

논문 2007-44SP-4-2

H.264/AVC 고속 매크로블록 모드 결정 알고리즘

(H.264/AVC Fast Macroblock Mode Decision Algorithm)

김지웅*, 김용관*

(Ji Woong Kim and Yong Kwan Kim)

요약

H.264/AVC 부호화 표준은 부호화 효율을 향상시키기 위하여 기존의 부호화 표준들과는 다른 새로운 부호화 기법들을 사용한다. 그러나 새로이 채택된 여러 기법들로 인해 H.264/AVC 표준 부호기 및 복호기의 복잡도는 극단적으로 증가하게 되었다. 특히 울-왜곡 최적화 기법에 의한 H.264/AVC의 인터/인트라 모드 결정 방법은 부호기의 복잡도를 증가시키는 가장 큰 원인 중 하나이다. 본 논문에서는 매크로블록 모드 결정 과정의 복잡도 감소에 주안점을 두며, 이에 대한 고속 매크로블록 모드 결정 알고리즘을 제안한다. 제안하는 방식에서는 간단한 구조의 4x4 정방형 필터와 블록 간 공간적 상관도를 이용하여 Intra4x4 모드 결정에 따른 울-왜곡 계산량을 줄이며, Inter8x8 모드 내 서브 매크로블록의 최적모드를 통해 현재 매크로블록에서 인트라 모드 결정 과정을 선택적으로 생략하도록 하는 알고리즘을 제안하였다. 또한, 선택 가능한 매크로블록 모드 중 상대적으로 복잡도와 발생 비트율이 낮은 SKIP, Inter16x16, Intra16x16 모드에 대한 발생 빈도수를 높여 발생 비트율을 낮추도록 하였다. 제안한 알고리즘을 적용한 실험 결과 최대 83%의 부호화 시간을 단축시킬 수 있었으며, 미미한 PSNR의 변화량에 비해 발생 비트율을 평균 8%~10% 감소시킴으로써 전체 부호화 효율을 향상시킬 수 있었다.

Abstract

For the improvement of coding efficiency, the H.264/AVC video coding standard employs new coding tools compared with existing coding standards. However, due to these new coding tools, the complexity of H.264/AVC standard encoder is greatly increased. Specifically, the inter/intra mode decision method using RDO(rate-distortion optimization) technique is one of the most complex parts in H.264/AVC. In this paper, we focus on the complexity reduction in macroblock mode decision. In the proposed method, we reduce the complexity of the Intra4x4 mode decision process using 4x4 simple square filters, and using spatial block correlation method. Additionally, exploiting the best mode of sub_macroblock in Inter8x8 mode, we proposed an algorithm to eliminate some intra modes in current macroblock mode decision process. In addition, we employed a method to raise the probability to select SKIP, Inter16x16, and Intra16x16 modes which usually show low complexity and low bitrate compared with other modes. From the simulation results, the proposed algorithm reduce the encoding time by maximum 83% of total, and reduce the bitrate of the overall sequences by 8~10% on the average compared with existing coding methods.

Keywords: H.264/AVC, rate-distortion optimization, macroblock mode, fast mode decision.

I. 서론

2003년 국제 동영상 부호화 표준으로 제정된 H.264/AVC 표준은 이미 표준화 개발 초기 무엇보다도 최대의 압축 효율을 최우선 목표로 하는 개발 방침을 내세웠다^[1,2]. 이 과정에서 보다 더 부호화 효율을 높이기

기 위하여 복잡도(Complexity)를 제한하지 않고 부호화 하고자 하는 검토가 진행되었고, 높은 복잡도로 인해 지금까지 채택되지 않았던 기술들을 적극 도입하게 되었다. 그 결과 H.264/AVC는 MPEG-2보다 약 40%~50%, MPEG-4 ASP(advanced simple profile)에 비해서는 약 30%~40% 성능이 향상된 것으로 보고되어 있다^[3]. 특히 최근에 울-왜곡 최적화(RDO : rate-distortion optimization) 방식이 제안되었는데, 이로 인해 H.26L의 압축 효율을 크게 향상시킬 수 있었다^[2,4].

반면, H.264/AVC에서는 복잡도의 증가를 고려하지

* 정회원, 호서대학교 정보통신공학과
(Dept. of Information and Telecommunication,
Hoseo University)
접수일자: 2006년12월22일, 수정완료일: 2007년6월8일

않고 압축 효율 향상을 최우선 목표로 삼았기 때문에 부호기 및 복호기의 계산량은 극단적으로 증가하게 되었다. 이는 결국 제한된 계산능력을 가지는 모바일, PDA 기기 및 실시간성을 요구하는 어플리케이션에 있어 H.264/AVC를 사용하기 어렵게 만드는 원인이 되었으며, 표준안이 제정된 이래 H.264/AVC의 복잡도를 감소시키기 위한 연구가 다각적으로 진행되어 왔다. 이러한 연구는 크게 고속 움직임 예측/보상 알고리즘과 고속 인터/인트라 모드 결정 알고리즘에 관한 연구 분야로 나눌 수 있다.

H.264/AVC는 기존의 동영상 부호화 표준에 비해 다양한 예측 모드를 가지고 있으며 최적의 인터/인트라 예측 모드를 선택하기 위해 다양한 인터/인트라 모드의 조합에 대해 울-왜곡 비용(RDO_cost)을 계산해야 한다. 울-왜곡 비용함수를 위해 사용되는 발생 비트량을 계산하기 위해서는 각 모드에 대한 DCT/Hadamard 변환 및 양자화/역양자화, 역 DCT/Hadamard 변환과정, 그리고 엔트로피 부호화 과정을 거치게 되므로 부호기의 복잡도를 크게 증가시키게 된다^[4]. 따라서 본 논문에서는 이러한 울-왜곡 계산에 따른 복잡도를 감소시키기 위하여 울-왜곡 계산을 필요로 하는 후보 모드의 수를 효율적으로 줄일 수 있는 알고리즘을 제안한다.

