Efficient Conversion of DCT Coefficients to H.264 Transform Coefficients Using Lapped Transforms

Chang-Woo Lee* ***Regular Member***

ABSTRACT

In order to convert MPEG-2 into H.264 format in ubiquitous communication environments, the efficient conversion of the discrete cosine transform (DCT) coefficients to H.264 transform coefficients is essential. In this paper, two efficient conversion systems are proposed. The proposed systems are composed of two parts. In the first part, the DCT coefficients are denoised using the lapped transform (LT) to reduce the quantization noises and blocking effects. In the second part, the denoised DCT coefficients are converted into the integer transform (IT) coefficients of H.264. Simulation results show that the proposed methods provide visually fine images. Moreover, the computational complexity of the proposed method is reduced compared with the conventional method, since the number of the DCT coefficients, which should be converted, is reduced in the first part.

Key Words : Lapped Transform, H.264, MPEG-2, Efficient Conversion, DCT

Ⅰ. Introduction

Digital images and videos often require conversion in ubiquitous communication environments[1,2]. For example, it is necessary to change the format of the original video, in order to adapt them to the decoding capability of the user's terminal. Since many video contents have been encoded using MPEG(Moving Picture Experts Group)-2 standard and H.264 is the new standard which shows the superior coding efficiency, the efficient conversion from MPEG-2 to H.264 is important[3-5]. However, decoding and then reencoding for the conversion are computationally too complex for real time processing. Thus, the direct conversion methods from MPEG-2 to H.264 have been proposed to reduce the computational complexity[3-5]. A lot of new coding tools, such as 4×4 integer transform (IT) and variable block size motion estimations, are adopted in H.264 to improve the coding efficiency. Since the discrete cosine transform (DCT) is used as a block transform in MPEG-2, the efficient conversion of DCT coefficients to IT coefficients is essential.

The conventional conversion method from DCT coefficients to IT coefficients focused their attention on improving the computational efficiency of the conversion[6-8]. In order to improve the quality of converted videos as well as the computational efficiency, we propose two conversion methods of the DCT coefficients to the IT coefficients using the lapped transform (LT). The method using lapped transforms to resize compressed images was first introduced in [9]. In this paper, the denoising concept before the conversion of the DCT coefficients to IT coefficients is proposed to improve the performance and efficiency of the conversion system. The proposed conversion system consists of two parts. The DCT coefficients are denoised using the LT

※ This study was supported by the Research Fund, 2007 of The Catholic University of Korea.
* School of Information, Communications and Electronics Engineering, The Catholic University of Korea (changwoo@catholic.ac.kr)
논문번호 : KICS2007-04-159, 접수일자 : 2007년 4월 3일, 최종논문접수일자 : 2007년 6월 26일

in the first part and the denoised DCT coefficients are converted to the IT coefficients in the second part. In the first part, the quantization noises and blocking effects are alleviated by removing high frequency components in the LT domain after converting the DCT coefficients to the LT coefficients. The loss caused by removing high frequency components in the LT domain is comparably small due to the high energy compaction of the LT. Moreover, the computational complexity of the proposed method is reduced compared with the conventional method, since the number of the DCT coefficients, which should be converted to the IT coefficients, has been reduced during the denoising process.

Ⅱ. Conversion of DCT to IT using lapped transforms

The lapped transform (LT) was proposed, in order to reduce the blocking effect and provide higher coding performance than the DCT[10,11]. Since multiple blocks in the spatial domain are transformed into one block in the LT domain, the blocking artifacts, which are observed in images compressed at low bit rates, are reduced to a very low level. Among the various lapped transforms, the LiftLT is one of the linear phase lapped transforms with fast and VLSI-friendly implementations via lifting steps[10]. The structure of the LiftLT is simpler than that of the other LTs, since the lifting steps in the LiftLT are composed of only additions and shift operations. It also shows the excellent energy compaction property.

