

일반논문-10-15-4-11

H.264/AVC에서 DCT계수 분포를 이용한 고속 인트라 모드 결정 방법

홍성욱^{a)}, 이영렬^{a)†}

Fast intra mode decision using DCT coefficient distribution in H.264/AVC

Sung-Wook Hong^{a)} and Yung-Lyul Lee^{a)†}

요약

H.264/AVC 부호기에서 각 모드 결정에 사용되는 방법인 윌-왜곡 최적화(RDO(Rate-distortion optimization))는 부호기의 높은 성능 향상을 보이지만 많은 계산량을 요구하는 문제점을 가진다. 이러한 계산량을 줄이기 위해 원본 영상의 각 블록에 DCT(Discrete Cosine Transform) 계수 분포를 통해 각 블록의 방향성을 예측하여 부호기의 RDO 수행에 참여하는 후보 모드 수를 줄이는 방법을 제안한다. 제안하는 방법은 H.264/AVC에 부호화 방법에 비해 적은 PSNR 손실 및 비트의 증가에도 불구하고 평균 68.40%의 부호기 속도 향상을 가진다.

Abstract

The rate-distortion optimization (RDO) method in the H.264/AVC encoder is a technology that improves the coding efficiency, but increases the computational complexity. In this paper, a fast Intra mode decision algorithm using DCT (Discrete Cosine Transform) coefficients distribution is proposed to reduce the H.264 encoder complexity. The proposed method reduces the encoder complexity on average 68.40%, while the coding efficiency is slightly decreased compared with the H.264/AVC encoder.

Keyword: Video coding, H.264/AVC, Encoder, Fast mode decision, Intra mode decision

I. 서론

동영상 압축 표준인 H.264/AVC(Advanced Video Coding)는 ITU-T VCEG과 ISO/IEC MPEG의 비디오 전문가들이 구성한 JVT(Joint Video Team)으로부터 개발되었다.^{[1][2]} 이 표준은 4x4 정수형 DCT, 다양한 블록 크기에

대한 움직임 보상, Hierarchical-B 구조, CABAC(Context Adaptive Binary Arithmetic Coding), CAVLC(Context-adaptive variable-length coding), 시간적/공간적(인tra/인트라) 예측 방법, 윌-왜곡 최적화 방법(RDO(Rate Distortion Optimization)), 1/4 화소 예측 방법 등 많은 기술을 사용함으로써 높은 성능 향상을 가진다. 하지만, 여러 기술의 사용에 따른 많은 계산량이 필요하다는 문제점을 가진다. 특히, RDO(Rate-Distortion Optimization)과정에서 각 블록 크기에 대한 최적의 모드를 결정하는 단계에서 많은 계산량을 요구한다. 제안하는 방법은 공간적 정보를 사용하는 인트라 부호화에 RDO 계산량을 줄이기 위한 방법으로 인트라

a) 세종대학교 컴퓨터공학과 DMS 연구실

DMS Lab. Dept. of Computer Engineering, Sejong University

† 교신저자 : 이영렬(ylee@sejong.ac.kr)

* 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임(2010-0016136)

·접수일(2010년5월17일),수정일(2010년7월12일),게재확정일(2010년7월13일)

부호화는 H.264/AVC에서 제공하는 Baseline 프로파일, Main 프로파일, High 프로파일 등 모든 프로파일에서 사용하는 부호화 방법이다.

인트라 부호화 과정은 프로파일에 따라 다르지만 각 매크로블록의 밝기신호(Luma)는 Intra16×16, Intra4×4, Intra8×8 (Baseline 프로파일에서는 사용안되고 Main, High 프로파일에서 사용) 예측방법을 사용하고 색차신호(Chroma)는 Chroma8×8 예측방법을 사용한다. 각 매크로블록(MacroBlock, MB)의 Intra 모드에 따라, Intra4×4 예측방법은 9가지 방향, Intra16×16과 Chroma8×8은 4가지 방향에 대한 예측 모드를 가진다. 그림 1은 Intra4×4 예측방법에 대한 주변 화소들의 분포와 9가지 예측방향을 나타낸다.