제안하는 알고리즘에서는 인터/인트라 예측 모드 결정에 있어 상대적으로 낮은 복잡도와 비트율을 가지는 SKIP, Inter16x16, Intra16x16 인터/인트라모드에 대한 가중치를 높인다. 그리고 P8x8 인터모드 내 서브 매크로블록(sub macro-block) 모드의 최적 모드 분포를 이용하여 Intra4x4, Intra16x16 모드에 대한 울-왜곡 계산 여부를 결정한다. 또한 Intra4x4 모드 결정의 계산량을 최소화하기 위하여 주변 블록의 모드와 현재 블록의 화소 분포를 이용하였다. II장에서는 H.264/AVC의 최적 모드 결정 방법에 대해 기술하고, III장에서는 제안하는 매크로블록 모드 결정법에 대해 기술한다. IV장에서 기존 선행 연구와의 비교 실험 결과를 통해 제안하는 방식의 성능을 살펴보고 V장에서 결론을 맺는다.

II. 울-왜곡 최적화에 기반 하는 H.264/AVC의 모드 결정 방법

H.264/AVC의 예측 부호화 기법은 화면 내 예측(intra prediction)과 화면 간 예측(inter prediction)으로 구성된다. 화면 내 예측은 입력 영상의 공간적 중복성을 제거하는 부호화 기법이며, 화면 간 예측은 움직임

추정 및 보상을 통해 입력 영상의 시간적 중복성을 제거하는 부호화 기법이다^[3]. H.264/AVC에서는 기존 동영상 부호화 표준에 비해 다양한 블록 크기로 구성되는 예측 모드에 대하여 최적의 모드를 결정하기 위해 왜곡도(SNR)와 함께 발생 비트율을 고려하는 울-왜곡 최적화 기법을 도입하였다. 울-왜곡 최적화 기법은 라그랑지안(Lagrangian) 함수에 기반을 둔 비용함수를 사용하며 식 (1)의 비용함수가 최소값을 가지는 모드를 최적의 모드로 결정하게 된다^[3].

$$J_{\text{mode}}(s, r, M | \lambda_{\text{mode}}) = SSD(s, r, M) + \lambda_{\text{mode}} \times R(s, r, M) \quad (1)$$

s, r 은 각각 원 매크로블록과 복호된 매크로블록의 화소값이며, M 은 매크로블록 모드, 예측 방향 모드 또는 서브 매크로블록 모드이다. $R(s, r, M)$ 은 M 에 해당하는 모드로 부호화했을 때 실제 발생하는 비트수이고, $SSD(s, r, M)$ (sum of squared difference)는 원 영상과 복호된 영상과의 차이에 대한 제곱의 합으로 정의된다. 또한 λ 는 라그랑지안(Lagrangian) 계수로 울-왜곡선에서의 접점의 기울기이다.

그림 1에서 볼 수 있듯이 H.264/AVC에서는 블록의 특성에 따라 휘도성분에 대해 Intra4x4 또는 Intra16x16 단위로 인트라 예측 부호화를 수행한다. 울-왜곡 최적화에 따른 최적의 예측 방향을 결정한 후 예측 신호와 현재신호와의 차이, 즉 잔여신호만을 변환/양자화 하고, 엔트로피 부호화 과정을 거쳐 최종 비트스트림을 구성하게 된다. 최적의 예측 모드를 결정하기 위해 각 블록 단위별로 보다 복잡한 영역으로 구성된 Intra4x4 휘도블록의 경우 9가지의 예측 방향에 대해 울-왜곡 최적화를 적용한다. 또한 비교적 평탄한 영역으로 구성된 Intra16x16 휘도블록에 대해서는 4가지의 예측 방향에 대해 울-왜곡 최적화를 적용한다^[1,2]. 이 때 예측을 수행하는 참조블록은 복호(decoding)된 블록의 경계화소를 이용하며, 각 모드별로 울-왜곡 비용함수가 최소가 되는 모드를 최적모드로 선택하게 된다.

인터 예측 부호화의 경우 총 5개의 모드를 지원하며 Inter8x8 모드의 경우 내부 4개의 서브 매크로블록(sub macroblock)모드를 지원하여 보다 세밀한 움직임까지 표현할 수 있도록 하였다.

H.264/AVC에서는 다양한 예측 모드에 대하여 울-왜곡 최적화를 이용하여 최적 모드를 결정함으로써 부호화 효율을 높일 수 있었지만, 하나의 매크로블록에 대하여 최적의 예측 모드를 결정하기 위한 울-왜곡 비용 계산 횟수는 식 (2)와 식 (3)에 의해 총 720회가 되어

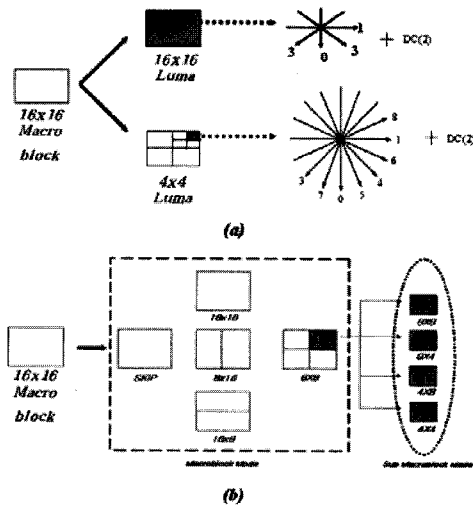


그림 1. H.264/AVC의 블록 크기별 예측 모드
(a) 인트라 예측 모드 (b) 인터 예측 모드

Fig. 1. Prediction mode of H.264/AVC
(a) Intra prediction mode (b) Inter prediction mode.

