

논문 2006-43SP-6-2

H.264 표준에서 양자화 계수를 고려한 고속 인터모드 결정 방법

(A Fast Inter Mode Decision Algorithm Considering Quantization Parameter in H.264)

김 근 용*, 호 요 성**

(Geun-Yong Kim and Yo-Sung Ho)

요 약

최근에 만들어진 비디오 압축 표준인 H.264는 매크로블록당 최적의 부호화 모드를 결정하기 위해 비트율-왜곡 (rate-distortion) 기법을 사용하지만, 그 복잡성으로 인해 부호화하는 데 많은 시간이 걸린다. 따라서, H.264의 부호화 시간을 단축하기 위해 고속 모드결정 방법이 필요하다. 본 논문에서는 양자화 계수에 따라 발생 모드가 변하는 특성에 기반하여 불필요한 움직임 예측 및 모드결정 과정을 생략하는 새로운 고속 모드결정 방법을 제안한다. 양자화 계수에 따라 최적모드의 발생 빈도가 변하게 되는데, 제안한 방법에서는 매크로블록의 양자화된 이산 여현변환 계수들이 모두 0일 때, 0의 값을 가지는 CBP (coded block pattern)를 고려하여 이러한 특성을 반영하며, 조기 SKIP 모드 결정방법과 조기 16×16 모드 결정방법을 이용한다. 컴퓨터 모의실험을 통해, 본 논문에서 제안한 고속 인터모드 결정방법이 H.264의 참조 소프트웨어에 구현된 고속 인터모드 결정방법에 비해, Baseline 프로파일의 경우 19.6%, Main 프로파일의 경우 18.8%의 부호화 시간을 감소시키는 것을 확인했다.

Abstract

The recent video coding standard H.264 employs the rate-distortion optimization (RDO) method for choosing the best coding mode; however, it causes a large amount of encoding time. Thus, in order to reduce the encoding time, we need a fast mode decision algorithm. In this paper, we propose a fast inter mode decision algorithm considering quantization parameter (QP). The occurrence of best modes depends on QP. In order to reflect these characteristics, we consider the coded block pattern (CBP) which has 0 value when all quantized discrete cosine transform (DCT) coefficients are zero. We also use the early SKIP mode decision and early 16×16 mode decision methods. By computer simulations, we have verified that the proposed algorithm requires less encoding time than the fast inter mode decision method of the H.264 reference software for the Baseline and Main profiles by 19.6% and 18.8%, respectively.

Keywords : Fast Inter Mode Decision, Quantization Parameter, H.264

I. 서 론

H.264 동영상 압축 표준은 기존의 MPEG-2/4, H.263

방식 등에 비해 높은 부호화 효율을 제공하는데, 이는 가변블록 움직임 보상, 복수의 참조영상, 1/4 화소 단위의 부화소 움직임 예측과 같은 다양한 기술들이 사용되기 때문이다^[1]. 특히, 가변블록을 사용한 움직임 보상 기술은 매크로블록의 크기를 16×16에서 4×4까지 나누어 영상의 특성을 잘 반영하므로 부호화 효율을 증가시키지만, 동시에 높은 복잡도를 요구한다^[2]. 이러한 복잡도의 증가는 가변블록 움직임을 위한 매크로블록의 모드를 결정하기 위해 많은 계산량을 요구하는 비트율-왜곡 최적화 기법을 사용하는 데에 기인한다.

* 학생회원, ** 평생회원, 광주과학기술원 정보통신공학과 (Gwangju Institute of Science and Technology, Department of Information and Communications)

※ 본 논문은 광주과학기술원(GIST) 실감방송연구센터 (RBRC)를 통한 정보통신부 대학 IT연구센터(ITRC)와 교육인적자원부 두뇌한국21(BK21) 정보기술사업의 지원으로 수행되었다.