Two systems are proposed for the conversion from the DCT coefficients to the IT coefficients using the LT, as is shown in Fig. 1. In the first system, eight point DCT coefficients are converted into the LT coefficients, and by removing two highest frequency coefficients in the LT domain and reconverting the LT coefficients into the DCT coefficients, the DCT coefficients are denoised. Then, six point DCT block are converted into two blocks of four point IT. In the second system, one block of eight point DCT is converted to four blocks of four point DCT and then four point DCT block is converted to four point LT block. Then, the coefficients are denoised in the LT domain by removing the highest frequency component in the LT domain. The LT block is reconverted to the DCT block. The first system is simpler than the second system, while the second system reduces blocking effects more efficiently by using smaller blocks. The detailed block diagram for the denoising part is depicted in Fig. 2, where high frequency components are removed after converting the DCT block to the LiftLT block and then the LiftLT block is reconverted to the DCT block. By removing high frequency components in the LiftLT domain, the quantization noises and blocking effects are reduced. The mathematical analysis for the performance improvement is given in Section 3. Due to the high energy compaction of the LiftLT, the performance degradation caused by removing the high frequency components in the LiftLT domain is comparably small. The denoising system proposed in this paper can also be used alone to reduce the quantization noises and blocking effects.

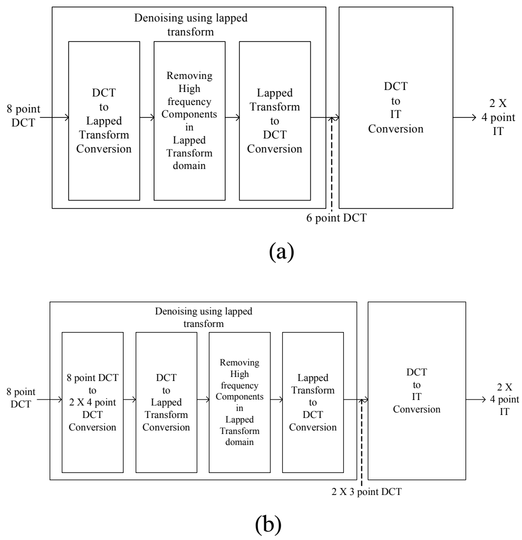
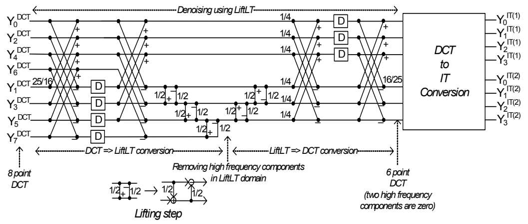
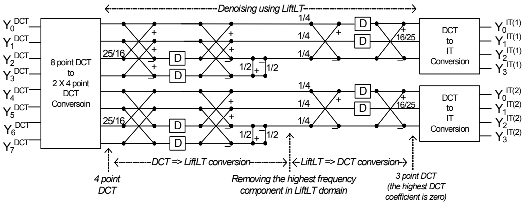


Fig. 1. Proposed systems for the conversion of the DCT coefficients to the IT coefficients (a) Proposed system 1 (b) Proposed system 2



(a)



(b)

Fig. 2. Denoising in the LiftLT domain (a) Proposed system 1 (b) Proposed system 2

The conversion from the DCT to the IT can be carried using the conversion matrix, $T_{DCT\text{-}ITM}$, which can be calculated using the product of T_{ITM} and T^t_{DCTM}, where T_{ITM} and T^t_{DCTM} are M point IT matrix and the transpose of M point DCT matrix, respectively[6]. The value of M is eight or four for the first and the second systems, respectively. Since the number of the DCT coefficients is reduced in the first part of the proposed system, the computational complexity for the conversion of DCT to IT is reduced. The number of multiplications required in the proposed system including the denoising part and conversion part is smaller than the conventional conversion system, as is analyzed in Section 4.