부호기의 복잡도를 크게 증가시키게 되었다.

$$RD_{Intra} = MB_{by8} \times (MB_{by4} \times 16 + M_{6by16}) = 4 \times ((9 \times 16) + 4) = 592 \quad (2)$$

$$RD_{Inter} = PSKIP + P_{16by16} + P_{16by8} + P_{8by16} + (P_{8by8} \times 4) = 16 + 16 + 16 + 16 + 64 = 128 \quad (3)$$

M과 P는 각각 인트라/인터 모드별 윌-왜곡 계산횟수이다.

III. 제안하는 매크로블록 모드 결정법

1. Intra4x4 모드 결정과정에 따른 복잡도 감소

H.264/AVC의 인트라 모드에서 Intra4x4 모드의 경우 총 9개의 예측 방향에 따라 최적 모드를 결정하고 가장 세분화된 블록에 대한 예측 부호화를 수행하기 때문에 상대적으로 높은 복잡도를 가진다. 그러나 그림 2와 같이 일반적으로 영상에서 주변 배경과 같은 영상신호의 변화가 적은 영역에서는 현재 블록과 주변 블록의 화소 정보가 유사하다. 또한 Intra4x4 인트라 모드에서 현재 블록의 최적 예측 방향이 주변 블록의 예측 방향과 같다면 현재 블록과 주변 블록의 화소 정보 또한 비슷할 확률이 높다.

블록 간 유사성이 높을 경우 현재 Intra4x4 블록의 최적 모드를 결정하기 위하여 9가지 후보모드 각각에 대한 윌-왜곡 비용을 계산하기보다는 이미 결정된 주변 모드 정보만을 이용하는 것이 복잡도 측면에서 유리하다. 본 논

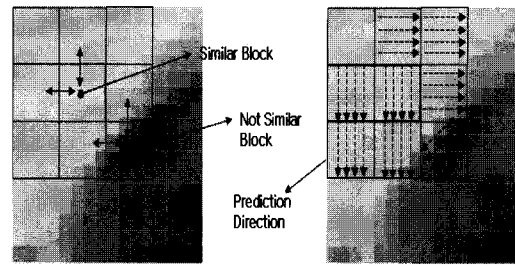


그림 2. Intra4x4 모드에서 화소 정보가 유사한 블록 간 최적 예측 방향

Fig. 2. In Intra4x4 mode, best prediction direction by similar block.

문에서 제안하는 방식은 우선 현재 블록의 화소 분포와 주변 블록과의 화소 분포의 차이에 따른 블록 간 유사성을 판별하는 전처리 과정을 통해 후보모드의 수를 줄이는 것이다.

현재 블록 내 화소 분포도를 C_{dis} 라 하고 현재 블록과 좌측 블록과의 화소 분포의 차이를 L_{cor} , 현재 블록과 상측 블록과의 화소 분포의 차이를 U_{cor} 이라고 할 때 $C_{dis}, L_{cor}, U_{cor}$ 은 다음과 같이 계산된다.

$$C_{dis} = \sum_{x=0}^{T-15} |C_{avr} - CP_x| \quad (4)$$

$$L_{cor} = \sum_{x=0}^{T-15} |C_{avr} - LP_x| \quad (5)$$

$$U_{cor} = \sum_{x=0}^{T-15} |C_{avr} - UP_x| \quad (6)$$

이 때, C_{avr} 은 현재 블록 내 화소 평균값 이고, CP_x, LP_x, UP_x 는 각각 현재, 좌측, 상측 블록의 x 인덱스에 해당하는 화소값이다. 평균값을 이용하여 블록 간 화소 분포 차이를 계산하는 이유는 Intra4x4 블록 내부의 화소 각각에 대하여 주변 블록과의 명도 차이를 계산한다면 화소 하나의 변화량에 블록 간 유사성이 결정될 확률이 높기 때문이다. 그림 3은 블록 간 화소 분포를 계산하기 위한 현재 블록과 주변 블록(좌측, 상측)의 화소 분포를 나타낸다. $C_{dis}, L_{cor}, U_{cor}$ 을 계산한 후 현재 블록과 주변 블록과의 화소 분포에 따른 블록 간 유사성을 판별하기 위한 L_{dif} 과 U_{dif} 은 식 (7)과 식 (8)에 의해 계산된다.

$$L_{dif} = |C_{dis} - L_{cor}| \quad (7)$$

$$U_{dif} = |C_{dis} - U_{cor}| \quad (8)$$

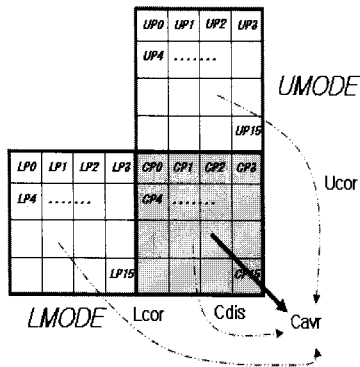


그림 3. 현재 블록과 주변 블록의 화소 분포 관계
Fig. 3. Pixel distribution of current and neighboring blocks.

제안하는 Intra4x4 모드 내 복잡도 감소 알고리즘은 다음과 같다.

첫째, 만약 $L_{dif} < T$ AND $U_{dif} < T$ 인 경우,

즉, L_{dif} 와 U_{dif} 가 모두 임계값 T 보다 작을 경우 현재 블록의 화소 분포와 주변 블록과의 화소 분포는 상관도가 매우 높다고 판단되므로 현재 블록의 최적 모드는 이미 결정된 주변 블록의 모드로부터 예측 가능하다. 따라서 이 경우 Intra4x4 블록의 최적 모드를 결정하기 위한 선택 가능한 후보 모드는 그림 3의 LMODE와 UMODE로 압축된다. 또한 LMODE와 UMODE가 동일할 경우 선택 가능한 후보 모드의 수는 하나가 된다. 본 논문에서 임계값 T는 현재 블록과 주변 블록의 평균 명도차로 정의된다.

