

H.264/AVC 부호화기에 대한 효과적인 모드 결정 알고리즘

정희원 문정미*, 김재호*, 문용호**

An Effective Mode Decision Algorithm in H.264/AVC Encoder

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

요 약

본 논문에서는 H.264/AVC의 부호화기에서 RDO 모드 결정을 위한 효율적인 방식을 제안한다. DCT 계수와 RDO 모드 결정 과정의 특징에 기반하여 모든 DCT 계수들이 양자화 후에 '0'이 되는 오차 블록(AZCB)을 검출하는 새로운 조건이 유도된다. 제안 알고리즘에서는 AZCB에 대한 (I)DCT, (역)양자화, 엔트로피 부호화 과정의 생략이 이루어진다. 이것은 RDO 모드 결정 과정에 요구되는 계산량을 감소시킨다. 모의 실험 결과는 기존 방식에 비하여 약 40% 이상의 계산량 감소가 제안 알고리즘에서 이루어짐을 보여준다.

Key Words : (I)DCT, (I)Q, H.264/AVC, SSD, RDO mode decision

ABSTRACT

In this paper, we propose an efficient algorithm for the RDO mode decision in H.264/AVC encoder. Based on the properties of DCT coefficients and the RDO mode decision processing, we derive a new condition for detecting an error block having all-zero DCT coefficient (AZCB). (I)DCT, (I)Q, and entropy coding are skipped for AZCBs in the proposed algorithm. It makes the reduction of the computational complexity for the RDO mode decision. Simulation results show that the proposed algorithm achieves computational saving over 40% compared to the conventional method.

1. 서 론

MPEG-1/2/4, H.261/H.263 등의 동영상 압축 표준안은 지금까지 동영상 압축 기술 및 관련 응용 분야들의 고속 발전에 토대가 되었다. 최근에 ITU-T Video Coding Experts Group과 ISO/IEC Moving Picture Experts Group은 새로운 압축 표준으로 H.264/AVC^[1]를 공동으로 제정하였다. H.264/AVC는 이전의 동영상 압축 표준에 비해 50%까지 향상된 성능을 보여준다^[2]. 그러나 H.264/AVC는 MPEG-4 SP 보다 10배 많은 부호화 계산량을 필요로 하기

때문에 실시간 응용에 적용될 수 없는 단점을 가진다^[2]. 따라서 오늘날 H.264/AVC의 부호화 계산량을 감소시키기 위한 연구가 활발히 진행되고 있다.

H.264/AVC의 부호화에 있어서 전체 계산량의 약 80%정도가 움직임 추정과 윌-웨곡 최적화(Rate-Distortion Optimization(RDO))에 의한 모드 결정 과정에서 소모된다고 알려져 있다^[3]. 따라서 이 두 과정에서의 계산량 감소방식을 개발하는 것이 무엇보다 중요하다. 그런데 움직임 추정 과정의 계산량은 지금까지 발표된 여러 압축 표준들에서도 심각한 문제로 인식되어 개선 방식이 폭넓게 연구되어

* 부산대학교 컴퓨터 및 정보통신연구소 ({moonjme, jhkim}@pusan.ac.kr),

** 부산외국어대학교 디지털미디어학부 (yhmoon5@pufs.ac.kr)

논문번호 : KICS2005-07-312, 접수일자 : 2005년 7월 30일, 최종논문접수일자 : 2006년 2월 24일

왔다. 현재 움직임 추정 과정에 대한 다양한 계산량 감소 방식들^[4-7]이 제안되어 있다. H.264/AVC의 경우, 고속 움직임 추정 방식^[6]이 채택되었으며 더욱 개선된 성능을 지닌 Yang의 방식^[7]이 제시되었다.

기존 움직임 추정은 주어진 왜곡 척도 하에서 탐색 과정을 통하여 각 매크로 블록에 대한 정합 블록을 결정하였다. 반면에 H.264/AVC에서 채택된 RDO에 의한 모드 결정 방식에서는 왜곡 정도와 비트 유효 측면 모두에서 가장 적합한 정합 블록의 위치 및 형태(모드)를 결정한다. 이를 위하여 H.264/AVC에서는 9가지 모드에 대하여 각각 탐색 과정을 거친 후 각 모드에서의 왜곡 정도 및 비트 유효를 구한다. 그리고 이들 중 최소의 왜곡 정도 및 비트 유효를 지닌 모드를 구한 후 이때의 정합 블록을 최적의 예측 블록으로 결정한다. 이러한 모드 결정 방식은 우수한 압축 성능을 가져온다. 그러나 왜곡 정도를 알기 위해서 항상 (I)DCT 및 (역)양자화가 수행되어야 하고 비트 유효를 얻기 위해서는 엔트로피 부호화가 9가지 모드의 전체 탐색 과정에서 수행되어야만 한다. 그런데 이것은 결코 무시할 수 없는 매우 큰 계산량을 필요로 한다.

