

H.264에서 고속의 모드 결정을 위한 율-왜곡 예측 방식

정회원 문정미*, 김재호**

A Rate-Distortion Estimation Method for Fast Mode Decision in H.264

Jeong-mee Moon*, Jae-ho Kim** *Regular Members*

요약

본 논문에서는 H.264 부호화기에서 고속의 율-왜곡 최적화된 모드 결정을 위한 효과적인 율-왜곡 예측 방식을 제안한다. 비트율과 왜곡치가 최소가 되는 율-왜곡 최적화된 모드를 결정하기 위해서는 각 매크로 블록 모드를 부호화하고 복호화 하여 비트율과 왜곡치를 계산해야 한다. 이 과정에서 요구되는 계산량을 줄이고자 DCT 영역에서 비트율과 왜곡치를 효과적으로 예측하는 방식을 제안한다. 모의실험 결과는 제안 방식을 이용하여 모드 결정을 하였을 때 율-왜곡 성능은 거의 유지되면서 H.264의 율-왜곡 최적화된 모드 결정 계산량의 약 81%를 감소시키고 전체 부호화 계산량의 약 26%를 감소시킴을 보여준다.

Key Words : H.264, Distortion, Bit-rate, Estimation, Mode decision, RDO

ABSTRACT

In this paper, a new rate-distortion estimation method is proposed for fast rate-distortion optimized(RDO) mode decision in H.264 encoder. A bit-rate and distortion should be calculated by encoding and reconstruction of each macroblock mode in order to select the optimal RDO mode. To reduce the computations for getting them, a bit-rate and distortion is efficiently estimated in DCT domain. Simulation results show that the proposed method achieves computational saving about 81% of RDO mode decision and saves about 26% of total encoding time without noticeable degradation in coding performance.

I. 서론

H.264^[1]는 가장 최근에 제정된 동영상 압축 표준으로서 이전의 압축 표준들인 MPEG-1/2/4와 H.263에 비해서 높은 압축률의 향상을 가져왔다. H.264의 압축률 향상을 가져온 주요 요인들 중 하나로 16x16에서 4x4까지 다양한 블록 크기들로 이루어진 매크로 블록 모드를 들 수 있다. 하나의 매크로 블록 모드를 구성하는 각각의 다양한 크기의 블록에 대해 움직임 예측을 한 후에 움직임 예측 오차와 움직임

벡터로 이루어진 움직임 비용이 가장 작은 매크로 블록 모드를 선택한다. 이 방식으로 압축률은 16x16 블록의 모드만 사용하는 경우에 비해서 약 16%의 비트율이 감소된 반면 약 100%의 전체 부호화 계산량이 증가하였다^[2]. 또 다른 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 매크로 블록 모드

* 부산대학교 일반대학원 전자공학과 영상통신 연구실(moonjmee@pusan.ac.kr)

** 부산대학교 컴퓨터 및 정보통신 연구소(Research Institute of Computer Information and Communication) (jhkim@pusan.ac.kr)

논문번호 : KICS2007-07-313, 접수일자 : 2005년 7월 20일, 최종논문접수일자 : 2007년 12월 24일

를 선택하는 것이다³⁾. 이 방식은 약 7%의 비트율 감소를 가져왔지만 전체 부호화 시간의 약 50%를 증가시켰다^{2,4)}. 이와 같이 H.264의 압축률을 높이기 위한 방식들은 부호화 계산량의 증가를 야기시켰다. 따라서 H.264의 성공적인 활용을 위해서는 이들에 대한 성능은 유지시키면서도 요구되는 계산량을 감소시키는 것이 필수적이다.

지금까지 각 매크로 블록 모드의 움직임 예측과 모드 결정에 대한 계산량을 줄이기 위한 많은 연구가 이루어져왔다⁵⁻⁷⁾. Lim⁵⁾은 경계의 크기 값으로부터 블록의 균질성을 측정하여 이에 따라 선택될 가능성이 높은 모드에 대해서만 움직임 예측과 모드 결정을 하는 방식을 제안하였다. Kim⁶⁾은 모드 결정 단계에서 양자화된 DCT 계수들로 후보 블록들의 선택 여부를 결정하였다. Jing⁷⁾은 현재 영상과 이전 영상의 차이로 균질성을 측정하여 움직임 예측과 모드 결정 여부를 판단하였다. 이러한 방식들은 약 50% 정도의 매크로 블록 모드들을 생략시킴으로서 남아있는 모드에 대해서만 최적의 모드를 선택하게 한다. 이 때 RDO 모드 결정 방식을 사용하게 되면 비트율과 왜곡치를 얻기 위해 생략되지 않은 각 모드를 부호화, 복원화 하여야 한다. 즉, 원 블록과 움직임 예측 블록의 차이 예측 오차 블록에 대한 DCT, 양자화, 엔트로피 부호화, 역 양자화, 역 DCT, 그리고 복원 과정이 차례로 이루어져야 한다. 따라서 이 부분에 대한 고속 부호화를 위하여 효율적인 윌-왜곡 계산에 대한 연구가 필요하다.

최근 Tu⁸⁾가 윌-왜곡 계산량을 감소시키는 방식을 제안하였다. DCT 영역에서 양자화 계수들로부터 비트율을 예측하고 DCT 계수와 역 양자화된 DCT 계수들로부터 왜곡치를 예측함으로써 엔트로피 부호화, 역 DCT, 그리고 복원 과정에 요구되는 계산량을 감소시켰다. 그러나 이 방식은 여전히 DCT 계수들의 양자화, 역 양자화 과정을 필요로 한다. 따라서 RDO 모드 결정의 계산량을 보다 더 감소시키기 위해서 효과적인 윌-왜곡 예측 방식이 요구된다.

본 논문에서는 DCT 영역에서 비트율과 왜곡치를 예측한다. 그러나 Tu⁸⁾의 방식과는 달리 양자화와 역 양자화를 하지 않고 양자화 오차의 특성을 분석하여 왜곡치를 예측하고 DCT 계수값으로부터 비트율을 효과적으로 예측한다. 제안 방식은 DCT 계수로부터 비트율과 왜곡치를 예측함으로써 RDO 모드 결정의 윌-왜곡 계산에 요구되었던 양자화, 엔트로피 부호화, 역 양자화, 역 DCT, 그리고 복원 과정

을 생략시킨다. 모의실험 결과는 제안하는 방식을 사용하여 성능은 H.264의 RDO 모드 결정 방식과 거의 유사 하면서도 계산량이 효과적으로 감소됨을 보인다.

