

논문 2004-41SP-6-19

# H.264 동영상 압축을 위한 고속 부호화 모드 결정 방법

## (Fast Coding Mode Decision for H.264 Video Coding)

이 제 윤\*, 전 병 우\*\*

(Jeyun Lee and Byeungwoo Jeon)

### 요 약

최근에 표준화가 완료된 H.264 비디오 부호화 표준은 기존 표준에 비해 높은 부호화 효율을 제공한다. H.264에는 움직임 보상 부호화를 위해 블록 크기에 따른 7가지 모드가 있으며, 인트라 부호화를 위해서는 인트라 예측 방향에 따라 다양한 모드를 지원한다. 이렇게 다양한 모드 중에서 가장 부호화 효율이 높은 모드를 결정하기 위하여, H.264 참조 모델은 복잡도가 높은 비트율-왜곡(Rate distortion) 최적화 기법을 채택하고 있다. 본 논문에서 제안한 '조기 SKIP 모드 결정법'과 '선택적 인트라 모드 생략법'은, 부호화 효율의 큰 감소 없이 H.264의 모드 결정에 소요되는 비트율-왜곡치(Rate distortion cost) 계산수를 평균 72% 감소시켰으며, 전체 부호화 시간도 평균 30% 감소시켰다.

### Abstract

H.264 is the newest international video coding standard that provides high coding efficiency. A macroblock in H.264 has 7 different motion-compensation block sizes in the Inter mode, and several different prediction directions in the Intra mode. In order to achieve as highest coding efficiency as possible, H.264 reference model employs complex mode decision technique based on rate-distortion (RD) optimization which requires high computational complexity. In this paper, we propose two techniques - 'early SKIP mode decision' and 'selective intra mode decision' - which can further reduce the computational complexity. Simulation results show that without considerable performance degradation, the proposed methods reduce encoding time by 30% on average and save the number of computing rate-distortion cost by 72%.

**Keywords :** encoder optimization, rate distortion optimization, SKIP mode, skip intra mode, fast mode decision

## I. 서 론

H.264의 근간이 되는 부호화 기술로는 4x4 블록 단위의 정수 변환 부호화, 가변 블록 움직임 보상, 방향성을 고려한 인트라 부호화 기법, 다중 참조 영상 부호화 기법, 그리고 1/4 화소 단위 움직임 벡터 추정 기법이 있다. 이러한 부호화 기법은 기존의 비디오 부호화 표준과 매우 상이한 기술로서, H.264만의 독특한 기술이

다. 이러한 H.264만의 독특한 부호화 기법으로 인해, 현재 H.264의 압축 효율은 기존의 MPEG-4 SP(Simple Profile)와 비교하여 동일한 PSNR에서 두 배 이상 좋으며, 동일한 비트율에서는 2dB 이상 화질이 개선되었다<sup>[1,2]</sup>. 하지만, 동시에 H.264의 복잡도는 14배 이상 높다는 것도 보고되었다<sup>[3]</sup>.

기존 H.263/MPEG-4의 부호화 기법에서는 일반적으로 16x16 매크로블록 또는 8x8 블록(H.263 Annex F/MPEG-4) 단위로 움직임 보상을 수행하고, 8x8 단위로 변환 부호화 하였다. 반면에 H.264는 4x4 단위 정수 변환 부호화 기법과 함께, 그림 1과 같이 변환 부호화 크기에 독립적인 움직임 보상 블록 크기를 제공하고 있다. 이 블록의 크기는 4x4에서부터 16x16에 이르기까지 다양하게 제공되고 있다. H.264 참조 모델인 JM6.1d 부호화기는 각각의 움직임 보상 블록들의 부호화 효율을

\* 정희원, LG 전자 Digital Media 연구소  
(LG Electronics Inc, Digital Media research laboratory)

\*\* 정희원 성균관대학교 정보통신공학부  
(School of Information and Computer Engineering, Sungkyunkwan University)

※ 이 논문은 2003년도 한국학술진흥재단의 지원에 의하여 연구되었음. (KRF-2003-041-D20405)  
접수일자: 2004년1월13일, 수정완료일: 2004년11월5일

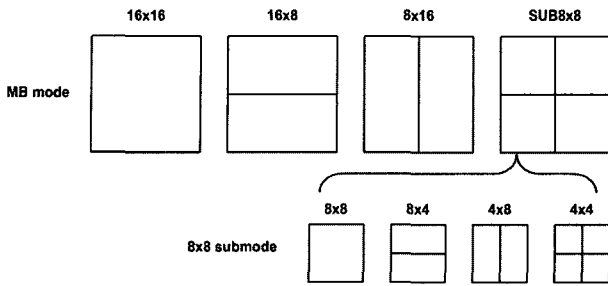


그림 1. 움직임 보상을 위한 매크로블록 분할  
Fig. 1. Macroblock segmentation for motion compensation.

계산하고, 가장 부호화 효율이 높은 것을 최종 움직임 보상 블록으로 결정하고 있다. 부호화 효율은 비트율-왜곡 최적화 기법에 근거하여 측정된다.