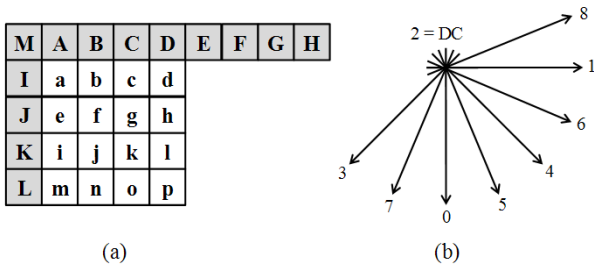


그림 1. Intra4×4에 예측 모드 (a) Intra4×4 블록 및 주변 블록 (b) 9가지 예측모드 방향

Fig. 1. Intra4×4 prediction modes (a) Boundary pixels and inside pixels for Intra4×4 prediction. (b) Nine directional prediction modes

그림 1-(a)의 a~p는 현재 부호화를 수행할 4×4 블록에 대한 원본 영상의 화소를 의미하고 A~M은 이미 부호화를 마치고 복원된 화소를 의미한다. 예를 들면, 그림 1-(b)의 모드 0 (Vertical prediction)인 경우는 a,e,i,m의 값을 수직 방향 상단에 존재하는 복원된 화소 A를 예측화소로 하여 차분 값을 각각 구하고, 나머지 화소들(b,c,d,f,g,h,j,k,l,n,o,p)은 마찬가지로 수직 방향에 위치하는 각 B,C,D를 예측화소로 하여 각각의 차분 값으로 하나의 4×4 차분 블록(residual block)을 생성한다. 이와 유사한 방법으로 9개의 예측방향에 대한 차분 블록을 생성한다. 그림 1-(b)는 8가지 예측방향과 주변 블록의 평균값을 이용하는 DC모드(mode=2)로 총 9가지 예측 가능한 방향을 나타낸다. Intra16×16과 Chroma8×8은 그림 1-(b)의 0~2번 모드와 45도 방향의 모드로 총 4가지 예측모드를 가진다.

율-왜곡 최적화 방법(RDO)은 각 블록의 예측 가능한 모드 중 가장 부호화 효율이 좋은 최적의 모드를 결정하기 위한 방법이다. 그림 1의 Intra4×4 예측방법에 대한 RDO 수행은 아래 식으로 표현된다.

$$J(s, c, IMODE | QP, \lambda_{MODE}) = D(s, c, IMODE | QP) + \lambda_{MODE} \cdot R(s, c, IMODE | QP) \quad (1)$$

$IMODE \in \{\text{all nine Intra4} \times 4 \text{ modes}\}$

식(1)에서 QP는 양자화 파라미터, λ_{MODE} 는 Lagrange Multiplier를 의미한다. D는 원래 블록과 복원된 블록의 SSD(Sum of Squared Difference), IMODE 는 Intra4×4에서 사용하는 9개의 예측모드를 의미한다. 좀 더 자세히 언급하면 s는 부호화할 현재 블록, c는 복원된 블록을 의미하며 마지막으로 R(s,c,IMODE|QP)는 모드에 따른 엔트로피 코딩을 마친 비트량을 의미한다. 위 식을 통해 Intra4×4에서 가능한 9가지 방향에 대해서 D(왜곡값, Distortion)와 R(부호화를 했을 때 필요한 비트량)을 모두 고려해 가장 부호화 효율이 좋은 모드를 최적의 모드로 선택한다. 이와 같이 RDO과정은 Intra4x4 및 Intra16x16 모드에서 가능한 예측방향을 모두 계산하여 가장 부호화 효율이 좋은 블록모드 및 이에 대한 예측방향을 결정하는 기술로 많은 계산이 요구되는 문제점을 가진다.

이러한 RDO의 계산량을 줄이기 위한 여러 방법들이 제안되었다^{[3][4][5]}. 이전에 제안된 방법들은 대부분 원본 영상의 화소 값을 이용해 각 블록이 가지는 방향성을 예측하고 예측된 방향과 유사한 방향의 예측 모드만 RDO과정을 수행함으로써 전체 RDO 모드 수를 줄이는 방법들이다. 제안하는 방법은 화소 도메인이 아닌 DCT 도메인 상의 계수 분포를 통해서 원 영상의 예측 방향을 결정하고 예측 후보의 그룹화 방법을 통한 속도 향상 방법을 제안한다. 또한, 제안하는 방법과 유사한 DCT 도메인을 이용한 방법인 Lin의 방법^[4]보다 PSNR, 비트증가량, 속도 측면에서 모두 좋은 성능을 가진다.

본 논문의 구성은 다음과 같다. 2장에서 DCT 계수 분포 특성에 대한 설명을 하고 3장의 각 계수 특성에 맞추어 RDO에 수행되는 후보 모드 그룹을 결정하는 방법을 설명

한다. 이어서 4장에서는 실험 및 결과, 5장에서 결론으로 마친다.