II. H.264의 RDO 모드 결정 방식

H.264에서는 그림 1과 같이 다양한 블록 크기를 가지는 7개의 모드들 중에서 최적의 인터 매크로 블록 모드를 결정한다. 이 때 RDO 모드 결정 방식을 사용하지 않고 움직임 예측 과정에서 얻은 움직임 비용이 가장 작은 모드를 선택할 수 있다. 그러나 보다 최적화된 성능을 얻기 위해서는 RDO 모드 결정 방식을 사용하여 모드를 결정한다.

RDO 모드 결정 방식에서는 움직임 예측 후에 각 모드에 대해서 식 (1)과 같은 윌-왜곡 비용이 최소가 되는 모드를 선택한다.

$$RD_Cost_{MODE} = Distortion_{MODE} + \lambda_{MODE} \cdot Rate_{MODE} \quad (1)$$

여기서 λ_{MODE} , $Rate_{MODE}$ 는 각각 라그랑지 계수, 해당 모드에 대한 비트수를 나타낸다. $Distortion_{MODE}$ 는 해당 모드에 대한 왜곡치로 원 블록과 복원 블록의 차의 제곱의 합(SSD:Sum of Squared Differences)으로 정의된다³⁾. 복원 블록을 얻기 위해서는 일반적인 부호화 과정인 원 블록과 움직임 예측 블록의 차이 예측 오차 블록에 DCT, 양자화, 역 양자화, 역 DCT를 차례로 적용한 후 움직임 예측 블록과 더하는 복원 과정이 수행되어야 한다. $Rate_{MODE}$ 는 헤더 정보에 해당하는 모드, 움직임 벡터, 양자화 변수 정보와 예측 오차 블록에 대한 비트 수를 포함한다. 따라서 $Rate_{MODE}$ 를 구하기 위해서는 헤더 정보에 대한 엔트로피 부호화와 예측 오차 블록에 대한 DCT, 양자화, 엔트로피 부호화가 수행되어야 한다. 그림 2는

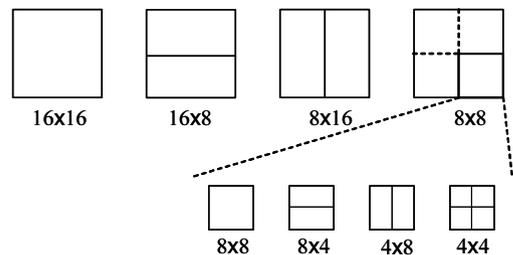


그림 1. H.264의 인터 매크로 블록 모드

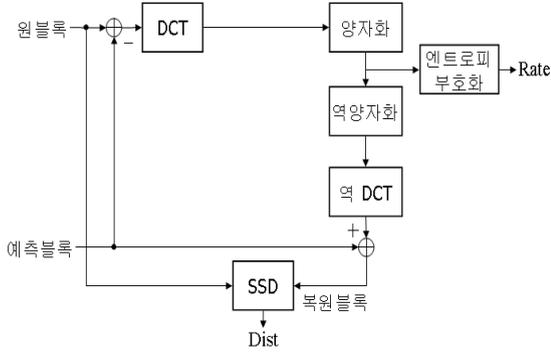


그림 2. H.264의 참조 프로그램[14]에서 RDO 모드 결정을 위한 비트율-왜곡치 계산 방식

H.264의 참조 프로그램[14]에서 비트율과 왜곡치를 계산하는 과정을 나타낸 것이다.

III. 기존의 고속 울-왜곡 예측 방식

일반적으로 공간 영역에서 복원 오차로부터 계산되는 왜곡치는 DCT 영역에서 양자화 오차로 표현된다. Tu^[8]의 방식을 설명하기에 앞서 H.264에서의 DCT, 양자화 과정에 대해서 설명한다. H.264에서 4x4 크기의 오차 블록 \mathbf{U} 를 DCT한 계수 블록 \mathbf{V} 를 구하는 과정은 다음 식과 같다^[4].

$$\mathbf{V} = (\mathbf{B}_F \mathbf{U} \mathbf{B}_F^T) \otimes \mathbf{P}_F = \mathbf{Z} \otimes \mathbf{P}_F. \quad (2)$$

여기서 \mathbf{B}_F , \mathbf{P}_F , \otimes 는 각각 정수 DCT 행렬, 후크기 변환 행렬, 그리고 각 요소에 대한 곱하기 연산을 나타낸다. 위의 식에서 후크기 변환을 하지 않은 정수 DCT 계수 행렬을 \mathbf{Z} 라고 정의한다. k, l 번째 위치의 양자화된 DCT 계수 $V_q[k, l]$ 은 다음 식으로 나타낼 수 있다^[4].

$$V_q[k, l] = \begin{cases} \frac{Z[k, l] \cdot P_F[k, l] + d}{\Delta}, & \text{if } Z[k, l] \geq 0, \\ \frac{Z[k, l] \cdot P_F[k, l] - d}{\Delta}, & \text{else,} \end{cases}$$

where $\Delta = 2^{(QP-4)/6}$. (3)

$[\cdot]$, Δ , d 는 각각 소수점 버림 연산, 양자화 크기(step-size), 그리고 오프셋을 나타낸다. 여기서 Δ

와 d 는 QP(Quantization Parameter) 값에 의해 정의되고 d 는 인트라 블록일 때는 $\Delta/3$, 인터 블록일 때는 $\Delta/6$ 의 값을 가진다. $V_q[k, l]$ 을 역 양자화한 계수를 $\hat{V}[k, l]$ 이라 할 때 역 DCT 후의 오차 블록 $\hat{\mathbf{U}}$ 은 식 (4)로 표현된다.

$$\hat{\mathbf{U}} = \frac{1}{64} \mathbf{B}_I^T \hat{\mathbf{Z}} \mathbf{B}_I, \quad \text{where } \hat{\mathbf{Z}} = \hat{\mathbf{V}} \otimes \mathbf{P}_I. \quad (4)$$

여기서 \mathbf{B}_I 와 \mathbf{P}_I 는 역 정수 DCT 행렬과 전크기 변환 행렬을 나타낸다.