H.263/MPEG-4는 움직임 보상에 바로 이전 참조 영상만을 이용했지만, H.264는 과거 여러 개의 참조 영상들을 버퍼에 저장해 놓고 그 가운데 부호화 효율이 가장 좋은 참조 영상을 이용해서 움직임 보상을 하고 있다. 이때 부호화 효율도 역시 비트율-왜곡 최적화 기법에 근거하여 측정되어 진다. 최대로 참조할 수 있는 영상 수는 프로파일과 레벨에 따라 정의 되어 있다<sup>[4]</sup>. H.264의 움직임 보상 블록들 중 크기가 8x8 보다 크거나 같은 블록들은 서로 다른 참조 영상으로부터 움직임 보상될 수 있다. 즉, 참조 영상 정보를 최소 8x8 단위로 전송해 주어야 하는 것이다. 다중 영상 참조 부호화 기법은 반복적인 움직임 특성을 갖고 있는 영상에서 큰 실효를 얻을 수 있다. 실제로 mobile, tempete와 같이 회전이나 반복적인 움직임이 많은 영상에서, 다중 참조 영상 부호화 기법은 최고 1 dB의 PSNR 이득을 제공한다<sup>[5]</sup>. 하지만, H.264의 다중 영상 참조 부호화 기법은 H.264의 복잡도를 증가시키는 주요 요인 중에 하나로서, 움직임 벡터 추정에 걸리는 시간이 참조 영상 수에 비례하여 증가하게 된다.

H.263의 'Annex I'에 해당되는 Advanced 인트라 모드 부호화 기법은 DCT 영역에서 일부 DCT 계수들만을 예측 부호화하기 때문에 효율적으로 잉여 데이터의 정보량을 줄이기 어려웠다. 하지만, H.264의 인트라 부호화는 공간 영역에서 이루어지며, 휘도 신호뿐만 아니라 색차 신호에 대해서도 인트라 부호화 기법을 적용하고 있다. 또한 휘도 신호에 대하여, 복잡도가 높은 영상은 방향에 따라 9개 모드를 제공하는 인트라 4x4 모드로 부호화 되고, 배경과 같이 평활한 영역에서는 방향에 따른 4가지 모드를 갖는 인트라 16x16 모드로 부호화 된다. 다음으로 색차 신호를 위하여 인트라 16x16과

방법은 같지만 단지 블록의 크기가 8x8로 줄어든 모드가 적용된다.

7개의 움직임 보상 블록과 다양한 인트라 모드 중에서 최고의 부호화 효율을 갖는 매크로블록 모드 선택을 위해, H.264는 비트율-왜곡 최적화 기법을 채용하고 있다. 비트율-왜곡 최적화 기법이 적용된 모드 결정법은 가능한 모든 모드의 비트율-왜곡치(Rate-distortion cost)를 계산하고, 이중 가장 작은 값을 갖는 모드를 선택하는 방법이다. 그리고 H.264는 높은 부호화 효율을 달성하기 위하여, P 슬라이스에서도 인트라 모드로 부호화 되는 것을 허용하고 있다. 결과적으로 P 슬라이스에서 매크로블록 모드 결정시, 그림 1의 인트라 모드뿐만 아니라 인트라 모드의 비트율-왜곡치까지도 계산해야 한다.

비트율-왜곡 최적화 기법은 Lagrangian 최적화 기법을 바탕으로 정리, 발전된 방법이다. 비트율-왜곡 최적화 기법은 움직임 벡터 선택, 참조 영상 선택, 인트라 및 인트라 모드 결정에 적용되고 있다. H.264의 비트율-왜곡 최적화 기법은 약 10%의 압축 효율<sup>[6]</sup>을 제공하는 반면, 많은 계산량을 필요로 한다. 비트율-왜곡 최적화 기법은 부호화기에서 효율적인 부호화를 위하여 고안된 방법으로, 표준 그 자체로서의 강제성은 없다. 즉, 더 좋은 최적화 기법이 있다면 그것을 적용하더라도 표준상의 어떠한 문제도 발생되지 않는다.

본 논문은 비트율-왜곡치 기반 최적 모드를 결정하면서 화질과 압축 성능의 저하 없이 계산량을 줄일 수 있는 방법을 제안한다. 본 논문은 II장에서 현재 H.264 참조 소프트웨어 JM6.1d<sup>[7]</sup>에 구현되어 있는 비트율-왜곡 최적화 기반 모드 결정법을 설명하고, III장과 IV장에서는 본 논문에서 제안한 고속 부호화 모드 결정법을 설명할 것이다. 마지막으로 V장과 VI장에서는 실험 방법과 결과, 그리고 결론을 제시할 것이다.

## II. 비트율-왜곡(Rate distortion) 최적화 기법

매크로블록 모드는 {SKIP, MB16x16, MB16x8, MB8x16, SUB8x8, INT4x4, INT16x16}로 총 7가지가 있다. 여기서 SKIP 모드는 움직임 보상 블록 크기가 16x16이면서 움직임 벡터가 예측 움직임 벡터와 같고, CBP(Coded Block Pattern)=0일 때 결정된다. 그리고 MBm<sub>x</sub>n 모드는 가로 길이가 m, 세로 길이가 n인 움직임 보상 블록을 나타낸다. 또한 SUB8x8의 경우, 각 8x8 블록 마다 독립적으로 {8x8, 8x4, 4x8, 4x4} 모드들

중에 하나로 부호화 될 수 있다. 마지막으로 INT4x4, INT16x16은 각각 인트라 4x4, 인트라 16x16 모드를 의미한다. 최종 부호화 모드 결정을 위해 각 모드별 부호화 효율을 측정하는데, 참조 모델인 H.264 JM6.1d는 각 모드별 비트율-왜곡치로 부호화 효율을 측정한다. 모드 결정을 위한 비트율-왜곡치 계산식은 다음과 같다.

$$J_{Mode} = SSD_{Mode} + \lambda_{Mode} \{R(Header) + R(residualdata)\} \quad (1)$$

- $J_{Mode}$ : 비트율-왜곡치
- $SSD_{Mode}$ : Sum of Squared Difference
- $\lambda_{Mode}$ : Lagrangian 계수(=0.85 × 2<sup>Qp/3</sup>)
- $R(x)$ : "x"를 부호화하는데 필요한 비트량
- $Header$ : 모드 정보와 차분 움직임 벡터, CBP(Coded Block Pattern) 등의 정보

최종적으로 결정되는 최적 모드는 비트율-왜곡치가 가장 작은 모드가 된다. 이는 왜곡치와 비트율이 작을 수록 부호화 효율은 높아지고, 비트율-왜곡치는 작아지기 때문이다. 모드 결정을 위한 비트율-왜곡치 계산식은 잉여 데이터의 비트량도 필요로 한다. 따라서 실제로 잉여 데이터를 부호화하기 위해서 뿐만 아니라, 모드 결정을 위해 4x4 정수 변환 부호화, 양자화, 가변길이 부호화가 반복적으로 수행되어진다.