H.264/AVC의 부호화기에서 고속으로 모드 결정을 수행하기 위한 다양한 방식들이 연구되어 왔다. 그중에서 대표적으로 경계(edge)의 크기로부터 얻어지는 블록의 균질성(homogeneity)을 이용한 고속 모드 결정 방식^[6]과 DCT 계수를 이용한 적응적 모드 결정 방식^[8]이 있다. 이 방식들은 모드 결정 과정 중에 얻어진 정보를 이용하여 불필요한 모드의 존재 여부를 판단한다. 그리고 불필요한 모드에 대하여 움직임 추정 과정 및 모드 결정 과정을 생략함으로써 부호화 계산량을 감소시킨다. 이와 같은 기법은 전체 부호화 계산량을 약 70% 정도 감소시키는 성과를 낳았다. 그러나 이 같은 계산량 감소는 왜곡 정도와 비트 유효 계산량의 비중을 점차 증가시키는 결과를 가져왔다. 현재 기존 고속 모드 결정 방식에서 왜곡 정도와 비트 유효에 대한 계산은 여전히 수행되어지고 있다. 따라서 왜곡 정도 및 비트 유효 계산량 감소에 관한 연구는 H.264/AVC의 고속 부호화에 있어서 매우 중요하다.

본 논문에서는 (I)DCT 및 (역)양자화, 그리고 DCT 계수 부호화 생략을 통한 효율적인 모드 결정 방식을 제안한다. Pao는 H.263 동영상 압축 표준 환경 하에서 DCT 및 양자화 과정을 생략하는 방식^[9]을 제시하였다. 기존의 방식은 통계적 분석을 토대로 움직임 추정 과정에서 얻어진 왜곡 척도 값을

이용하여 오차 블록의 분산을 추정 한 후 DCT 계수들이 '0'으로 양자화 될 조건을 제시하였다. 그런데 이전의 압축 표준과는 달리 H.264/AVC 부호화기에서는 4×4 블록 단위로 DCT가 수행되며 RDO 기법에 기초한 새로운 방식에 의하여 움직임 추정 정보가 얻어진다. 따라서 기존 방식을 H.264/AVC에 직접 적용하는 것은 불가능하다. 이에 본 논문은 H.264/AVC의 RDO에 의한 모드 결정 과정의 특성을 이용하여 오차 블록의 분산을 새롭게 추정한다. 그리고 양자화된 DCT 계수가 모두 '0'인 오차 블록을 효과적으로 검출할 수 있는 조건을 제안한다. 모의 실험 결과는 제안 방식으로 RDO에 의한 모드 결정 과정의 계산량이 효과적으로 감소됨을 보인다.

II. RDO 기법에 의한 H.264/AVC에서의 모드 결정

기존의 동영상 압축 표준에서는 시간적 중복성을 제거하기 위하여 움직임 추정을 이용하였다. 기존 움직임 추정에서는 16x16 또는 8x8 크기의 원 블록 각각에 대하여 참조 프레임의 일정 영역 내에 존재하는 후보 블록들을 탐색한다. 이때 왜곡 정도를 나타내는 척도로서 SAD(Sum of Absolute Differences)가 일반적으로 사용되었으며 SAD를 최소화 하는 후보 블록이 정합 블록으로 결정되었다. 식(1)은 M×N 크기의 원 블록 S에 대한 SAD 정의 식이다.

$$SAD(mv) = \sum_{x=1}^M \sum_{y=1}^N |S[x, y] - C[x, y]|$$

$$= \sum_{x=1}^M \sum_{y=1}^N |\varepsilon[x, y]| \quad (1)$$

여기서 c와 ε는 후보 블록과 오차 블록을 각각 나타낸다.

H.264/AVC에서는 매크로 블록 모드라 불리는 다양한 형태의 매크로 블록들에 대하여 움직임 추정 개념을 적용한다. 그림 1은 H.264/AVC에서 INTER 매크로 블록에 대한 7가지 모드를 나타낸 것이다.

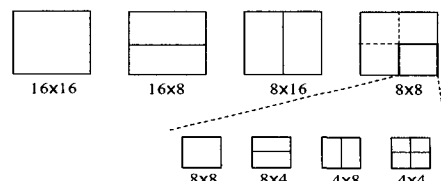


그림 1. H.264/AVC의 매크로 블록 모드

다양한 매크로 블록 유형들에 대한 시간적 중복성의 고려는 H.264/AVC의 압축 성능 향상에 큰 역할을 한다.

