

# H.264/AVC에서 모드 결정을 위한 개선된 율-왜곡 예측

## Improved Rate-Distortion Estimation for Mode Decision in H.264/AVC

박기홍\*, 김윤호\*\*

Ki-Hong Park\*, Yoon-Ho Kim\*\*

### 요 약

본 논문은 H.264/AVC의 효율적 모드 결정을 위한 율-왜곡 예측 방법을 제안하였다. 제안한 방법에서 모드 결정은 DCT 계수의 라플라시안 분포 모델링을 적용하여 비트율을 예측함으로써 양자화와 엔트로피 부호화의 과정이 불필요하다. 실험결과, 실제 율-왜곡치와 제안한 방식과의 율-왜곡치가 거의 일치하였고, 기존의 방식보다 0.02 dB의 PSNR 이득을 보였다.

### Abstract

This paper presented a rate-distortion estimation method for effective mode decision in H.264/AVC. In this approach, in order to decide a mode, laplacian distribution modeling of DCT coefficients is utilized, which do not need to such process as quantization, entropy coding. From the simulation results, proposed a method showed that rate-distortion between proposed scheme and practical value was almost the same and performed 0.02dB of PSNR gain.

Key words : H.264/AVC, Bit-Rate, Distortion, Estimation, RDO

### I. 서 론

H.264/AVC [1]는 ITU-T의 Video Coding Expert Group (VCEG)과 ISO/IEC의 MPEG의 활동에 의해 2003년 5월에 표준이 완성되었고, 이전의 압축 표준들인 MPEG-1/2/4 및 H.263에 비해서 높은 압축률의 향상을 가져왔다. 즉, 기존 영상 압축 부호화 표준인 MPEG-4, H.263, MPEG-2와 비교하여 동일 화질에 39%, 49%, 64%의 비트율 감소를 보인다 [2]. 이러한 H.264/AVC의 높은 압축 성능은 새롭게 추가된 예측 기법들과 매크로블록 모드를 구성하는 다양한 크기

의 블록에 대해 움직임 예측 후의 움직임 예측 오차와 움직임 벡터로 이루어진 움직임 비용이 가장 작은 매크로블록 모드를 선택하여 비트율 감소를 가져왔다. 또한 압축률 향상의 요인으로 율-왜곡 최적화(RDO:Rate-Distortion Optimization)된 모드 결정 방식을 들 수 있다.

RDO 모드 결정 방식은 움직임 예측 과정에서 계산된 움직임 비용으로 모드를 결정하지 않고 비트율과 왜곡치를 계산하여 율-왜곡 비용을 최소로 가지는 매크로블록 모드를 선택하는 것이다 [3]. 이 방식은 약 7%의 비트율 감소를 가져왔지만 전체 부호화 시간의 약 50%를 증가시켰다 [4][5]. 따라서 압축률을

\* 목원대학교 일반대학원 IT공학과 박사과정

\*\* 목원대학교 컴퓨터공학부 교수

· 제1저자 (First Author) : 박기홍 · 교신저자 (Corresponding Author) : 김윤호

· 투고일자 : 2010년 2월 8일

· 심사(수정)일자 : 2010년 2월 9일 (수정일자 : 2010년 2월 15일)