그림 2는 비트율-왜곡치 계산 과정을 도식화한 것으로, 각 모드에 따른 잉여 데이터가 입력되면, 4x4 정수 변환 부호화, 양자화가 이루어져, 헤더 정보를 비롯한 잉여 데이터의 비트량을 산출 할 수 있게 된다. 그리고 양자화된 잉여 데이터를 역 양자화, 역 정수 변환 부호화하면, 복원 영상을 얻을 수 있게 되어, 원 영상과의 차이 값인 왜곡치를 얻을 수 있는 것이다. 하지만, 비트율-왜곡치를 계산하는 과정은 단지 모드 결정을 위해 반복적으로 수행되며, 이는 부호화기의 복잡도를 높이는 원인이 된다.

다음은 H.264 참조 모델의 비트율-왜곡치 계산을 정량화하는 과정이다. H.264는 4x4 단위의 정수 변환 부호화를 채용하고 있기 때문에, 4x4 블록을 비트율-왜곡

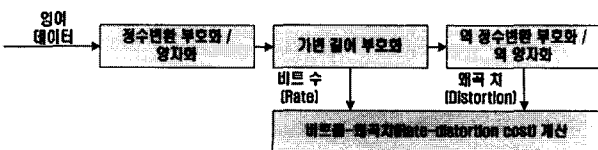


그림 2. 비트율-왜곡치 계산 과정  
Fig. 2. Computation of RDCost.

치 계산의 최소 단위로 가정했을 때, 16x16 블록의 비트율-왜곡치 계산은 16번 수행되어진다. 그림 3은 JM6.1d 부호화기의 최대 비트율-왜곡치 계산수를 나타내고 있다.

그림 3(a)의 SKIP, MB16x16의 비트율-왜곡치 계산수는 4x4 블록을 비트율-왜곡치 계산의 최소 단위로 했을 때 각각 16번의 계산이 수행된다. 또한 MB16x8, MB8x16들도 각각의 하부 블록들이 독립적인 움직임 벡터와 참조 영상을 갖지만, 각 하부 블록들이 독립적인 모드로 부호화 되는 것이 아니기 때문에 단순히 16번의 비트율-왜곡치가 계산된다. 하지만 SUB8x8은 각각의 하부 블록들이 독립적인 모드로 움직임 보상이 가능하기 때문에, 각 8x8 블록 별로 최적의 움직임 보상 블록을 결정하기 위한 별도의 비트율-왜곡치 계산이 수행된다. 따라서 4 (4개의 8x8 블록) x 4 (4개의 4x4 블록) x 4 (4가지 모드) = 64 번의 비트율-왜곡치가 계산되는 것이다.

그리고 그림 3(b)에서 인트라 4x4(INT4x4) 모드로

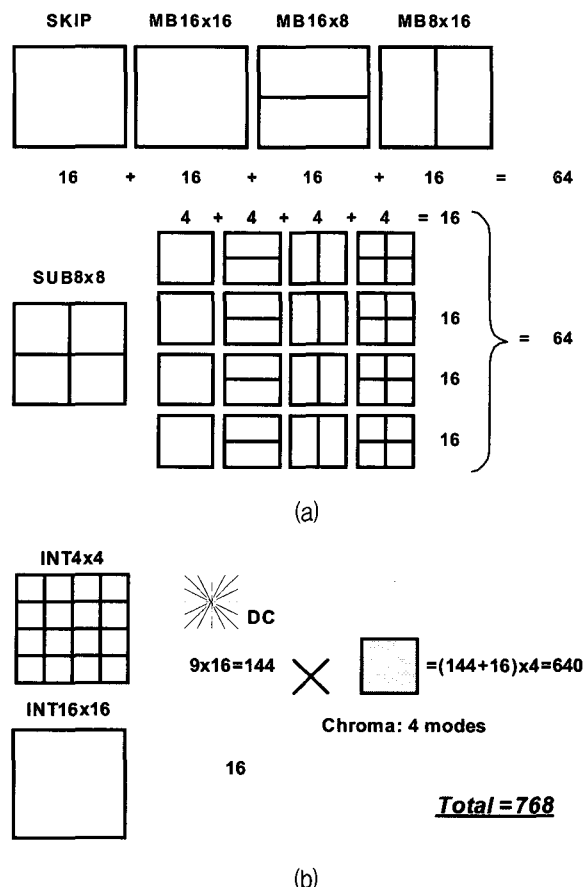


그림 3. 매크로블록 모드 결정의 최대 비트율-왜곡치 계산 수; (a) 인터 부호화 (b) 인트라 부호화  
Fig. 3. Maximum number of computing Rdcost in mode decision; (a) Inter coding (b) Intra coding.

부호화 될 경우, 16 (16개의 4x4 블록) x 9 (9개의 인트라 4x4 모드) =144 번의 비트율-왜곡치가 계산된다. 그런데, 현재의 JM6.1d에 구현되어 있는 인트라 16x16 (INT16x16)의 모드 결정법은 비트율-왜곡치를 이용하는 것이 아니라, 각 모드별 잉여 데이터에 대한 SAD (Sum of Absolute Difference) 또는 SATD(Sum of Absolute Transformed Difference)를 구하여 최적의 인트라 16x16 모드를 결정한다. 그리고 다른 매크로블록 모드들의 비트율-왜곡치와 비교하기 위하여, 이미 결정된 최적 인트라 16x16 모드의 비트율-왜곡치를 계산하기 때문에 16번의 계산이 수행된다.