II. DCT 계수 분포의 특성

제안하는 방법은 H.264/AVC에서 현재 사용하는 정수형 4×4 DCT 및 양자화를 이용한다. 이 장에서는 4×4 DCT 및 양자화를 통한 계수의 분포 특성을 알아본다. 먼저 DCT 변환은 아래 식으로 표현된다.

$$Y = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{pmatrix} X \begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{pmatrix} \otimes \begin{pmatrix} a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \\ a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \end{pmatrix} \quad (2)$$

위 식에서 X는 부호화 과정에서 블록이 가지는 각 방향에 따른 예측 잔차 신호에 대해 DCT를 수행하는 4×4 행렬을 의미하고 \otimes 는 행렬에서 동일 (i,j) 위치에 해당하는 값의 곱셈연산 의미한다. 또한, 위 식에서 오른쪽 4×4 행렬은 PF로 정의 되고 a는 0.5, b는 $\sqrt{2/5}$ 이다.

식(2)의 DCT를 마친 행렬 Y에 대해 양자화 과정을 수행한다. 여기서 Qstep은 양자화 파라미터(QP(Quantization parameter))를 결정하는 양자화 값(Quantization step size)을 의미한다. 양자화는 균등 양자화기를 사용하며 양자화 값 수행하는 식은 아래와 같다.

$$Z_{ij} = Y_{ij} / Qstep \quad (3)$$

식(3)의 Yij는 식(2)에서 Y 행렬을 의미하고 Qstep은 위와 같이 양자화 파라미터에 의해 결정된다. 식(3)은 식(2)의 조합으로 다음과 같이 표현된다.

$$Z_{ij} = W_{ij} \cdot \frac{PF}{Qstep} \quad (4)$$

식(4)에서 Wij는 식(2)에서 PF의 곱셈을 수행하기 전 단계를 의미하고 식(4)는 H.264/AVC에서 소수점을 없애기 위해 아래 식으로 표현된다.

$$Z_{ij} = W_{ij} \cdot \frac{MF}{2^{qbits}} \quad (5)$$

$$\frac{MF}{2^{qbits}} = \frac{PF}{Qstep} \quad (6)$$

$$Z_{ij} = (W_{ij} \cdot MF + f) \gg qbits \quad (7)$$

H.264/AVC에서 MF를 이용한 쉬프트 연산은 식(5)로 표현되고, 결과적으로 식(6)으로 표현이 가능하다. 식(5)로부터 얻은 식(7)은 쉬프트 연산으로 소수점이 필요 없는 정수형 계산이 가능하게 되고, 식(7)은 f의 값에 의해서 쉬프트 연산을 통한 Zij의 반올림이 결정된다고 할 수 있다.

식(7)은 소숫점의 값이 모두 제거되므로 식(3)과 비교하여 소수점을 표현하지 못한다. 더 정확한 계수분포를 위해 제안하는 방법은 소숫점을 가지는 Qstep의 나눗셈을 이용하고 소수점을 포함한 값을 구한다. 결과적으로 식(3)을 이용한 소수점계산이 식(7)보다 더 정확한 계수 분포를 확인할 수 있었다. 또, 제안하는 방법은 소수부분의 값을 밝기신호(Luma)에서 각 MB에 대한 Intra4x4, Intra16x16 예측모드 및 색차신호(Chroma)의 Intra8x8 예측모드 별로 임계값을 정하고 그 값보다 큰 경우만 계수를 가진다고 판단하는 방법을 추가적으로 사용한다.

제안하는 방법에서 DCT 계수 분포는 식(2)에서 X 행렬을 잔차 블록이 아닌 원영상의 4×4 블록에 대한 계산 후, 식(3)에 Qstep에 따른 양자화를 계수를 구한다. 본 논문에서는 DCT 계수분포로부터 예측된 화소들의 유사성을 크게 4가지로 분포 형태로 분류하였다.

그림 2는 4×4 블록 크기의 원본 영상의 한 블록에 대한 식(2), 식(3)을 이용한 DCT 및 양자화가 수행된 계수 분포를 나타낸다. 예를 들어 그림 2-(a)에서 DCT 및 양자화 과정을 거친 계수 Zij가 DC 성분만을 갖는 경우, 4×4 복원된 블록은 같은 화소 값을 갖는다. 따라서 DC 계수만을 갖는

블록은 DC 모드, 수직예측모드(mode 0), 수평예측모드(mode 1)로부터 예측될 수 있다.

그림 2-(b)에서 DCT 및 양자화 된 Z_{ij} 가 첫 번째 열(row)에서만 0이 아닌 계수분포를 보이는 경우, 화소영역에서 각 수직(column)방향의 4개의 화소는 같은 화소 값을 갖는다. 이와 같은 블록은 수직예측모드(mode 0) 및 DC 모드로부터 예측 가능하다.