Ⅲ. Performance analysis of the proposed system

The performance of the proposed system is analyzed using the fact that two dimensional transform can be implemented by applying the following one dimensional transform in the row and column directions[10,11]. The $M\times 2M$ LiftLT matrix, T_{LiftLT_M}, is composed of two $M\times M$ matrices, $T_{POST_M(1)}\cdot T_{DCT_M}$ and $T_{POST_M(2)}\cdot T_{DCT_M}$, where T_{DCT_M} is the $M\times M$ DCT matrix whose rows are the bases of the DCT and $T_{POST_M(i)}$ is the $M\times M$ post-processing matrix for the LiftLT. Thus, the vector of the LiftLT coefficients, $\underline{y}_i^{LiftLT_M}$, is given by

$$\underline{y}_i^{LiftLT_M} = T_{POST_M(1)}\cdot \underline{y}_i^{DCT_M} + T_{POST_M(2)}\cdot \underline{y}_{i+1}^{DCT_M} \tag{1}$$

where $\underline{y}_i^{DCT_M}$ is the vector of size M in the DCT domain. The inverse LiftLT is performed by the preprocessing operation using the matrix, $T_{PRE_M(i)}$ $(i=1 \text{ or } 2)$, and the inverse DCT. Thus, the denoised DCT coefficients, $\hat{\underline{y}}_{i+1}^{DCT_M^Q}$, which is obtained after removing high frequency components in the LiftLT domain and reconverting them to the DCT domain, can be expressed as

$$\begin{aligned}\hat{\underline{y}}_{i+1}^{DCT_M^Q} &= T'_{PRE_M(1)}\cdot(R_L\cdot T_{POST_M(1)}\cdot \underline{y}_i^{DCT_M^Q} + R_L\cdot T_{POST_M(2)}\cdot \underline{y}_{i+1}^{DCT_M^Q})\\ &\quad + T'_{PRE_M(2)}\cdot(R_L\cdot T_{POST_M(1)}\cdot \underline{y}_{i+1}^{DCT_M^Q} + R_L\cdot T_{POST_M(2)}\cdot \underline{y}_{i+2}^{DCT_M^Q})\end{aligned} \tag{2}$$

where $T'_{PRE_M(i)}$ represents the preprocessing matrix of the inverse LiftLT for the block whose high frequency coefficients are removed. The matrix R_L is defined as

$$R_L = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & 0 & \cdots & \cdots & 0 \\ \vdots & 0 & 1 & 0 & \cdots & \vdots \\ \vdots & \vdots & 0 & 0 & 0 & \vdots \\ \vdots & \vdots & & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 0 \end{pmatrix} \tag{3}$$

where the number of zeros in the diagonal element represents the number of removed high frequency components in the LiftLT domain. Since the quantized DCT vector, $\underline{y}_i^{DCT_M^Q}$, is the sum of the DCT vector and the quantization noise vector, the formula (2) can be rewritten as

$$\hat{\underline{y}}_{i+1}^{DCT_M} + \hat{\underline{n}}_{Q_{i+1}}^{DCT_M} = \begin{pmatrix} T_{D-L-D(1)} & T_{D-L-D(2)} & T_{D-L-D(3)} \end{pmatrix}\cdot \begin{pmatrix} \underline{y}_i^{DCT_M} + \underline{n}_{Q_i}^{DCT_M} \\ \underline{y}_{i+1}^{DCT_M} + \underline{n}_{Q_{i+1}}^{DCT_M} \\ \underline{y}_{i+2}^{DCT_M} + \underline{n}_{Q_{i+2}}^{DCT_M} \end{pmatrix} \tag{4}$$

where $T_{D-L-D(1)}$, $T_{D-L-D(2)}$, and $T_{D-L-D(3)}$ are given by

$$T_{D-L-D(1)} = T'_{PRE_M(1)} \cdot R_L \cdot T_{POST_M(1)} \quad (5)$$

$$T_{D-L-D(2)} = T'_{PRE_M(1)} R_L \cdot T_{POST_M(2)} + T'_{PRE_M(2)} \cdot R_L \cdot T_{POST_M(1)} \quad (6)$$

$$T_{D-L-D(3)} = T'_{PRE_M(2)} \cdot R_L \cdot T_{POST_M(2)} . \quad (7)$$

In the second system, before denoising the DCT coefficients, eight point DCT block should be converted to four blocks of four point DCT using the following relation

$$\begin{pmatrix} \underline{y}_{2i}^{DCT_4^Q} \\ \underline{y}_{2i+1}^{DCT_4^Q} \end{pmatrix} = \begin{pmatrix} T_{DCT8\_4(1)} \\ T_{DCT8\_4(2)} \end{pmatrix} \cdot \underline{y}_{2i}^{DCT_8^Q} \quad (8)$$