둘째, 만약 $L_{dif} > T$ OR $U_{dif} > T$ 인 경우,

Intra4x4 블록의 최적 모드는 현재 블록 내 방향성과 보다 밀접한 관계를 가진다. 최근에 소벨 마스크(Sobel Mask)와 에지 히스토그램(Edge Histogram)을 이용한 전처리 과정을 통해 Intra4x4 모드의 선택 가능한 후보 모드의 수를 4개로 줄임으로써 Intra4x4 모드 결정에 따른 복잡도를 감소시키는 알고리즘이 제안되었다^[1]. 그러나 이러한 방식은 블록 내 정확한 방향성을 계산하기 위해 다소 복잡한 전처리 과정을 필요로 하기 때문에 복잡도 감소 효과를 충분히 반영하지 못하는 단점이 있다. 즉, Intra4x4 모드의 8가지 예측 방향 외에 불필요한 방향 성분을 모두 검출하기 때문에 계산 과정이 복잡해진다. 따라서 본 논문에서는 다소 복잡한 연산을 필요로 하는 소벨 마스크를 이용하는 대신에 그림 4의 (a)와 같은 보다 간단한 구조의 4x4 정방형 마스크를 이용하도록 하였다. 또한, 그림 4의 (b)에서 볼 수 있듯이 Intra4x4 모드의 8가지 예측 방향 중 실제 선택 비율

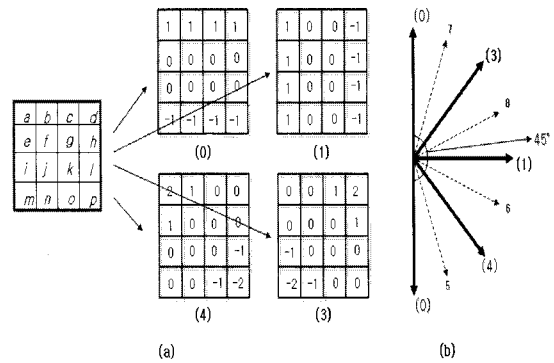


그림 4. (a) 4x4 정방형 마스크 구조 (b) 검출 방향
Fig. 4. (a) 4x4 square mask structure (b) Direction detection.

이 높은 주요 4가지 방향 성분에 대해서만 마스크 연산을 수행하도록 하였다.

최종적으로 Intra4x4 모드의 최적 모드를 결정하기 위한 후보 모드는 마스크 연산 결과가 최소가 되는 모드와 그림 4의 (b)에서 점선으로 표시되는 인접한 두 개의 모드 및 평균치 예측(DC) 모드로서 총 4개의 후보 모드를 가지게 된다.

$$Mode(0) = |a - m| + |b - n| + |c - o| + |d - p| \quad (9)$$

$$Mode(1) = |a - d| + |e - h| + |i - l| + |m - p| \quad (10)$$

$$Mode(3) = |c - i| + 2|d - m| + |h - n| \quad (11)$$

$$Mode(4) = |b - l| + 2|a - p| + |e - o| \quad (12)$$

H.264/AVC의 Intra4x4 블록에서 최적의 예측 모드를 결정하기 위한 율-왜곡 계산 횟수는 식 (2)에서와 같이 총 144회이며, 최근에 제안된 방식^[1]의 경우 $4 \times 16 = 64$ 회가 된다. 반면, 본 논문에서 제안하는 방식의 경우 선택 가능한 후보 모드의 수가 최소 1개에서 통상 4개 이하이므로 Intra4x4 블록의 율-왜곡 계산 횟수를 보다 감소시킬 수 있다.

2. 선택적 인트라 모드 결정법

H.264/AVC 부호기에서 I-picture 는 화면 내 예측(Intra Prediction)만을 수행한다. 반면 P-picture의 경우 인터/인트라 예측을 모두 수행하는데, 실제 P-picture의 경우 최적의 예측 모드가 인트라 모드로 결정될 확률은 4%~5% 정도로 미미한 수준이다. 따라서 H.264/AVC의 복잡도를 고려할 때 P-picture의 경우 인트라 모드 결정 과정을 꼭 필요한 경우에 한해 수행하도록 하는

것이 시간 측면에서 유리하다.

[5]와 [7]등에서 제안하는 방식은 모두 특정 임계값에 기반 하여 인트라 모드 결정 과정을 생략하도록 하였다. [5]에서는 최적 인터 모드의 울-왜곡 비용에 비례상수를 곱한 값과 Intra16x16 모드의 울-왜곡 비용을 비교하여 Intra4x4 모드 결정 과정을 생략하도록 하였으며, [7]에서 제안한 방식 또한 특정 임계값 이상일 경우에 한해 인트라 모드 결정 과정을 수행토록 하였다. 그러나 이러한 방식은 특정 임계값의 변화에 민감하여 최적의 임계값을 설정하는 데 어려움이 있다. 본 논문에서는 특정 임계값을 이용하지 않고 Inter8x8 모드 내 서브 매크로블록(sub macroblock)의 최적 모드 분포를 이용하여 선택적으로 인트라 모드 검색을 수행하는 방식을 제안한다.

일반적으로 영상에서 배경과 같이 움직임이 작고 매크로블록 내 화소 분포가 비교적 균일한 경우 최적의 인트라 모드는 대부분 Intra16x16 모드에서 결정된다. 또한 사물영역과 같이 복잡한 매크로블록에서는 최적의 인트라 모드는 Intra4x4 모드에서 결정될 확률이 높다. 따라서 현재 매크로블록의 특성이 움직임이 작고 화소 분포가 비교적 균일하다고 판단될 경우 Intra4x4 모드에 대한 방향 결정 및 모드 결정 과정을 전부 수행하는 것은 복잡도 측면에서 비합리적이다.