이때, 색차 신호의 인트라 모드는 휘도 신호의 인트라 모드의 영향을 받지 않고 독립적으로 결정되기 때문에, 비트율-왜곡치 계산에 색차 신호의 4개 인트라 모드와 휘도 신호의 인트라 모드가 조합되어 비트율-왜곡치가 계산된다. 이는 색차 신호의 인트라 모드와 휘도 신호의 인트라 모드 조합에 따라서, CBP(Coded Block Pattern), 인트라 16x16의 매크로블록 모드 정보(CBP 정보와 인트라 16x16 모드가 결합되어 있음), CABAC (Context-based Adaptive Binary Arithmetic Coding) 사용 시 통계적인 정보 등이 달라질 수 있기 때문이다. 따라서 인트라 모드의 비트율-왜곡치 계산은 (144+16) x4 (4개의 색차 신호에 대한 인트라 모드) = 640 번이 수행된다.

H.264 부호화기 복잡도 분석<sup>[8]</sup>에 따르면 비트율-왜곡 최적화 기법을 사용하지 않았을 때 복잡도는 30% 감소하나, 이때 PSNR은 0.64dB 감소한다. 이는 비트율-왜곡 최적화 기법이 전체 부호화기 복잡도에서 많은 비중을 차지하고 있는 동시에 높은 부호화 효율을 제공하는 중요한 요소임을 의미한다.

### III. 초기 SKIP 모드 결정법

동영상에서 배경과 같은 부분은 매 프레임마다 반복되어 나타난다. 이렇게 변화 없이 반복되는 부분을 효

표 1. SKIP 모드의 통계적 데이터  
Table 1. Statistical data of SKIP mode.

	container	foreman	news	mobile	silent	paris	tempe	average
검출율	0.67	0.48	0.84	0.39	0.82	0.61	0.49	0.69
에러율	0.02	0.08	0.02	0.13	0.01	0.01	0.09	0.05

실험조건: Baseline profile, JM6.1d, Qp=28

과적으로 부호화하기 위하여, H.264는 SKIP 모드를 제공하고 있다. H.264 표준에서 매크로블록이 SKIP 모드로 결정 되려면 다음의 4가지 조건을 만족해야 한다<sup>[4]</sup>.

1. 움직임 보상 블록이 16x16이어야 한다.
2. 참조 영상이 바로 이전 영상이어야 한다.
3. 움직임 벡터가 예측 움직임 벡터와 같아야 한다.  
(주변 블록의 움직임 벡터를 이용할 수 없는 경우 (0,0)으로 예측 움직임 벡터를 결정)
4. 잉여 데이터의 비트량이 0 이어야 한다.

앞서 설명하였듯이, 매크로블록 모드 결정은 각 모드별 비트율-왜곡치를 구하고, 그 중 가장 작은 비트율-왜곡치를 갖는 모드를 선택함으로써 이루어진다. 그런데 SKIP 모드의 비트율-왜곡치 계산은 실제로 움직임 벡터 추정, 참조 영상 결정, 잉여 데이터의 비트량 산출 과정 없이 단순히 움직임 벡터를 예측 움직임 벡터로 가정하고, 참조 영상도 바로 이전영상으로 가정하며, 잉여 데이터의 비트량도 0으로 가정하여 비트율-왜곡치를 계산한다. 이렇게 가정된 값으로 계산된 SKIP 모드의 비트율-왜곡치를 다른 모드의 비트율-왜곡치와 비교하는 것이다. 즉, SKIP 모드 조건이 만족되지 않더라도, 비트율-왜곡치를 계산하는 것이다. 이는 비록 SKIP 모드 조건이 만족되지 않더라도, 비트율-왜곡 관점에서 SKIP 모드의 비트율-왜곡치가 다른 어떤 매크로블록 모드 보다 작다면, 최종 매크로블록 모드로 결정될 수 있기 때문이다. 실제 비디오 데이터로 예비 실험한 결과 SKIP 조건이 만족되는 경우, 최종 매크로블록 모드가 SKIP 모드로 결정될 가능성이 매우 높다는 것을 관찰할 수 있었다.

표 1은 이를 뒷받침해 주는 결과로써, '검출율'은 SKIP 모드 조건이 만족될 경우 최종 매크로블록 모드가 SKIP으로 결정되는 확률을 나타내며, '에러율'은 SKIP 모드 조건을 만족하지만, 최종 매크로블록 모드가 SKIP 모드로 결정되지 않는 확률을 의미한다. 위 실험결과로부터 SKIP 모드 조건이 만족되면, SKIP 모드가 최적 매크로블록 모드로 결정될 확률이 약 0.7 로 매우 높다는 것을 알 수 있다. 반면, SKIP 모드 조건이 만족되더라도, 최종 매크로블록 모드로 결정되지 않을 확률은 0.05 정도에 불과하다. 그러므로 표 1의 실험을 바탕으로 다음과 같은 방법을 제안한다.

초기에 SKIP 모드 조건을 확인하기 위하여, 가장 먼저 MB16x16의 움직임 벡터, 참조 영상을 결정하여, 비

트율-왜곡치 계산을 수행한다. 이후 SKIP 모드 조건을 검사해서 조건을 만족할 경우, 최종 매크로블록 모드를 SKIP 모드로 결정한다. 조건을 만족하지 않으면, 기존 방법대로 모든 매크로블록 모드의 비트율-왜곡치를 계산해서 최종 매크로블록 모드를 결정하도록 한다.

