

논문 2008-45SP-3-10

잔여 신호의 상관성에 기반한 선택 변환

(Alternative Transform Based on the Correlation of the Residual Signal)

임성창*, 김대연**, 이영렬***

(Sung-Chang Lim, Dae-Yeon Kim, and Yung-Lyul Lee)

요약

기존의 영상 부호화 표준들보다 높은 압축 성능을 얻기 위해 부호화 효율 측면에서 우수한 기술들이 가장 최근에 완성된 영상 부호화 표준인 H.264/AVC에 채택되었다. 가변 블록 단위의 움직임 예측/보상과 다양한 방향성을 통한 화면 내 예측 방법 등의 영상 부호화 기술들의 발달에도 불구하고, 이산 여현 변환은 초기 영상 부호화 표준에서부터 계속적으로 사용되고 있다. 일반적으로 실제 영상 신호가 부호화될 때 생성되는 잔여 신호의 상관 계수 값은 0.5 미만이다. 하지만, 이러한 상관 계수 값의 범위는 이산 여현 변환이 최적의 성능을 나타내는 범위가 아니며, 상관 계수의 범위가 -0.5에서 0.5일 경우 차선의 변환인 이산 정현 변환이 이산 여현 변환과 함께 영상 부호화에 사용될 수가 있다. 본 논문에서는 이산 정현 변환과 H.264/AVC에서의 정수 여현 변환 중 최적의 변환을 율-왜곡 최적화 과정을 이용하여 선택적으로 사용하는 선택 변환을 제안한다. 제안하는 방법을 통해 H.264/AVC의 JM 10.2와 비교하여 상대적으로 높은 비트율에서 최대 0.71 dB의 BD-PSNR 향상을 얻을 수가 있었다.

Abstract

Many predominant video coding tools in terms of coding efficiency were adopted in the latest video coding standard, H.264/AVC. Regardless of development of these predominant video coding tools such as the variable block-size motion estimation/compensation, intra prediction based on various directions, and so on, the discrete cosine transform has been continuously used starting from the early video coding standards. Generally, the correlation coefficient of the residual signal is usually less than 0.5 when this residual signal is actually encoded. In this interval of correlation coefficient, the discrete cosine transform does not show the optimal coding gain, and the discrete sine transform which is a sub-optimal transform when the correlation coefficient is in the interval from -0.5 to 0.5 can be used in conjunction with the discrete cosine transform in the video coding. In this paper, an alternative transform that alternatively uses the discrete sine transform and integer cosine transform in H.264/AVC by using rate-distortion optimization is proposed. The proposed method achieves a BD-PSNR gain of up to 0.71 dB compared to H.264/AVC JM 10.2 at relatively high bitrates.

Keywords: Alternative Transform, Discrete Cosine Transform, Discrete Sine Transform, H.264/AVC

I. 서론

영상 부호화 표준들 중 MPEG (Moving Picture Experts Group)과 VCEG (Video Coding Experts Group)의 공동 개발로 가장 최근에 완성된 H.264/AVC (Advanced Video Coding)에는 기존의 영상 부호화 표

준들보다 높은 압축 성능을 얻기 위해 가변 블록 단위의 움직임 예측/보상 (Variable Block-size Motion Estimation/Compensation), 다양한 방향성을 통한 화면 내 예측 (Intra Prediction), 다중 참조 프레임 (Multiple Reference Frame), 이산 여현 변환 (Discrete Cosine Transform)에 기반한 4x4 정수 여현 변환 (Integer Cosine Transform), 그리고 표준에 직접적으로 포함되지 않았지만 사용이 권고되고 있는 율-왜곡 최적화 (Rate-Distortion Optimization) 등과 같은 기술들이 채택되었다^[1]. 또한, 보다 좋은 화질, 보다 큰 영상과 같은 전문적인 영상 환경을 지원하기 위해 H.264/AVC 개정 (Amendment) 1: FRExt (Fidelity Range Extensions)가

* 정회원, ** 학생회원, *** 평생회원, 세종대학교 컴퓨터공학과 DMS 연구실 (DMS Lab., Department of Computer Engineering, Sejong University)

※ 본 논문은 서울시 산학연 협력사업(11098)의 지원으로 수행되었음.

접수일자: 2007년10월5일, 수정완료일: 2008년4월14일

2004년 7월에 완성되었다. FRExt에서는 다양한 컬러 형식(Color Format), 8 비트 크기 이상의 영상 표본(Sample)을 지원하며, 추가적으로 전문적 영상 환경에서의 부호화 효율 향상을 위해 8x8 단위의 화면 내 예측, 8x8 정수 여현 변환 등을 제공하고 있다^[2].

하지만, 이러한 영상 부호화 기술들의 발달에도 불구하고, 이산 여현 변환에 기반한 정수 여현 변환은 초기 영상 부호화 표준에서부터 계속적으로 사용되고 있다. 이산 여현 변환은 영상이 1 차 고정 마르코프 수열(First-order Stationary Markov Sequence)의 형태를 가지며, 상관성의 척도인 상관 계수(Correlation Coefficient)의 값이 0.5 보다 크고 1.0 보다 작을 경우 통계적으로 최적의 공간적 압축 성능을 보이는 K-L 변환(Karhunen-Loeve Transform)과 비슷한 성능을 보인다^[3~4]. 하지만, H.264/AVC에서의 가변 블록 단위의 움직임 예측/보상, 화면 내 예측과 같은 기술들이 최신 영상 부호화 표준에 쓰임에 따라, 잔여 신호(Residual Signal)의 상관 계수가 -1.0 에서 0.5 사이의 값을 가지는 경우가 발생할 수 있다. 이 때, 영상 신호의 상관 계수가 -0.5 에서 0.5 사이일 경우 이산 여현 변환보다 더 뛰어난 공간적 압축 성능을 나타내는 이산 정현 변환(Discrete Sine Transform)이 이산 여현 변환의 대체 방법으로 영상 부호화 방법에 고려될 수가 있다^[4~6].

이산 정현 변환은 다른 유니터리 변환(Unitary Transform)과 비슷하게 대칭성(Symmetry)과 직교성(Orthogonality)의 특성을 지니고 있으며, 이산 여현 변환과 비슷하게 총 8 가지의 형태로 존재하고 있다. 그리고 이산 정현 변환은 주변 블록들 간의 중복성을 제거하거나^[7], 영상 내에서 블록킹 현상(Blocking Artifact)를 줄이기 위해 사용되고 있지만^[8], 완벽한 정수 형태의 기저 벡터(Basis Vector)를 가지는 형태가 아직 존재하지 않아^[9], 실제 영상 부호화 방법이나 다양한 응용 분야에 쓰이는데 어려움이 있으며, 기존 영상 부호화 표준에서의 완벽하지 않은 화면 간/내 예측 방법으로 인해 주목할 만한 부호화 효율을 보여주지 못했다^[10].

따라서 본 논문에서는 다양하고 세밀한 화면 간/내 예측 방법과 율-왜곡 최적화를 통해서 우수한 부호화 효율을 보이는 H.264/AVC에 잔여 신호의 상관성을 고려하여 정수 여현 변환과 이산 정현 변환을 율-왜곡 최적화 내에서 선택적으로 사용하는 방법을 제안한다.

다음 II 장에서는 기존의 H.264/AVC에서의 4x4 정수 여현 변환과 그에 따르는 양자화(Quantization)에

대해서 소개하며, III 장에서는 H.264/AVC에서 사용하기 위해서 4x4/8x8 블록 단위로 설계된 이산 정현 변환이 제시된다. IV 장에서는 정수 여현 변환과 이산 정현 변환을 H.264/AVC 부호화기(Encoder)의 율-왜곡 최적화에서 선택적으로 사용하기 위한 방법이 제시되며, V 장에서는 제안하는 선택 변환(Alternative Transform)방법과 기존 H.264/AVC 참조 소프트웨어(Reference Software)인 JM (Joint Model) 10.2^[11]의 방법이 비교 실험되며, 마지막으로 VI 장의 결론과 함께 본 논문의 끝을 맺는다.

II. H.264/AVC의 4x4 정수 여현 변환과 양자화

H.264/AVC에 채택한 4x4 정수 여현 변환^[12]이 기반한 이산 여현 변환은 영상이 1차 고정 마르코프 수열의 형태를 가지며, 상관 계수의 값이 0.5 와 1.0 사이에 존재할 경우, 통계적으로 최적의 공간적 압축 성능을 보이는 K-L 변환과 비슷한 성능을 보이며, 이러한 특징으로 인해 기존의 많은 영상 부호화 표준에서부터 사용되고 있다.

H.264/AVC 표준에 따르면, 현재 부호화하는 블록에 대해 화면 간/내 예측이 수행되고 난 뒤, 블록 내 잔여 신호의 공간적 상관성(Spatial Correlation)을 감소시키고, 블록 내의 잔여 신호 에너지의 압축률을 높이기 위한 방법으로 이산 여현 변환에 기반한 4x4 정수 여현 변환이 사용된다. 이 때, 기존의 8x8 블록 단위의 실수(Real Number)형 이산 여현 변환과 달리 더 낮은 계산 복잡도를 위해 4x4 단위로 변환 기저 벡터들이 정수로 근사화된 방법이 쓰이고 있다. 또한, 정수 여현 변환 이후 변환된 계수(Transformed Coefficient)에 적용되는 양자화 과정에서의 곱셈 연산을 더욱 줄이기 위한 방법으로 정수 여현 변환과 양자화가 통합되어 사용된다.

다음 식 (1)은 블록 내에서 공간적인 상관성을 감소하고 4x4 정수 여현 변환된 계수 행렬 Y 를 얻기 위해, 화면 간/내 예측이 수행된 4x4 블록 단위의 잔여 신호 X 에 대한 정수 여현 변환을 행렬식으로 나타내고 있다.

$$Y = C_f X C_f^T = \begin{pmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} [X] \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \end{pmatrix} \quad (1)$$

여기서 C_f 는 정수로 근사화된 1차원 정수 여현 변환 행렬을 나타내며, C_f^T 는 행렬 C_f 가 전치(Transposition)

표 1. 양자화에서의 곱셈 인자
Table 1. Multiplication factor for quantization in H.264/AVC.

QP%6	(0,0),(2,0),(2,2),(0,2)의 위치	(1,1),(1,3),(3,1),(3,3)의 위치	그 밖의 4x4 행렬에서의 위치
0	13107	5243	8066
1	11916	4660	7490
2	10082	4194	6554
3	9362	3647	5825
4	8192	3355	5243
5	7282	2893	4559

표 2. 양자화에서의 크기조정 인자
Table 2. Scaling factor for dequantization in H.264/AVC.

QP%6	(0,0),(2,0),(2,2),(0,2)의 위치	(1,1),(1,3),(3,1),(3,3)의 위치	그 밖의 4x4 행렬에서의 위치
0	10	16	13
1	11	18	14
2	13	20	16
3	14	23	18
4	16	25	20
5	18	29	23

된 행렬을 나타낸다.

위와 같은 행렬식을 이용해 정수 여현 변환 수행 후, 4x4 블록 단위로 정수 여현 변환된 계수 행렬 Y 는 다음 식 (2)와 같은 양자화를 통해서 양자화된 변환 계수 $Z_{(i,j)}$ 를 생성하게 된다.

$$Z_{(i,j)} = \text{sgn}(Y_{(i,j)}) \cdot (Y_{(i,j)} \cdot MF_{(i,j)} + DZ) \gg \left(15 + \left\lfloor \frac{QP}{6} \right\rfloor \right) \quad (2)$$

$Y_{(i,j)}$ 는 4x4 블록 단위의 행렬 (i,j)위치에서의 정수 여현 변환된 계수를 나타내며, QP 는 양자화 매개변수 (Quantization Parameter)를 나타내며, $MF_{(i,j)}$ 는 곱셈 인자 (Multiplication Factor)를 나타낸다. 그리고 $\text{sgn}()$ 는 부호 함수 (Sign Function), DZ 는 데드 존 (Dead Zone)을 조정하는 역할을 한다. 표 1은 양자화에서의 곱셈 인자를 나타내며, (0,0), (1,0), ..., (3,3)은 4x4 행렬의 위치 (i,j)를 나타낸다^[11].