본 논문에서 제안하는 방식은 그림 5와 같은 Inter8x8 모드 내 서브 매크로블록의 최적 모드 분포를 이용하여 다음과 같이 수행된다.

첫째, 만약 $Mode(Block1, Block2, Block3, Block4) = 8 \times 8 Mode$ 인 경우, 즉, Inter8x8 모드 내 서브 매크로블록의 최적 모드가 모두 8x8 블록 크기로 결정된 경우 현재 매크로블록은 8x8 블록 크기 이하의 세분화 과정을 필요로 하지 않는다. 왜냐하면 8x8 블록 크기 이하의 세분화 과

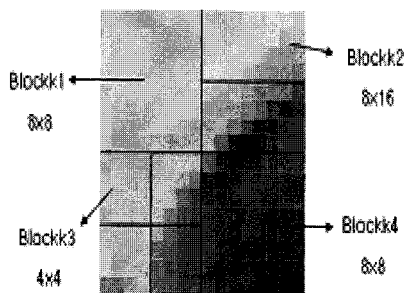


그림 5. Inter8x8 모드 내 서브 매크로블록 모드 분포
Fig. 5. The sub macroblock mode distribution in Inter8x8 mode.

정을 통해 왜곡 오차를 줄이더라도, 세분화된 블록 각각에 대한 비트 정보가 많아지기 때문에 부호화 효율 측면에서 향상된 성능을 기대하기 어렵기 때문이다. 즉, 이 경우 인트라 예측 부호화에서도 마찬가지로 블록 세분화 과정을 통해 울-왜곡 측면에서 효율 향상을 기대하기 어렵다고 판단할 수 있기 때문에 Intra4x4 모드 결정 과정을 생략하도록 한다.

둘째, 만약, $Mode(Block1, Block2, Block3, Block4) = 4 \times 4 Mode$ 인 경우, 즉, Inter8x8 모드 내 서브 매크로블록의 최적 모드 분포가 모두 4x4 블록 크기로 결정된 경우 현재 매크로블록은 H.264/AVC에서 지원하는 가장 작은 블록 크기까지 세분화된다. 이는 결국 현재 매크로블록 특성이 영상에서 사물영역과 같이 복잡하고 움직임이 많은 영역으로 판단되기 때문이다. 따라서 이 경우 영상의 배경과 같은 비교적 평탄한 영역에 대한 인트라 예측 부호화를 수행하는 Intra16x16 모드 결정 과정을 생략한다.

3. 조기 모드 결정을 통한 부호기 복잡도 감소

H.264/AVC에서는 앞서 기술한 바와 같이 식 (1)의 비용함수를 이용하여 최적의 예측 모드를 결정한다. 그러나 인터 모드 결정 과정에서는 최적의 인터 모드를 결정하기에 앞서 움직임 벡터와 참조 프레임을 먼저 결정해야 한다. 특히 H.264/AVC의 인터 모드 중 Inter8x8 모드는 나머지 4가지 모드에 비해 높은 복잡도를 지닌다. 움직임 벡터와 참조 프레임을 결정하기 위한 계산 과정은 Inter16x16, Inter16x8, Inter8x16, Inter8x8 모두 필요로 하지만 Inter8x8 모드의 경우 다른 인터 모드에서는 없는 4개의 서브 매크로블록 단위의 울-왜곡 비용 함수 계산 과정을 포함하기 때문이다^[7].

본 논문에서 제안하는 방식은 조기 SKIP 및 Inter16x16 모드에 대한 가중치를 높임으로써 이하 모드 결정 과정을 생략하는 것이다. 즉 Inter16x16 모드가 최적 모드 조건을 만족할 경우 이하 Inter8x8 모드 결정 과정을 생략토록 하며, SKIP 모드가 최적 모드 조건을 만족할 경우 이하 인터/인트라 모드 결정 과정을 생략하도록 한다. 조기 SKIP 모드 결정의 경우 H.264/AVC 참조 소프트웨어인 JM(Joint Model)9.4 버전 이후부터 이미 도입된 방식이지만^[8], 본 논문에서는 모드 결정 과정 초기에 SKIP 모드와 Inter16x16 모드의 식 (1)에 따른 울-왜곡 비용만을 비교하는 조건을 통해 SKIP 모드의 울-왜곡 비용이 Inter16x16 모드의 울-왜곡 비용보다 작을 경우 현재 매크로블록의 최적 모드를 SKIP

모드로 결정하며 이하 모든 모드 결정 과정을 생략하도록 하였다.

현재 매크로블록이 조기 SKIP 모드 조건을 만족하지 않을 경우 Inter16x16, Inter16x8, Inter8x16 모드에 대한 움직임 벡터와 참조 프레임 결정을 위한 계산과정을 수행한다. 일반적으로 영상에서 매크로블록이 배경이나 움직임이 느린 부분에 속해 있을 경우에는 최적 모드가 SKIP이나 Inter16x16 모드로 결정되는 경우가 많다. 이는 앞서 기술한 바와 같이 큰 블록을 사용하여 매크로블록을 부호화해도 왜곡도가 충분히 작기 때문에 작은 블록으로 부호화할 필요가 없기 때문이다. 따라서 Inter16x16 모드의 움직임 벡터 비용이 Inter16x8, Inter8x16 모드의 움직임 벡터 비용보다 작을 경우 현재 매크로블록의 최적 인터 모드가 Inter8x8 모드로 결정될 경우가 적다고 판단할 수 있다. 따라서 이 경우 본 논문에서는 최적의 인터 모드를 Inter16x16 모드로 결

정하고 Inter16x8, Inter8x16, Inter8x8 모드 결정 과정을 생략하도록 하였다.