Tu^[8]는 H.264의 울-왜곡 계산량을 감소시키기 위해서 엔트로피 부호화와 복원 과정을 생략하고 DCT 영역에서 울-왜곡을 예측하였다. Tu^[8]가 제안한 DCT 영역에서 k, l 번째 위치의 왜곡치 $D_{Tu}[k, l]$ 은 식 (5)와 같이 계산하였다.

$$D_{Tu}[k, l] = \left(Z[k, l] - \frac{1}{64} [\mathbf{P} \hat{\mathbf{Z}} \mathbf{P}]_{k, l} \right)^2 \times P_F[k, l]^2,$$

$$\mathbf{P} = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}. \quad (5)$$

여기서 $[\mathbf{A}]_{k, l}$ 은 \mathbf{A} 행렬의 k, l 번째 원소를 나타낸다. 그리고 4x4 블록에서 $V_q[k, l]$ 이 '0'이 아닌 개수를 N 이라고 할 때 비트율 R_{Tu} 는 다음 식과 같이 예측하였다.

$$R_{Tu} = \alpha \cdot N + \beta \cdot \sum_{k=0}^3 \sum_{l=0}^3 |V_q[k, l]|. \quad (6)$$

여기서 α , β 는 비트율 예측을 위한 상수 값들을 나타낸다. 기존 방식에서는 DCT 영역에서 왜곡치와 비트율을 예측함으로써 엔트로피 부호화, 역 DCT, 복원 과정에 대한 계산량을 생략시키고 실제 왜곡치와 비트율에 근사한 값을 얻을 수 있었다. 그러나 비트율과 왜곡치를 예측하는데 있어서 예측 오차 블록에 대한 DCT, 양자화, 역 양자화를 항상 필요로 한다. 그림 3은 Tu^[8]의 방식에서 제안된 비트율-왜곡치 예측 방식의 블록도를 나타낸다.

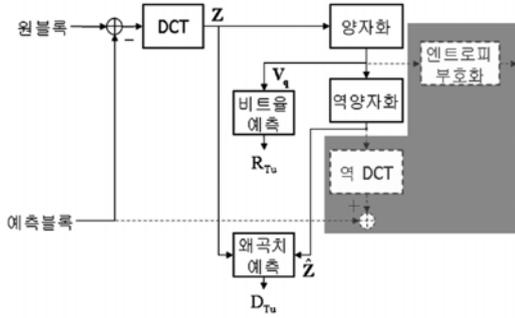


그림 3. Tu[8]가 제안한 비트율-왜곡치 예측 방식

그림 3에서 회색 영역은 Tu의 방식으로 인해 H.264의 비트율-왜곡치 계산에서 생략되어지는 과정을 나타낸 것이다.

IV. 제안하는 울-왜곡 예측 방식

본 논문에서는 H.264의 울-왜곡 계산에 요구되었던 엔트로피 부호화, 역 DCT, 복원 과정을 포함하여 Tu^[8]의 방식에서 남아있던 양자화와 역 양자화 과정을 생략시키기 위하여 정수 DCT 계수 블록 Z로부터 효과적으로 비트율과 왜곡치를 예측하는 방식을 제안한다. 그림 4는 제안 방식의 블록도를 나타낸다.

그림 4에서 회색 영역은 제안 방식으로 인해 H.264의 비트율-왜곡치 계산에서 생략되어지는 과정을 나타낸 것이다.

양자화와 역 양자화 과정 없이 왜곡치를 얻기 위해서 양자화 오차의 특성에 대해 알아볼 필요가 있다. 먼저 H.264의 양자화 방식인 식 (3)을 이용하면 정수 DCT 계수 $Z[k, l]$ 로부터 양자화된 DCT 계수 $V_q[k, l]$ 의 값이 '0'인지 아닌지를 인터 블록에 대해서 다음 식으로 판별할 수 있다.

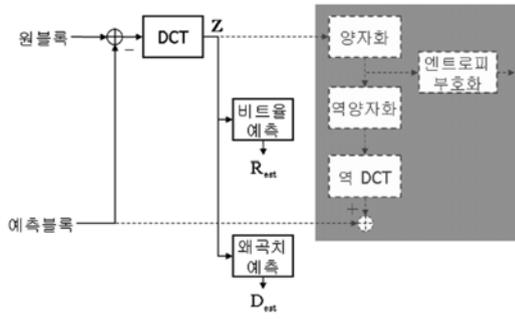


그림 4. 제안하는 비트율-왜곡치 예측 방식

$$nz[k, l] = \begin{cases} 0, & \text{when } |Z[k, l]| < \frac{5}{6} \cdot \frac{2^{(QP-4)/6}}{P_f[k, l]} = Th[QP, k, l] \\ 1, & \text{otherwise} \end{cases} \quad (7)$$

$Th[QP, k, l]$ 은 각 QP와 k, l 위치에 따른 임계값을 나타내고 이 값은 테이블에 저장하여 사용할 수 있다. 식 (7)의 결과인 $nz[k, l]$ 으로 각 계수들에 대한 양자화 오차 $\varepsilon[k, l]$ 을 다음 식과 같이 두 경우로 나누어 표현한다.

$$\varepsilon[k, l] = \begin{cases} V[k, l] - Z[k, l] \cdot P_f[k, l], & \text{when } nz[k, l] = 0 \\ V[k, l] - \hat{V}[k, l] = \varepsilon_{nc}[k, l], & \text{otherwise} \end{cases} \quad (8)$$

$nz[k, l]$ 이 '0'인 경우의 왜곡치는 위의 식과 같이 DCT 계수 값으로 쉽게 얻을 수 있기 때문에 $nz[k, l]$ 이 '0'이 아닌 경우의 양자화 오차를 $\varepsilon_{nc}[k, l]$ 로 정의하고 이 경우에만 양자화 오차를 추정한다.