양자화된 (Quantized) 정수 여현 변환 계수 $Z_{(i,j)}$ 는 zigzag 스캐닝 (Zigzag Scanning) 과정과 엔트로피 부호화 (Entropy Coding)를 거쳐 비트스트림 (Bitstream)으로 변환되어 전송되거나, 저장된다. 부호화 과정과 반대로 복호화 (Decoding) 과정에서는 변환된 비트스트림에

대해 엔트로피 복호화를 수행한 후 역양자화 (Dequantization)와 이산 여현 변환에 기반한 4x4 정수 여현 역변환 (Inverse Integer Cosine Transform)을 수행하게 된다. 엔트로피 복호화 후 양자화된 정수 여현 변환 계수들에 다음 식 (3)과 같은 역양자화가 수행된다.

$$Y'_{(i,j)} = Z_{(i,j)} \cdot SF_{(i,j)} \cdot 2^{\left\lfloor \frac{QP}{6} \right\rfloor} \quad (3)$$

복호화 과정에서 $Y'_{(i,j)}$ 는 4x4 블록 단위의 행렬 (i,j) 위치에서 역양자화가 수행된 정수 여현 변환된 계수를 나타내며, 이 신호는 부호화 과정에서의 신호와는 달리 양자화와 역양자화 과정을 통해 신호가 왜곡되어 있으며, $SF_{(i,j)}$ 는 역양자화에서의 크기조정 인자 (Scaling Factor)를 나타낸다.

표 2는 역양자화 과정에서의 크기조정 인자를 나타내며, (0,0), (1,0), ..., (3,3)은 표 1의 곱셈 인자와 마찬가지로 4x4 행렬에서의 위치를 나타낸다^[11].

4x4 정수 여현 변환된 계수 행렬 Y' 는 식 (4)와 같은 정수 여현 역변환 과정을 거쳐 4x4 블록 단위의 복원된 잔여 신호 행렬 X' 로 표현된다.

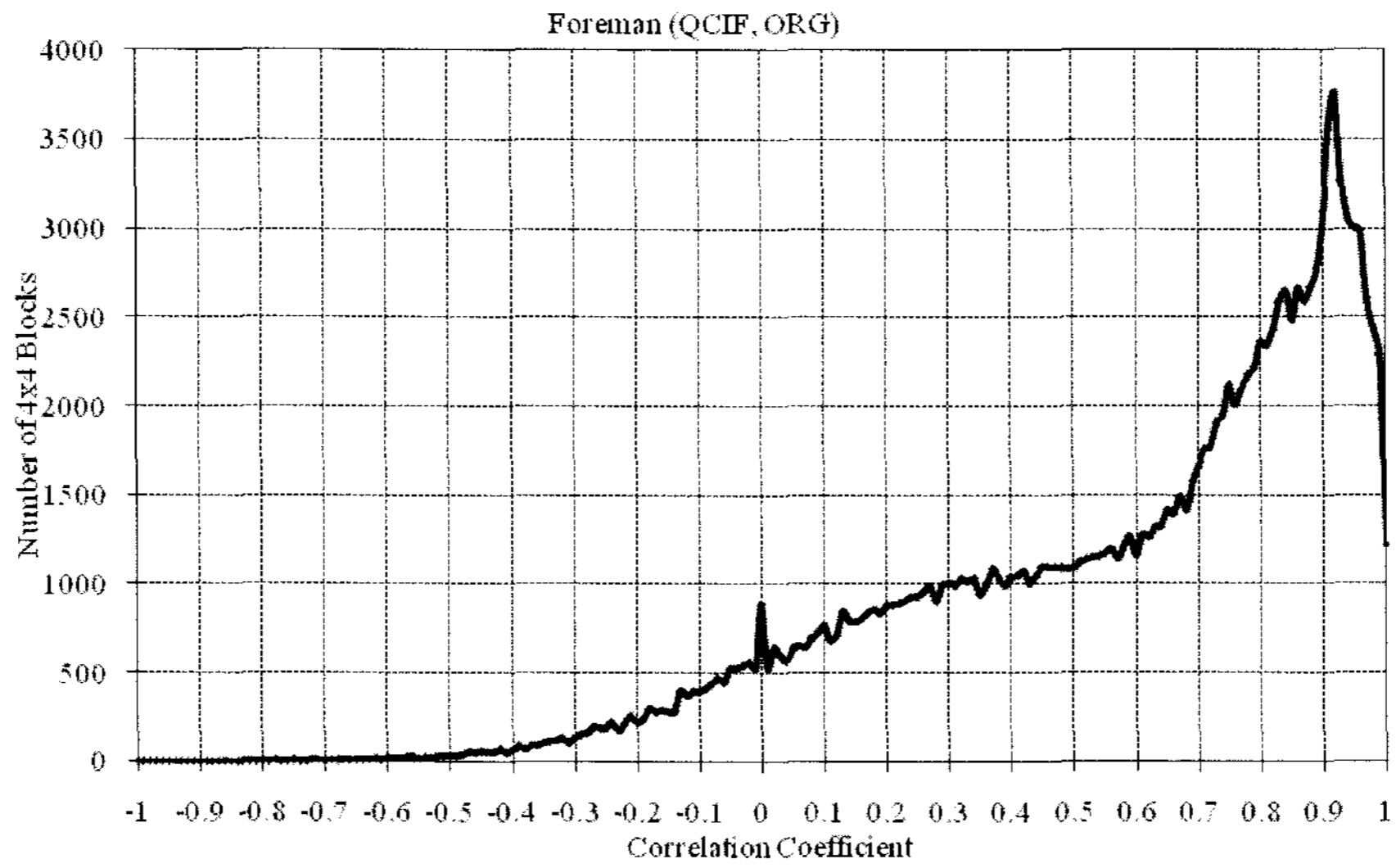
$$X' = C_i^T Y' C_i = \begin{pmatrix} \begin{bmatrix} 1 & 1 & 1 & \frac{1}{2} \\ 1 & \frac{1}{2} & -1 & -1 \\ 1 & -\frac{1}{2} & -1 & 1 \\ 1 & -1 & 1 & -\frac{1}{2} \end{bmatrix} & [Y'] & \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \frac{1}{2} & -\frac{1}{2} & -1 \\ 1 & -1 & -1 & 1 \\ \frac{1}{2} & -1 & 1 & -\frac{1}{2} \end{bmatrix} \end{pmatrix} \quad (4)$$

여기서 C_i 는 이산 여현 역변환에 기반하며 정수로 근사화된 1 차원 정수 여현 역변환 행렬을 나타내며, 정수 여현 변환과 마찬가지로 C_i^T 는 행렬 C_i 가 전치된 행렬을 나타낸다.

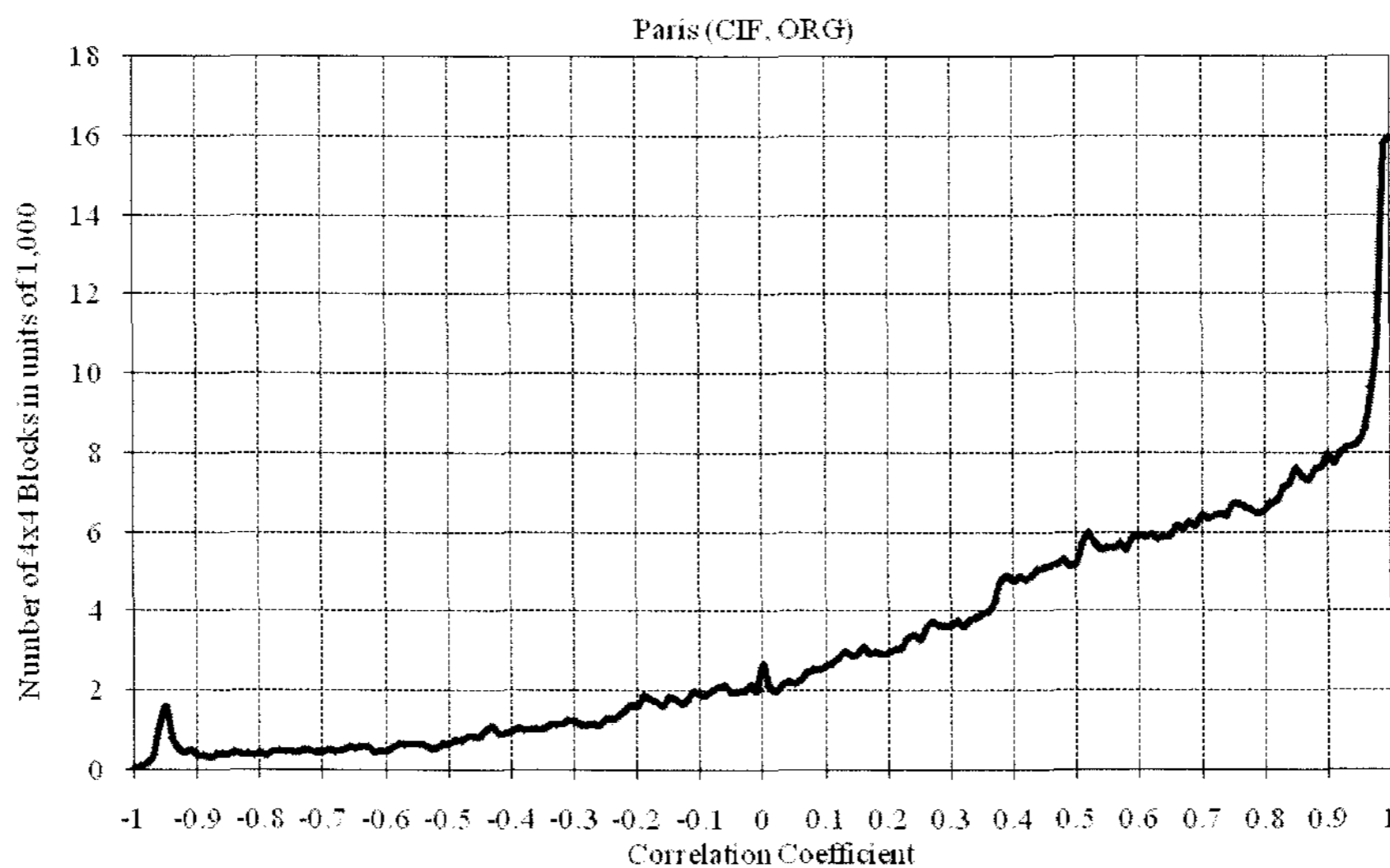
4x4 블록 단위의 복원된 잔여 신호 행렬 $X'_{(i,j)}$ 는 다음 식 (5)와 같은 후-크기조정 (Post-scaling) 과정을 거쳐 4x4 블록 단위의 최종 복원된 잔여 신호 행렬 $X''_{(i,j)}$ 로 표현된다.

$$X''_{(i,j)} = \left\lfloor \frac{X'_{(i,j)}}{64} + 0.5 \right\rfloor \quad (5)$$

한편, H.264/AVC에서는 가변 블록 단위의 움직임 예측/보상, 다양한 방향성을 통한 화면 내 예측과 같은 기술들이 쓰임에 따라, 영상 잔여 신호의 상관 계수가 -1.0 에서 0.5 사이의 값을 가지는 경우가 발생할 수 있



(a) "Foreman" 영상



(b) "Paris" 영상

그림 1. "Foreman"과 "Paris"에서 원 영상 신호의 수평 상관 계수 분포

Fig. 1. 4x4 block distribution of correlation coefficient for original signal in "Foreman" and "Paris" sequences.

다. 이 때, 영상 신호의 상관 계수가 -0.5 에서 0.5 사이 일 경우 이산 여현 변환보다 더 뛰어난 공간적 압축 성능을 나타내는 이산 정현 변환이 이산 여현 변환의 대체 방법으로 H.264/AVC에 고려될 수가 있다.