#### IV. 선택적 인트라 모드 생략법

P 슬라이스에서 인트라로 부호화 되는 매크로블록의 발생 빈도는 매우 낮다. 이는 많은 경우 프레임간의 시간 상관도가 프레임내의 공간 상관도보다 더 높아서, 인트라 매크로블록 모드의 부호화 효율이 더 높기 때문이다. H.264 매크로블록 모드의 통계적 데이터<sup>[9]</sup>에 따르면, P 슬라이스에서 인트라 매크로블록의 개수는 3%로 그 개수는 매우 미비하다. 하지만, P 슬라이스에서 인트라 매크로블록이 발생하지 않도록 강제로 제한할 경우, 평균 0.18 dB BDPSNR(Bjonteggard Delta PSNR) 감소와, 3.5% BDBR(Bjonteggard Delta BitRate)<sup>[11]</sup> 증가가 발생하여, 이에 따른 부호화 성능의 감소 정도는 매우 크다. 여기서 BDPSNR, BDBR은 JVT에서 채택된 객관적 부호화 효율 측정치이다<sup>[11]</sup>. 인트라 매크로블록은 전체적인 부호화 효율에 매우 큰 공헌을 하고 있으므로, 그 발생하는 빈도가 낮다고해서 무조건 생략하는 것은 부호화 효율의 큰 감소를 초래한다.

그림 3에서 볼 수 있듯이 최적 인트라 모드를 결정하기 위한 비트율-왜곡치 계산은 인트라 매크로블록 모드의 총 비트율-왜곡치 계산 수 보다 5배 많다. 또한, H.264는 높은 부호화 효율을 얻기 위해, P 슬라이스에서도 인트라 매크로블록 모드의 비트율-왜곡치를 계산한다. 즉, P 슬라이스의 매 매크로블록마다 768번의 비트율-왜곡치가 계산되는 것이다. 따라서 인트라 모드의 부호화 과정을 간략화하거나 생략할 수 있다면, H.264의 부호화 복잡도를 크게 낮출 수 있을 것이다.

매크로블록 모드 결정시, 프레임간 시간 상관도가 높다면 인트라 모드가 결정될 것이고, 프레임내 공간 상관도가 높다면 인트라 모드로 결정될 것이다. 따라서 P 슬라이스에서 인트라 모드 부호화 이전에 프레임간 시간 상관도와 프레임내 공간 상관도를 비교해서, 프레임간 시간 상관도가 높다면, 인트라 모드로 결정될 확률이 높으므로 인트라 모드에 대한 부호화 과정을 생략할 수 있을 것이다. 본 논문에서 프레임간 시간 상관도는 최적 인트라 모드의 평균 비트율 AR(Average Rate)로 측정하였고, 프레임내 공간 상관도는 평균 블록 경계 에러

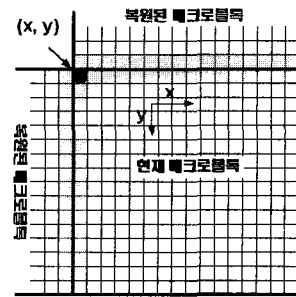


그림 4. ABE(Average Boundary Error) 계산  
Fig. 4. Computing average boundary error.

ABE(Average Boundary Error)로 측정 하였다. 인트라 모드는 이미 부호화된 주변 블록을 이용하여 예측 부호화하기 때문에, 평균 블록 경계 에러는 이전에 부호화된 주변 블록과 현재 블록의 경계에 있는 화소 값들의 평균 블록 경계 에러 ABE를 계산한다. 최종적으로 ABE를 최적 인트라 모드의 평균 비트율 AR과 비교하여 인트라 모드의 부호화 여부를 결정한다. 이에 대한 구체적인 방법은 그림 4와 같다.

ABE 계산에 사용되는 화소들은 그림 4에서와 같이, 이미 부호화 되어 복원된 좌측, 상위 블록의 블록 경계에 있는 화소 열을 이용한다. ABE 계산은 휘도 신호뿐만 아니라, 색차 신호도 포함한다. ABE 계산식은 다음과 같다.

$$\begin{aligned}
 SBE(\text{Sum of Boundary Error}) &= \\
 &\sum_{i=0}^{15} Y_{Orig}(x, y+i) - Y_{Rec}(x-1, y+i) \\
 &+ \sum_{i=0}^{15} Y_{Orig}(x+i, y) - Y_{Rec}(x+i, y-1) \\
 &+ \sum_{i=0}^7 Cb_{Orig}(cx, cy+i) - Cb_{Rec}(cx-1, cy+i) \\
 &+ \sum_{i=0}^7 Cb_{Orig}(cx+i, cy) - Cb_{Rec}(cx+i, cy-1) \\
 &+ \sum_{i=0}^7 Cr_{Orig}(cx, cy+i) - Cr_{Rec}(cx-1, cy+i) \\
 &+ \sum_{i=0}^7 Cr_{Orig}(cx+i, cy) - Cr_{Rec}(cx+i, cy-1) \\
 \therefore ABE &= \frac{1}{64} SBE \\
 &, \text{ where } 64 = (16+16)_{Luma} + (8+8)_{Chroma} \times 2
 \end{aligned}
 \tag{2}$$

$-Y_{Orig}, Cb_{Orig}, Cr_{Orig}$ : 원 영상의 휘도 신호 및 색차 신호  
 $-Y_{Rec}, Cb_{Rec}, Cr_{Rec}$ : 복원 영상의 휘도 신호 및 색차 신호  
 $-(x, y), (cx, cy)$ : 부호화될 매크로블록의 휘도 및 색차 신호의 시작점 (왼쪽 상위 지점)

그리고 최적 인터 모드의 평균 비트율 AR(Average Rate)은 다음과 같이 정의 된다.

$$AR = \frac{1}{384} (\#bits\ of\ texture\ data) \times \lambda \quad (3)$$

$$where\ \lambda = 0.85 \times 2^{Qp/3},\ 384 = (16 \times 26)_{Luma} + (8 \times 8)_{Chroma} \times 2$$

평균 비트율 계산식은 최적 인터 모드의 헤더 정보 비트량을 제외한 순수한 휘도 신호와 색차 신호의 잉여 데이터 비트량만을 고려한다.