4.1 양자화 오차의 확률 분포에 대한 해석

양자화 오차에 대한 분포를 알기 위해서는 양자화기의 입력 분포에 대한 정의가 필요하다. 이전의 많은 연구들에서 예측 오차 블록의 DCT 계수는 영-평균 라플라시안 분포(Zero-mean Laplacian Distribution)로 모델링 되었다^{9,10)}. 따라서 본 논문에서는 양자화기의 입력인 DCT 계수 $V[k, l]$ 을 다음 식과 같이 영-평균 라플라시안 분포를 가지는 임의 변수(Random Variable) V 로 가정하였다.

$$f_V(v) = \frac{\tau}{2} e^{-\tau|v|}, \quad -\infty < v < \infty, \quad \tau > 0. \quad (9)$$

여기서 τ 는 라플라시안 계수이다. 그림 5는 식 (3)을 근거로 하여 $nz[k, l]$ 이 '0'이 아닌 경우, 즉 DCT 계수가 '0'으로 양자화 되지 않는 경우에 대해서 식 (9)로 가정한 입력과 양자화 후의 출력, 그리고 양자화 오차의 분포를 도식화 한 것이다.

일반적으로 오프셋이 $\Delta/2$ 인 양자화기의 양자화 오차 ε 는 모두 $-\Delta/2$ 에서 $\Delta/2$ 구간 사이에 존재하고 양자화 오차의 PDF는 다음 식과 같이 모델링 된다^{11,12)}.

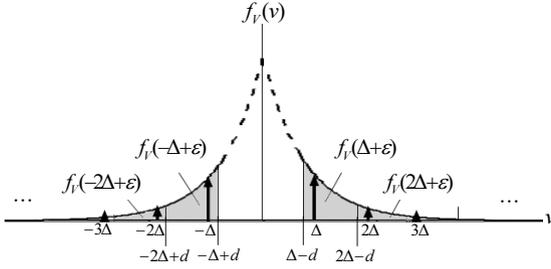


그림 5. '0'이 아닌 수로 양자화 되는 경우의 양자화기 입력, 출력, 양자화 오차의 분포 ; 실선은 라플라시안 분포로 가정한 입력 DCT 계수의 확률을 나타내고 화살표는 출력 양자화 계수의 확률, 그리고 회색 영역은 양자화 오차의 확률을 나타낸다.

$$f_{\varepsilon}(\varepsilon) = \sum_{n=-\infty}^{\infty} f_v(n\Delta + \varepsilon), \quad -\Delta/2 \leq \varepsilon < \Delta/2 \quad (10)$$

그림 5에서 '0'보다 작은 수로 양자화 되는 경우의 양자화 오차는 $-\Delta+d$ 에서 d 구간 사이에 존재하고 '0'보다 큰 수로 양자화 되는 경우의 양자화 오차는 $-d$ 에서 $\Delta-d$ 구간 사이에 존재하는 것을 알 수 있다. 이전의 방식들^{[11],[12]}과 같이 양자화 오차의 PDF는 양자화 오차 구간 안의 확률들의 합으로 모델링 할 수 있다. 양자화된 계수 값이 '0'보다 작을 경우와 클 경우에 대해서 정규화 이전의 양자화 오차 분포를 나타내면 식 (11)과 같다.

$$\tilde{f}_{\varepsilon_{nz}}(\varepsilon) = \begin{cases} \sum_{n=-1}^{-\infty} f_v(n\Delta + \varepsilon), & \text{for } -\Delta+d < \varepsilon < d \\ \sum_{n=1}^{\infty} f_v(n\Delta + \varepsilon), & \text{for } -d \leq \varepsilon < \Delta-d. \end{cases} \quad (11)$$

식 (12)는 위의 식을 양자화 오차 ε 의 각 구간 별로 표현한 것이다.

$$\tilde{f}_{\varepsilon_{nz}}(\varepsilon) = \begin{cases} \sum_{n=-1}^{-\infty} f_v(n\Delta + \varepsilon), & \text{for } -\Delta+d < \varepsilon < -d \\ \sum_{n=-1}^{-\infty} f_v(n\Delta + \varepsilon) + \sum_{n=1}^{\infty} f_v(n\Delta + \varepsilon), & \text{for } -d \leq \varepsilon \leq d \\ \sum_{n=1}^{\infty} f_v(n\Delta + \varepsilon), & \text{for } d < \varepsilon < \Delta-d. \end{cases} \quad (12)$$

식 (12)에 식 (9)의 라플라시안 확률 밀도 함수 (PDF : Probability Density Function)를 대입하고 정규화 하면 그림 5의 회색 영역에 대한 양자화 오

차의 PDF는 식 (13)과 같이 얻어진다.

$$f_{\varepsilon_{nz}}(\varepsilon) = \begin{cases} \frac{\tau}{2} e^{\tau\varepsilon} \frac{1}{(e^{\tau d} - e^{\tau(-\Delta+d)})}, & \text{for } -\Delta+d < \varepsilon < -d, \\ \frac{\tau}{2} (e^{\tau\varepsilon} + e^{-\tau\varepsilon}) \frac{1}{(e^{\tau d} - e^{\tau(-\Delta+d)})}, & \text{for } -d \leq \varepsilon \leq d, \\ \frac{\tau}{2} e^{-\tau\varepsilon} \frac{1}{(e^{\tau d} - e^{\tau(-\Delta+d)})}, & \text{for } d < \varepsilon < \Delta-d. \end{cases} \quad (13)$$

식 (13)의 PDF로부터 '0'으로 양자화 되지 않는 경우의 양자화 오차 ε_{nz} 의 평균과 분산을 다음 식과 같이 구할 수 있다.

$$E[\varepsilon_{nz}] = 0, \\ E[\varepsilon_{nz}^2] = \frac{\left\{ (\tau d - 1)^2 + 1 \right\} - e^{-\tau\Delta} \left\{ (\tau(-\Delta+d) - 1)^2 + 1 \right\}}{\tau^2 (1 - e^{-\tau\Delta})}. \quad (14)$$