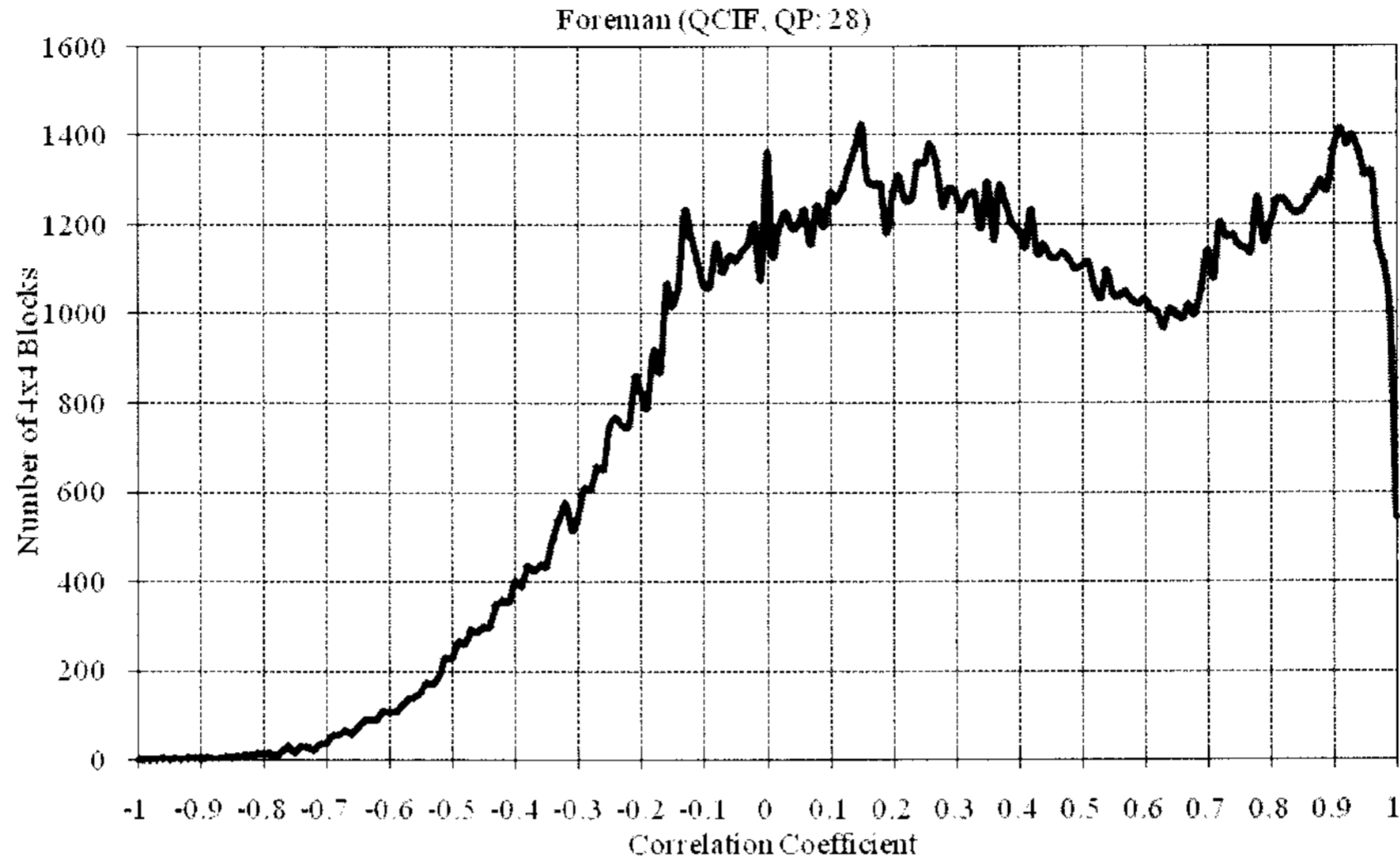
III. H.264/AVC에서의 이산 정현 변환

H.264/AVC에서의 이산 정현 변환 방법이 제시되기 앞서 원 영상 (Original Video Sequence)과 원 영상을 부호화할 때 생성되는 잔여 신호들에서의 상관 계수를 측정하고 이들을 각각 그림 1과 2에 나타내었다. 상관 계수는 다음 식 (6)과 같이 계산된다.

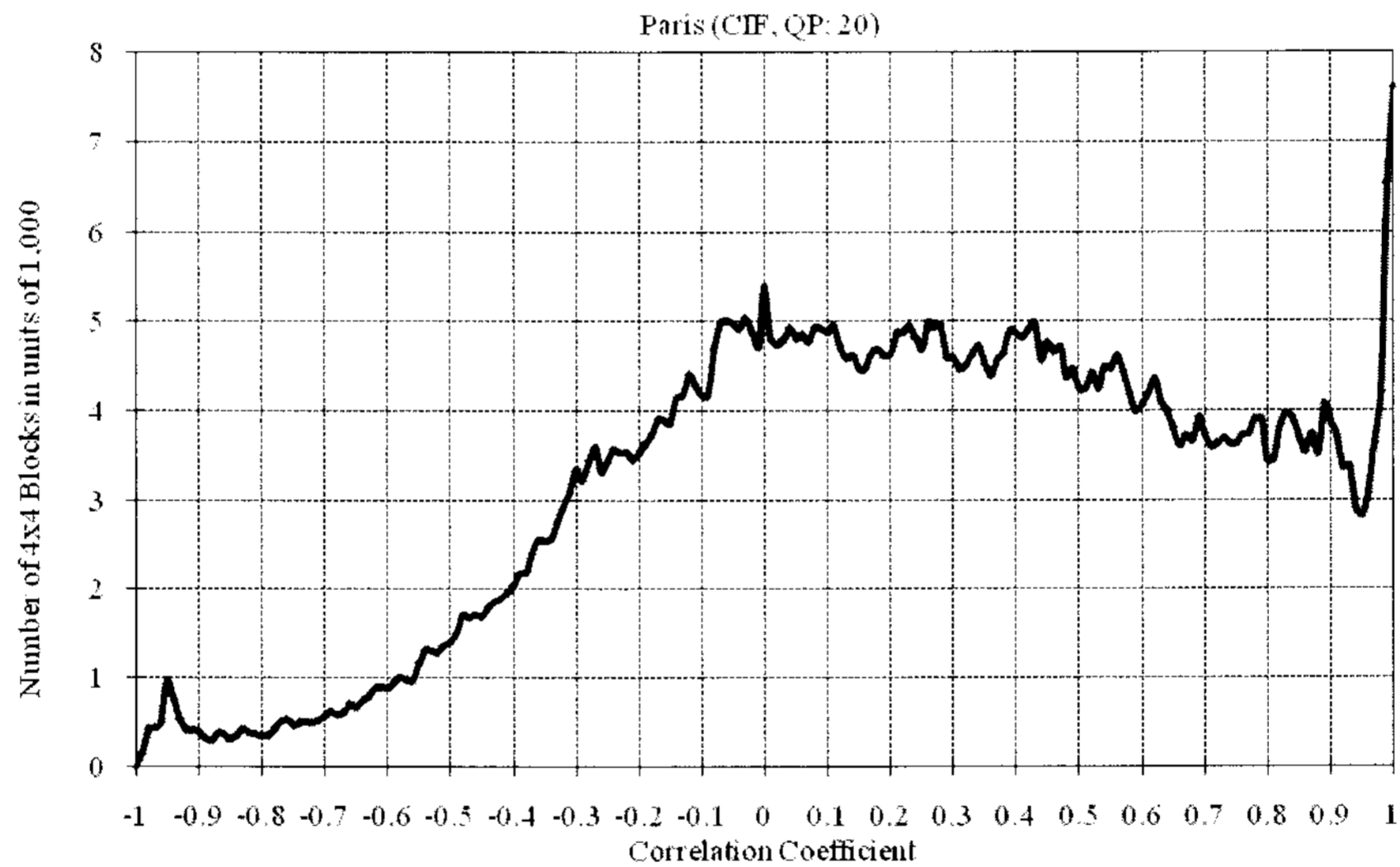
$$\rho_{B^C, B^S} = \frac{\sum_{x=0}^3 \sum_{y=0}^3 (B_{x,y}^C - \bar{B}^C)(B_{x,y}^{LS} - \bar{B}^{LS})}{\sqrt{\sum_{x=0}^3 \sum_{y=0}^3 (B_{x,y}^C - \bar{B}^C)^2} \sqrt{\sum_{x=0}^3 \sum_{y=0}^3 (B_{x,y}^{LS} - \bar{B}^{LS})^2}} \quad (6)$$

$$\rho_{B^C, B^S} = \frac{\sum_{x=0}^3 \sum_{y=0}^3 (B_{x,y}^C - \bar{B}^C)(B_{x,y}^{RS} - \bar{B}^{RS})}{\sqrt{\sum_{x=0}^3 \sum_{y=0}^3 (B_{x,y}^C - \bar{B}^C)^2} \sqrt{\sum_{x=0}^3 \sum_{y=0}^3 (B_{x,y}^{RS} - \bar{B}^{RS})^2}}$$

여기서, B^C 와 B^S 는 각각 현재 4x4 블록과 현재 블록을 수평 방향으로 1 화소 (Pixel) 위치 이동한 블록을



(a) "Foreman" 영상



(b) "Paris" 영상

그림 2. "Foreman"과 "Paris" 영상에서 잔여 신호의 수평 상관 계수 분포

Fig. 2. 4x4 block distribution of correlation coefficient for residual signal in "Foreman" and "Paris" sequences.

나타낸다. $B_{x,y}^C$ 는 현재 4x4 블록 내의 (x,y) 위치를 나타내며, $B_{x,y}^{LS}$ 와 $B_{x,y}^{RS}$ 는 각각 현재 4x4 블록에서 왼쪽과 오른쪽으로 1 화소 위치 이동된 4x4 블록 내에서의 (x,y) 위치를 나타낸다. \bar{B}^C 와 \bar{B}^{LS} , 그리고 \bar{B}^{RS} 는 각각 4x4 블록에서의 화소의 평균값을 나타낸다. 현재 4x4 블록이 매크로블록 (Macroblock) 내에서 가장 오른쪽에 위치한 블록이라면, 식 (6)에서 위의 식이 이용되고, 그렇지 않은 경우 아래의 식이 상관 계수 계산에 이용된다.

그림 1은 "Foreman"과 "Paris" 원 영상 중 첫 번째

장을 제외하고 100 장에 대한 4x4 블록 단위의 수평 상관 계수 분포를 나타낸다. 또한, 그림 2는 "Foreman"과 "Paris" 영상 중 첫 번째 I-프레임 (I-frame)을 제외하고 100장의 P-프레임 (P-frame)들에 대해서 각각의 매크로블록들이 울-왜곡 최적화를 통해서 최적의 매크로블록 방식 (Mode)으로 선택될 때의 잔여 신호의 수평 상관 계수 분포를 나타낸다. 이를 통해서 실제 영상을 부호화할 때 발생하는 잔여 신호의 상관 계수 분포는 원 영상에서의 상관 계수의 분포와 크게 다르며, -0.5 와 0.5 사이의 값을 가지는 상관 계수가 많이 발생하는 것을 알 수 있다.