최적의 블록 유형 및 정합 블록을 결정하기 위하여 H.264/AVC에서는 왜곡 정도와 비트 율을 동시에 고려할 수 있는 RDO 기법을 이용한다. 먼저 각각의 모드에 대하여 움직임 추정이 수행된다. H.264/AVC에서 움직임 추정은 여러 참조 프레임에 대해서 수행되고 식 (2)가 최소가 되는 움직임 벡터와 참조 프레임을 결정한다.

$$J_{motion}(mv) = SA(T)D(mv) + \lambda_{motion} \cdot R(mv) \quad (2)$$

여기서 λ_{motion} 는 라그랑지 계수를 나타낸다. 그리고 $R(mv)$ 는 움직임 벡터와 참조 프레임 부호화에 대한 비트 율이다. 식 (2)는 왜곡 정도와 비트 율 측면 모두에서 가장 우수한 정합 블록을 결정하는 것을 의미한다. 각각의 모드에서 결정된 참조 프레임과 움직임 벡터들에 대하여 식 (3)을 최소로 하는 모드를 구한다.

$$J_{MODE}(MODE) = D(MODE) + \lambda_{MODE} \cdot R(MODE) \quad (3)$$

$$where \ mode \in \left\{ \begin{array}{l} SKIP, 16 \times 16, 16 \times 8, 8 \times 16, \\ 8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4, \\ INTRA4 \times 4, INTRA16 \times 16 \end{array} \right\}.$$

여기서 $R(MODE)$ 는 원 매크로 블록 전체를 부호화 하는데 필요한 전체 비트 수를 나타낸다. 그리고 $D(MODE)$ 는 복원 오차를 나타내는 왜곡 척도로써 식 (4)와 같이 정의된다.

$$D(MODE) = \sum_{x=1}^M \sum_{y=1}^N |S[x, y] - \hat{S}[x, y]|^2 \quad (4)$$

식 (4)의 \hat{S} 는 참조 프레임과 양자화 된 오차 블록으로부터 얻어진 복원 블록을 의미한다.

RDO 기법의 활용으로 인하여 PSNR은 RDO 기법을 사용하지 않았을 때에 비해 0.35dB 증가하고, 비트율은 9% 감소한다^[2]. 그러나 각각의 모드에 대해서 $D(MODE)$ 와 비트 율 계산은 50%의 계산량 증가를 야기시킨다^[3]. 이를 개선하기 위하여 지금까지 효과적인 모드 선택 방식들이 제시되어 왔다. 한편 식 (3)과 식 (4)로부터 왜곡치 $D(MODE)$ 를 얻기

위해서는 오차 블록에 DCT 및 양자화, 역양자화 및 IDCT를 차례로 적용한 후 예측 블록과 더하는 복원 과정이 필요함을 알 수 있다. 그리고 각 모드의 비트 율 $R(MODE)$ 를 알기 위해서는 엔트로피 부호화가 수행되어야 함을 알 수 있다. 따라서 고속 부호화를 위하여 왜곡치와 비트 율을 구하기 위한 계산량 감소에 대한 연구가 이루어져야 한다.

III. 기존의 DCT 및 양자화 생략 방식

일반적으로 오차 블록의 공분산(covariance)이 주어지면 각각의 DCT 계수들에 대한 분산이 얻어진다는 사실이 알려져 있다^[10]. 2차원 이산 확률장(discrete random field)에서 가로, 세로 방향에 대한 픽셀 간 거리를 각각 m, n 이라 할 때 공분산 함수 $r(m, n)$ 는 식 (5)와 같이 표현된다.

$$r(m, n) = r_h(m)r_v(n) = \sigma_e^2 \rho_h^{|m|} \rho_v^{|n|}, \quad where \ |\rho_h| < 1 \ and \ |\rho_v| < 1 \quad (5)$$

여기서 $r_h(m)$, $r_v(n)$, 그리고 σ_e^2 는 가로, 세로 방향의 공분산과 오차 블록의 분산을 나타낸다. 그리고 ρ_h 와 ρ_v 는 가로, 세로 방향에 대한 상관 계수(correlation coefficient)를 의미한다. 식 (5)에 DCT가 유니터리(unitary) 변환이라는 특성을 적용하면 (k, l) 위치의 DCT 계수에 대한 분산 $\sigma_f^2(k, l)$ 은 식 (6)과 같이 유도된다^[10].

$$\sigma_f^2(k, l) = \sigma_e^2 [AR_h A^T]_{k,k} [AR_v A^T]_{l,l}$$

$$where \ R_h = \begin{bmatrix} 1 & \rho_h & \rho_h^2 & \cdots \\ \rho_h & 1 & \rho_h & \\ \rho_h^2 & \rho_h & 1 & \\ \vdots & & & \ddots \\ & & & & 1 \end{bmatrix}$$

$$and \ R_v = \begin{bmatrix} 1 & \rho_v & \rho_v^2 & \cdots \\ \rho_v & 1 & \rho_v & \\ \rho_v^2 & \rho_v & 1 & \\ \vdots & & & \ddots \\ & & & & 1 \end{bmatrix} \quad (6)$$

여기서 $[\cdot]_{i,j}$ 는 행렬의 (i, j) 번째 요소이다. 그리고 식 (6)의 A , R_h , 그리고 R_v 는 DCT의 기저 행렬과 가로, 세로 방향에 대한 상관 행렬(correlation matrix)을 각각 나타낸다. 식 (6)은 오차 블록의 분

산과 상관 계수가 주어질 경우 DCT 계수의 분산이 어떻게 구해지는지를 보여준다.