· 게재일자 : 2010년 2월 28일

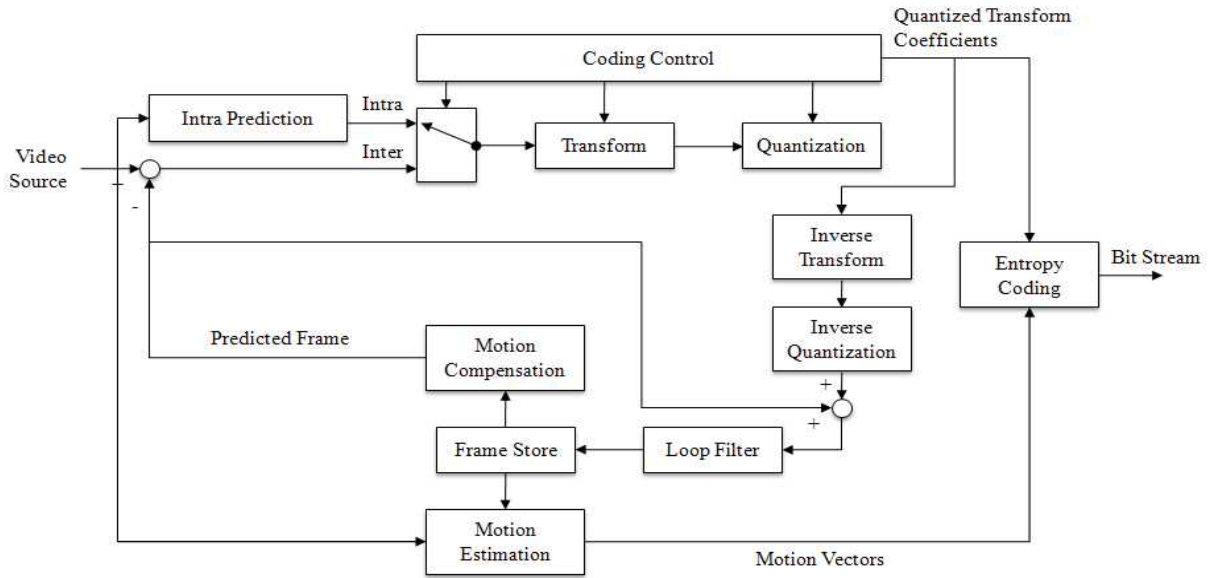


그림 1. H.264 블록도  
Fig. 1. H.264 Block Diagram

높이기 위한 방식들은 부호화 계산량이 증가되었고, 이에 매크로블록 모드의 움직임 예측과 모드 결정에 대한 계산량을 줄이기 위한 연구가 활발히 이루어지고 있다.

본 논문에서는 기존의 율-왜곡 예측 결정을 위해 모드 제거 방식을 살펴보고, DCT 영역과 양자화 단계에서의 율-왜곡 예측 방식을 제안한다.

## II. H.264/AVC의 RDO 모드 결정

H.264/AVC에서는 그림 2와 같은 블록 모드에서 움직임 예측 과정 후의 움직임 비용이 가장 작은 모드를 선택할 수 있지만, 보다 최적화된 성능을 얻기 위해 RDO 모드 결정 방식을 사용하여 모드를 결정한다.

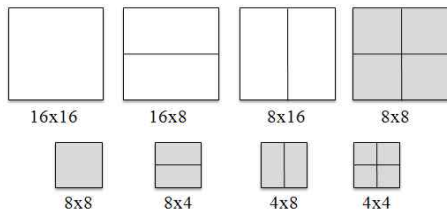


그림 2. 인터 매크로블록 모드  
Fig. 2. Inter Macro block Mode

RDO 모드 결정 방식에서는 움직임 예측 후에 각 모드에 대해서 식 1과 같은 율-왜곡 비용이 최소화 되는 모드를 선택한다.

$$\min_I D(MODE | QP), \quad (1)$$

$$R(MODE | QP) \leq R_T$$

여기서  $MODE$ 는 매크로블록 부호화 모드,  $QP$ 는 주어진 양자화 파라미터,  $R_T$ 는 매크로블록 할당 비트,  $D$ 와  $R$ 은 각각 왜곡과 생성 비트이다. 최소가 되는 해를 구할 때 라그랑지 승수(Lagrange multiplier)  $\lambda$ 를 이용한 최적화 기법을 사용하여 식 2와 식 3에 적용할 수 있다.

$$MODE^* = \operatorname{argmin}_{MODE} J(MODE | QP, \lambda_{MODE}) \quad (2)$$

$$J(MODE | QP, \lambda) = D(MODE | QP) + \lambda_{MODE} \cdot R(MODE | QP) \quad (3)$$

$\lambda_{MODE}$ 는 해당 모드에 대한 왜곡치로 원 블록과 복원 블록의 차의 제곱의 합(SSD: Sum of Squared Differences)이다.  $R(MODE | QP)$ 는 해당 모드에 대한 비트수를 나타내고, 헤더 정보, 변환과 양자화

처리 후의 변환 계수부분의 비트수를 포함한다. 따라서, 복원 블록을 얻기 위해서는 일반적인 부호화 과정인 원 블록과 움직임 예측 블록의 차인 예측 오차 블록에 DCT, 양자화, 역양자화, IDCT를 차례로 적용한 후 움직임 예측 블록과 더하는 과정이 수행되어야 한다. 그림 3은 H.264/AVC의 참조 프로그램 [6]에서 비트율과 왜곡치를 계산하는 과정을 나타낸다.