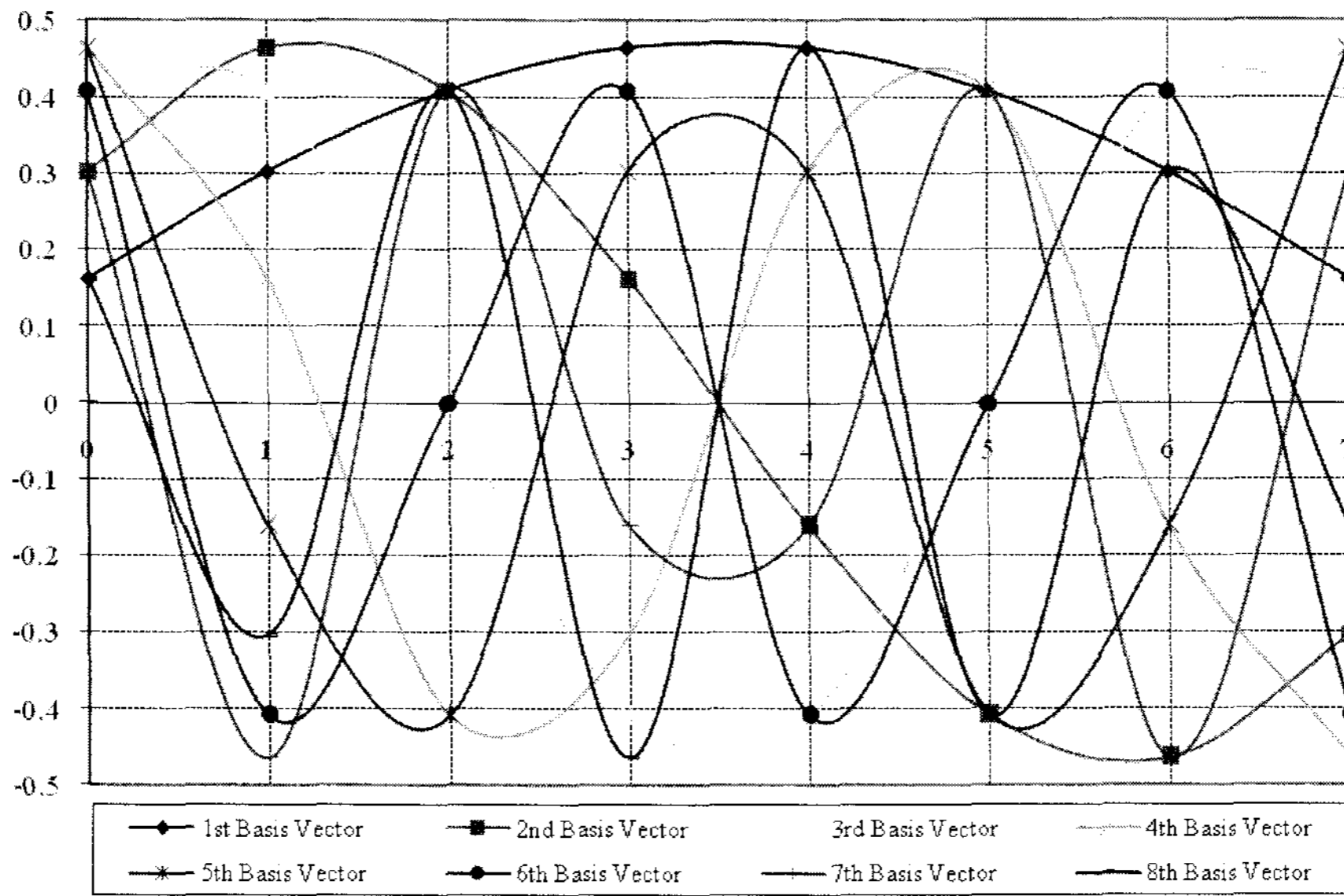


그림 3. 8x8 이산 정현 변환에서의 기저 벡터
 Fig. 3. Basis vector for 8x8 discrete sine transform.

그림 1과 2에서 나타난 H.264/AVC에서 잔여 신호의 특징과 II 장에서 언급된 이산 정현 변환의 장점을 고려하기 위해 4x4와 8x8 블록 단위의 이산 정현 변환들이 본 장에서 소개된다. 이산 정현 변환은 다른 사인 곡선(Sinusoid) 계통 유니터리 변환과 비슷하게 대칭성과 직교성의 특성을 지니고 있으며, 이산 역현 변환과 같이 총 8 가지의 형태로 존재하고 있다. 이들 8 가지의 이산 정현 변환 형태 중 이산 정현 변환-I (DST-I)의 4x4와 8x8 블록 단위의 실수형 이산 정현 변환 방법을 선택 변환에서 사용하기 위해 제시한다. 다음 식 (7)과 식 (8)은 각각 1차원 이산 정현 변환과 1차원 이산 정현 역변환(Inverse Discrete Sine Transform)에 관한 식을 나타낸다^[6].

$$Y(k) = \sqrt{\frac{2}{N+1}} \sum_{n=0}^{N-1} X(n) \sin \frac{\pi(k+1)(n+1)}{N+1}, \quad 0 \leq k \leq N-1 \quad (7)$$

$$X(n) = \sqrt{\frac{2}{N+1}} \sum_{k=0}^{N-1} Y(k) \sin \frac{\pi(k+1)(n+1)}{N+1}, \quad 0 \leq n \leq N-1 \quad (8)$$

$X(n)$ 는 이산 정현 변환될 잔여 신호를 나타내고, $Y(k)$ 는 이산 정현 변환된 계수를 나타낸다. 그리고 N 은 이산 정현 변환의 단위 크기를 나타낸다. 본 논문에서의 단위 크기는 각각 4와 8이 고려되었다. 4x4 블록과 8x8 블록 중 4x4 블록 단위에 대한 예로, 식 (7)과 식 (8)은 실제 영상 부호화 과정에서 사용되기 위해 식 (9)과 식 (10)과 같이 4x4 블록 단위의 이산 정현 변환 행

렬로 나타낼 수 있다.

$$Y = SXS = \begin{pmatrix} a & b & b & a \\ b & a & -a & -b \\ b & -a & -a & b \\ a & -b & b & -a \end{pmatrix} [X] \begin{pmatrix} a & b & b & a \\ b & a & -a & -b \\ b & -a & -a & b \\ a & -b & b & -a \end{pmatrix} \quad (9)$$

$$X' = SYS = \begin{pmatrix} a & b & b & a \\ b & a & -a & -b \\ b & -a & -a & b \\ a & -b & b & -a \end{pmatrix} [Y'] \begin{pmatrix} a & b & b & a \\ b & a & -a & -b \\ b & -a & -a & b \\ a & -b & b & -a \end{pmatrix} \quad (10)$$

식 (9)에서 S 는 X 에 대한 4x4 이산 정현 변환 행렬을 나타내며, 이산 역현 변환과 달리 이산 정현 변환은 대칭적인 특징을 가지고 있기 때문에 $S = S^T$ 가 성립하게 된다. 식 (10)에서의 S 역시 식 (9)에서의 S 와 동일하다. 그리고 행렬 S 내의 원소들 a 와 b 는 각각 상수

$\sqrt{\frac{2}{5}} \sin(\frac{\pi}{5})$ 와 $\sqrt{\frac{2}{5}} \sin(\frac{2}{5}\pi)$ 를 나타낸다.

식 (9)을 이용해 잔여 신호 행렬 X 에 대해서 이산 정현 변환을 수행한 후, 다음 식 (11)을 이용하여 양자화를 수행한다.

$$Z_{(i,j)} = \left\lfloor \frac{Y_{(i,j)}}{QStep} + 0.5 \right\rfloor \quad (11)$$

여기서, 기존의 H.264/AVC에서의 정수 역현 변환과 양자화와 달리 이산 정현 변환에서의 변환 기저 벡터는 실수로 나타나며, 양자화 후 엔트로피 부호화 과정

전에서 실수 형식으로 계산된 변환된 계수를 정수로 근사화하여 복호화기 (Decoder)에 전송해야 한다. 식 (11)에서 $QStep$ 은 양자화의 간격 크기 (Step Size)를 나타낸다.

양자화된 이산 정현 변환 계수 $Z_{(i,j)}$ 에 대해서 다음 식 (12)을 이용해 역양자화를 수행 후, 식 (10)와 식 (13)를 이용해 이산 정현 역변환과 복원된 잔여 신호에

대한 최종적인 근사화를 수행하여 복원된 잔여 신호 행렬 $X''_{(i,j)}$ 를 얻을 수가 있다.

$$Y'_{(i,j)} = Z_{(i,j)} \cdot QStep \tag{12}$$

$$X''_{(i,j)} = \lfloor X'_{(i,j)} + 0.5 \rfloor \tag{13}$$

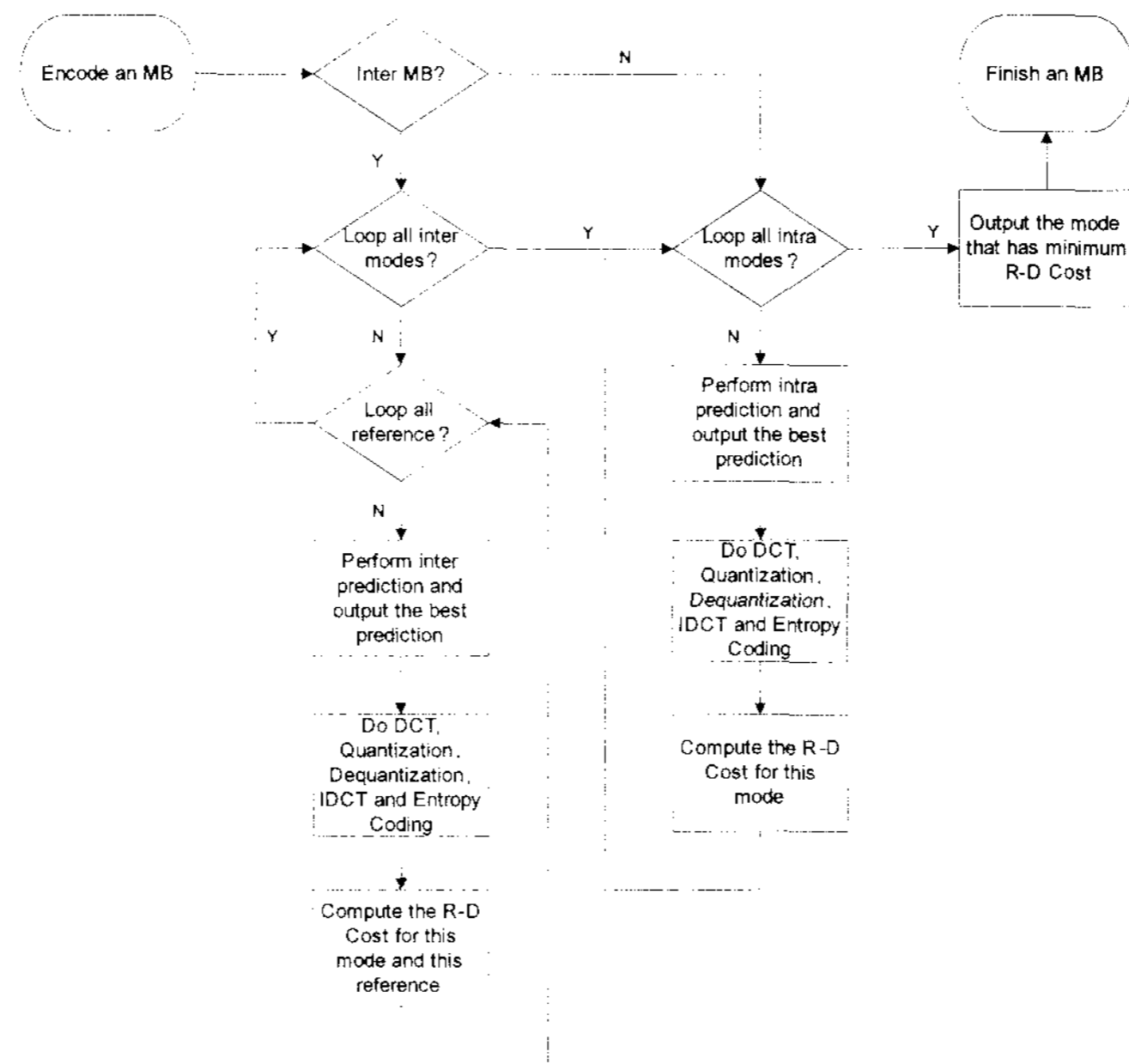


그림 4 H.264/AVC에서의 율-왜곡 최적화 과정
Fig. 4. Rate-distortion optimization in H.264/AVC.