이 같은 DCT 계수의 통계적 특성을 바탕으로 H.263 동영상 압축 표준에 대한 DCT 및 양자화 과정 생략 방식이 제안되었다⁹⁾. 기존 방식에서는 오차 블록의 픽셀을 영 평균 라플라시안 분포(zero-mean Laplacian distribution)를 가지는 이산 확률장이라 가정하고 매크로 블록의 예측 오차에 대한 기대치를 식 (7)과 같이 근사화하였다.

$$\mu_\epsilon \approx \frac{SAD}{256} \quad (7)$$

그런데 일반적으로 영 평균 라플라시안 분포에서 양의 부분에 대한 기대치 μ_ϵ 는 식 (8)과 같다.

$$\mu_\epsilon = \frac{\sigma_\epsilon}{\sqrt{2}} \quad (8)$$

따라서 식 (7)과 식 (8)로부터 표준 편차는 식 (9)와 같이 추정된다.

$$\sigma_\epsilon \approx \sqrt{2} \cdot \frac{SAD}{256} \quad (9)$$

H.263 부호화기에서 QUANT와 DZ를 양자화 파라미터와 데드 존(dead-zone)이라 할 때 DCT 계수가 $2 \cdot QUANT + DZ$ 보다 작을 경우 항상 '0'으로 양자화 된다. 일반적으로 영 평균 라플라시안 분포에서 각 화소의 절대치가 $(0, 3\sigma_\epsilon)$ 영역 내에 99% 존재한다고 알려져 있다. 그런데 DCT의 유니터리 특성과 DCT 계수의 분산이 화소 값의 분산으로 유도되어지는 식 (6)으로부터 DCT 계수 역시 영 평균 라플라시안 분포를 지닌다고 볼 수 있다⁹⁾. 그러므로 DCT 계수의 절대치 역시 $(0, 3\sigma_f)$ 안에 99% 존재한다. 따라서 DCT 계수가 양자화 후에 '0'이 될 조건은 식 (10)과 같이 표현될 수 있다.

$$3\sigma_f < 2 \cdot QUANT + DZ = 2.5 \cdot QUANT \quad (10)$$

그리고 식 (6)을 이용하여 식 (10)을 정리하면 다음과 같다.

$$3\sigma_f = 3\sigma_\epsilon \sqrt{[A_{8 \times 8} R_h A_{8 \times 8}^T]_{k,k} [A_{8 \times 8} R_v A_{8 \times 8}^T]_{l,l}} < 2.5 \cdot QUANT \quad (11)$$

여기서 $A_{8 \times 8}$ 은 8×8 블록에 대한 DCT의 기

저행렬을 나타낸다. 그리고 식 (11)의 표준편차를 식 (9)로 치환하면 DCT 계수가 '0'이 될 조건식이 식 (12)와 같이 유도된다.

$$SAD < \frac{K}{n \cdot \sqrt{[A_{8 \times 8} R_h A_{8 \times 8}^T]_{k,k} [A_{8 \times 8} R_v A_{8 \times 8}^T]_{l,l}}} \cdot QUANT = T(k,l) \cdot QUANT \quad (12)$$

여기서 K 는 상수, n 은 3 이다. 기존 방식에서는 R_h 와 R_v 를 구성하기 위한 상관 계수 ρ_h, ρ_v 의 값을 여러 영상에 대한 실험을 통하여 0.6으로 설정하였다. 그리고 $T(k,l)$ 을 테이블 화하여 모든 영상에 동일하게 적용하였다. 식 (12)에서 알 수 있듯이 기존 방식은 움직임 추정에서 얻어지는 SAD 값과 8×8 DCT에 기반하여 DCT 및 양자화 과정의 생략을 수행한다.

IV. 제안 알고리즘

기존 방식⁹⁾에서는 H.263 동영상 부호화기에서 SAD를 이용하여 오차 블록의 분산을 추정하였다. 그러나 H.264/AVC의 RDO 모드 결정 과정에서는 SSD가 왜곡 척도로서 사용되기 때문에 SAD를 사용할 수 없다. 따라서 기존 방식의 적용은 SAD 계산에 따른 부가 계산량이 야기된다. 또한 기존 방식은 8×8 DCT를 기반으로 있기 때문에 4×4 DCT를 기본으로 하고 있는 H.264/AVC 부호화기에는 적합하지 않다. 따라서 본 논문은 4×4 DCT와 RDO에 의한 모드 결정 과정의 특징을 이용하여 새로운 오차 블록의 분산을 추정한다. 그리고 (I)DCT 및 (역)양자화를 생략하기 위한 효과적인 임계값을 제시한다.