where $T_{DCT8\_4(i)}$ is the 4×8 matrix which converts the eight point DCT block to the upper ($i=1$) or lower ($i=2$) four point DCT block. Since the quantized DCT block, $\underline{y}_i^{DCT_8^Q}$, is the sum of the DCT coefficients, $\underline{y}_{i+1}^{DCT_8}$, and quantized noises, $\underline{n}_{Q_{i+1}}^{DCT_8}$, the following relation holds

$$\underline{\hat{y}}_{2i+1}^{DCT_4} + \underline{\hat{n}}_{Q2i+1}^{DCT_4} = \begin{pmatrix} T_{conv(1)} & T_{conv(2)} \end{pmatrix} \cdot \begin{pmatrix} \underline{y}_{2i}^{DCT} + \underline{n}_{Q2i}^{DCT} \\ \underline{y}_{2i+2}^{DCT} + \underline{n}_{Q2i+2}^{DCT} \end{pmatrix} \quad (9)$$

where $T_{conv(1)}$ and $T_{conv(2)}$ are given by

$$\begin{aligned} T_{conv(1)} &= T'_{PRE_4(1)} \cdot R_L \cdot T_{POST_4(1)} \cdot T_{DCT8\_4(1)} \\ &+ T'_{PRE_4(1)} \cdot R_L \cdot T_{POST_4(2)} \cdot T_{DCT8\_4(2)} \\ &+ T'_{PRE_4(2)} \cdot R_L \cdot T_{POST_4(1)} \cdot T_{DCT8\_4(2)} \end{aligned} \quad (10)$$

and

$$T_{conv(2)} = T'_{PRE_M(2)} \cdot R_L \cdot T_{POST_M(2)} \cdot T_{DCT8\_4(2)} \quad (11)$$

The reconstructed vector in the spatial domain, $\underline{\hat{x}}_{i+1}$, can be obtained by

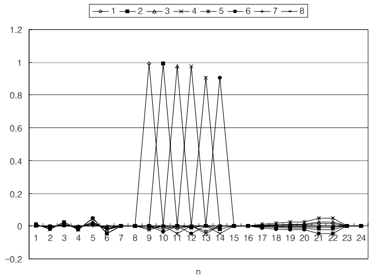
$$\underline{\hat{x}}_{2i+1} + \underline{\hat{n}}_{Q2i+1} = \begin{pmatrix} T_{DCT_4}^t \cdot T_{conv(1)} & T_{DCT_4}^t \cdot T_{conv(2)} \end{pmatrix} \cdot \begin{pmatrix} \underline{y}_{2i}^{DCT} + \underline{n}_{Q2i}^{DCT} \\ \underline{y}_{2i+2}^{DCT} + \underline{n}_{Q2i+2}^{DCT} \end{pmatrix} \quad (12)$$

In order to analyze the performance of the proposed conversion system, the rows of $((T_{D-L-D(1)}\ T_{D-L-D(2)}\ T_{D-L-D(3)}))$ in formula (4), which are the impulse responses for the conversion from the quantized DCT coefficients to the denoised DCT coefficients, are calculated using formulas (5)~(7) and are depicted in Fig. 3. Since the length of the impulse response is longer than that of the input, the overlap occurs in the neighboring blocks. Thus, the quantization noises are mixed, reducing blocking effects. However, the loss caused by removing the high frequency components in the LiftLT domain is small due to the high energy compaction of the LiftLT. Assuming that the quantization noise for each DCT coefficient is independent to each other and the mean value of the quantization noise is zero, the mean square value of quantization noises is given by

$$\sum_i E(n_{Q_i}^2) = \sum_i E\{(\sum_j a_{ij} \cdot n_{Q_j})^2\} = \sum_i \sum_j a_{ij}^2 \cdot E(n_{Q_j}^2) \quad (13)$$