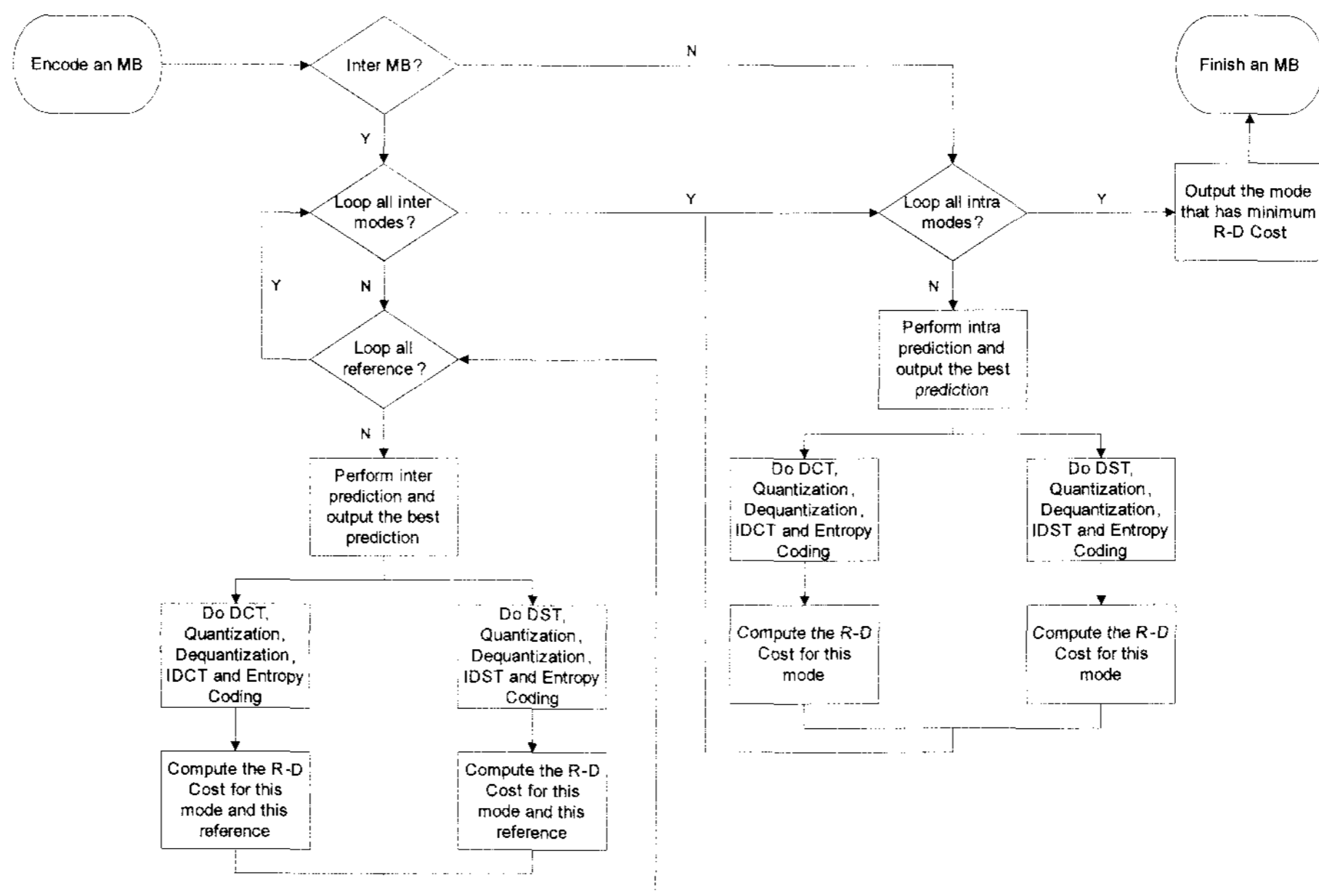


그림 5. 제안하는 율-왜곡 최적화 과정
Fig. 5. Proposed rate-distortion optimization.

8x8 블록 단위의 이산 정현 변환 역시 위에서 제시된 4x4 이산 정현 변환과 동일하게 유도될 수가 있으며, 8x8 블록 단위의 이산 정현 변환에 사용되는 기저 벡터는 그림 3과 같이 표현된다.

IV. 율-왜곡 최적화 내에서의 선택 변환

II 장에서 설명된 H.264/AVC에서의 정수 여현 변환과 III 장에서 설계된 이산 정현 변환과 그에 따른 양자화를 H.264/AVC에서 표준으로 제정되진 않았지만 사용이 권고되는 율-왜곡 최적화 과정에서 선택적으로 수행하기 위해서, 그림 4와 같은 H.264/AVC에서의 율-왜곡 최적화의 구성을 그림 5와 같이 본 논문에서 제안하

는 방법으로 수정할 수가 있다. 실제로 식 (6)에서의 상관 계수 계산 방법은 제안하는 율-왜곡 최적화 과정 내에서 수행되지 않으며, 율-왜곡 최적화 과정 내에서 2가지의 변환 중 최적의 율-왜곡 값 (R-D Cost)을 나타내는 변환을 선택하여 부호화한다.

부호화기에서 선택된 변환의 종류를 복호화기에 전송하기 위해서, 매크로블록 단위로 추가적인 플래그 비트 (Flag Bit)를 1 비트 두어 복호화기가 올바른 방법으로 복호화하게 할 수 있다. 다만, 잔여 신호를 부호화하지 않는 SKIP 모드에서는 플래그 비트를 전송하지 않으며, INTER8x8 방식일 경우에는 8x8 블록 단위로 총 4 개의 플래그 비트를 복호화기에 전송한다. 원칙적으로 제안하는 선택변환은 매크로블록 내의 모든 4x4 블

표 3. 전송되는 플래그 비트의 비율과 매크로블록 당 평균 플래그 비트수

Table 3. Transmitted percentage of flag bits for whole bitstream and average flag bits per a macroblock.

Sequence	플래그 비트의 비율 (Percentage of flag bits for whole bitstream)	매크로블록 당 평균 플래그 비트수 (Average flag bits per a macroblock)
Stefan (QCIF)	0.49%	1.76
News (QCIF)	1.24%	0.93
Foreman (CIF)	0.93%	1.18
Mobile & Calendar (CIF)	0.52%	1.79

표 5. 실험 조건

Table 5. Experimental condition.

GOP Structure	IBPBP..
Encoded Frames	299 Frames
Intra Period	Every 30th Frame
Quantization Parameters	12, 16, 20, 24
Search Range	±16 for QCIF, ±32 for CIF
Multiple Reference Frames	5
Entropy Coding Method	CABAC
Rate-Distortion Optimization	On
Intra 8x8 Prediction, 8x8 Integer Transform	Enabled

표 4. 정수 여현 변환과 이산 정현 변환의 계산 복잡도 비교

Table 4. Comparison of computational complexity of integer cosine transform and discrete sine transform.

Operation	4x4 ICT in H.264/AVC	8x8 ICT in H.264/AVC	Proposed 4x4 DST	Proposed 8x8 DST	4x4 DST with fast implementation [13]	8x8 DST with fast implementation [13]
Addition	256	256	1536	3584	128	176
Shift	64	80	0		0	0
Multiplication	0	0	2048	4096	64	64

록에 대해서 적용이 되어 총 16 개의 플래그 비트가 전송될 수 있지만, 너무 많은 플래그 비트의 수는 전체 부호화 효율을 감소할 수가 있다. 실제로 전송되는 플래그 비트수가 전체 비트율에서 차지하는 비율과 매크로블록 당 평균 플래그 비트수를 표 3에 나타내었다. 표 3은 QP 20에서 “Stefan” 영상, “News” 영상, “Foreman”

영상, “Mobile & Calendar” 영상에 대하여 표 5의 실험 조건 하에서 실험한 결과이다. 표 3에서 확인할 수 있듯이, QP가 20일 때, 플래그 비트는 전체 비트율에서 1% 이하를 차지하며, 매크로블록 당 약 1.5 bit가 평균적으로 전송되는 것을 알 수 있다.

한편으로 울-왜곡 최적화 과정 내에서의 2 가지 변

표 6. H.264/AVC와 제안하는 방법의 실험 결과

Table 6. Experimental result of H.264/AVC and proposed method.

Sequence	QP	H.264/AVC		Proposed Method		
		Bitrate (kbps)	PSNR (dB)	Bitrate (kbps)	PSNR (dB)	BD-PSNR (dB)
Coastguard (QCIF)	12	1485.93	47.85	1574.18	49.18	0.39
	16	984.46	44.40	1038.61	45.40	
	20	587.61	40.85	640.17	41.70	
	24	335.51	37.63	375.76	38.22	
Mother & Daughter (QCIF)	12	664.70	48.65	712.83	49.38	0.13
	16	353.90	45.99	378.02	46.48	
	20	194.03	43.02	204.04	43.32	
	24	110.02	40.07	114.31	40.22	
News (QCIF)	12	470.22	49.10	493.73	49.62	0.03
	16	312.28	46.41	328.87	46.84	
	20	205.86	43.43	218.25	43.84	
	24	135.97	40.42	144.17	40.76	
Stefan (QCIF)	12	2160.91	48.11	2276.06	49.69	0.71
	16	1539.18	44.91	1603.66	46.08	
	20	1018.36	41.34	1061.89	42.32	
	24	651.36	37.88	682.37	38.77	
Carphone (CIF)	12	6222.82	48.07	6944.26	49.32	0.04
	16	3334.12	44.69	3668.41	45.18	
	20	1564.47	41.93	1664.47	42.12	
	24	875.91	39.84	933.28	39.99	
Foreman (CIF)	12	4915.37	48.26	5546.85	49.38	0.04
	16	2825.00	45.09	3088.64	45.59	
	20	1413.58	41.99	1503.03	42.26	
	24	746.16	39.30	775.12	39.46	
Mobile & Calendar (CIF)	12	8679.70	47.85	9229.66	49.52	0.58
	16	6051.85	44.55	6414.66	45.79	
	20	3858.43	40.93	4110.97	41.91	
	24	2322.93	37.51	2508.42	38.25	
Paris (CIF)	12	3668.24	47.99	4088.74	48.80	-0.03
	16	2151.36	45.07	2286.12	45.40	
	20	1316.52	42.08	1392.71	42.36	
	24	838.90	39.10	895.04	39.40	

환 모두 사용되기 때문에, 부동 소수점과 곱셈 연산이 필요한 이산 정현 변환에 의해서 제안하는 선택 변환의 계산 복잡도가 크게 증가할 수 있다. 표 4는 정수 여현 변환과 이산 정현 변환을 복호화기에서 하나의 매크로 블록에 대해 수행할 때 필요한 계산 복잡도를 나타낸다. 이산 정현 변환의 경우 많은 양의 계산 복잡도가 필요하지만, [13] 에서 제안된 이산 정현 변환의 빠른 구현 방법을 통해 계산 복잡도를 크게 감소할 수 있다.