그림 6에 본 논문에서 제안하는 H.264/AVC 고속 매크로블록 모드 결정을 위한 전체 순서도를 도시하였다.

IV. 실험 및 결과

1. 실험 환경

제안하는 알고리즘의 성능을 평가하기 위하여 본 논문에서는 H.264/AVC 참조 소프트웨어 JM9.5를 사용하였고, 테스트 방법은 JVT에서 권고하는 시뮬레이션 조건^[9]을 따르도록 하였다. 또한 기존에 제안된 방식과 본 논문에서 제안하는 알고리즘의 비교를 통해 제안하는 알고리즘의 성능을 분석하도록 하였다. 실험 결과에 따른 성능 분석을 위해 사용되는 PSNR의 변화량($\Delta PSNR$), 비트율 변화량(ΔBit) 및 전체 부호화 시간의 변화량($\Delta Time$)은 다음의 식 (13), (14), (15) 의해 결정된다.

$$\Delta PSNR = PSNR_{Algorithm} - PSNR_{original} \quad (13)$$

$$\Delta Bit = \frac{Bitrate_{Algorithm} - Bitrate_{original}}{Bitrate_{original}} \quad (14)$$

$$\Delta Time = \frac{Time_{Algorithm} - Time_{Original}}{Time_{Original}} \quad (15)$$

$\Delta PSNR$, ΔBit , $\Delta Time$ 의 값이 음수일 경우 감소량을 나타내며 양수일 경우 증가량을 나타낸다.

2. 실험 결과 및 고찰

본 논문에서 제안하는 알고리즘 중 III-3의 조기 모드 결정법을 적용하기 전의 방식(A)과 [1]에서 제안한 고속 인터 모드 결정 방식을 비교 테스트한 결과를 표 1에 나타내었다. 또한 실험에 사용된 구체적인 부호화 옵션은 다음과 같으며, 테스트 영상으로는 QCIF(176x144) 크기의 영상 3개와 CIF(352x288) 크기의 영상 3개를 사용하였다.

- 1) Search range = ± 32
- 2) Rate Distortion Optimization = Use
- 3) # of reference frames = 1
- 4) Symbol mode = CABAC
- 5) GOP structure = IPPP... (I_frame_rate=100)
- 6) QP = 28, 32, 36, 40

표 1로부터 제안하는 방식(A)에 대한 실험 결과

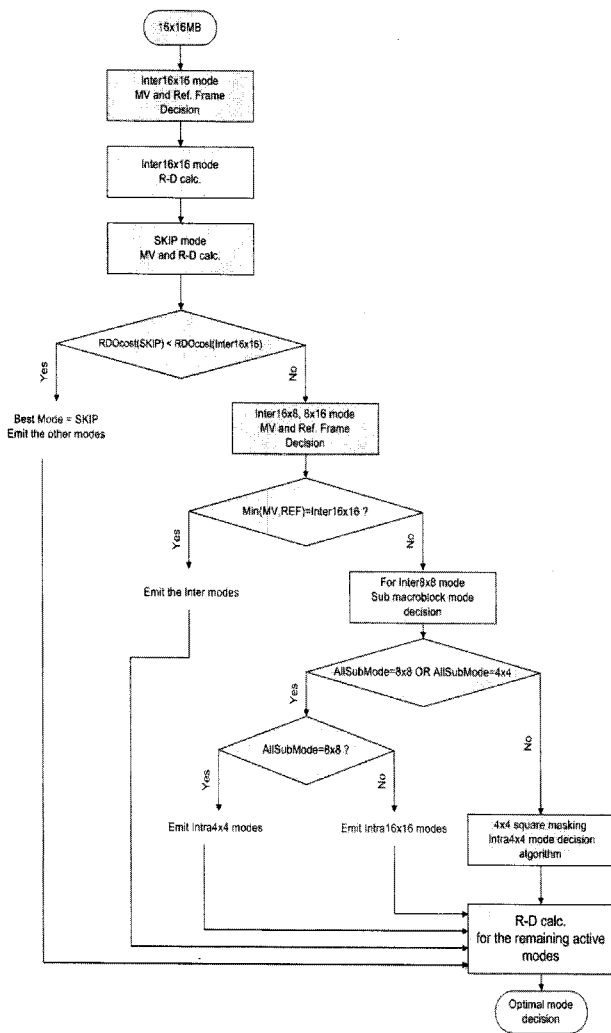


그림 6. 제안하는 알고리즘 전체 순서도
Fig. 6. Flow chart of the proposed algorithm.

표 1. 제안하는 방식(A)과 [1]의 알고리즘(F. Pan 등의 제안 방식)과의 실험 결과 비교

Table 1. Simulation results of the proposed method(A) and the existing algorithm(F. Pan's method)[1].

Sequences (QCIF)	F. Pan 등의 제안한 방식			제안하는 (A)의 방식								
	$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)	QP=28			QP=32			QP=40		
				$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)	$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)	$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)
Foreman	-0.077	1.536	-21.800	0.025	-5.602	-24.080	0.062	-3.307	-20.706	0.107	-0.574	-17.604
Container	-0.061	1.803	-20.764	-0.015	-3.094	-23.517	0.005	-1.727	-21.485	0.068	1.104	-18.065
Akiyo	-0.067	1.226	-23.108	0.027	-4.395	-21.216	0.006	-0.968	-19.615	0.061	1.336	-17.543
Mobile	-0.018	0.451	-27.565	-0.018	-4.182	-31.008	0.002	-5.360	-27.851	0.021	-3.088	-21.306
Paris	-0.023	0.504	-26.901	0.023	-6.584	-27.607	0.013	-7.276	-24.979	0.090	-5.871	-22.618
Tempo	-0.029	0.812	-26.723	-0.016	-4.526	-27.553	-0.010	-4.279	-24.754	0.016	-2.382	-22.010

