of the energy distibutions in the DCT domain for the sub-bands on the "MODEL" sequence;
(a) LL band; (b) LH band; (c) HL band; (d) HH band.



(a) (b) (c) (d)

Fig. 10. The proposed scanning methods for the DCT coefficients in the
(a) LL band; (b) LH band; (c) HL band; (d) HH band.

corner in the DCT domain for the LH, HL, and HH sub-bands, respectively. By carefully examining the energy distributions in each sub-band, we propose efficient scanning methods for the DCT coefficients in each sub-band as shown in Fig. 10. Note that the proposed scanning methods yield the consecutive zeros more probable than the zig-zag scanning, especially for the higher sub-bands.

In order to improve the efficiency of the VWL coding further, we design separate 2-D Huffman code tables for each sub-band. Fig. 11 shows the histogram of the quantized DCT coefficients in each sub-band on the "MODEL" sequence. Note that the quantizer step size is set to 8. It is observed in Fig. 11 that the DCT coefficients of all sub-bands show sharply peaked distributions rapidly falling off as the quantized value increases, implying that the VWL coding would be very efficient to these signals. However, the distributions in each sub-band are observed to be quite different. For example, the distributions of the higher sub-bands are even more highly peaked than that of the LL sub-band. For this reason, we have chosen to use separate 2-D Huffman code tables for each sub-band. Each 2-D Huffman code table is designed according to the probability distributions of the DCT coefficients of each sub-band of the "MODEL" sequence.

Fig. 12 shows the performance evaluation of the adaptive scanning method described so far. In Fig. 12, it is observed that the MCSBC-Ⅲ with adaptive scanning technique provides a coding gain of about 0.3 dB over the MCSBC-Ⅲ with the zig-zag scanning in all sub-bands. Moreover, by employing the separate Huffman code tables optimized to the statistics of each sub-band, the performance of the MCSBC-Ⅲ is further improved by about 0.5 dB. From the simulation results presented so far, it seems that the adaptive scanning methods, together with separate Huffman code tables for each sub-band, indeed improve the performance of the MCSBC-Ⅲ.

C. The proposed coding technique

Now, let us summarize the proposed coding

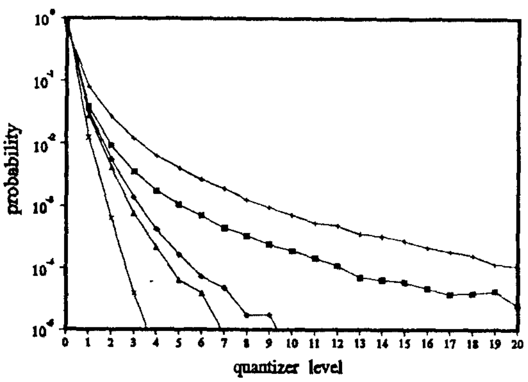


Fig. 11. An example of the histogram of the quantized DCT coefficients in each sub-band on the "MODEL" sequence.

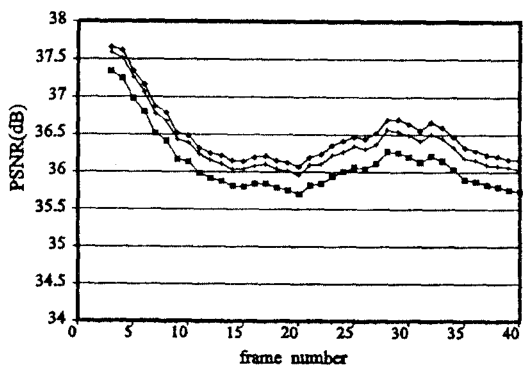


Fig. 12. Performance of the adaptive DCT coding technique for the MCSBC-Ⅲ on the "MODEL" sequence.

technique described so far. In the proposed coding technique, after the sub-band decomposition, the motion estimation and compensation is performed only on the LL sub-band. But, the motion compensation is not employed on the rest of sub-bands. Then, after applying the DCT to all sub-bands, the DCT coefficients are scanned differently in each sub-band as shown in Fig. 10, and coded into the VWL codes using separately prepared 2-D Huffman code tables for each sub-band. Note that in the MCSBC-I, the DCT coefficients are scanned in a zig-zag manner and coded into the VWL codes using a single 2-D Huffman code table.