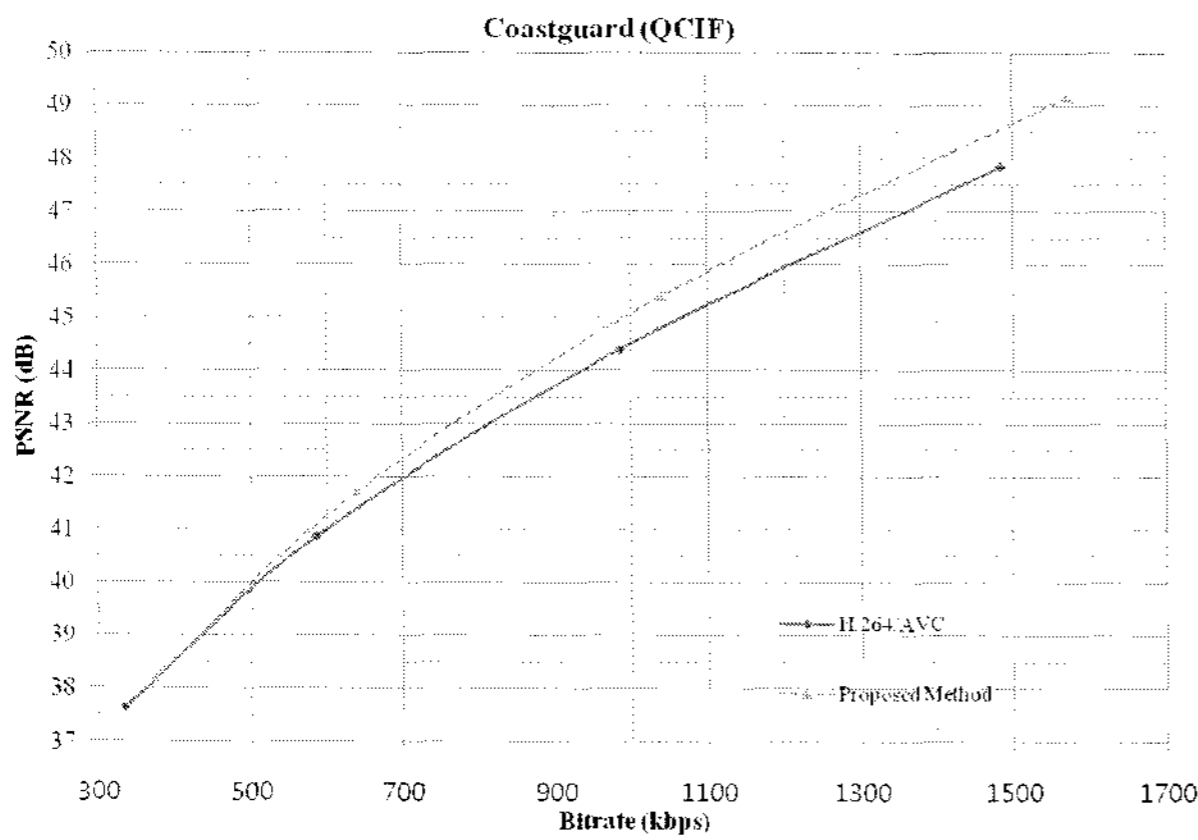
V. 실험 결과

본 장에서는 제안된 잔여 신호의 상관성을 고려한 선택 변환 방법의 성능을 여러 영상에 대하여 H.264/AVC JM 10.2의 방법과 비교 실험한 결과를 제시한다. 실험 영상으로는 각각 4 개의 QCIF (Quarter Common Intermediate Format) 영상과 4 개의 CIF (Common Intermediate Format) 영상을 이용했으며, 상

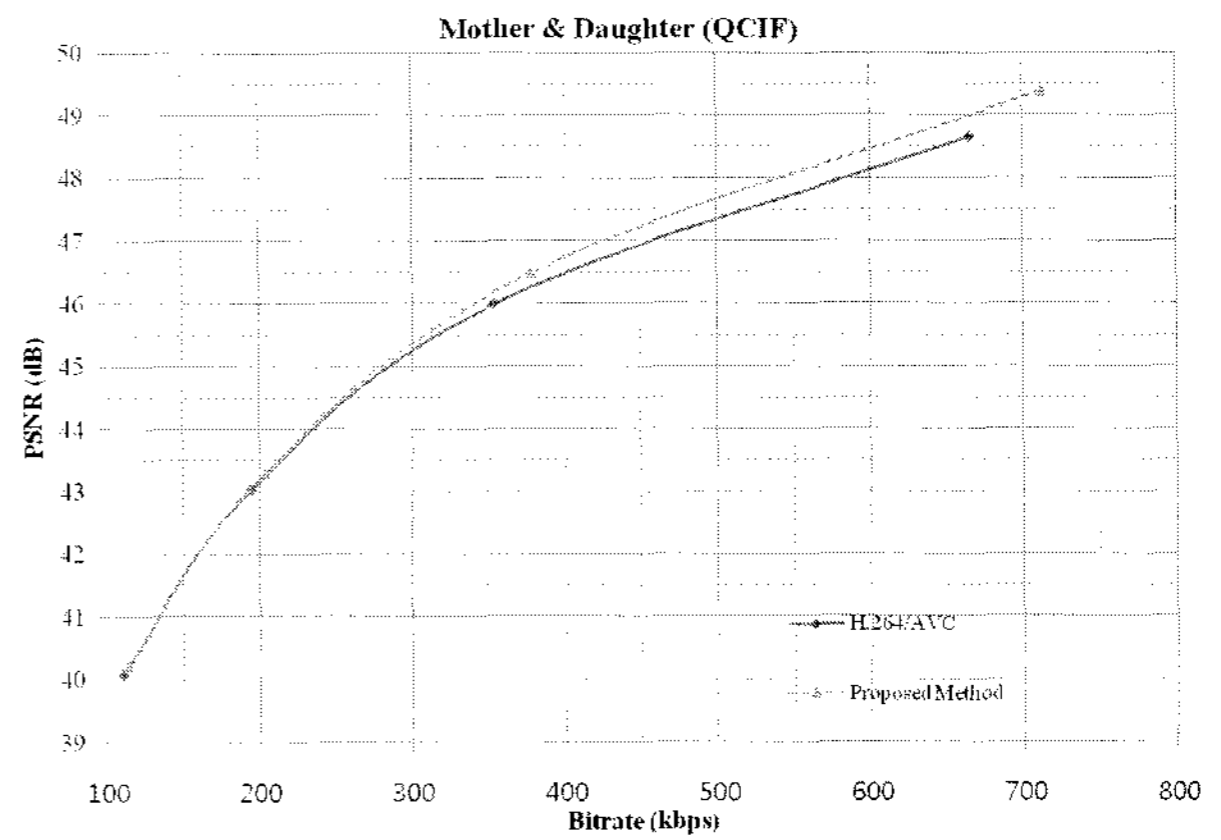
세한 실험 조건은 표 5에 나타내었다.

다음 표 6은 표 5의 실험 조건으로 기존 H.264/AVC의 정수 여현 변환 방법과 본 논문에서 제안하는 방법 간의 비교 실험 결과이며, 그림 6과 7은 표 6에서의 각 영상들에 대한 율-왜곡 곡선 (Rate-Distortion Curves)을 나타낸다. 제안하는 선택 변환은 상대적으로 높은 비트율 (High Bitrate)에서 부호화 성능을 나타내며, 특히 영상 내의 움직임이 상대적으로 큰 “Stefan” 영상이나 “Mobile & Calendar” 영상에서 높은 부호화 성능을 나타낸다. 특히 “Stefan” 영상에서는 H.264/AVC보다 최대 0.71 dB에 가까운 BD-PSNR (Bjontegaard Delta Peak Signal-to-Noise Ratio)^[14] 성능 향상을 보인다. 하지만, “News” 영상이나 “Carphone” 영상, “Foreman” 영상에서는 미미한 부호화 효율 향상을 보여주며, “Paris” 영상에서는 오히려 부호화 효율이 저하되는 것을 확인할 수 있다.