RDO 모드 결정 과정에서 왜곡치 $D(MODE)$ 는 예측 블록과 역양자화 후의 오차 블록으로 구성되는 복원 블록을 필요로 한다. 그런데 이때 양자화된 DCT 계수 값이 모두 '0'인 오차 블록(all zero coefficient block(AZCB))의 경우, 복원 블록은 예측 블록과 동일하다. 즉, AZCB에 대하여 4×4 블록에 대한 왜곡치는 식 (13)과 같이 표현된다.

$$\begin{aligned} D_{4 \times 4} &= \sum_{x=1}^4 \sum_{y=1}^4 |S[x,y] - \hat{S}[x,y]|^2 \\ &= \sum_{x=1}^4 \sum_{y=1}^4 |S[x,y] - C[x,y] - \hat{\epsilon}[x,y]|^2 \\ &= \sum_{x=1}^4 \sum_{y=1}^4 |S[x,y] - C[x,y]|^2 \end{aligned} \quad (13)$$

여기서 $\hat{\varepsilon}$ 은 (I)DCT 및 (역)양자화 후에 복원된 오차 블록을 의미한다. 본 논문은 식 (13)의 원 블록과 예측 블록간의 왜곡을 식 (14)와 같이 새롭게 정의한다.

$$\hat{D}_{4 \times 4} = \sum_{x=1}^4 \sum_{y=1}^4 |S[x,y] - C[x,y]|^2 \quad (14)$$

식 (1)과 식 (13)로부터 $\hat{D}_{4 \times 4}$ 는 (I)DCT 및 (역)양자화를 거치지 않고 얻어짐을 알 수 있다. 따라서 AZCB인 경우 모드 결정 과정에서 $D_{4 \times 4}$ 대신 $\hat{D}_{4 \times 4}$ 를 사용하여 (I)DCT 및 (역)양자화를 생략하더라도 성능이 저하되지 않는다. 그러나 이것은 AZCB가 사전에 검출된 경우에 유효하다. 따라서 AZCB를 효과적으로 검출할 수 있는 방법이 필요하다.

본 논문에서는 기존 방식의 접근 기법에 기초하여 AZCB를 검출하고자 한다. 먼저 오차 블록의 화소가 영 평균 라플라시안 분포라고 가정하면 오차 블록의 분산 σ_ε^2 은 식 (15)와 같다.

$$\begin{aligned} \sigma_\varepsilon^2 &= E[|e[x,y]|^2] \quad (15) \\ &= \frac{1}{16} \sum_{x=1}^4 \sum_{y=1}^4 |e[x,y]|^2 \\ &= \frac{1}{16} \sum_{x=1}^4 \sum_{y=1}^4 |S[x,y] - C[x,y]|^2 \end{aligned}$$

식 (14)와 식 (15)를 결합하여 오차 블록의 분산을 식 (16)과 같이 정의한다.

$$\sigma_\varepsilon^2 = \frac{\hat{D}_{4 \times 4}}{16} \quad (16)$$

H.264/AVC에서 k,l 번째 위치의 DCT 계수 $F_{k,l}$ 는 다음과 같이 양자화된다.

$$Q(F_{k,l}) = \left\lceil \frac{|F_{k,l}| + d}{\Delta} \right\rceil, \quad (17)$$

where $\Delta = \text{step}(QP\%6) \ll \text{floor}(QP/6)$.

여기서 $\lceil \cdot \rceil$, Δ , d 는 각각 버림 연산, 양자화 크기, 그리고 오프셋을 의미한다. $\text{step}(QP\%6)$ 은 초기 양자화 크기로서 표 1과 같이 설정되어 있다.

표 1. H.264/AVC의 초기 양자화 크기

QP%6	0	1	2	3	4	5
step (QP%6)	0.625	0.6875	0.8125	0.875	1	1.125

3장에서 언급한 바와 같이 영 평균 라플라시안 분포에서 DCT 계수의 절대치 $|F_{k,l}|$ 는 $(0, 3\sigma_f)$ 안에 99% 존재한다고 알려져 있다. 그러므로 식 (17)로부터 양자화 계수가 '0'이 될 조건은 식 (18)과 같이 정의할 수 있다.

$$\frac{3\sigma_f}{\Delta} + d < 1 \quad (18)$$

식 (18)을 분산 σ_f^2 에 대하여 정리하면 식 (19)가 얻어진다.

$$9\sigma_f^2 < (1-d)^2 \cdot \Delta^2 \quad (19)$$

식 (6)를 식 (19)에 적용하면 식 (20)과 같다.

$$\begin{aligned} 9\sigma_f^2 &= 9\sigma_\varepsilon^2 [A_{4 \times 4} R_h A_{4 \times 4}^T]_{k,k} [A_{4 \times 4} R_v A_{4 \times 4}^T]_{l,l} \\ &< (1-d)^2 \cdot \Delta^2 \quad (20) \end{aligned}$$

여기서 $A_{4 \times 4}$ 는 4×4 DCT의 기저행렬을 나타낸다. 그리고 식 (16)을 식 (20)에 대입하고 d 에 1/6을 대입하면 Inter 블록의 DCT 계수값이 '0'으로 양자화 될 조건식이 얻어진다. 식 (21)은 이를 나타낸다.

$$\begin{aligned} \hat{D}_{4 \times 4} \quad (21) \\ < \frac{1.235}{[A_{4 \times 4} R_h A_{4 \times 4}^T]_{k,k} [A_{4 \times 4} R_v A_{4 \times 4}^T]_{l,l}} \cdot Qstep^2 \\ = Th(k,l) \end{aligned}$$

식 (21)에서 R_h 와 R_v 는 H.264/AVC의 압축 방식에서 여러 테스트 영상들을 실험을 토대로 하여 0.19로 결정된 ρ_h 와 ρ_v 에 의하여 계산되었다.