In the next section, we shall present the computer simulation results to compare the performance of various coding techniques discussed so far.

Ⅴ. Simulations

The results of computer simulation is provided to demonstrate the performance of the MCSBC-Ⅲ over the MCTC on the "MODEL", "FRUITS" and "CATHEDRAL" sequences. However it should be noted that in the computer simulation, the adaptive schemes in each sub-band obtained from the "MODEL" sequence are used to encode the "FRUITS" and "CATHEDRAL" sequences without any modification. For a fair comparison, we also include a coding technique, in which the PCM is employed to encode the higher sub-bands, to demonstrate the employment of the DCT to encode each sub-band. In the case of using PCM, the coding strategy for the LL sub-band is the same as the MCSBC-Ⅲ, but a different strategy is employed for the higher sub-bands. Similar to the technique of Gharavi[13], the 8×8 blocks in the higher sub-bands, on which neither the motion compensation nor the DCT is performed, are scanned in the horizontal, vertical, or zig-zag direction, according to the characteristics of each sub-band. Then, those are quantized by uniform quantizer with dead zone, and coded by using separate 2-D Huffman code tables optimized to the statistics of each sub-band. Here, it is worth to note that the MCSBC-Ⅲ

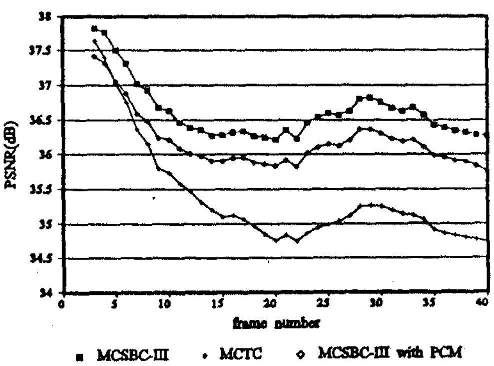


Fig. 13. Performance of the MCSBC-Ⅲ technique on the "MODEL" sequence.

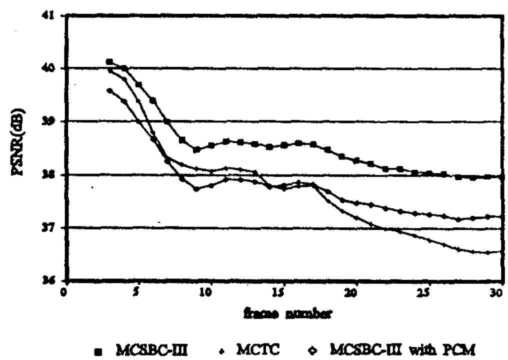


Fig. 14. Performance of the proposed MCSBC-Ⅲ technique on the "FRUITS" sequence.

differs from the techniques by Irie and Kishimoto[12] and Gharavi[13] in that the coding technique for the higher sub-bands in this paper is based on the DCT, while others are based on the PCM or DPCM.

The performance of the proposed MCSBC-Ⅲ is compared with those of the MCTC and the MCSBC-Ⅲ with PCM for the higher sub-bands, in terms of the PSNR, as shown in Fig. 13, Fig. 14 and Fig. 15, respectively.

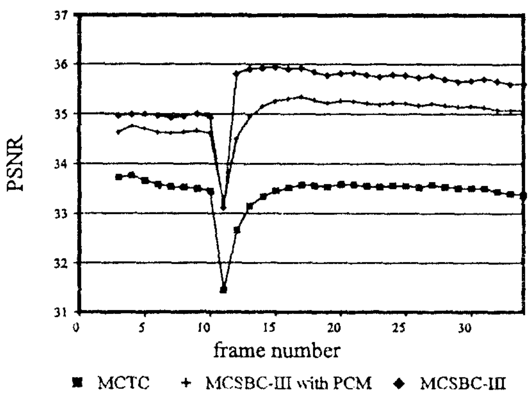
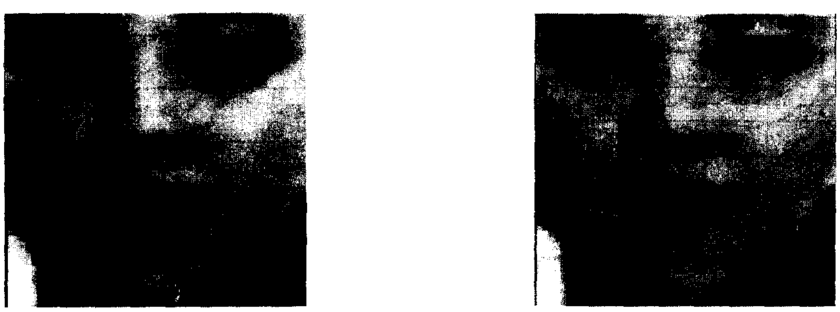