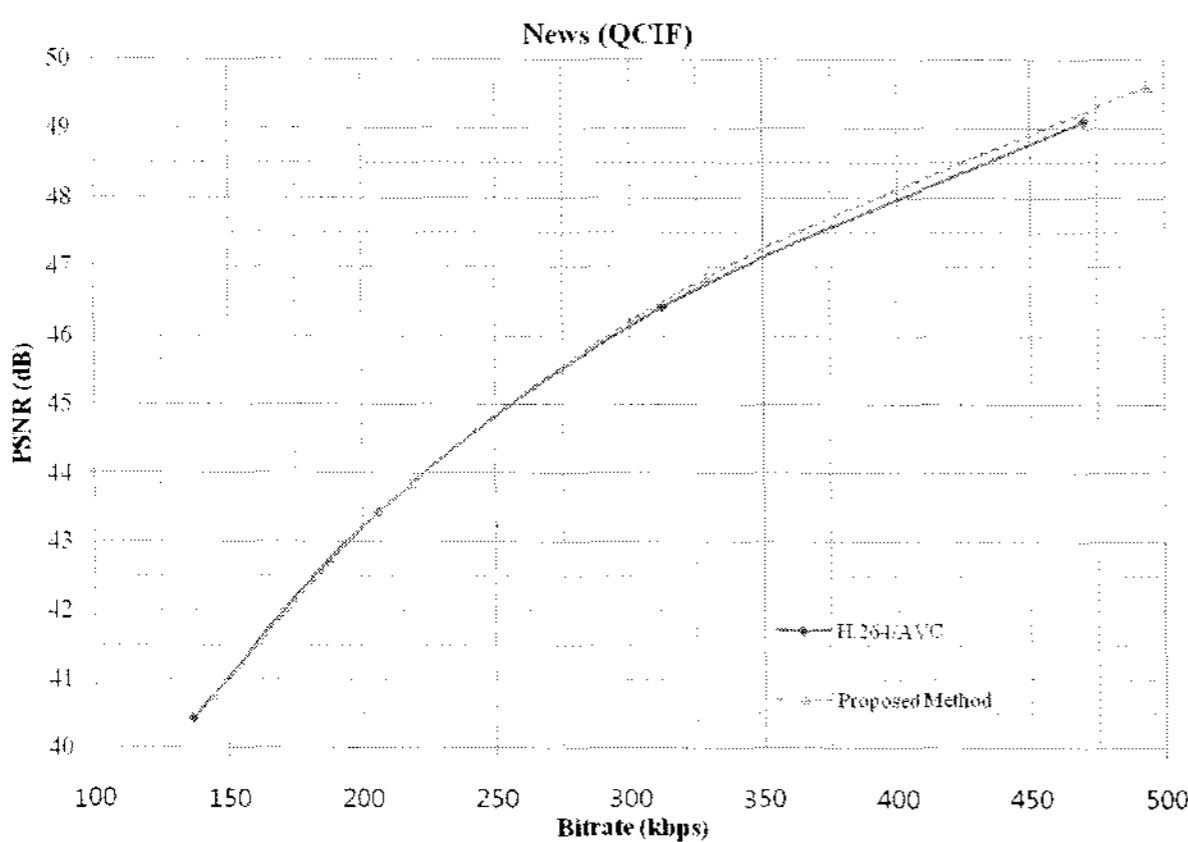
표 7은 제안하는 선택 변환 방법을 통해서 실제 부호



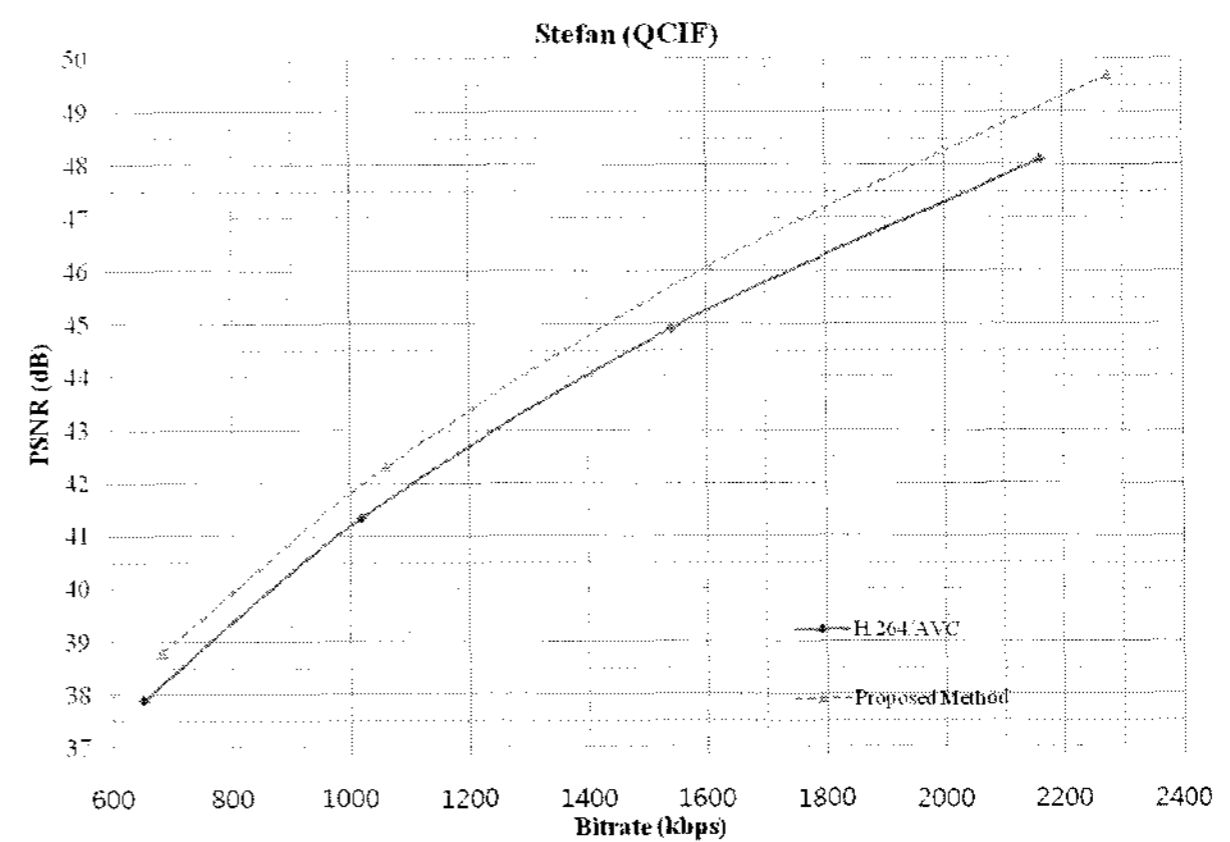
(a) “Coastguard” 영상



(b) “Mother & Daughter” 영상



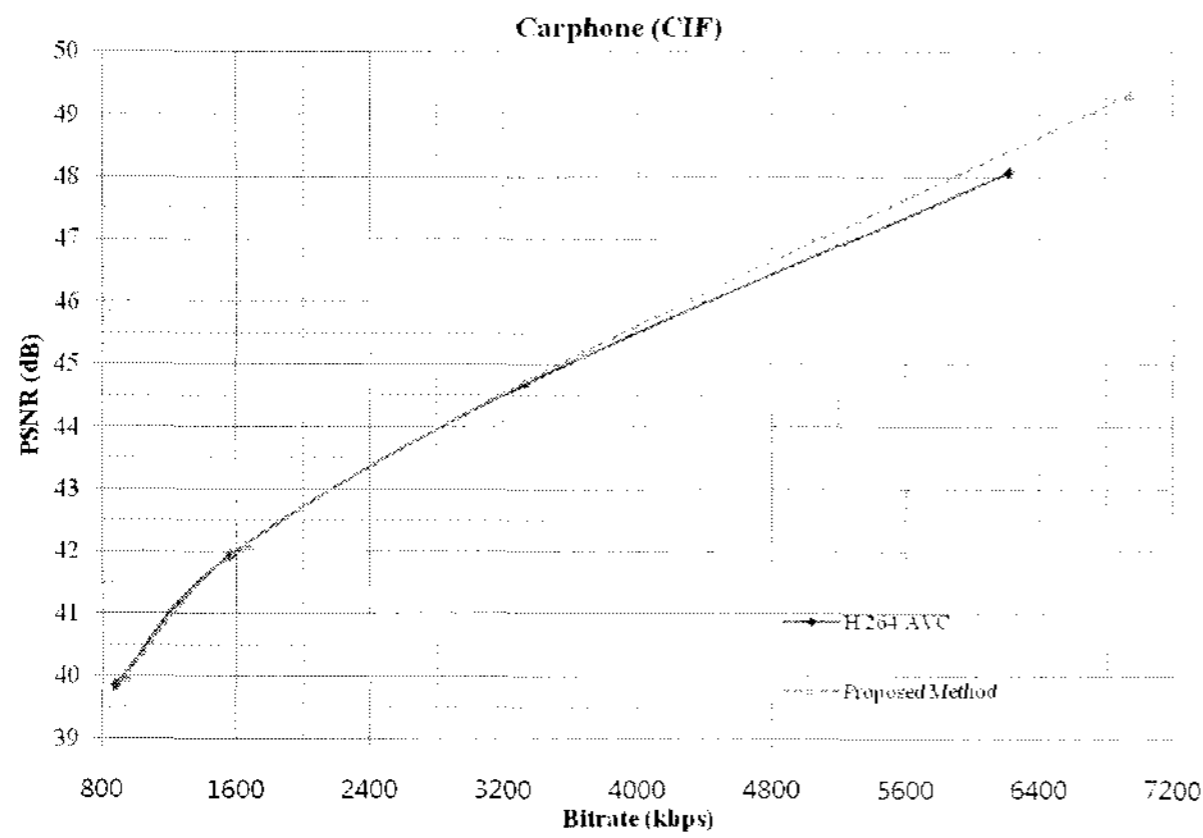
(c) “News” 영상



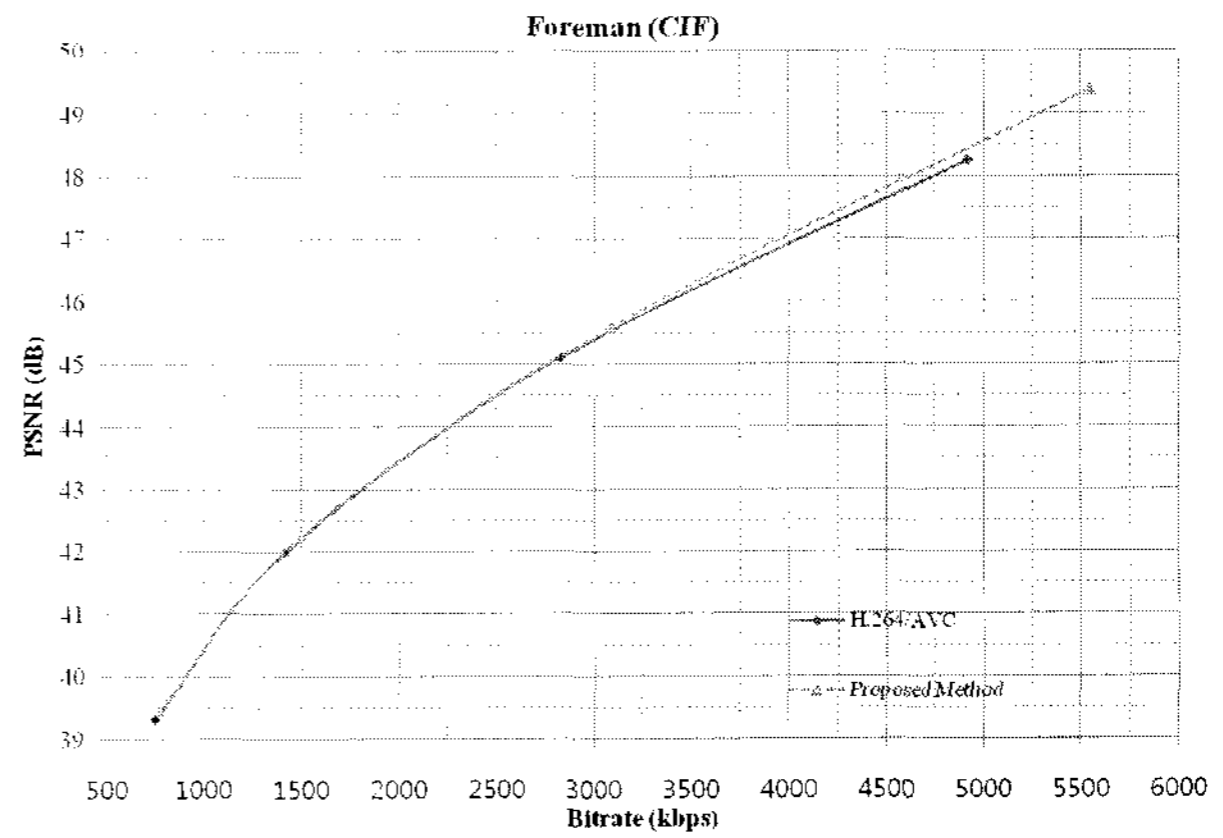
(d) “Stefan” 영상

그림 6. H.264/AVC와 제안하는 방법의 율-왜곡 곡선 (QCIF)

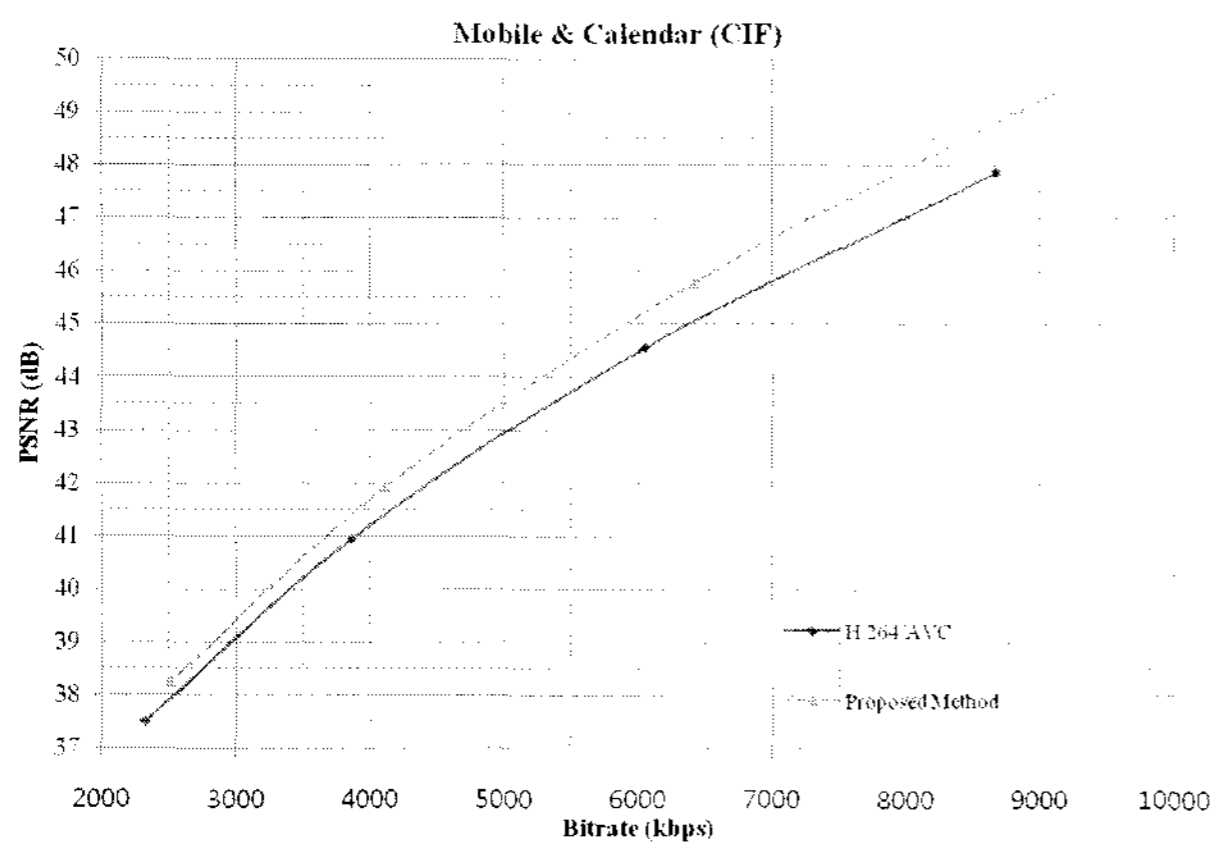
Fig. 6. Rate-distortion curves of H.264/AVC and proposed method (QCIF).