양자화 변수를 28로 설정할 경우 양자화된 DCT 계수 각각이 '0'임을 판별하기 위한 $Th(k,l)$ 은 식 (22)과 같다.

$$Th(k,l) = \begin{bmatrix} 180.25 & 215.96 & 273.73 & 322.28 \\ 215.96 & 258.72 & 327.97 & 386.03 \\ 273.73 & 327.97 & 415.45 & 489.41 \\ 322.28 & 386.03 & 489.41 & 575.88 \end{bmatrix} \quad (22)$$

상관관계 계수가 양의 값을 가지기 때문에 식

(22)에서 보는 것처럼 16개의 $Th(k,l)$ 중에서 $Th(0,0)$ 이 최소값이 된다. 그러므로 $\hat{D}_{4 \times 4}$ 값이 $Th(0,0)$ 보다 작으면 주어진 오차 블록은 AZCB이라고 판단 할 수 있다. 이러한 사실에 기초하여 주어진 양자화 변수에 대한 임계값 TH 는 식 (23)과 같이 결정된다.

$$TH = \min Th(k,l) \text{ for all } k,l \quad (23)$$

임계값 TH 는 양자화 변수가 주어지면 사전에 계산되어 모든 동영상에 적용되어진다. 표 2는 대표적인 양자화 변수들에서 본 논문에서 제시한 방식에 의하여 설정된 AZCB 검출을 위한 임계값이다.

표 2. 양자화 변수에 대한 AZCB 검출 임계값

QP	28	32	36	40
TH	180.25	475.97	1126.57	2884.01

본 논문에서 제안하는 방식을 요약하면 그림 1과 같다. 제안 알고리즘에서는 각 4×4 블록에 대해서 $\hat{D}_{4 \times 4}$ 를 계산하고 임계값 TH 와의 비교를 통해 AZCB 여부를 판정한다. 만약 AZCB로 판단되면 $\hat{D}_{4 \times 4}$ 를 복원 오차로 사용하고 (I)DCT, (역)양자화, DCT 부호화 과정을 생략한다. 그러나 AZCB가 아니라고 판단된 경우에는 DCT 및 양자화 후에 AZCB 여부를 실제로 판정한다. 실제 AZCB라고 판단된 블록에 대해서 역시 $\hat{D}_{4 \times 4}$ 를 복원 오차로

사용하고 역양자화 및 IDCT 과정을 생략한다. 한편 AZCB가 아닌 블록에 대해서는 (I)DCT 및 (역)양자화, 그리고 $D_{4 \times 4}$ 를 계산하고 DCT 계수 부호화를 수행하여 비트 율을 구한다. 따라서 이 경우 사전에 이루어진 $\hat{D}_{4 \times 4}$ 계산은 부가 계산량이 된다. 그러나 AZCB로 검출되는 블록의 수가 많을 경우 부가 계산량의 비중은 무시할 수 있을 정도로 작을 것이다.

V. 실험 결과 및 고찰

제안 방식의 우수성을 확인하기 위하여 본 논문에서는 H.264/AVC의 압축 방식에 대하여 기존의 방식과 제안 방식의 성능을 비교, 평가하였다. 모의 실험은 QCIF 형식의 'silent', 'foreman', 'container', 그리고 'news' 시퀀스에 대하여 JM 9.6에 의해서 수행되었다. 실험에서 사용된 양자화 변수는 28, 32, 36, 40으로 하였다. 그리고 움직임 추정 및 보상에 사용된 참조 프레임은 5 프레임, 탐색 범위는 16, 프레임 비율은 10Hz와 15Hz로 100프레임을 실험하였다. 제안 알고리즘은 P 프레임의 INTER 모드에만 적용되었다.

앞에서 설명한 바와 같이 기존 방식을 H.264/AVC에 직접 적용할 수는 없다. 그리하여 본 논문에서는 객관적 성능 비교를 위하여 기존 방식을 H.264/AVC의 환경에 맞게 변형하였다. 즉, 모든 조건을 제안 방식과 동일하게 하고 오차 블록의 분산만 기존 방식대로 설정하였다. 식 (24)는 기존 방식의 변형된 조건을 나타낸다.

$$SAD(MODE) < \frac{3.14}{\sqrt{[A_{4 \times 4} R_1 A_{4 \times 4}^T]_{k,k} [A_{4 \times 4} R_2 A_{4 \times 4}^T]_{l,l}}} \cdot \Delta \quad (24)$$

표 3은 전체 Inter 모드 블록에 대하여 기존 방식과 제안 방식에 의하여 검출된 AZCB의 수를 정리한 것이다. 표 3의 검출 오류는 AZCB이 아닌 블록을 AZCB이라고 제안 방식에서 잘못 검출하는 경우를 의미한다. 표 3은 기존 방식의 경우 평균 55%의 블록에서 (I)DCT 및 (역)양자화, DCT 계수 부호화 과정이 생략됨을 보인다.