Fig. 15. Performance of the proposed MCSBC-Ⅲ technique on the "CATHEDRAL" sequence

From Fig. 13, it is observed that the proposed MCSBC-Ⅲ yields about 1.5 dB improvement over the MCTC and about 0.5 dB improvement over the MCSBC-Ⅲ with PCM on the "MODEL" sequence. In the case of the "FRUITS" sequence, as shown in Fig. 14, it is also observed that the proposed MCSBC-Ⅲ provides about 0.8 dB improvement over the MCSBC-Ⅲ with PCM, while the MCSBC-Ⅲ with PCM provides only slightly better performance than the MCTC. Also, in the case of the "CATHEDRAL" sequence which has a scene change in 10-th frame, the proposed MCSBC-Ⅲ yields the better performance than the MCTC by about 2 dB and MCSBC-Ⅲ with PCM by about 0.6 dB, as shown in Fig. 15.

Moreover, in the subjective quality, it is shown that the proposed MCSBC-Ⅲ outperforms the MCTC and the MCSBC-Ⅲ with PCM. For an easy comparison of the subjective quality, we also present, in Fig. 16, the enlarged photographs for the 20-th frame of the "MODEL" sequence reconstructed by the proposed MCSBC-Ⅲ and the MCTC. Fig. 16 reveals that the visually objectionable blocking effect, which is prominent by the MCTC, is significantly alleviated in the reconstructed



(a) (b)

Fig. 16. The enlarged photographs of the reconstruted "MODEL" sequence by (a) MCSBC-Ⅲ ; (b) MCTC.

image by the proposed MCSBC-Ⅲ. Hence, it is believed that the proposed MCSBC-Ⅲ outperforms the MCTC and the techniques in [12,13,14], in terms of both the PSNR and the subjective quality.

Ⅵ. Conclusions

In this paper, we described the MCSBC techniques by combining the SBC with the motion compensation in order to alleviate the blocking effect, which is prominent in the reconstructed image by the MCTC at low bit rate. The performances, according to the position of the band splitting, were compared, and it was found that the MCSBC-I provides better performance than the MCSBC-Ⅱ, in terms of the motion compensation and energy packing efficiency.

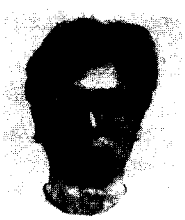
In addition, we proposed an efficient MCSBC technique called the MCSBC-Ⅲ, in which after sub-band decomposition, the motion estimation and compensation were performed only on the LL sub-band, and after applying the DCT to all sub-bands, the adaptive scanning methods together with separate 2-D Huffman code tables were employed for encoding each sub-band. The performance of the proposed MCSBC-Ⅲ was examined intensively by computer simulations on the HDTV image sequences. It was found that the MCSBC-Ⅲ proposed in this paper outperformed the MCTC by about 1.5 dB, in terms of the average PSNR. Moreover, the visually objectionable blocking effect was significantly alleviated by the MCSBC-Ⅲ, providing a good subjective quality. Hence, it is believed that the MCSBC-Ⅲ is a good candidate for the HDTV image data compression.

In this paper, we only dealed with the MCSBC technique based on the DCT coding technique. However, it is worth while to note that by employing the lapped orthogonal transform[8] or the vector quantization[2], further performance improvements would be possible. But these problems need a further study.