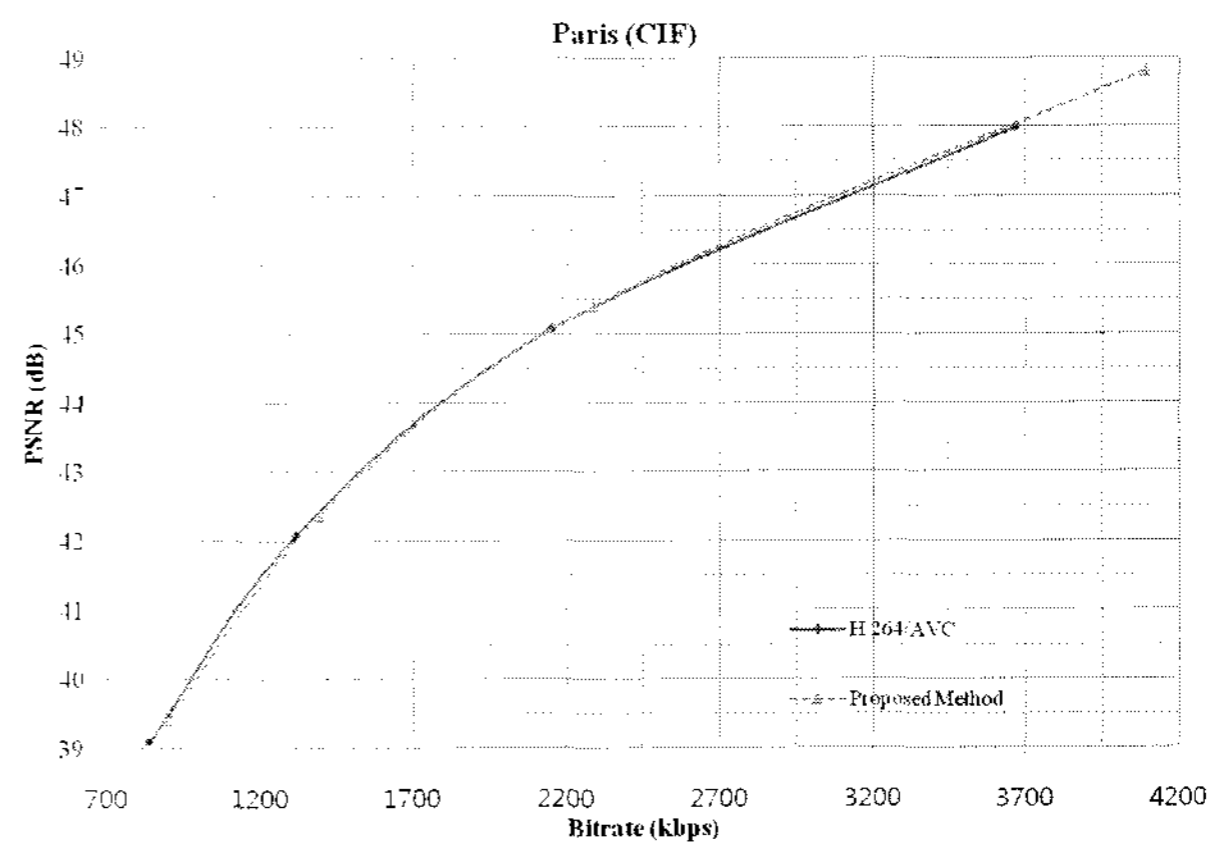
(a) "Carphone" 영상



(b) "Foreman" 영상



(c) "Mobile & Calendar" 영상



(d) "Paris" 영상

그림 7. H.264/AVC와 제안하는 방법의 율-왜곡 곡선 (CIF)

Fig. 7. Rate-distortion curves of H.264/AVC and proposed method (CIF).

표 7. QCIF 영상들과 CIF 영상들에서 선택된 변환 방법의 통계

Table 7. Statistic of selected transform method in QCIF sequences and CIF sequences.

Sequence	QP	ICT (%)	DST (%)
Coastguard (QCIF)	12	46.51	53.49
	16	58.94	41.06
	20	67.06	32.94
	24	78.65	21.35
Mother & Daughter (QCIF)	12	76.23	23.77
	16	88.02	11.98
	20	92.54	7.46
Mobile & Calendar (CIF)	24	96.43	3.57
	12	31.73	68.27
	16	45.44	54.56
Paris (CIF)	20	57.41	42.59
	24	69.83	30.17
	12	58.36	41.64
	16	80.16	19.84
Paris (CIF)	20	86.49	13.51
	24	90.7	9.30

화 과정 시 제안하는 이산 정현 변환이 어느 정도의 빈도수로 선택되었는지에 대한 통계를 나타낸다. 표 5의 실험 조건을 기반으로 하여 표 7은 양자화 매개변수 12, 16, 20, 24에 대해서 2 개의 QCIF 영상과 2 개의 CIF 영상을 이용하여 모든 화면 간/내 매크로블록 모드들에 대해서 이산 정현 변환으로 선택된 전체 비율을 나타낸다. INTER8x8 모드인 경우 4 개의 8x8 블록들 중 하나의 블록이라도 이산 정현 변환으로 선택된 경우, 이산 정현 변환이 사용되었다고 고려했다.

표 6에서 높은 부호화 성능 향상을 보여준 영상과 상대적으로 낮은 부호화 성능의 향상을 보여 준 영상들의 통계를 표 7에 나타내었으며, 높은 부호화 성능 보여주는 "Coastguard" 영상과 "Mobile & Calendar" 영상에서는 각각 최대 53.49 %와 68.27 %의 매크로블록에서 이산 정현 변환을 선택하였으며, 상대적으로 낮은 부호화 성능을 보여주는 "Mother & Daughter" 영상과 "Paris" 영상에서도 최대 23.77 %와 41.64 %의 매크로블록에서 이산 정현 변환을 선택하였다. 또한, 부호화

성능의 향상이 상대적으로 작은 낮은 비트율 (Low Bitrate)보다 부호화 성능의 향상이 큰 높은 비트율에서 보다 많은 매크로블록에서 이산 정현 변환을 변환 방법으로 선택한 것을 확인할 수 있다.

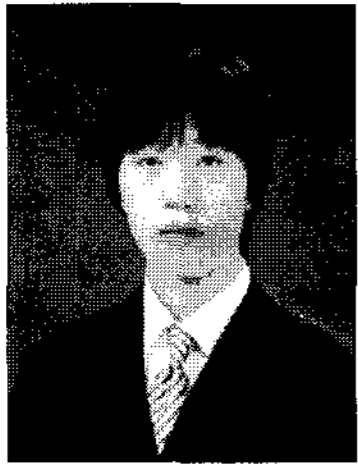
VI. 결 론

본 논문에서 제안한 선택 방법을 통해서, 실제 영상이 부호화될 때 생성되는 잔여 신호의 상관 계수 값의 범위가 이산 여현 변환이 최적의 성능을 나타내는 상관 계수의 범위가 아닌 경우, 이산 여현 변환보다 더 뛰어난 공간적 압축 성능을 나타내는 이산 정현 변환이 이산 여현 변환의 대체 방법으로 영상 부호화 방법에 고려되었다. 또한, 제안된 방법으로 이산 정현 변환과 H.264/AVC에서의 정수 여현 변환을 올-왜곡 최적화 과정에서 선택적으로 사용 가능하며, 기존의 H.264/AVC의 JM 10.2와의 비교 실험을 통해 상대적으로 높은 비트율에서 최대 0.71 dB의 BD-PSNR 향상을 얻을 수 있었다.

참 고 문 헌

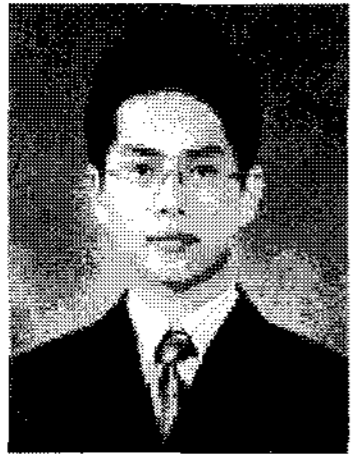
- [1] ITU-T Recommendation H.264 and ISO/IEC 14496-10, *Advanced Video Coding for Generic Audiovisual Services*, May 2003.
- [2] G. J. Sullivan, T. McMahan, T. Wiegand, and A. Luthra, Eds., Draft Text of H.264/AVC Fidelity Range Extensions Amendment to ITU-T Rec. H.264 ISO/IEC 14496-10 AVC, *ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/SG16 Joint Video Team document JVT-L047*, Jul. 2004.
- [3] N. Ahmed, T. Natarajin, and K. R. Rao, "Discrete cosine transform," *IEEE Transactions on Computers*, Vol. C-23, pp. 88-93, Jan. 1974.
- [4] A. K. Jain, "A sinusoidal family of unitary transforms," *IEEE Transactions on Pattern Analysis. Machine Intelligence*, Vol. PAMI-1, pp. 356-365, Oct. 1979.
- [5] A. K. Jain, "Fast Karhunen-Loeve transform for a class of stochastic processes," *IEEE Transactions on Communications*, Vol. COM-24, pp. 1023-1029, Sep. 1976.
- [6] A. K. Jain, "Fundamentals of Digital Image Processing," *Pearson Education Inc., Prentice Hall*, pp. 154-155, 1989.
- [7] K. Rose, A. Heiman, and I. Dinstein, "DCT/DST alternate-transform image coding," *IEEE Transactions on Communications*, Vol.38, No.1, Jan. 1990.
- [8] E. G. Pelaes, Y. Iano, "Image coding with low blocking effects using discrete sine transform with axis rotation," *Telecommunications Symposium, 1998. ITS'98 Proceedings. SBT/IEEE International*, Vol. 2, pp. 588-593, 9-13 Aug. 1998.
- [9] S.-C. Pei and J.-J. Ding, "The integer transforms analogous to discrete trigonometric transforms," *IEEE Transactions on Signal Processing*, Vol. 48, No. 12, pp. 3345-3364, Dec. 2000.
- [10] M. Helsingius, P. Kuosmanen, J. Astola, "Image compression using multiple transforms," *Signal Processing: Image Communication 15 (2000)*, pp. 513-529.
- [11] H.264/AVC Reference Software Version 10.2, <http://iphome.hhi.de/suehring/tml/download/>
- [12] H. S. Malvar, A. Hallapuro, M. Karczewicz, L. Kerofsky, "Low-complexity transform and quantization in H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 7, Jul. 2003.
- [13] P. Yip and K. R. Rao, "A Fast Computational Algorithm for the Discrete Sine Transform," *IEEE Transactions on Communications*, Vol. COM-28, No. 2, Feb. 1980.
- [14] G. Bjontgaard, "Calculation of average PSNR differences between RD-curves", *ITU-T SG16 Q.6 VCEG, Doc. VCEG-M33*, Mar. 2001.

저 자 소 개



임 성 창(정회원)
 2006년 세종대학교 컴퓨터공학과
 학사 졸업.
 2008년 세종대학교 컴퓨터공학과
 석사 졸업.
 2008년~현재 한국전자통신
 연구원 방송통신융합연구
 부문 방통미디어연구부
 연구원

<주관심분야 : 영상압축, 영상처리>



김 대 연 (학생회원)
 2005년 세종대학교 인터넷학과
 학사 졸업.
 2005년~현재 세종대학교 컴퓨터
 공학과 석박통합 과정
 <주관심분야 : H.264/AVC, MPEG
 video compression, Future video
 coding, Entropy coding,
 Transcoding>



이 영 렬(평생회원)
 1985년 서강대학교 전자공학과
 학사 졸업.
 1987년 서강대학교 전자공학과
 석사 졸업.
 1999년 한국과학기술원 전기전자
 공학과 박사 졸업.

1987년~2001년 삼성전자 R&D센터 Digital
 Media Lab. 수석연구원

2001년~현재 세종대학교 컴퓨터공학과 부교수
 <주관심분야 : 비디오코딩, 영상처리, 영상통신시
 스템, 멀티미디어전송>