반면 제안 방식의 경우 1단계에서 평균 64%의 블록에 대해 (I)DCT 및 (역)양자화, DCT 계수 부호화 과정의 생략이 이루어진다. 그리고 제안 방식 전체에서는 2단계에서 평균 94%의 블록에 대해 IDCT 및 역양자화, DCT 계수 부호화 과정이 생략

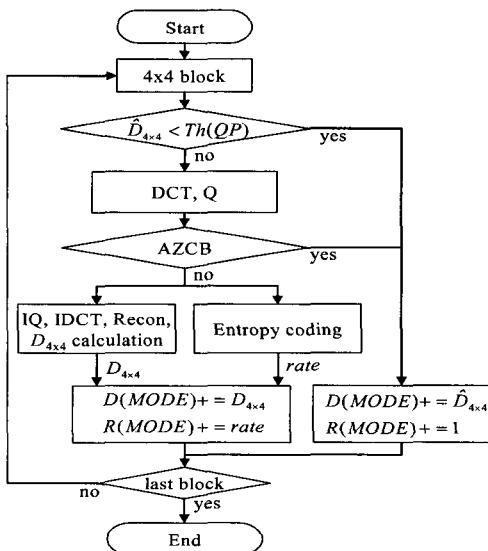


그림 1. 제안 알고리즘의 순서도

표 3. 각 방식에서 검출된 AZCB 수 비교

seq.	QP	Total	기존방식	제안방식 1단계	제안방식 1,2단계	검출 오류
foreman 10Hz	28	1097712	354213	442430	878206	0
	32	1097712	547933	637091	977337	0
	36	1097712	684058	783631	1034888	0
	40	1097712	826719	917939	1068292	0
news 10Hz	28	1097712	541532	613153	984233	0
	32	1097712	647184	713581	1029876	0
	36	1097712	729034	797946	1058699	0
	40	1097712	817893	837377	1075975	0
container 10Hz	28	1097712	416641	521784	1020090	0
	32	1097712	610920	702827	1058541	0
	36	1097712	707457	777284	1077290	0
	40	1097712	793202	867726	1086775	0
silent 15Hz	28	1097712	284184	375237	983123	0
	32	1097712	437159	563863	1032185	0
	36	1097712	598524	743386	1063704	0
	40	1097712	787938	913179	1081501	0
평균		1097712 (100%)	611537 (55%)	703652 (64%)	1031920 (94%)	0 (0%)

표 4. 각 경우별 사용되는 계산량

			DCT, Q	IQ, IDCT	SSD 계산	부가 계산량
			기존방식	생략	ADD	0
		SHF	0	0	0	0
		MUL	0	0	16	0
		ABS	0	0	0	16
		CMP	0	0	0	1
	수행	ADD	80	80	47	15
		SHF	32	48	0	0
		MUL	16	16	16	0
		ABS	0	0	0	16
		CMP	0	0	0	1
제안방식	1단계 생략	ADD	0	0	31	0
		SHF	0	0	0	0
		MUL	0	0	16	0
		ABS	0	0	0	0
		CMP	0	0	0	1
	2단계 생략	ADD	80	0	31	0
		SHF	32	0	0	0
		MUL	16	0	16	0
		ABS	0	0	0	0
		CMP	0	0	0	2
	수행	ADD	80	80	47	31
		SHF	32	48	0	0
MUL		16	16	16	16	
ABS		0	0	0	0	
	CMP	0	0	0	2	

됨을 보여준다. 또한 검출 오류가 모두 '0'이므로 제안 방식에 의한 성능 저하가 크지 않을 것이라는 사실을 알 수 있다.

다음의 표 4는 기존 방식과 제안 방식에 있어서 각각의 경우에 따라 사용되는 계산량을 정리한 것

이다. 일반적으로 DCT는 덧셈 연산 64회, 자리 옮김 연산 16회만으로 구현되며, 16개의 DCT 계수에 대한 양자화 과정은 덧셈 16회, 곱셈 16회, 자리 옮김 16회가 각각 사용된다.

역양자화 과정에서는 곱셈 16회와 자리 옮김 16회, IDCT 과정에서는 덧셈 80회와 자리 옮김 32회가 사용된다. 그리고 표 4의 부가 계산량은 각 영역에서 새로운 임계값의 계산 및 그 적용을 위하여 요구되는 계산량을 나타낸 것이다. 즉, 기존 방식의 경우 SAD 계산과 임계값 비교의 양이고, 제안 방식의 경우 $\hat{D}_{4 \times 4}$ 계산과 임계값 비교의 양이다.