그림 2-(c)에서 DCT 및 양자화 된 Z_{ij} 가 첫 번째 행(column)에서만 0이 아닌 계수분포를 보이는 경우, 화소영역에서 각 열(row)방향의 4개의 화소는 같은 화소 값을 갖는다. 이와 같은 블록은 수평예측모드(mode 1) 및 DC 모드로부터 예측 가능하다.

그림 2-(d)에서 DCT 및 양자화 된 Z_{ij} 가 $\pm 45^\circ$ 방향으로 대칭을 이루는 경우, 이러한 블록은 모드3, 모드4 대각선(diagonal) 예측모드 및 DC 모드로부터 예측이 가능하다.

하지만 실제 영상은 균질의 영역(homogeneous region)에서도 4x4 블록내의 화소들 간의 아주 작은 화소 값의 변화로 인하여 정확한 수평방향성, 수직방향성, DC, $\pm 45^\circ$ 방향성을 갖지 않을 수 있다. 따라서 제안하는 방법에서 Z_{ij} 계수분포 정보는 9개의 방향성 모드들을 유사 방향성에 따라서 그룹화 하여 RDO를 수행할 후보를 결정하는 방법에 이용한다.

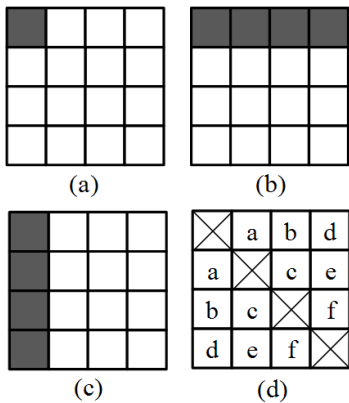


그림 2. DCT계수 분포에 따른 블록의 예지 방향 (a) DC, (b) 수직 방향 (모드 0), (c) 수평방향, (d) $\pm 45^\circ$ 대각선방향
 Fig. 2. Four directions on distribution of 4x4 Block Quantized DCT values (a) DC, (b) Vertical direction, (c) Horizontal direction, (d) $\pm 45^\circ$ Diagonal direction

III. DCT 계수 분포를 이용한 모드 결정 방법

위 식(2), 식(3)을 통해 원 영상의 각 블록 별 DCT 및 양자화 변환으로 얻은 계수 분포의 특성을 이용하여 RDO에 포함되는 후보모드를 줄이는 방법을 제안한다. 각 블록의 변환된 계수들 분포 특성에 맞추어 RDO과정에 수행되는 각 후보 모드에 대해 각각 그룹으로 나누고 각 그룹에 해당하는 모드만을 선택적으로 RDO 과정의 후보 모드로 사용함으로써 전체적인 부호화 속도를 향상시키는 방법이다.

Intra4x4, Intra16x16, Chroma8x8에 해당하는 밝기 신호 및 색차 신호에 해당하는 인트라 블록에 대해 각 그룹으로 후보 모드를 나누는 과정을 설명한다.

1. Intra4x4의 모드 결정 방법

아래 식(8)에 Z_{ij} 는 식(3)에서 계산된 블록의 값을 의미한다. 임계값 w 에 따라 0 또는 값이 존재하는 영역으로 나뉜다. 이 값을 통해 최종적으로 값의 분포를 확인한다.

$$\begin{cases} \hat{Z}_{ij} = Z_{ij}, & \text{if } |Z_{ij}| > |w| \\ \hat{Z}_{ij} = 0, & \text{otherwise} \end{cases} \quad (8)$$

위 식 (8)에서의 w 는 0.1로 정의 하였다.

- 그룹 G0: 만약 의 계수 분포가 그림 2-(a)와 같이 분포한다면 원영상은 화소간 유사성이 큰 영상으로 판단하여 가장 많이 선택되는 모드인 모드 0, 모드 1, 모드 2를 후보 모드로 결정한다. 여기서 3~8 모드를 RDO과정에서 제외시킨다.
- 그룹 G1: 만약 의 계수 분포가 그림 2-(b)와 같이 분포한다면 원영상은 수직방향에 화소간 유사성을 가지고 있다고 가정 하고 모드 0, 모드 2를 선택한다.
- 그룹 G2: 같은 방법으로 의 계수 분포가 그림 2-(c)와 같이 분포 한다면 원영상은 수평방향에 화소간 유사성을 가지고 있다고 가정고 모드 1과 모드 2를 선택한다.
- 그룹 G3: 그림 2-(d)와 같은 분포를 가지면 $+45^\circ$, -45° 의 각도에 화소간 유사성을 가지고 있다는 가정을 하고 모드 2, 3, 4를 선택한다.