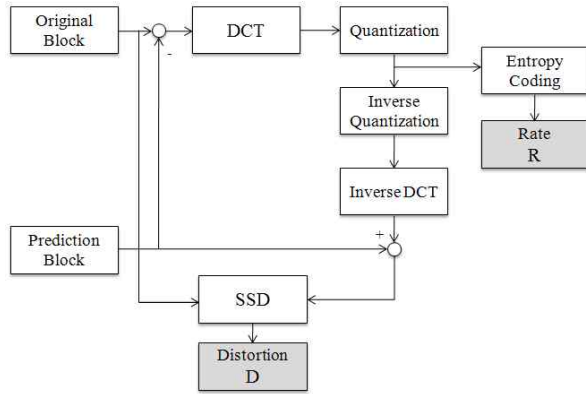


그림 3. H.264/AVC 참조 프로그램에서의 RDO 모드 결정을 위한 율-왜곡 비용 계산  
Fig. 3. Cost calculation for RDO mode decision in H.264/AVC reference software

H.264/AVC에서 복원 오차로부터 계산되는 왜곡치는 DCT 영역에서 양자화 오차로 표현된다. 4x4 크기의 잔여 영상 X를 DCT한 계수 블록 Y를 구하는 과정은 다음 식과 같다.

$$Y = (C_f X C_f^T) \otimes E_f = W \otimes E_f \quad (4)$$

$C_f$ ,  $E_f$ ,  $\otimes$ 는 각각 정수 DCT 행렬, 스케일링 행렬 및 각 요소에 대한 스칼라 곱하기 연산을 나타낸다.  $k$ ,  $l$ 번째 위치의 양자화된 DCT 계수  $Y_Q[k, l]$ 은 다음 식과 같다.

$$Y_Q[k, l] = \begin{cases} \left\lceil \frac{W[k, l] \cdot E_f[k, l] + d}{\Delta} \right\rceil, & \text{if } W[k, l] \geq 0 \\ \left\lfloor \frac{W[k, l] \cdot E_f[k, l] - d}{\Delta} \right\rfloor, & \text{else} \end{cases} \quad (5)$$

here,  $\Delta = 2^{(QP-4)/6}$ .

$\lceil \cdot \rceil$ ,  $\Delta$ ,  $d$ 는 각각 소수점 버림 연산, 양자화 크기 및 오프셋이다.  $\Delta$ 와  $d$ 는 QP 값에 의해 정의되고, 인터 블록일 때  $d$ 는  $\Delta/6$ 의 값을 갖는다.  $Y_Q[k, l]$ 를 역양자화 한 계수를  $\hat{Y}_Q[k, l]$ 라 할 때 오차 블록  $\hat{X}$ 는 다음 식으로 표현되고,  $C_i$ 와  $E_i$ 는 역 정수 DCT 행렬과 역 스케일링 행렬을 나타낸다.

$$\hat{X} = \frac{1}{64} C_i^T (\hat{Y} \otimes E_i) C_i \quad (6)$$

Tu [7]는 H.264/AVC의 율-왜곡 계산량을 감소시키기 위해서 DCT 영역에서 율-왜곡을 예측하였다. Tu [7]가 제안한 DCT 영역에서  $k$ ,  $l$ 번째 위치의 왜곡치  $D_{Tu}[k, l]$ 는 식 7과 같이 계산하였다.

$$D_{Tu}[k, l] = \left( W[k, l] - \frac{1}{64} [EWE]_{k, l} \right)^2 \times E_f[k, l]^2 \quad (7)$$

here,  $E = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$ .