표 5는 표 4에 근거하여 RDO를 사용한 모드 결정에 있어서 기존 방식과 제안 방식간의 실제 계산량을 비교한 것이다. 표 5는 기존 방식에 비하여 제안 방식이 약 50%의 계산량 이득을 가져옴을 보여준다. 한편 계산량 비교에서 DCT 계수 부호화에 요구되는 계산량은 고려하지 않았다. 그러나 표 3에서 기존 방식에 비해 제안 방식에 의하여 검출되는 AZCB가 더욱 더 많다는 사실을 알 수 있다. 이것은 DCT 계수 부호화 계산량을 고려 할 경우 제안 방식에서 더욱 계산량 감소가 발생할 것이다. 따라서 실제 계산량 이득은 제안 방식에서 더욱 더 클 것으로 예상된다.

표 5. 기존 방식과 제안 방식의 실제 계산량

Seq.	QP	총 계산량	기존 방식의 계산량	제안 방식의 계산량
foreman 10Hz	28	367733520	300846960	182659296
	32	367733520	245055600	137027910
	36	367733520	205851600	106211193
	40	367733520	164765232	81970833
news 10Hz	28	367733520	246899088	138688440
	32	367733520	216471312	116285127
	36	367733520	192898512	99435681
	40	367733520	167307120	84322950
container 10Hz	28	367733520	282867696	143052642
	32	367733520	226915344	111738738
	36	367733520	199112688	98252742
	40	367733520	174418128	84622329
silent 15Hz	28	367733520	321015312	169609374
	32	367733520	276958512	135120786
	36	367733520	230485392	105437886
	40	367733520	175934160	79850610
Average		367733520 (100%)	226737666 (61.66%)	117142909 (31.86%)

VI. 결론

지금까지 H.264/AVC의 고속 부호화를 위하여 수행된 많은 연구는 RDO 모드 결정에 있어서 불필요한 모드를 제거하는 기법을 개발하였다. 반면에 RDO 모드 결정에서 요구되는 왜곡 정도 및 비트 율에 대한 계산량에 관한 개선은 시도되지 않았다. 그런데 현재 왜곡 정도 및 비트 율에 대한 계산량의 비중은 점점 증가하고 있다. 이에 본 논문에서는 H.264/AVC의 부호화기에서 왜곡 정도 및 비트 율을 효율적으로 계산하는 새로운 방식을 제안한다.

본 논문은 H.264/AVC의 RDO에 의한 모드 결정 과정의 특성과 DCT 계수의 특성을 이용하여 오차 블록의 분산을 새롭게 추정하고 양자화된 DCT 계수가 모두 '0'인 오차 블록에 대한 검출 조건을 제안한다. 제안된 검출 조건에 의하여 RDO 모드 결정 과정에 요구되는 (IDCT, (역)양자화, 엔트로피 부호화 과정이 효과적으로 생략되어 계산량 감소가 이루어진다.

제안 방식의 우수성을 검증하기 위하여 수행된 모의 실험은 제안 방식이 기존 방식에 비해 약 50%의 계산량 절감을 가져옴을 보여준다. 한편 제안 방식은 기존 고속 모드 결정 방법들과 결합하여 사용할 경우 더욱 더 계산량을 절감할 수 있는 장점을 지니고 있다.

참고 문헌

[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 Curcuits and Systems Magazine*, First Quarter, pp. 7-28, 2004.

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

[4] S. Zhu and K.-K. Ma, "A New Diamond Search Algorithm for Fast Block-Matching

Motion Estimation", *IEEE Transactions on Image Processing*, Vol. 9, No. 2, pp.287-290, Feb. 2000.

[5] C. Zhu, X. Lin, and L.-P. Chau, "Hexagon-Based Search Pattern for Fast Block Motion Estimation", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 12, No. 5, pp.349-355, May 2002.

[6] "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.

[7] L. Yang, K. Yu, J. Li, and S. Li, "An Effective Variable Block-Size Early Termination Algorithm for H.264 Video Coding", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 15, No. 6, pp.784-788, Jun. 2005.

[8] 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, 16th Sep. 2004.

[9] I.-M. Pao and M.-T. Sun, "Modeling DCT Coefficients for Fast Video Encoding", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 9, No. 4, pp.608-616, Jun. 1999.

[10] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, pp. 132-140, 1986

문 정 미 (Jeong-mee Moon)

정회원



2001년 2월 홍익대학교 전자공학과 (공학사)

2001년 9월~현재 부산대학교 전자공학과 석박사통합과정 <관심분야> 영상처리, 영상부호화, 신호처리

김 재 호 (Jae-ho Kim)

정회원

한국통신학회 논문지 제 29권, 제 1C호 참조

문 용 호 (Yong-ho Moon)

정회원

한국통신학회 논문지 제 29권, 제 1C호 참조