•그룹 G4: 계수 분포가 그림 2에 어떤 경우에도 속하지 않는 경우는 계수 분포를 통해서 유사성을 찾지 못했다고 판단하고 9가지 모든 후보 모드를 모두 선택해 RDO 과정을 수행한다.

2. Intra16×16의 모드 결정 방법

Intra16×16은 우선 16×16에 해당하는 원 영상의 블록을 sub-sampling 방법을 통해 4×4 블록을 생성한다. 앞에 설명한 Intra4×4 블록과 같은 방법으로 계수분포에 따른 분류 과정을 수행 한다. 이때, 식(8)의 임계값 w 는 0.3으로 정의 한다. Intra16×16도 같은 방법으로 그림 2-(a)에 해당하는 경우는 G0로 그룹을 결정하고 그림 2-(b)에 해당하는 경우는 그룹 G1, 그림 2-(c)에 해당하는 경우는 그룹 G2로 구분 한다. 그 이외의 경우는 Intra4×4와 같이 계수 분포를 통해서 유사성을 찾지 못했다고 판단하며 G3로 분류한다.

3. Chroma8×8의 모드 결정 방법

Intra16×16과 같은 방법으로 Chroma8×8의 블록을 4×4 크기로 블록으로 sub-sampling하여 4×4 블록을 생성하여 계산 및 그룹으로 나누는 동일 과정을 수행한다. 이때 임계값 w 는 0.8로 정의 하였다.

표 1. 각 그룹별 예측 후보 모드
Table 1. Group of fast Intra prediction

Prediction Modes	Group	Candidate Modes
Intra4×4	G0	0, 1, 2
	G1	0, 2
	G2	1, 2
	G3	2, 3, 4
	G4	0, 1, 2, 3, 4, 5, 6, 7, 8
Intra16×16	G0	0, 1, 2
	G1	0, 2
	G2	1, 2
	G3	0, 1, 2, 3
Chroma8×8	G0	0
	G1	2, 0
	G2	1, 0
	G3	0, 1, 2, 3

표 1은 Intra4×4, Intra16×16, Chroma8×8에 해당하는 모든 그룹별 후보 모드를 보여준다.

IV. 실험 및 결과

제안하는 DCT 계수의 분포를 이용한 인트라 고속 모드 결정 방법은 JM11.0[6]을 기반으로 구현 되었다. Intel(R) Core(TM) 2.66GHz, 4GB RAM을 가진 동일한 환경의 컴퓨터에 Windows7 64bits의 운영체제 환경과 인트라 예측에 관한 실험조건으로 H.264/AVC의 Baseline 프로파일에서 모든 프레임은 I-picture로 설정하여 실험을 하였다. 또, 실험을 수행하는 영상은 현재 개발이 진행되고 있는 HEVC(High Efficiency Video Coding)에서 사용되어지는 테스트 영상들로 구성되어 있다. 또, QP는 28, 32, 36, 40, 44로 설정하여 각 300장씩 부호화 하였다. 표 2는 실험 영상 및 조건에 대해 자세히 보여준다.

표 2 실험 영상 및 조건
Table 2 Experimental Conditions

Resolution	Test Sequence	Experimental Conditions
1920x1080 (Full HD)	ParkScene	Baseline profile All I-frame 300 frames RDO enable QP(28,32,36,40,44) CAVLC
	Kimono	
	BasketballDrive	
	BQTerrace	
832x480 (WVGA)	BasketballDrill	
	RaceHorses	
	PartyScene	
	BQMall	
416x240 (WQVGA)	BasketballPass	
	BQSquare	
	BlowingBubbles	
	RaceHorses	

제안하는 고속 모드 결정방법은 JM 11.0에서 Pan의 방법^[3]과 DCT 도메인을 이용하는 유사방법인 Lin의 방법^[4]과 비교하여 성능 비교를 수행하였다. 비교 대상인 Pan의 방법은 원영상의 각 블록에 Sobel 에지 연산자를 통해 블록의 에지를 분석하고 에지에 방향에 따라 각 예측 후보 모드를 그룹화 하여 유사한 방향의 예측 방향만 선택하여 사용하

는 방법이다. Lin의 방법은 제안하는 방법과 유사한 방법으로 정수형 변환 도메인 상에서 계수의 분포를 이용한 후보모드를 그룹화 하는 방법이다. Lin의 방법은 양자화 이전의 DCT 변환 값을 이용한 방법으로 제안하는 방법과 차이가 있다. 이 방법에서는 식(2)에서 PF를 곱하는 과정에서 생기는 나눗셈을 없애기 위해 각 PF계수에 100을 곱한 근사화(Approximation)과정을 거친 값을 사용하여 나눗셈을 제거했다. 제안하는 방법에서는 양자화 파라미터(QP)에 따라 변화된 계수 값을 이용하는 방법으로써 양자화를 고려하지 않는 Lin의 방법과 차이를 가진 QP에 따라 적응적으로 변화된 계수 값을 사용할 수 있는 장점을 가진다. 또, 임계값을 이용한 후보모드 분류과정을 통해 Lin의 방법보다 RDO를 수행하는 후보모드를 결정할 때 더 정확하고 적은 수의 후보 모드의 RDO 수행을 통해 부호화 속도 및 성능 향상을 가진다.