참고문헌

1. J. R. Jain and A. K. Jain, "Displacement measurement and its application in interframe image coding," *IEEE Trans. Commun.*, Vol. 29, pp.1799-1808, Dec. 1981.
2. K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications*, San Diego, CA: Academic Press, Inc., 1990.
3. *DIGICIPHER HDTV SYSTEM*, submitted by General Instrument, Co., VideoCipher Division, June 1990.
4. ITU/TS Working Party XV/4, "Description of Ref. Model 8," *Specialist Group On Coding for Visual Telephony*, June 1989.
5. ITU/TS Working Party XV/1, "Video Codec for Audiovisual Services at $p\times 64$ Kbits/s," Recommendation H.261, July 1990.
6. ISO-IEC/JTC1/SC2/WG11, "Test Model 2," *Test Model Editing Committee*, July 1992.
7. B. Ramamurthi and A. Gersho, "Nonlinear space-variant processing of block coded images," *IEEE Trans. Acoust., Speech, and Signal processing*, Vol. 34, pp.1258-1268, Oct. 1986.
8. H. S. Malva and D. H. Staelin, "The LOT: Transform coding without blocking effects," *IEEE Trans. Acoust., Speech, and Signal processing*, Vol. 37, pp.553-559, Apr. 1989.
9. J. W. Woods and S. D. O'Neil, "Sub-band coding of image," *IEEE Trans. Acoust., Speech, and Signal Processing*, Vol. 34, pp.1278-1288, Oct. 1986.

10. J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," in *Proc. 1980 Int. Conf. Acoust., Speech, and Signal Processing*, Denver, Colorado, pp.291-294, Apr. 1980.
11. Truong Q. Nguyen and P. P. Vaidyanathan, "Two-channel perfect-reconstruction FIR QMF structures which yield linear-phase analysis and synthesis filters," *IEEE Trans. on Acoust., Speech, and Signal Processing*, Vol. 37, No. 5, May 1989.
12. K. Irie and R. Kishimoto, "Adaptive sub-band DCT coding for HDTV signal transmission," *Signal Processing: Image Communication*, Vol. 2, pp.333-341, Oct. 1990.
13. H. Gharavi, "Subband coding algorithms for video applications: Videophone to HDTV-conferencing," *IEEE Trans. Cir. and Syst. for Video Technology*, Vol. 1, pp.174-183, June 1991.
14. H. -M. Hang, et. al, "Digital HDTV compression using parallel motion-compensated transform coders," *IEEE Trans. Cir. and Syst. for Video Technology*, Vol. 1, pp.210-221, June 1991.
15. N. Farvardin and J. W. Modestino, "Optimum quantizer performance for a class of non-Gaussian memoryless source," *IEEE Trans. Inform. Theory*, Vol. 30, pp.485-497, May 1984.
16. Herbert Gish and John N. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. on Inform. Theory*, Vol. 14, No. 5, Sep. 1968.
17. N. S. Jayant and P. Noll, *Digital coding of waveform*, Englewood Cliffs, NJ: Prentice-Hall, 1984.



白 勳(Hoon Paek) 정회원

1990년 2월 : 서울대학교 제어계측공학과 졸업(공학사)
1992년 2월 : 서울대학교 제어계측공학과 졸업(공학석사)
1992년 3월~현재 : 서울대학교 제어계측공학과 박사과정
※주관심 분야 : 영상 신호처리, 동영상부호화 및 해석, 영상후처리기법 등

金 麟 澈(Rin Chul Kim) 정회원

1985년 2월 : 서울대학교 제어계측공학과 졸업(공학사)
1987년 2월 : 서울대학교 제어계측공학과 졸업(공학석사)
1992년 2월 : 서울대학교 제어계측공학과 졸업(공학박사)
1992년 10월~1993년 10월 : 서울대학교 자동화시스템 공동연구소(특별연구원)
1992년 8월~1994년 2월 : 대우전자(주) 영상연구소(선임연구원)
1994년 3월~현재 : 한성대학교 정보공학과(전임강사)
※주관심 분야 : 영상신호처리, 실시간 신호처리, 영상부호화 등

李 商 郁(Sang Uk Lee) 정회원

현재: 서울대학교 제어계측공학과 교수
한국통신학회 논문지 제19권 제10호 참조