그리고 비트량에 Lagrangian 계수인  $\lambda$ 를 곱하고, 매크로블록의 전체 화소수인 384로 나누어 평균 비트율인 AR을 얻게 된다. 최종적으로 최적 인터 모드의 평균 비트율 AR이 ABE보다 작다면, 인트라 모드를 생략한다. 그러나, ABE 계산은 이전에 복원된 주변 블록을 이용해야 하기 때문에, 영상 경계나 슬라이스 경계에서는 이 방법을 적용할 수 없는 상황이 발생한다. 그래서 주변 블록의 존재 여부에 따라 인트라 모드 생략법을 다음과 같이 세분화하였다.

(1) 좌측/상위 블록 사용 불가:

AR=0이면, 인트라 모드 생략

(2) 좌측블록 사용 불가:

$$\begin{aligned} SBE(\text{Sum of Boundary Error}) = & \sum_{i=0}^{15} Y_{Orig}(x+i, y) - Y_{Rec}(x+i, y-1) \\ & + \sum_{i=0}^7 Cb_{Orig}(cx+i, cy) - Cb_{Rec}(cx+i, cy-1) \\ & + \sum_{i=0}^7 Cr_{Orig}(cx+i, cy) - Cr_{Rec}(cx+i, cy-1) \end{aligned} \quad (4)$$

$$\therefore ABE = \frac{1}{32} SBE$$

그리고 AR<ABE 이면, 인트라 모드 생략

(3) 상위블록 사용 불가:

$$\begin{aligned} SBE(\text{Sum of Boundary Error}) = & \sum_{i=0}^{15} Y_{Orig}(x, y+i) - Y_{Rec}(x-1, y+i) \\ & + \sum_{i=0}^7 Cb_{Orig}(cx, cy+i) - Cb_{Rec}(cx-1, cy+i) \\ & + \sum_{i=0}^7 Cr_{Orig}(cx, cy+i) - Cr_{Rec}(cx-1, cy+i) \end{aligned} \quad (5)$$

$$\therefore ABE = \frac{1}{32} SBE$$

그리고 AR<ABE 이면, 인트라 모드 생략

(4) 좌측/상위 블록이 모두 사용 가능:

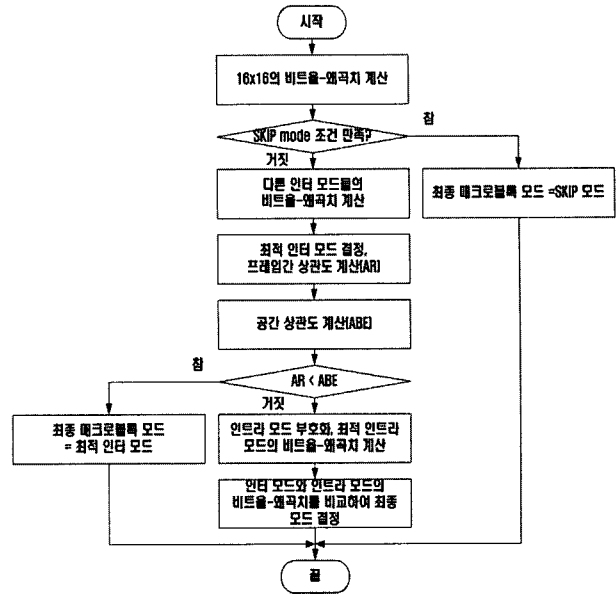


그림 5. 제안된 방법의 흐름도  
Fig. 5. Flowchart of the proposed method.

그림 4와 같이 SBE를 구하고, 이를 64로 나누어 ABE를 계산한다. 그리고 AR<ABE 이면, 인트라 모드 생략.

그림 5는 본 논문에서 제안한 '조기 SKIP 모드 결정법'과 '선택적 인트라 모드 생략법'을 단계별로 도식화한 흐름도이다. SKIP 모드 조건을 검사하기 위해서, 16x16 블록의 비트율-왜곡치를 계산한다. 이때 비트율-왜곡치 계산에 앞서, 16x16 블록의 움직임 벡터, 참조 영상 결정이 선행되어야 한다. SKIP 모드 조건이 만족되면, 최종 매크로블록 모드를 SKIP 모드로 결정하고, 남아 있는 인터 모드 및 인트라 모드의 부호화 과정을 생략한다. 그리고 SKIP 모드가 만족되지 않을 경우 남아 있는 인터 모드의 움직임 벡터와 참조 영상을 탐색해서 각 모드별 비트율-왜곡치를 계산한다. 최소 비트율-왜곡치를 갖는 인터 모드를 최적 인터 모드로 결정한다. 그리고 최적 인터 모드의 잉여 데이터에 대한 비트량을 이용해서 프레임간 상관도를 나타내는 평균 비트율, AR을 계산한다. 다음으로 공간 상관도를 나타내는 평균 블록 경계 에러, ABE를 계산한다. AR과 ABE를 비교해서, AR이 ABE보다 작다면, 인트라 모드의 부호화 과정을 생략하고 최종 매크로블록 모드를 최적 인터 모드로 결정한다. 만약 AR이 ABE보다 크거나 같다면, 인트라 모드의 부호화 과정을 수행한 후 인트라 모드들의 비트율-왜곡치를 계산한다. 인터 모드의 경우와 마찬가지로 가장 작은 비트율-왜곡치를 갖는 인트라 모

드를 최적 인트라 모드로 결정한다. 마지막 단계에서 최적 인터 모드와 최적 인트라 모드의 비트율-왜곡치를 비교해서 최종 매크로블록 모드를 결정한다.