제안하는 방법과 비교 대상인 방법은 모두 RDO를 수행하기 전 일정한 계산과정을 통해 RDO를 수행하는 후보 모드를 결정하여 기존 H.264/AVC 보다 적은 후보 모드 수를 통해 속도 향상을 가져온다. H.264/AVC Baseline Intra 코딩의 경우 (Intra16×16, Intra4×4, Chroma8×8), RDO 횟수는 매크로블록 당 총 592회 (4회(Chroma8×8모드수)×(4회(Intra16×16모드수)+9회(Intra4×4모드수)×16회(MB당Intra4×4블록수)))를 수행한다. 이는 현재 블록의 주변 복원블록이 모두 예측에 사용 가능할 때의 수를 의미하며 주변 복원 블록이 존재하지 않는 경우는 더 적은 숫자를 수행 할 수도 있다. 표 3은 제안하는 방법과 JM11.0을 표 2와 같은 실험 조건을 통한 WQVGA영상의 RDO 모드 수를 나타낸다.

위 표 3과 같이 70%이상의 후보 모드수가 들어 들었고 이는 아래 표 4의 속도 향상 결과와 유사한 결과를 보인다. 부호화 속도의 계산과 후보 모드 수는 아래와 같은 방법으

표 3. RDO의 후보 모드 수 비교

Table 3. Compared to the number of candidate modes for RDO

Resolution	Test Sequences	JM11.0	Proposed Method	Total candidate modes saving(%)
WQVGA (416x240)	BasketballPass	328338000	75521207	-76.99
	BQSquare	328338000	80506050	-75.48
	BlowingBubbles	328338000	103313046	-68.53
	RaceHorses	328338000	97937955	-70.17

표 4. 제안하는 방법의 성능 비교

Table 4. Performance of the proposed fast mode decision method compared with JM11.0 and other methods

Resolution	Test Sequences	BD-PSNR			BD-Rate			Total Encoding Time saving(%)		
		JM vs Pro	JM vs [3]	JM vs [4]	JM vs Pro	JM vs [3]	JM vs [4]	JM vs Pro	JM vs [3]	JM vs [4]
Full HD (1920x1080)	ParkScene	-0.09	-0.16	-0.67	2.05	3.74	16.93	-70.67	-59.07	-34.49
	Kimono	-0.14	-0.16	-0.82	3.12	3.86	20.97	-74.02	-61.99	-32.98
	BasketballDrive	-0.27	-0.17	-1.65	6.55	4.02	48.31	-72.08	-60.83	-41.31
	BQTerrace	-0.07	-0.14	-1.23	1.48	2.99	28.53	-72.09	-58.55	-42.46
WVGA (832x480)	BasketballDrill	-0.15	-0.17	-0.85	3.09	3.38	18.28	-65.83	-61.59	-41.16
	RaceHorses	-0.14	-0.16	-0.73	2.90	3.17	15.61	-66.62	-59.77	-40.63
	PartyScene	-0.10	-0.20	-0.96	1.67	3.39	17.38	-64.25	-57.38	-42.85
	BQMall	-0.18	-0.25	-1.69	3.20	4.32	33.18	-71.07	-59.54	-44.23
WQVGA (416x240)	BasketballPass	-0.19	-0.22	-1.09	3.48	4.12	22.51	-68.75	-58.86	-49.78
	BQSquare	-0.07	-0.24	-0.85	1.03	3.30	12.20	-69.87	-59.16	-41.39
	BlowingBubbles	-0.10	-0.18	-0.75	1.97	3.42	15.41	-62.82	-57.18	-42.75
	RaceHorses	-0.14	-0.18	-0.84	2.65	3.47	17.31	-62.79	-59.68	-42.02
Total average		-0.14	-0.18	-1.01	2.76	3.60	22.22	-68.40	-59.47	-41.34

로 표현된다.

$$S(\%) = \frac{T_{method} - T_{JM}}{T_{JM}} \times 100 \quad (9)$$

위 식(9)에서 S(%)는 절약된 %를 나타내며 TJM 은 동일한 실험 조건에서 JM11.0 부호기에 전체 부호화 시간 또는 후보 모드 수를 의미하고 TMethod 는 제안하는 방법 및 다른 제안 방법에 부호화 시간 또는 후보 모드 수를 의미한다.