4.2 왜곡치와 비트율 예측

본 논문에서는 식 (14)의 분산 값을 '0'으로 양자화 되지 않는 DCT 계수에 대한 왜곡치로 근사화한다. 식 (14)의 분산 값에서 지수 항은 τ 또는 Δ 의 값이 커질수록 무시할 수 있을 정도로 작아지므로 지수 항을 없애고, d 를 식 (3)을 이용하여 QP에 대해서 정리하면 왜곡치는 인터 블록에서 다음 식과 같이 QP와 τ 의 함수로 표현된다.

$$E[\varepsilon_{nz}^2] \approx \frac{(\tau \cdot (2^{(QP-4)/6}) / 6 - 1)^2 + 1}{\tau^2} = D_{nz}[\text{QP}, \tau]. \quad (15)$$

이 값을 미리 계산하여 각 QP와 τ 에 대해서 테이블에 저장하여 사용함으로써 계산량을 줄인다. 제안하는 DCT 영역에서의 왜곡치 예측 방식을 식 (7,8,15)을 이용하여 정리하면 다음 식과 같다.

$$D_{est} = \sum_{k=0}^3 \sum_{l=0}^3 \varepsilon^2[k, l],$$

$$\varepsilon^2[k, l] = \begin{cases} Z^2[k, l] \cdot P_F^2[k, l], & \text{when } nz[k, l] = 0, \\ D_{nz}[\text{QP}, \tau], & \text{otherwise} \end{cases} \quad (16)$$

$nz[k,l]$ 값은 식 (7)에서 k,l 위치의 정수 DCT 계수 $Z[k,l]$ 을 임계값 $Th[QP,k,l]$ 과 비교하여 구하였다. 비트율은 식 (17)과 같이 블록 내의 '0'으로 양자화 되지 않는 계수들의 개수로부터 예측한다.

$$R_{est} = \begin{cases} 1, & \text{when } N = 0 \\ a \cdot N + b, & \text{otherwise,} \end{cases}$$

$$N = \sum_{k=0}^3 \sum_{l=0}^3 nz[k,l]. \tag{17}$$

a 와 b 는 상수로서 M 개의 표본 블록에 대해서 다음 식과 같은 일반적인 선형 회기 기법으로 구할 수 있다^[13].

$$a = \frac{M \cdot \sum_{i=0}^M N_i R_i - \sum_{i=0}^M N_i \sum_{i=0}^M R_i}{M \cdot \sum_{i=0}^M N_i^2 - \left(\sum_{i=0}^M N_i\right)^2}, b = \frac{\sum_{i=0}^M R_i - a \sum_{i=0}^M N_i}{M}. \tag{18}$$

N_i 와 R_i 는 각각 i 번째 표본 블록에서의 '0'이 아닌 계수들의 개수와 실제 비트율을 나타낸다. 본 논문에서 여러 테스트 동영상들을 표본으로 하여 얻은 a 와 b 값은 표 1과 같다.

제안 방식으로 비트율과 왜곡치를 예측함으로써 H.264의 RDO 모드 결정 방식에서 요구되었던 양자화, 엔트로피 부호화, 역 양자화, 역 DCT, 그리고 복원 과정의 생략이 가능하다.

표 1. 비트율 예측 상수 a, b

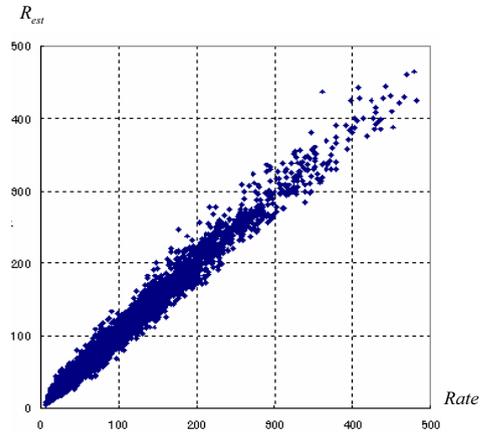
QP	a	b
20	4.7564	2.9142
24	4.8288	2.435
28	4.8211	1.6336
32	4.932	0.5913
36	4.878	0.1953
40	4.5875	0.271

V. 실험 결과 및 고찰

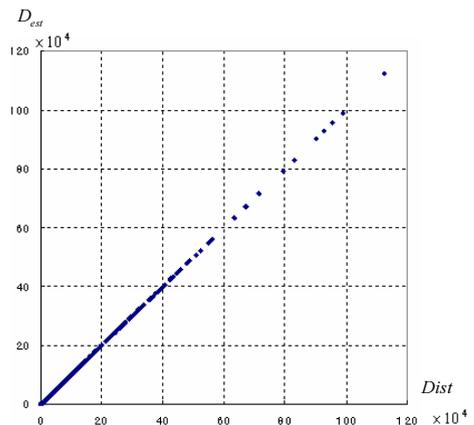
제안 방식의 성능을 알아보기 위하여 H.264의 참조 프로그램 JM9.6^[14]에서 모의실험 하였다. 동영상은 QCIF 형식의 'silent', 'foreman', 'container',

'news', 'grandma'와 CIF 형식의 'mobile', 'akiyo', 'news', 'coastguard', 'salesman'이고 10Hz, 15Hz, 30Hz로 각각 100 프레임을 실험하였다. 사용된 QP 값은 28, 32, 36, 40이고 IPPP 형식으로 P 프레임에서 인트라 모드는 사용하지 않았다. 움직임 추정 및 보상에 사용된 참조 프레임 수는 1 프레임, 탐색 범위는 16으로 설정하였고 고속 움직임 예측 방식^[14]을 사용하였다. 제안 알고리즘은 인터 모드에만 적용되었고, 왜곡치 계산에 필요한 τ 는 계산량 절감을 위해서 미리 여러 동영상에 대한 실험을 통하여 10으로 고정하였다.

그림 6은 실험에서 얻은 실제 비트율, 왜곡치와 제안 방식으로 계산한 비트율, 왜곡치를 비교한 그림이다.