### V. 실험 방법 및 결과

실험 조건은 H.264 표준화 그룹에서 권고하고 있는 공통 실험 조건<sup>[10]</sup>을 따라 표 2와 같이 설정하였다. 실험에 사용된 부호화기는 JM6.1d이며, 실험은 B 슬라이스 CABAC을 사용하지 않는 H.264의 베이스라인 프로파일에 맞추어 수행되었다.

제안된 방법의 비트율-왜곡치 계산 수 감소를 측정하기 위하여, S를 다음과 같이 정의하였다.

$$S = \frac{\#RDcost[reference] - \#RDcost[proposed]}{\#RDcost[reference]} \times 100 [\%] \quad (6)$$

#RDcost[reference]는 기존 H.264의 참조 모델인 JM6.1d에 의한 비트율-왜곡치 계산 수를 나타내고, #RDcost[proposed]는 제안된 방법의 비트율-왜곡치 계산 수를 나타낸다. 그리고 부호화 시간의 변화량은 다음과 같이 계산하였으며, 그 단위는 백분율로 나타내었다.

$$\Delta T = \frac{Time[reference] - Time[proposed]}{Time[reference]} \times 100 [\%] \quad (7)$$

부호화 성능 평가를 위한 지표로 BDBR(Bjonteggard Delta BitRate)과 BDPSNR (Bjonteggard Delta PSNR)<sup>[11]</sup>을 사용하였는데, 이들은 각각 비교하고자 하는 두 방법 간의 비트율과 PSNR 차이의 평균을 의미한다. 그리고 BDBR의 (+) 부호와 BDPSNR의 (-) 부호는 부호화 성능의 손실을 나타낸다.

표 3의 결과로부터 제안된 SKIP 모드 조건을 이용한 고속 모드 결정법은 BDBR, BDPSNR의 감소 없이 비트율-왜곡치 계산 수를 평균 31.8% 감소시켰다. 특히 BDBR은 오히려 0.003% 감소되었는데, 이는 SKIP 모드의 비트량이 0인 것에서 기인한다. 즉, 기존 모드 결정법에서는 SKIP 모드의 왜곡치를 모드 결정에 반영했으나, 조기 SKIP 모드 결정법은 단순히 비트량이 0이 되는 것만을 검사했기 때문이다.

표4에서 인트라 모드 생략법을 사용했을 경우 비트율-왜곡치 계산 수 감소량은 SKIP 모드 조건을 이용한 모드 결정법보다 많다. 인트라 모드의 비트율-왜곡치 계산 수는 인터 모드보다 약 5배 많기 때문에, 인트라 모드 생략에 의한 계산 수 감소가 더 큰 것이다. 제안된

표 2. 실험 조건

Table 2. Simulation condition.

영 상	news (qcif)	container (qcif)	foreman (qcif)	silent (qcif)	paris (cif)	mobile (cif)	tempe (cif)
주사율(Hz)	10	10	10	15	15	30	30
영상수	100	100	100	150	150	300	250
Qp	28, 32, 36, 40						
부호화 기법	비트율-최적화 기법, 하다마드 변환부호화, IPPP 구조, CAVLC(Context-based adaptive variable length coding), #REF=4, SR=+/-16						
부호화기	JM 6.1d 부호화기						

표 3. 조기 SKIP 모드 결정법

Table 3. SKIP mode decision at early stage.

영상	BDBR[%]	BDPSNR[dB]	S [%]	ΔT [%]
container	-0.302	0.017	47.92	23.40
foreman	-0.340	0.018	15.09	3.49
news	0.265	-0.016	49.68	25.49
mobile	-0.202	0.009	10.00	0.74
silent	0.234	-0.013	44.45	21.99
paris	0.090	-0.005	41.34	20.41
tempe	0.231	-0.010	14.14	3.04
평균	-0.003	0.000	31.80	14.08

표 4. 인트라 모드 생략법

Table 4. Selective intra mode decision.

영상	BDBR[%]	BDPSNR[dB]	S [%]	ΔT [%]
container	1.745	-0.087	76.19	22.45
foreman	1.090	-0.061	58.97	16.02
news	0.158	-0.010	65.79	20.16
mobile	0.219	-0.010	59.96	22.34
silent	0.333	-0.018	69.32	20.59
paris	0.711	-0.037	70.09	23.54
tempe	0.663	-0.028	59.78	19.19
평균	0.703	-0.036	65.73	20.61

표 5. 두 방법을 결합

Table 5. Combined two methods.

영상	BDBR[%]	BDPSNR[dB]	S [%]	ΔT [%]
container	1.309	-0.067	84.30	36.48
foreman	1.163	-0.065	62.04	20.24
news	0.004	0.000	77.55	36.11
mobile	0.065	-0.003	61.61	24.11
silent	0.498	-0.028	77.02	33.75
paris	0.923	-0.048	76.74	34.69
tempe	0.621	-0.026	63.22	22.83
평균	0.655	-0.034	71.78	29.74

방법에 의한 부호화 효율의 변화는 BDPSNR이 평균 0.036dB 감소하였고, 비트량은 0.7%로 소폭 증가하였다.