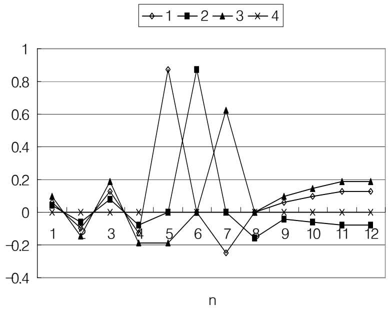
where a_{ij} is the element of each row i in $T_{D-L-D(k)}$. The value of $\sum_j a_{ij}^2$ for each i varies from 0.836 to 0.987 for the first proposed system and from 0.627 to 0.918 for the second proposed system, respectively. Thus, assuming that $E(n_{Q_i}^2)$ in (13) are equal for each i, the mean square value of quantization noises for the proposed systems is reduced than that for the conventional system.

The effect of the quantization noises to the reconstructed value in the spatial domain can be analyzed using formula (12). The impulse responses, which are the rows of $(T_{DCT_M}^t \cdot T_{conv(1)}\ \ T_{DCT_M}^t \cdot T_{conv(1)})$, are given in Fig. 4. Since the length of the impulse response is longer than that of the input, the overlap occurs in the neighboring blocks, thus reducing the blocking effects.



(a)

Fig. 3. Impulse responses for the denoising system (DCT to DCT domain) (a) Proposed system 1



(b)

Fig. 3. Impulse responses for the denoising system(DCT to DCT domain) (b) Proposed system 2

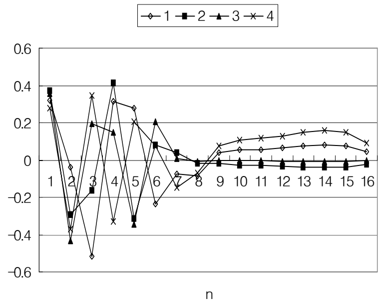


Fig. 4. Impulse responses for the denoising system (DCT to spatial domain, Proposed system 2)

Ⅳ. Complexity analysis and simulations results

The complexity of the proposed system is compared with the conventional conversion system by calculating the number of operations required for the conversion. The proposed conversion system is composed of the preprocessing for denoising and the conversion from the DCT to the IT, while the conventional conversion system only requires the conversion from the DCT to the IT. The conversion system proposed in [6] is used as a standard of the conventional conversion system, since it has the efficient structure and is widely used. The number of multiplications for real numbers to convert one block of eight point DCT to two blocks of four point IT is calculated as 22, and the number of additions is also 22[6]. For the conversion of 8×8 DCT blocks in the row and column directions, the number of operations (multiplication or addition) is $352(=22\times8+22\times8)$.

In the first proposed system, the number of the DCT coefficients for the conversion is reduced to six in the first part of denoising. Thus, the number of multiplications for the conversions is reduced to $224(=16\times6+16\times8)$. The number of additions is also 224. However, the proposed system requires the preprocessing for denoising. The operations for preprocessing are composed of only additions, shift operations and integer multiplications. The total numbers of operations are given in Table 1.

In the second proposed system, $T_{DCT8\_4(1)}$ and $T_{DCT8\_4(2)}$ in formula (8) can be decomposed into $(A+B)/\sqrt{2}$ and $(A-B)/\sqrt{2}$, respectively, where

$$A=\begin{pmatrix}1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0.416 & 0 & 0.791 & 0 & -0.352 & 0 & 0.278\\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0.023 & 0 & -0.098 & 0 & 0.490 & 0 & 0.866\end{pmatrix} \tag{14}$$

and

$$B=\begin{pmatrix}0 & 0.906 & 0 & -0.318 & 0 & 0.213 & 0 & -0.180\\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\ 0 & -0.075 & 0 & 0.513 & 0 & 0.768 & 0 & -0.375\\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\end{pmatrix} \tag{15}$$

Thus, the number of multiplications required to convert one block of eight point DCT to two blocks of four point DCT is 16 and the number of additions is 20. The denoising system using the LiftLT, which is shown in Fig. 2(b), requires 28 additions and 4 shift operations for real numbers and 2 multiplications for integers. Finally, four point DCT block is converted to four point IT block using the conversion matrix T_{DCT-_4}, which is the product of T_4 and $T^t_{DCT_4}$. Since the highest frequency component of the DCT block is removed, the numbers of multiplications, additions and shift operations are 2, 2 and 2, respectively. The total numbers of operations are given in Table 1. As is shown in Table 1, the number of multiplications is reduced in the proposed system, while additions and shift operations are added. The multiplications are much more complex than the additions and shift operations. Thus, the proposed system is not complex compared with the conventional conversion system.