표 4는 비교대상인 방법들과 제안하는 방법의 Bit-rate, PSNR, 및 부호화 시간의 성능을 나타낸다. 표 4에서 BD-PSNR과 BD-Rate는 각각 비트 증가율 및 PSNR의 계산 방법을 통하여 계산하였다^[7]. 표 4와 같이 제안하는 방법은 JM11.0과 비교 대상인 방법들과 비교하여 가장 적은 비트증가 및 PSNR 손실에도 불구하고 가장 좋은 부호화 시간 감소를 보인다.

아래 그림 3~5는 JM11.0, 제안하는 방법, Pan의 방법, Lin의 방법에 대한 RD 곡선을 나타낸다. 결과적으로 제안

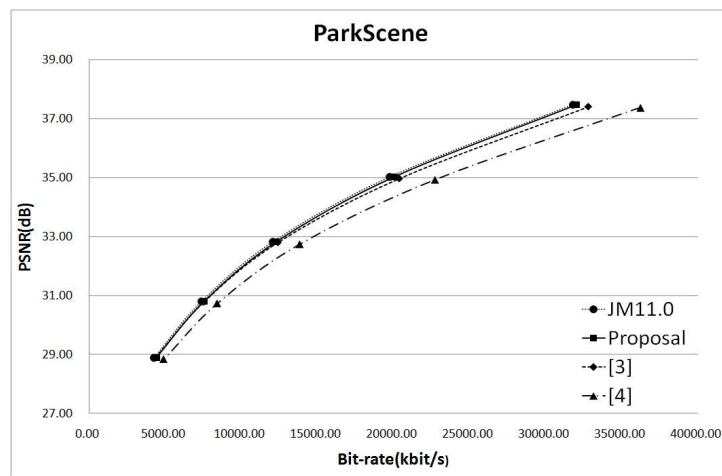


그림 3. JM 11.0에서 ParkScene 영상의 제안하는 방법 비교 RD곡선

Fig. 3. RD-curve for ParkScene with/without the proposed fast algorithm and with other algorithms in JM 11.0

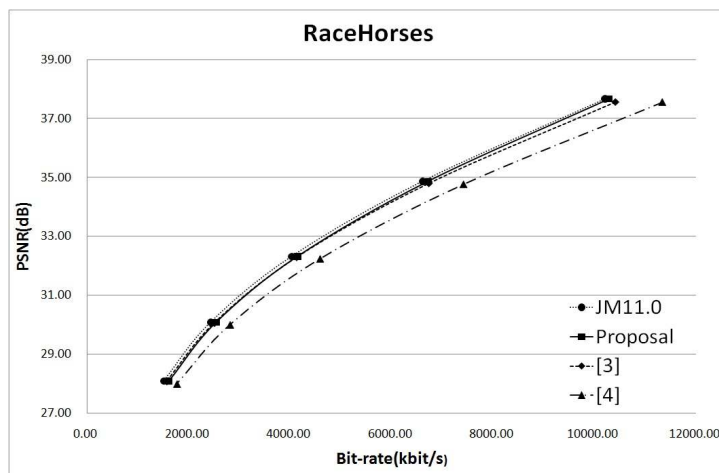


그림 4. JM 11.0에서 RaceHorses 영상의 제안하는 방법 비교 RD곡선

Fig. 4. RD-curve for RaceHorses with/without the proposed fast algorithm and with other algorithms in JM 11.0

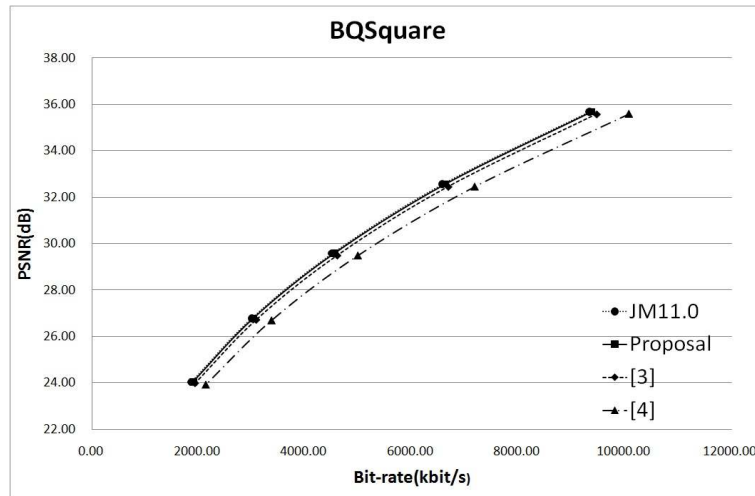


그림 5. JM 11.0에서 BQSquare 영상의 제안하는 방법 비교 RD곡선
 Fig. 5. RD-curve for BQSquare with/without the proposed fast algorithm and with other algorithms in JM 11.0