또한 4x4 블록에서  $Y_Q[k, l]$ 이 '0'이 아닌 개수를  $N$ 이라고 할 때, 비트율  $R_{Tu}$ 는 다음 식과 같이 예측하였다.

$$R_{Tu} = \alpha \cdot N + \beta \cdot \sum_{k=0}^3 \sum_{l=0}^3 |Y_Q[k, l]| \quad (8)$$

여기서  $\alpha$ ,  $\beta$ 는 비트율 예측을 위한 상수 값들을 나타낸다 [8]. Tu [7]의 방식에서는 DCT 영역에서 왜곡치와 비트율을 예측함으로써 엔트로피 부호화, 역 DCT, 복원 과정에 대한 계산량을 생략시키고 실제 왜곡치와 비트율에 근사한 값을 얻을 수 있었다. 또한 Moon [9]은 양자화 오차의 확률 분포 [9][10]를 이용하여 Tu [7]의 방식에서 남아있던 양자화와 역 양자화 과정을 생략시켜 비트율과 왜곡치를 예측하는 방식이 제안되었다.

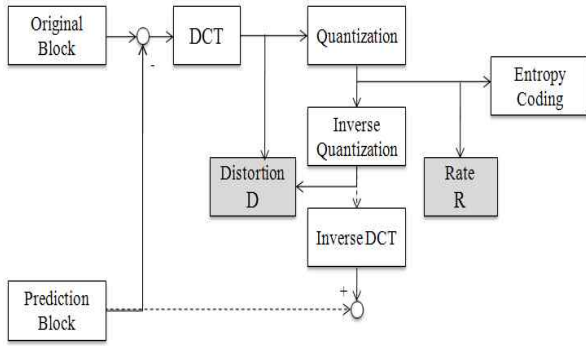


그림 4. Tu[7]가 제안한 율-왜곡 예측  
Fig. 4. Rate-Distortion method by Tu[7]

### III. 제안한 율-왜곡 예측

본 논문에서 제안하는 율-왜곡 예측 방식은 그림 5와 같고, 복잡도의 감소를 위하여 오차 블록의 DCT 계수의 영-평균 라플라시안 분포 모델링을 적용하여 비트율을 예측하였다.

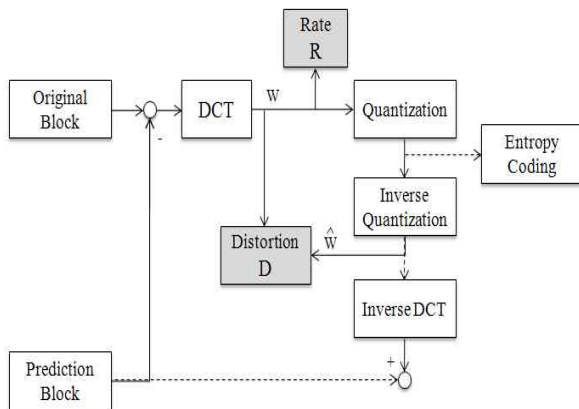


그림 5. 제안 방식의 율-왜곡 예측  
Fig. 5. Proposed Rate-Distortion method

라플라시안 확률 분포 함수(PDF: probability density function)에 기반한 '0'이 아닌 수로 양자화 되는 경우의 양자화기 입력, 출력 및 양자화 오차의 분포는 다음 식으로 나타낼 수 있다.

$$f_{\epsilon_{nz}}(\epsilon) = \begin{cases} \frac{\tau}{2} e^{\tau\epsilon} \frac{1}{(e^{\tau d} - e^{\tau(-\Delta+d)})}, & \text{for } -\Delta+d < \epsilon < -d, \\ \frac{\tau}{2} (e^{\tau\epsilon} + e^{-\tau\epsilon}) \frac{1}{(e^{\tau d} - e^{\tau(-\Delta+d)})}, & \text{for } -d < \epsilon < d, \\ \frac{\tau}{2} e^{-\tau\epsilon} \frac{1}{(e^{\tau d} - e^{\tau(-\Delta+d)})}, & \text{for } d < \epsilon < \Delta-d. \end{cases} \quad (9)$$