Table 1. The number of operations for the conversion

	Conventional system	Proposed system 1	Proposed system 2
The number of operations for the conversion from DCT to IT	$[(22\times8+22\times8)M + (22\times8+22\times8)A]$	$[(37\times8+37\times6)A + (10\times8+10\times6+36) + (2\times8+2\times6)IM] + [(16\times6+16\times8)M + (16\times6+16\times8)A]$	$[(16\times8+16\times8)M + (20\times8+20\times8)A] + [(14\times2\times8+14\times2\times6)A + (2\times2\times8+2\times2\times6+36)S + (1\times2\times8+1\times2\times6)IM] + [(2\times2\times6+2\times2\times8)M + (2\times2\times6+2\times2\times8)S]$
The total number of operations	$352M+352A$	$224M+742A+176S+28IM$	$312M+712A+148S+28IM$

(Conventional system: [DCT to IT onversion], proposed system 1: [Denoising]+[DCT to IT conversion], Proposed system 2: [8x8 to 4x4 conversion]+[Denoising]+[DCT to IT conversion], M: multiplications of real numbers, A: additions of real numbers, S: shift operations, IM: multiplications of integers)

Table 2. PSNR between the original frames and converted frames (in dB)

CIF sequences	Quantizer	Conventional system	Proposed system 1	Proposed system 2
Football	Q1	32.433	32.702	32.685
	Q2	30.378	30.673	30.913
Foreman	Q1	32.712	32.480	32.322
	Q2	31.007	31.010	31.152
Stefan	Q1	30.218	30.318	29.569
	Q2	27.812	28.108	27.917

(Conventional system: Quantized in the DCT domain & conversion from DCT to IT domain, Proposed system: Quantized in the DCT domain & denoised before conversion from DCT to IT domain))

In order to measure the performance of the proposed system, the CIF sequences are used as the test sequences. The simulations are focused on the conversion from the DCT to IT domain. The first frame of each CIF sequence is transformed into 8×8 DCT blocks and the DCT coefficients are quantized using the quantizer whose step size is proportional to the quantization table of MPEG.

Then the quantized DCT blocks are converted to four 4×4 IT blocks using the conventional

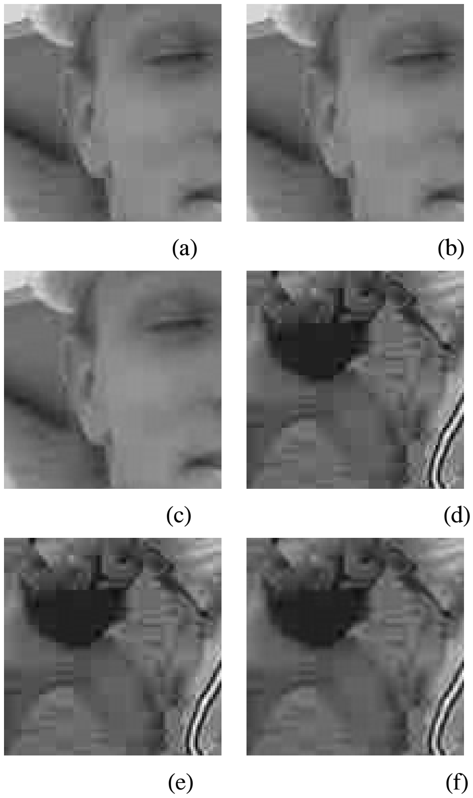


Fig. 5. Reconstructed frames
(a) Conventional system (Quantized in the DCT domain, conversion from DCT to IT domain, foreman, Q1)
(b) Proposed system 1 (Quantized in the DCT domain, denoised before conversion from DCT to IT domain, foreman, Q1)
(c) Proposed system 2 (Quantized in the DCT domain, denoised before conversion from DCT to IT domain, foreman, Q1)
(d) Conventional system (Quantized in the DCT domain, conversion from DCT to IT domain, football, Q2)
(e) Proposed system 1 (Quantized in the DCT domain, denoised before conversion from DCT to IT domain, football, Q2)
(f) Proposed system 2 (Quantized in the DCT domain, denoised before conversion from DCT to IT domain, football, Q2)