하는 DCT 계수를 이용한 인트라 고속 모드 결정 방법은 가장 좋은 비트 증가율 및 PSNR 손실률에도 평균 68.40%의 속도 향상으로 평균 59.47%의 Pan의 방법과 41.34%의 Lin의 방법과 비교해 가장 좋은 성능을 보였다. 또, 향후 개발될 코덱에서 사용될 고해상도 영상에서 더 좋은 성능을 가지는 실험 결과를 통해 향후 더욱 발전 가능성을 보여준다.

V. 결론

부호기의 율-왜곡 최적화(RDO)의 수행 방법의 개선을 통해 부호화 속도를 감소시키는 고속 인트라 모드 결정 방법을 제안하였다. 이 방법은 원 영상의 각 블록에 DCT 변환 계수를 통한 예측을 통해서 RDO과정을 수행할 후보 모드의 수를 그룹화 함으로써 전체 부호화의 속도향상을 가져오는 방법이다. 제안한 방법은 기존 고속 인트라 모드 결정 방법들과 비교해 성능 향상을 가지며, 적은 부호화 손실에도 불구하고 전체 부호화 시간을 평균 68.40% 부호화 시간의 성능 향상을 가진다.

참고 문헌

- [1] ITU-T Recommendation H.264 and ISO/IEC 14496-10, Advanced Video Coding for Generic Audio visual Services, May 2003 (and subsequent amendment [X] and corrigenda).
- [2] G. J. Sullivan, T. McMahon, T. Wiegand, and A. Luthra, "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 JVT Document JVT-L047, July 2004.
- [3] Feng Pan, Xiao Lin, Susanto Rahardja, Keng Pang Lim, Z. G. Li, Dajun Wu, and Si Wu, "Fast Mode Decision Algorithm for Intraprediction in H.264/AVC Video Coding", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 15, NO. 7, JULY 2005.
- [4] Heng-Yao Lin, Kuan-Hsien Wu, Bin-Da Liu and Jar-Ferr Yang, "Transformed-Based Mode Decision Algorithm for H.264/AVC Intraprediction", IEEE Asia Pacific Conference on Digital Object Identifier, Circuits and Systems, APCCAS 2008, pp. 1344-1347, DEC 2008.
- [5] Huanqiang Zeng, Canhui Cai, Senior Member, IEEE, and Kai-Kuang Ma, Senior Member, IEEE, "Fast Mode Decision for H.264/AVC Based on Macroblock Motion Activity", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 19, NO. 4, APRIL 2009.
- [6] http://iphome.hhi.de/suehring/tml/download/old_jm/11.0.zip
- [7] G.Bjontegaard, Calculation of average PSNR differences between RD-curves, ITU-T Q.6/SG16 VCEG, 15th Meeting, VCEG-M33, Austin, Texas, USA, 2-4 April, 2001.

저 자 소 개



홍 성 옥

- 2009년 2월 : 세종대학교 인터넷공학과 (학사)
- 2009년 3월 ~ 현재 : 세종대학교 컴퓨터공학과 (석사과정)
- 주관심분야 : 영상압축, H.264/AVC, HEVC(High Efficiency Video Coding)



이 영 렬

- 1985년 2월 : 서강대학교 전자공학과 (학사)
- 1987년 2월 : 서강대학교 전자공학과 (석사)
- 1999년 2월 : 한국 과학기술원 전기전자공학과 (박사)
- 1987년 1월 ~ 1994년 2월 : 삼성전자 R&D 센터 Digital Media Lab.
- 1999년 3월 ~ 2001년 8월 : 삼성전자 R&D 센터 Digital Media Lab. 수석연구원
- 2001년 9월 ~ 현재 : 세종대학교 전자정보대학 컴퓨터공학부 컴퓨터공학과 교수
- 주관심분야 : Multi-view video coding, MPEG video compression, Multimedia(MPEG) systems, H.264/AVC, Image compression, Image processing, Multimedia data(A/V/System) over IP, Embeded system for A/V transport, Transcoding, Watermarking, HEVC(High Efficiency Video Coding)