이 두 가지 방법을 동시에 적용했을 때, BDPSNR은 0.034dB 감소했으며, BDBR은 0.66% 증가했다. 이때, foreman, mobile, tempete 영상에 대한 비트율-왜곡치 감소는 다른 영상들에 비해 약 10% 정도 적다. 이는 이들 영상의 특징이 SKIP 모드의 개수가 적으며, 인트라 모드로 부호화 되는 매크로블록들이 많기 때문이다. 즉, 최적 인트라 모드의 평균 비트율인 AR이 공간상의 평균 블록 경계에러인 ABE보다 큰 경우가 많음을 나타낸다.

## VI. 결 론

본 논문은 H.264의 가변 블록 움직임 보상을 위한 고속 모드 결정법을 제안하였다. 고속 모드 결정법은 두 단계로 이루어져 있는데, 부호화 초기 단계에 SKIP 모드 조건을 이용하여 전체 모드 결정 과정을 간략화 하는 방법과 인트라 모드를 생략하는 방법이다.

먼저, SKIP 모드 결정법은, 모드 결정 초기 단계에 16x16 블록에 대한 비트율-왜곡치를 계산하여, SKIP 모드 조건이 만족하는지 확인하고 이를 만족한다면 SKIP 모드를 최종 매크로블록 모드로 결정하는 것이다. H.264 부호화기에서 대부분의 복잡도는 움직임 보상에 기인하는데, 조기에 SKIP 모드 조건이 만족되면 각 가변 블록들의 움직임 벡터 추정과 비트율-왜곡치 계산을 생략할 수 있어 복잡도를 감소시킬 수 있다. SKIP 모드 결정법은 평활하고 움직임이 적은 영상에서 효율적임을 알 수 있다.

인트라 모드 생략법은 최적 인트라 모드의 평균 비트량과 공간상의 평균 블록 경계 에러를 비교하여 인트라 모드를 생략하는 것이다. H.264는 높은 부호화 효율을 위하여 P 슬라이스에서 인트라 모드로 부호화 되는 것이 가능하다. 하지만, 인트라 모드의 발생 빈도는 매우 낮다. 따라서 모든 P 슬라이스의 매크로블록마다 인트라 모드의 비트율-왜곡치를 계산하는 것이 부호화기의 복잡도를 높이는 한 요인이다. 더욱이 인트라 모드는 인트라 모드 보다 많은 비트율-왜곡치 계산 수를 필요로 한다.

두 가지 방법을 동시에 적용했을 때 얻게 되는 비트율-왜곡치 계산 수 감소는 평균 72%이며, BDPSNR의 감소는 약 0.03dB로 매우 미비하다. 즉 제안된 고속 모드 결정법은 큰 부호화 효율의 감소 없이, H.264의 복잡도를 크게 낮추었으며, 부호화 시간도 평균 30% 이상

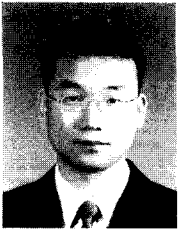
감소시켰다.

## 참 고 문 헌

- [1] P. Topiwala, G. Sullivan, A. Joch, and F. Kossentini, "Overview and Performance Evaluation of the Draft ITU-T H.26L Video Coding Standard," Proc. SPIE, Appl. Dig. Im. Proc, Aug 2001.
- [2] ITU-T SG16 Q6, "Performance Evaluation of H.26L, TML 8 vs. H.263++ and MPEG-4," Doc. VCEG-N18, Sep 2001.
- [3] ISO/IEC JTC1/SC29WG11, "A Computational Complexity Comparison of MPEG4 and JVT Codecs," Doc. M8696, Jul 2002.
- [4] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)," Doc. JVT-G050r1, Mar 2003.
- [5] W. Choi and B. Jeon, "Selective fast motion estimation with variable motion block sizes," International Workshop on Advanced Image Technology (IWAIT'03), pages.101-106, Nagasaki, 21-22, Jan 2003.
- [6] M. Zhou, "Evaluation and Simplification of H.26L Baseline Coding Tools," ITU-T Q.6/16, Doc. JVT-B030, Jan 2002.
- [7] <http://bs.hhi.de/~suehring/tml/download/jm61d.zip>.
- [8] 이제윤, 최웅일, 전병우, 석민수, "H.264의 가변 블록 움직임 보상을 위한 고속 움직임 벡터 탐색 및 모드 결정법," 전자공학회 논문지 제 40호 SP편 제 4호, 49-59쪽, 2003년 7월.
- [9] J. Lee and B. Jeon, "Pruned Mode Decision based on Variable Block Sizes Motion Compensation for H.264," Lecture Notes in Computer Science(LNCS), vol. 2899, pages. 410~418, Nov 2003.
- [10] G. Sullivan and G. Bjontegaard, "Recommended simulation common conditions for H.26L coding efficiency experiments on low-resolution progressive-scan source material," ITU-T Q.6/16, Doc. VCEG-N81, Sep 2001.
- [11] G. Bjontegaard, "Calculation of Average PSNR Differences between RD-curves," Doc. VCEG-M33, Apr 2001.



저 자 소 개



이 제 운(정회원)  
 2002년 성균관대학교 전자공학과  
 학사  
 2004년 성균관대학교 정보통신  
 공학부 석사  
 2004년 2월~현재 LG 전자 Digital  
 Media 연구소 근무 중

<주관심분야: 영상압축, 멀티미디어 응용>



전 병 우(정회원)  
 1985년 서울대학교 전자공학과  
 학사  
 1987년 서울대학교 전자공학과  
 석사  
 1992년 Purdue Univ, School of  
 Elec. 박사

1993년~1997년 8월 삼성전자 신호처리연구소  
수석연구원

1997년 9월~현재 성균관대학교 정보통신공학부  
부교수

<주관심분야: 멀티미디어, 영상압축, 영상인식, 신  
호처리>