(a) 실제 비트율(Rate)과 제안 방식으로 예측한 비트율(R_{est})

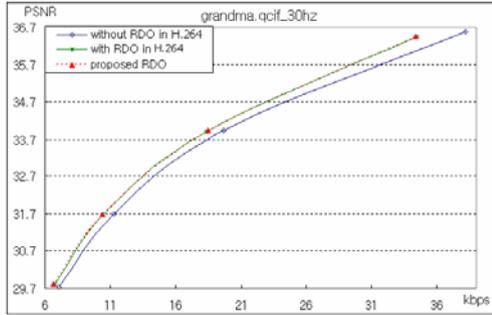


(b) 실제 왜곡치(Dist)와 제안 방식으로 예측한 왜곡치(D_{est})

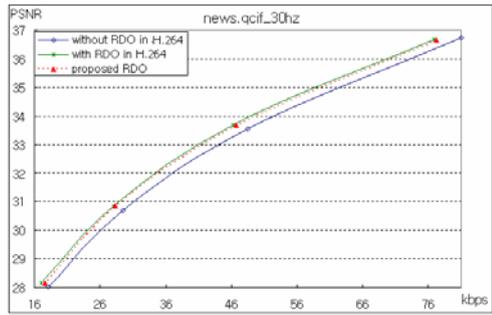
그림 6. Foreman QCIF 동영상, QP값 28일 때 얻은 실제 비트율, 왜곡치와 제안 방식으로 예측한 비트율, 왜곡치

위의 그림에서 제안 방식으로 예측된 비트율과 왜곡치가 실제 비트율, 왜곡치와 거의 일치하는 것을 알 수 있다.

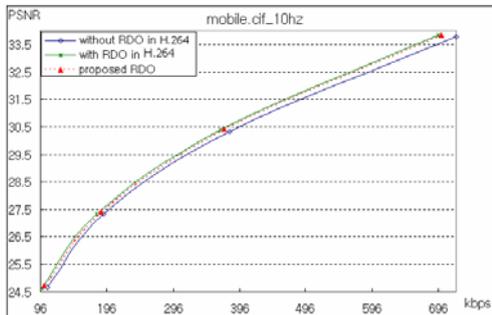
그림 7은 H.264의 RDO 모드 결정 방식을 사용하였을 경우와 RDO 모드 결정 방식을 사용하지 않았을 경우, 그리고 제안하는 윌-왜곡 예측 방식을 이용하여 모드를 결정한 경우의 R-D 곡선(Rate-Distortion Curve) 그래프를 나타낸 것이다.



(a) granma 동영상



(b) news 동영상



(c) mobile 동영상

그림 7. H.264의 RDO 모드 결정 방식을 사용한 경우(with RDO in H.264)와 사용하지 않은 경우(without RDO in H.264), 그리고 제안하는 윌-왜곡 예측 방식을 사용하여 모드를 결정한 경우(proposed RDO)의 R-D 곡선 그래프

표 2. H.264의 RDO 모드 결정 방식에 대한 Tu의 방식(Tu's)과 제안 방식(Proposed)의 평균 PSNR 감소량

QCIF		PSNR(dB)		CIF		PSNR(dB)	
		Tu's	Proposed			Tu's	Proposed
foreman	10hz	-0.055	-0.059	mobile	10hz	-0.057	-0.067
	30hz	-0.014	-0.012		30hz	-0.041	-0.052
news	10hz	-0.078	-0.076	akiyo	10hz	0.015	0.002
	30hz	-0.084	-0.034		30hz	0.051	0.057
container	10hz	-0.071	-0.068	news	10hz	-0.073	-0.098
	30hz	-0.096	-0.085		30hz	-0.093	-0.081
silent	15hz	-0.072	-0.051	coastguard	10hz	-0.044	-0.065
	30hz	-0.052	-0.045		30hz	-0.062	-0.066
grandma	10hz	0.007	0.029	salesman	10hz	-0.054	-0.057
	30hz	0.001	0.017		30hz	-0.061	-0.052
Average		-0.051	-0.038	Average		-0.042	-0.048

그림 7에서 제안하는 윌-왜곡 예측 방식을 사용하여 모드를 결정하였을 때 H.264의 RDO 모드 결정 방식과 R-D 곡선 관점에서 성능의 저하가 거의 없음을 알 수 있다. 표 2에는 H.264의 RDO 모드 결정 방식을 사용한 경우에 대해서 Tu의 방식과 제안 방식의 평균 PSNR의 차이를 나타내었다. 표 2로부터 H.264의 RDO 모드 결정 방식에 대한 QCIF와 CIF 동영상의 평균 PSNR 감소는 Tu의 방식은 0.047dB, 제안 방식은 0.043dB로 Tu의 방식과 성능이 거의 동일한 것을 알 수 있다.

다음은 H.264의 RDO 모드 결정 방식에 대한 계산량 감소 비율을 알아보기 위해서 사용한 식을 나타낸다.

$$\text{Saving}_{\text{RDO}}(\%) = \frac{\text{Time}_{\text{with RDO in H.264}} - \text{Time}_{\text{proposed RDO}}}{\text{Time}_{\text{with RDO in H.264}} - \text{Time}_{\text{without RDO in H.264}}} \times 100. \quad (19)$$

위의 식으로부터 얻은 H.264의 RDO 모드 결정 방식에 대한 Tu의 방식과 제안 방식의 계산량 감소 비율을 표 3에 나타내었다.

그림 8은 표 3을 기반으로 하여 H.264 RDO 모드 결정 방식에 대한 Tu와 제안 방식의 평균 계산량 감소 비율을 나타낸 것이다.