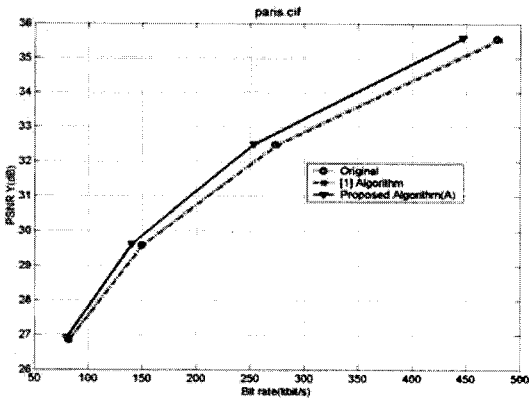


그림 7. Paris 영상에서 울-왜곡 곡선 성능 비교
Fig. 7. R-D performance comparison in Paris sequence.

PSNR의 변화량은 평균 0.05dB 내외로 미미한 수준에 그친 반면, 발생 비트율은 대부분의 테스트 영상에서 평균 3%이상 감소함을 볼 수 있다.

이는 Inter8x8 모드 내 서브 매크로블록의 최적 모드 분포를 이용하여 Intra4x4, Intra16x16 모드를 선택적으로 생략하는 과정에서 보다 적은 발생 비트수를 가지는 Intra16x16 모드의 선택 빈도수를 증가시켰기 때문이다. 즉 미미한 수준의 PSNR의 변화량에 비해 발생 비트율을 보다 큰 폭으로 낮출 수 있기 때문에 전체 부호화 효율은 그림 7에서 볼 수 있듯이 향상된 성능을 보인다. 부호화 시간 측면에서는 [1]에서 제안한 방식과 비슷한 성능을 보이며, QP가 증가할수록 부호화 시간 단축 효과가 감소하는데 이는 QP가 증가할수록 최적의 모드가 인터 모드에서 발생될 경우가 증가하기 때문이다. 그림 7은 Paris.CIF 영상에서 제안하는 방식(A)과 기존의 방식들에 의한 울-왜곡 곡선(R-D curve)을 비교하였다.

동일한 부호화 옵션을 사용한 경우, III-3의 조기 매크

표 2. 제안하는 최종 알고리즘(C)의 실험 결과

Table 2. Simulation results of the proposed final algorithm(C).

Sequences	본 논문에서 제안하는 최종 Algorithm(C)											
	QP=28			QP=32			QP=36			QP=40		
	$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)	$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)	$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)	$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)
Foreman	-0.004	-6.2	-26.73	0.05	-6.21	-27.71	0.097	-6.56	-32.01	0.238	-11.83	-28.73
Container	0.015	-11.28	-36.76	0.005	-14.88	-36.86	0.071	-16.64	-35.08	0.063	-18.53	-34.64
News	0.052	-14.23	-34.18	0.054	-12.80	-34.36	0.157	-10.34	-33.28	0.106	-14.61	-33.82
Siard	-0.015	-7.6	-34.73	-0.029	-7.23	-32.31	-0.018	-8.08	-32.14	0.055	-10.46	-32.86
Carphone	-0.008	-7.84	-32.27	0.086	-8.22	-32.01	0.123	-7.45	-29.97	0.059	-10.9	-30.8
Salesman	-0.047	-12.32	-36.38	0.019	-12.02	-34.79	0.063	-11.91	-34.13	0.005	-14.48	-34.37
Claire	0.193	-12.37	-38.12	0.189	-12.71	-38.33	0.944	-13.21	-33.38	0.286	-16.64	-33.69

로블록 모드 최종 결정 알고리즘(C)를 적용시킨 결과를 표 2에 나타내었다. 이 때 테스트 영상으로는 QCIF 크기의 영상 7개를 사용하였다.

실험 결과 표 2에서 볼 수 있듯이 전체 부호화 시간은 평균 33%이상 단축시킬 수 있었다. 또한 부호화 효율 측면에서 약 0.06dB 내외의 미미한 PSNR의 변화량을 가지는 반면 비트율은 평균 8% 이상 감소시킬 수 있었다. 특히 Container, News와 같이 움직임이 적고 배경 특성이 일정한 영상에서는 평균 13% 이상 큰 폭으로 비트율을 감소시킬 수 있었다. 이는 결국 앞서 기술한 바와 같이 큰 블록을 통해 매크로블록을 부호화하여도 왜곡 오차가 충분히 작기 때문에 상대적으로 큰 블록을 통해 부호화를 수행하는 SKIP, Inter16x16 모드에 대한 가중치를 높인 결과이다. 제안하는 최종 알고리즘의 성능을 [6]에서 제안하는 고속 모드 결정법과 비교 실험한 결과를 표 3에 나타내었으며, 이 때 사용된 부호화 옵션은 다음과 같다.

- 1) # of reference frames = 5
- 2) Fast motion estimation = Use
- 3) Rate Distortion Optimization = Use
- 4) Symbol mode = CAVLC
- 5) GOP structure = IPPP...
- 6) Search range = ± 32
- 7) QP = 28, 32, 36, 40

표 3의 결과로부터 전체 부호화 시간은 [6]에서 제안한 방식과 비교하여 평균 15%이상 단축시킬 수 있었다. 특히 고속 움직임 추정법과 결합하였기 때문에 부호화 시간을 최대 83%까지 단축시킬 수 있음을 알 수 있다. 또한 PSNR은 0.04dB 내외의 미미한 변화량을 가졌으며, 비트

표 3. [6]의 알고리즘과 본 논문에서 제안하는 고속 매크로블록 모드 결정 알고리즘과의 실험 결과 비교

Table 3. Simulation results of [6]'s algorithm and the proposed macroblock mode decision algorithm.