여기서  $\tau$ 는 라플라시안 계수이다. 식 5를 이용하여 DCT 계수  $Y_Q[k, l]$ 로부터 양자화된 DCT 계수의 값이 '0'인지 아닌지를 식 10에 의해 판별한다.

$$nz[k, l] = \begin{cases} 0, & \text{when } |W[k, l]| < \frac{\Delta}{E_f[k, l]}, \\ 1, & \text{otherwise.} \end{cases} \quad (10)$$

식 9와 식 10에 의해 식 11과 같이 블록 내의 '0'으로 양자화 되지 않는 계수들의 개수로부터 비트율이 예측된다.

$$R(est) = \begin{cases} 1, & \text{when } N = 0, \\ a \cdot N + b, & \text{otherwise.} \end{cases} \quad (11)$$

$$\text{here, } a = \frac{M \cdot \sum_{i=0}^M N_i R_i - \sum_{i=0}^M N_i \sum_{i=0}^M R_i}{M \cdot \sum_{i=0}^M N_i^2 - \left(\sum_{i=0}^M N_i\right)^2},$$

$$b = \frac{\sum_{i=0}^M R_i - a \sum_{i=0}^M N_i}{M}$$

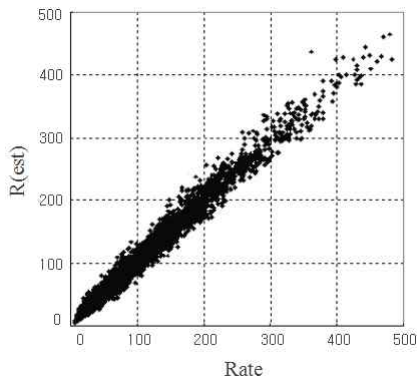
여기서  $a$ 와  $b$ 는 상수로서  $M$ 개의 표본 블록에 대해서 일반적인 선형 회기 기법으로 구할 수 있다. 왜곡치 예측은 양자화와 역 양자화를 하지 않고 오차 블록의 DCT 계수 영역에서 오차의 특성을 분석하여 예측하기도 하지만, QP 값에 대한 복원 왜곡치와 예측된 왜곡치의 평균 절대 차이(MAD)의 고른 분포를 위해 식 7을 적용하였다.

### IV. 실험 결과 및 고찰

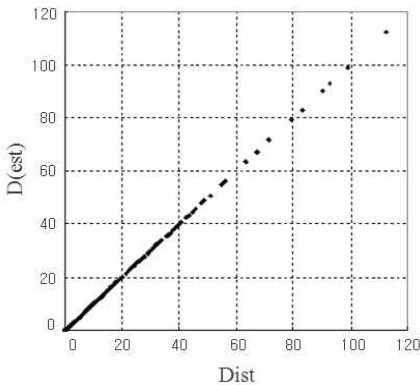
제안 방식의 타당성 평가를 위하여 기존의 방식과 성능을 비교, 평가하였다. H.264/AVC의 참조 프로그램

램인 JM에서 모의 실험하였고, 테스트 시퀀스로는 QCIF 형식의 'foreman', 'news', 'container' 그리고 'silent' 영상을 이용하였으며, 시퀀스는 각각 IPPP 형식의 100 프레임으로, QP 값은 28, 32, 36, 40이다. 또한 탐색 범위는 16이고, 인터 모드만 적용하였다.

그림 6은 'foreman.qcif' 영상의 QP값이 28일 때, 실험에서 얻은 실제 율-왜곡치와 제안 방식으로 계산된 율-왜곡을 비교한 실험 결과값으로 두 값이 대동소이함을 확인하였다.



(a) 실제 비트율(Rate)과 제안 방식의 비트율(R(est))



(b) 실제 왜곡치(Dist)와 제안 방식의 왜곡치(D(est))

그림 6. 실제 율-왜곡치와 제안 방식의 율-왜곡치  
Fig. 6. Actual rate-distortion versus proposed rate-distortion

표 1에서 PSNR의 음수는 PSNR의 감소를 의미하는 바, Tu [7]의 방식과 제안하는 방식의 성능이 평균 0.076dB 와 0.056dB로 거의 일치함을 알 수 있다.