표 3. H.264의 RDO 모드 결정 방식에 대한 Tu의 방식 (Tu's)과 제안 방식(Proposed)의 계산량 감소 비율(%)

Saving _{RDO} (%)						
QCIF	Hz	Method	QP 28	QP 32	QP 36	QP 40
foreman	10hz	Tu's	52.51	64.35	68.24	64.79
		Proposed	72.04	80.49	83.29	81.85
foreman	30hz	Tu's	54.76	56.36	65.82	64.54
		Proposed	71.24	81.59	78.24	86.22
news	10hz	Tu's	69.41	63.80	63.34	71.34
		Proposed	79.51	82.95	84.12	84.28
news	30hz	Tu's	69.77	62.10	59.95	57.29
		Proposed	84.74	80.59	83.15	79.30
container	10hz	Tu's	58.11	59.43	62.04	62.13
		Proposed	76.27	80.25	85.32	83.89
container	30hz	Tu's	67.17	67.40	67.31	66.60
		Proposed	86.86	85.83	78.28	79.18
silent	15hz	Tu's	66.45	65.50	65.99	67.86
		Proposed	81.59	83.25	87.28	86.00
silent	30hz	Tu's	69.51	65.80	68.36	57.39
		Proposed	79.74	82.95	87.74	81.83
grandma	10hz	Tu's	41.79	59.20	70.81	65.99
		Proposed	85.56	79.25	80.07	78.80
grandma	30hz	Tu's	58.92	57.97	66.27	66.63
		Proposed	79.39	84.89	80.16	85.21

(a) QCIF 동영상

Saving _{RDO} (%)						
CIF	Hz	Method	QP 28	QP 32	QP 36	QP 40
mobile	10hz	Tu's	63.27	63.33	56.87	67.25
		Proposed	77.08	80.11	76.70	84.92
mobile	30hz	Tu's	57.51	61.13	56.81	53.67
		Proposed	75.83	74.76	81.71	82.33
akiyo	10hz	Tu's	62.71	64.35	60.74	61.47
		Proposed	80.74	78.49	79.86	80.15
akiyo	30hz	Tu's	59.81	60.60	64.99	57.98
		Proposed	85.86	83.27	79.37	81.93
news	10hz	Tu's	68.20	59.43	67.49	58.25
		Proposed	80.00	78.49	80.65	82.72
news	30hz	Tu's	61.34	67.55	59.80	59.83
		Proposed	83.87	84.67	84.76	79.66
coastguard	15hz	Tu's	60.60	61.91	53.72	52.18
		Proposed	75.52	75.24	75.85	80.56
coastguard	30hz	Tu's	56.07	52.02	59.82	61.30
		Proposed	75.40	79.82	73.45	75.66
salesman	10hz	Tu's	64.18	65.45	61.78	65.52
		Proposed	87.64	79.79	85.74	81.83
salesman	30hz	Tu's	69.98	64.15	65.31	86.05
		Proposed	89.44	81.87	83.95	83.02

(b) CIF 동영상

위의 그림은 Tu의 방식은 평균 62.6%, 제안 방식은 평균 81.2%의 RDO 계산량을 감소시킴을 보여준다.

전체 부호화 계산량에 대한 Tu의 방식과 제안 방식의 계산량 감소 비율은 다음 식으로 계산되었고 이는 표 4에 나타내었다.

$$\text{Saving}_{\text{Total encoding}} (\%) = \frac{\text{Time}_{\text{with RDO in H.264}} - \text{Time}_{\text{proposed RDO}}}{\text{Time}_{\text{with RDO in H.264}}} \times 100. \quad (20)$$

표 4. Tu의 방식(Tu's) 제안 방식(Proposed)을 H.264의 RDO 모드 결정 방식에 적용하였을 경우 전체 부호화 시간 감소 비율(%)

		Saving Total encoding (%)				
QCIF	Hz	Method	QP 28	QP 32	QP 36	QP 40
foreman	10hz	Tu's	14.41	18.15	19.04	17.29
		Proposed	19.77	22.70	23.24	21.84
foreman	30hz	Tu's	16.84	16.54	19.01	18.06
		Proposed	21.90	23.94	22.60	24.13
news	10hz	Tu's	22.65	18.45	19.32	21.71
		Proposed	25.95	23.99	25.65	25.65
news	30hz	Tu's	25.39	20.01	19.36	17.68
		Proposed	30.84	25.96	26.85	24.47
container	10hz	Tu's	19.19	18.90	18.95	19.12
		Proposed	25.18	25.53	26.06	25.82
container	30hz	Tu's	22.60	22.36	22.42	21.69
		Proposed	29.22	28.47	26.08	25.79
silent	15hz	Tu's	24.15	22.85	22.46	21.75
		Proposed	29.65	29.04	29.71	27.56
silent	30hz	Tu's	23.65	19.97	20.95	18.00
		Proposed	27.13	25.17	26.89	25.67
grandma	10hz	Tu's	13.18	18.12	23.55	21.94
		Proposed	26.98	24.26	26.63	26.20
grandma	30hz	Tu's	18.77	18.07	21.98	21.60
		Proposed	25.29	26.46	26.59	27.63

(a) QCIF 동영상

		Saving Total encoding (%)				
CIF	Hz	Method	QP 28	QP 32	QP 36	QP 40
mobile	10hz	Tu's	20.49	19.51	16.06	18.92
		Proposed	24.97	24.68	21.66	23.89
mobile	30hz	Tu's	19.33	18.15	16.32	14.29
		Proposed	25.49	22.19	23.47	21.92
akiyo	10hz	Tu's	21.62	21.72	22.04	22.47
		Proposed	27.84	26.50	28.98	29.29
akiyo	30hz	Tu's	20.74	21.40	22.47	20.52
		Proposed	29.78	29.41	27.44	29.00
news	10hz	Tu's	22.10	19.18	23.15	19.49
		Proposed	25.92	25.33	27.66	27.68
news	30hz	Tu's	20.41	22.99	20.13	19.72
		Proposed	27.91	28.82	28.54	26.26
coastguard	15hz	Tu's	17.82	16.61	13.92	13.36
		Proposed	22.21	20.19	19.65	20.63
coastguard	30hz	Tu's	17.12	14.24	15.65	15.97
		Proposed	23.02	21.85	19.22	19.72
salesman	10hz	Tu's	19.22	20.39	18.74	19.49
		Proposed	26.25	24.85	26.01	24.34
salesman	30hz	Tu's	27.25	19.99	21.40	21.37
		Proposed	35.22	25.51	27.51	27.07

(b) CIF 동영상

H.264의 RDO 모드 결정 방식을 사용할 경우의 전체 부호화 시간의 평균 20%와 26%가 각각 Tu의 방식과 제안 방식을 사용하여 감소되었다.