Sequences (CIF, QCIF)	[6]에서 제안한 방식			본 논문에서 제안하는 최종 알고리즘(C)								
	$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)	QP=20			QP=32			QP=40		
				$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)	$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)	$\Delta PSNR$ (dB)	ΔBit (%)	$\Delta Time$ (%)
CIF container	0.002	1.05	-50.09	0.046	-13.03	-77.77	-0.005	-16.1	-80.10	0.05	-15.57	-89.34
	-0.056	1.08	-44.22	0.047	-5.93	-51.30	0.09	-5.79	-56.51	0.03	-8.09	-69.95
				0.04	-6.71	-49.2	0.052	-9.09	-57.32	0.045	-11.16	-71.01
CIF foreman	-0.066	1.07	-43.69	0.073	-4.49	-52.25	0.125	-5.08	-60.34	0.002	-10.34	-79.10
	-0.155	1.26	-41.55	-0.026	-0.81	-40.79	-0.045	-1.47	-45.6	0.016	-2.89	-56.08
				-0.032	-1.18	-40.16	-0.022	-2.34	-47.64	0.122	-6.72	-66.96

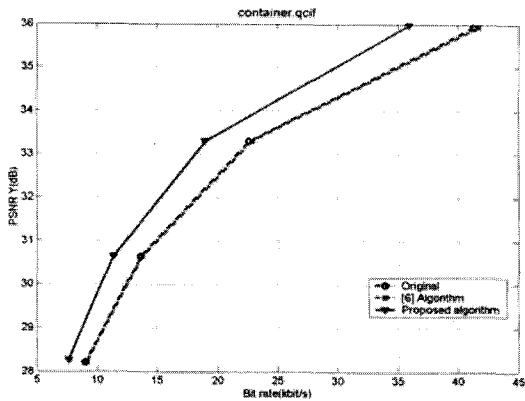


그림 8. Container 영상에서 울-왜곡 곡선 성능 비교
Fig. 8. R-D performance comparison in Container sequence.

울의 경우 평균 8%이상 감소시킬 수 있었다. [6]에서 제안한 방식과의 부호화 효율에 따른 RD 곡선을 그림 8에 나타내었다.

V. 결 론

본 논문에서는 기존의 R-D 최적화 기법에 따른 H.264/AVC 부호기의 복잡도를 감소시킬 수 있는 효율적인 알고리즘을 제안하였다. 제안 기법에서는 매크로블록 모드 결정과정에서 기존에 확보된 매크로블록 정보를 이용하여 후보 모드의 가지 수를 제한하는 기법을 적용하였다. 실험결과로부터 제안 알고리즘은 기존의 부호화기법에 비해 평균 30% 이상의 부호화 시간을 단축시켰으며, 비트율을 평균 8% 이상 크게 감소시킴으로써 전체 부호화 효율을 향상시킬 수 있었다. 앞으로 본 논문에서 제안한 고속 모드 결정방법을 고속 움직임

추정 알고리즘과 결합할 경우, 고속 부호기의 구현에 유용하게 사용될 수 있을 것으로 사료된다.

참 고 문 헌

- [1] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, D. Wu, and S. Wu, "Fast Mode Decision Algorithm for Intraprediction in H.264/AVC Video Coding", *IEEE Trans. Circuits and System for Video Technology*, vol.15, no.7, pp. 813-822, Jul. 2005.
- [2] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard", *IEEE Trans. Circuits and System for Video Technology*, vol.13, no.7, pp.560-576, Jul. 2003.
- [3] ITU-T Rec. H.264/ISO/IEC 11496-10, "Advanced Video Coding", *Final Committee Draft, Document JVT-E022*, September 2002.
- [4] J. W. Seok, B. H. Kim, and C. S. Cho, "The Research Trend of the H.264Technology", *ETRI*, vol.32, no1, pp.25-34, 2006.2.
- [5] K. Donghyung, J. Jechang, "A Fast MB Mode Selection Algorithm in the H.264 Standard", *Korea Institute Of Communication Sciences*, vol.30, no.1C, pp. 61-72, 2005.1.
- [6] L. Woongho, L. Jungho, C. Ikhwan, J. Dongseok, "Fast Mode Decision in H.264/AVC Using Adaptive Selection of Reference Frame and Selective Intra Mode", *Korea Institute Of Communication Sciences*, vol.31, no.3, pp. 271-278, 2006.3.
- [7] Y. Peng, C. T. Hyeyeon, M. T. Alexis, and B. Jill, "Fast Mode Decision and Motion Estimation for JVT/H.264", *Image Processing, 2003 International Conference on*, vol.3, pp.853-856, September 2003.
- [8] G. Sullivan, T. Wiegand, and K. P. Lim, "Joint model reference encoding methods and decoding concealment methods", presented at the 9th JVT Meeting(JVT-I049d0), San Diego, CA, Sep. 2003.
- [9] G. Sullivan, "Recommended simulation common conditions for H.26L coding efficiency experiments on low resolution progressive scan source material", presented at the 14th VCEG-N81 Meeting, Santa Barbara, CA, Sep. 2001.

저 자 소 개



김 지 웅(정회원)

2005년 호서대학교 정보통신
공학과 학사.

2007년 호서대학교 정보통신
공학과 석사.

2007~현재 (주)우리기술 연구원.

<주관심분야 : 신호처리, ARM 기반 H/W 설계>



김 용 관(정회원)

1988년 서울대학교 제어계측
공학과 학사.

1990년 서울대학교 제어계측
공학 석사.

1996년 서울대학교 제어계측
공학과 박사.

1994년 4월~1997년 2월 대우전자 영상연구소.

1997년 3월~현재 호서대학교 정보통신공학과
부교수.

<주관심분야 : 영상신호처리, 멀티미디어 통신,
신호처리 시스템 H/W 설계>