표 1. PSNR 비교

Table 1. Comparison of PSNR

QCIF		PSNR	
		Tu's	Proposed
foreman	10 Hz	-0.055	-0.061
	30 Hz	-0.104	-0.019
news	10 Hz	-0.078	-0.071
	30 Hz	-0.084	-0.042
container	10 Hz	-0.071	-0.069
	30 Hz	-0.096	-0.091
silent	10 Hz	-0.072	-0.049
	30 Hz	-0.052	-0.043

### V. 결 론

본 논문에서는 H.264/AVC 의 효율적 모드 결정을 위하여 개선된 율-왜곡 예측 방법을 제안하였다. 제안한 방법에서는 율-왜곡 예측 모드 결정을 위하여 오차 블록의 DCT 계수의 영-평균 라플라시안 분포 모델링을 적용하였는 바, 따라서 양자화와 엔트로피 부호화의 과정이 생략되었다. 실험결과, 실제 율-왜곡치와 제안 방식으로 계산된 율-왜곡은 거의 일치하였고, Tu [7]보다 평균 PSNR 0.02dB의 성능 이득을 얻었다. 향후, 계산량 감소를 위해 율-왜곡치 예측 모드 결정에 요구되는 양자화와 역 양자화 과정의 생략과 성능 개선 연구가 필요하다.

### 참 고 문 헌

[1] Joint Video Team of ITU-T and ISO/IEC JTC 1. "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H264 ISO/IEC 14496-10 AVC", *Joint Video Team of ISO/IEC MPEG and ITU-T VCEG. doc. JVT-Go5o. Mar. 2003.*

[2] R. Schafer, T. Wiegand, and H. Schwarz, "The emerging H.264 standard EBU Technical Review," <http://www.ebu.ch/>, January 2003

[3] "Text Description of Joint Model Reference Encoding Methods and Decoding Concealment Methods", *Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Document JVT-K049, Munich, Germany*, Mar. 2004.

[4] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC: Tools, Performance, and Complexity", *IEEE Circuits and Systems Magazine, First Quarter*, pp. 7-28, 2004.

[5] Lain E.G. Richardson, *H.264 and MPEG-4 Video Compression*, John Wiley & Sons Inc. 2003.

[6] Joint Video Team (JVT) Reference Software. Available: <http://bs.hhi.de/~suehring/tml/download/>

[7] Y. K. Tu, J. F. Yang, and M. T. Sun, "Efficient Rate-Distortion Estimation for H.264/AVC Coder", *IEEE Trans. Circuits Syst. Video Technol.*, vol 16, no. 5, pp. 600-611, May 2006.

[8] D. F. Mix, *Random signal processing*, Prentice Hall International, Inc. pp. 285-289.

[9] J. M. Moon and J. H. Kim, "A Rate-Distortion Estimation Method for Fast Mode Decision in H.264", *Journal of KICS*, vol. 33, no. 1, pp. 34-44, Jan. 2008.

[10] I. Pao and M. Sun, "Modeling DCT Coefficients for Fast Video Encoding", *IEEE Trans. Circuits Syst. Video Technol.*, vol 9, no. 4, pp. 608-616, Jun. 1999.

김 윤 호 (金允鎬)



2003 - 현재: 목원대학교 컴퓨터공학부 교수  
 2005-2006: University of Auckland, NZ, CITR Lab Research Fellow.  
 2008-현재: ISO/TC223 Societal Security Korea Delegate.  
 IEEE, 대한전자공학회, 한국통신학회 정회원/ 한국정보 기술학회, 한국해양정보통신학회, 한국 모바일학회, 한국항행학회, 사회안전학회 종신회원  
 관심분야: 영상처리, 컴퓨터비전, 사회안전 표준화, IT정책, 방재정보통신 등.

박 기 홍 (朴起弘)



2004년 2월: 목원대학교 컴퓨터공학과 (공학사)  
 2006년 2월 : 목원대학교 공학석사  
 2006년 8월~현재 : 목원대학교 IT공학부 박사과정  
 관심분야: 영상처리, 컴퓨터비전