일반적으로 H.264의 RDO 모드 결정 방식을 사용함으로써 RDO 모드 결정 방식을 사용하지 않았을 경우보다 평균 7%의 비트율이 감소하지만 전체

부호화 계산량은 평균 50%가 증가한다²⁴⁾. 모의실험 결과로부터 제안 방식은 H.264에서 RDO 모드 결정 방식을 사용하지 않는 경우의 전체 부호화 계산량의 약 10%의 추가 계산량을 가지면서 RDO 모드 결정 방식을 사용할 때와 거의 비슷한 성능을 얻을 수 있음을 알 수 있다.

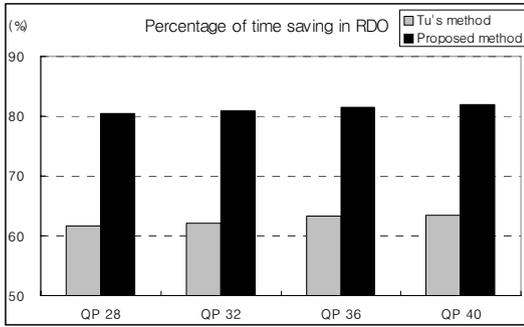


그림 8. Tu(Tu's method)와 제안 방식(Proposed method)의 H.264 RDO 모드 결정 방식에 대한 평균 계산량 감소 비율(%) 그래프

VI. 결 론

지금까지 고속 RDO 모드 결정을 위해서 수행된 많은 연구⁵⁻⁷⁾는 주로 모드 생략에 관한 것이었다. 이러한 방식들에서 제거되지 않은 모드에 대해서는 여전히 RDO 모드 결정을 위한 윌-왜곡 계산량을 필요로 한다. 따라서 고속 부호화를 위하여 효율적인 윌-왜곡 계산에 대한 연구를 할 필요가 있다. 윌-왜곡 계산량을 감소시키기 위해서 Tu^[8]는 DCT 영역에서 비트율과 왜곡치를 예측함으로써 엔트로피 부호화, 역 DCT, 그리고 복원 과정에 요구되는 계산량을 감소시켰지만 양자화와 역 양자화 과정이 여전히 남아 있게 되었다.

본 논문에서는 양자화와 역 양자화 과정 없이 왜곡치와 비트율을 예측하는 방식을 제안하였다. DCT 계수를 라플라시안 분포로 가정하고 양자화 오차를 추정하여 왜곡치를 예측하고, '0'으로 양자화 되지 않는 계수들의 개수로 비트율을 예측하였다. 제안 방식으로 H.264의 RDO 모드 결정 과정에 요구되었던 양자화, 엔트로피 부호화, 역 양자화, 역 DCT, 그리고 복원 과정이 생략되었다.

일반적으로 H.264의 RDO 모드 결정 방식을 사용함으로써 RDO 모드 결정 방식을 사용하지 않았을 경우보다 평균 7%의 비트율이 감소하지만 전체 부호화 계산량은 평균 50%가 증가한다^{2,4)}. 모의실험 결과 제안 방식을 사용함으로써 성능 저하는 거의 없이 H.264의 RDO 모드 결정 계산량의 약 81%가 감소되었다. 이는 제안 방식으로 전체 부호화 계산량의 약 10%의 계산량만 추가하면서도 RDO 모드 결정 방식을 사용할 때와 거의 비슷한 성능을 얻을 수 있음을 의미한다. 따라서 많은 성능 개선을 가져오면서도 불구하고 구현하기 어려웠던 RDO 모드 결

정 방식을 알고리즘을 개선하여 실제 하드웨어 구현에 사용할 수 있는 수준으로 감소시킨 것에 큰 의의가 있다.

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. H.264 ISO/IEC 14496-10 AVC)", *Joint Video Team(JVT) of ISO/IEC MPEG and ITU-T VCEG. doc. JVT-G050*. Mar. 2003.
- [2] 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.
- [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] Lain E.G. Richardson, *H.264 and MPEG-4 Video Compression*, John Wiley & Sons Inc. 2003
- [5] K. P. Lim, S. Wu, D. J. Wu, S. Rahardja, X. Lin, F. Pan, Z. G. Li "Fast Inter Mode Selection", *Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Document JVT-I020*, San Diego, US, 2-5 Sept. 2003
- [6] Y. H. Kim, J. W. Yoo, S. W. Lee, J. Shin, J. Paik, and H. K. Jung, "Adaptive mode decision for H.264 encoder," *IEE Electronics Letters*, vol. 40, no. 19, Sep. 2004.
- [7] X. Jing and L. P. Chau, "Fast approach for H.264 inter mode decision", *IEE Electronics Letters*, vol. 40, no. 17, pp.1050-1052, Aug. 2004.
- [8] Y. K. Tu, J. F. Yang, and M. T. Sun, "Efficient Rate-Distortion Estimation for H.264/AVC Coders", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 5, pp. 600-611, May 2006.

- [9] 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
- [10] M. Robertson and R. Stevenson, "DCT Quantization noise in compressed images", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 1, pp. 27-38, Jan. 2005
- [11] A. B. Sripad and D. L. Snyder, "A Necessary and Sufficient Condition for Quantization Errors to be Uniform and White", *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-25, no. 5, pp. 442-448, Oct. 1977
- [12] B. Widrow, I. Koll?r and M. C. Liu, "Statistical Theory of Quantization", *IEEE Trans. Instrumentation and Measurement*, vol. 45, no. 2, pp. 353-361, April 1996
- [13] D. F. Mix, *Random signal processing*, Prentice Hall International, Inc. pp 285-289.
- [14] Joint Video Team (JVT) Reference Software [Online]. Available: <http://bs.hhi.de/~suehring/tml/download/>

문 정 미 (Jeong-mee Moon)

정회원



2001년 2월 홍익대학교 전기 전자공학부 졸업
현재 부산대학교 전자공학과 석박사통합과정

<관심분야> 영상처리, 영상압축, VLSI, SOC 설계

김 재 호 (Jae-ho Kim)

정회원



1980년 2월 부산대학교 전기 기계공학과 졸업
1982년 2월 한국과학기술원 산업전자공학과 석사
1990년 2월 한국과학기술원 전기 및전자공학과 박사
현재 부산대학교 전자공학과정교수

<관심분야> 영상처리, 영상압축, VLSI, SOC 설계