conversion system and the proposed conversion system, respectively. The PSNR performance is calculated by reconverting the images in the IT domain to the spatial domain and comparing them with the original images. The magnified images, which are reconstructed from the IT domain to the spatial domain, are shown in Fig. 5 and the PSNR results are given in Table 2. In Table 2, Q1 and Q2 mean that the step sizes of quantizers are two or three times as large as the MPEG quantization table, respectively. As can be seen, the proposed system gives better results both subjectively and objectively. Especially, the blocking effects are reduced significantly in the proposed system 2 by using smaller block when the quantization noise is large.

Ⅴ. Conclusion

In this paper, two conversion methods for the conversion of DCT coefficients to H.264 transform coefficients were proposed. By denoising the DCT coefficients before the conversion from the DCT to IT, the proposed system showed the improved performance both objectively and subjectively. The denoising system is efficiently implemented using the LiftLT. Moreover, by reducing the number of the DCT coefficients to be converted, the number of multiplications is reduced in the proposed system. Thus, it is concluded that the proposed system can be used as the efficient system for the conversion of the DCT coefficients to H.264 transform coefficients.

References

〔1〕 S. F. Chang and A. Vetro, “Video adaptation: concepts, technologies and open issues,” *Proceedings of the IEEE*, Vol. 93, pp.148-158, 2003.

〔2〕 A. Vetro, C. Christopoulos and H. Sun, “Video transcoding architectures and techniques: An overview,” *IEEE Signal Processing Magazine*, Vol. 20, pp.18-29, 2003.

〔3〕 G. F. Escribano, P. Cuenca, L. O. Barbosa and A. Garrido, “A fast intra-frame prediction algorithm for MPEG-2/H.264 video transcoders,” *IEEE International Conference on Image Processing*, pp.III.684-III.687, 2005.

〔4〕 X. Lu, A. M. Tourapis, P. Yin and J. Boyce, “Fast mode decision and motion estimation for H.264 with a focus on MPEG-2/H.264 transcoding,” *IEEE Symposium on Circuits and Systems*, pp.II.1246-II.1249, 2005.

〔5〕 Y. Su, J. Xin, A. Vetro and H. Sun, “Efficient MPEG-2 to H.264/AVC intra transcoding in transform-domain,” *IEEE Symposium on Circuits and Systems*, pp.II.1234-II.1237, 2005.

〔6〕 J. Xin, A. Vetro and H. Sun, “Converting DCT coefficients to H.264/AVC transform coefficients,” *Lecture Notes in Computer Science*, Vol. 3332, pp.939-946, 2004.

〔7〕 C. Y. Park and N. I. Cho, “A fast algorithm for the conversion of DCT coefficients to H.264 transform coefficients,” *IEEE International Conference on Image Processing*, pp.III.664-III.667, 2005.

〔8〕 K. T. Fung and W. C. Siu, “Conversion between DCT coefficients and IT coefficients in the compressed domain for H.263 to H.264 video transcoding,” *IEEE International Conference on Image Processing*, pp.III.57-III.60, 2005.

〔9〕 C. W. Lee and H. Ko, “Arbitrary resizing of images in DCT domain using lapped transforms,” *Electronics Letters*, Vol. 41, pp.1319-1320, 2005.

〔10〕 T. D. Tran, “The LiftLT: Fast-lapped transforms via lifting steps,” *IEEE Trans. Signal Processing Letters*, Vol. 7, pp.145-148, 2000.

〔11〕 H. S. Malvar, “Lapped transforms for efficient transform/subband coding,” *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 38, pp. 969-978, 1990.

이 창 우 (Chang-Woo Lee) 정회원

현재 가톨릭대학교 정보통신전자공학부 교수

<관심분야> 영상 통신, 영상 압축