접수일자: 2006년1월31일, 수정완료일:2006년10월16일

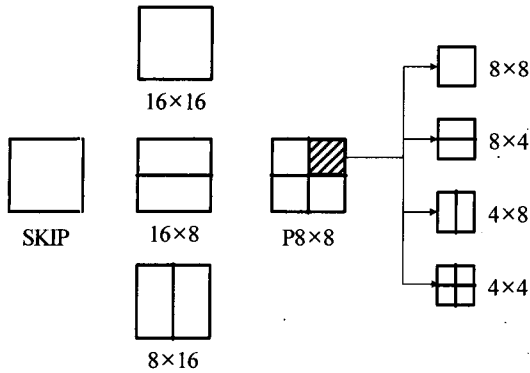


그림 1. H.264의 인터모드
Fig. 1. Inter Modes of H.264.

H.264는 매크로블록 모드를 위해 5개의 인터모드와 3개의 인트라모드를 지원한다. 이중 SKIP, 16×16, 16×8, 8×16과 P8×8은 인터모드에 해당되며, 움직임 예측을 통한 매크로블록 부호화에 이용된다. P8×8 모드는 각각의 8×8 블록 안에서 8×8, 8×4, 4×8, 4×4 중의 하나로 나뉘어진다. 그림 1은 H.264에서 사용되는 인터모드를 나타낸다. H.264는 인트라모드 예측을 위해 Intra4×4, Intra8×8, Intra16×16을 제공하며, 휘도 블록을 위한 Intra8×8 모드는 FRExt 프로파일에서만 제공된다.

인트라모드에서는 이전에 부호화된 유효한 이웃 매크로블록의 경계 화소를 이용해 현재 매크로블록을 예측한다. 휘도 블록의 Intra4×4와 Intra8×8은 각각 4×4와 8×8 블록 단위로 총 9가지 예측 방향을 이용하여 예측을 수행하고, Intra16×16은 16×16 블록 단위로 4가지 예측 방향을 이용해 현재 매크로블록을 예측한다. 이렇게 예측된 인터모드와 인트라모드들 중에서 비트율-왜곡 비용이 가장 작은 모드가 최종 매크로블록 모드로 결정된다.

인터모드에서 색차 신호는 휘도 신호의 움직임 벡터를 이용해 예측되지만, 인트라모드에서 색차 신호는 휘도 신호와 별도로 예측된다. 인트라모드에 대한 색차 신호 예측은 8×8 블록 단위로 이루어지며, 각 8×8 블록 당 4가지 예측 방향이 사용된다.

본 논문에서는 양자화 계수를 고려한 고속 인터모드 결정 방법을 제안한다. 양자화 계수가 증가함에 따라 SKIP과 16×16 모드의 발생 빈도가 증가하는데, 이것은 이산 여현변환과 양자화 과정으로 인해 최적모드를 결정하는 비용함수의 값이 변하기 때문이다. 이러한 특성을 반영하기 위해 매크로블록의 모든 양자화된 DCT 계수들이 0일 때, 0의 값을 가지는 CBP 값을 이용한다. 따라서, QP 값이 클 경우에는 필요치 않은 P8×8에 대한 움직임 예측 및 모드결정 과정을 생략한다. 뿐만 아니라, 16×16, 16×8, 8×16의 움직임 예측과

정의 결과를 기반으로 조기에 인터모드를 16×16 모드로 결정하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. II장에서 H.264의 매크로블록 모드 결정방법에 대해 설명하고, III장에서 제안한 양자화 계수를 고려한 고속 모드결정 방법을 설명한다. IV장에서 실험환경과 실험결과를 제시한 후에, 마지막으로 V장에서 결론을 맺는다.

II. H.264 매크로블록 모드결정 방법

매크로블록을 어떤 모드로 부호화할 지를 결정하는 것은 H.264 부호기의 핵심이다. 기존의 비디오 압축 부호화 표준과는 달리 H.264는 여러 개의 모드 중에서 최적의 모드를 결정하기 위해서 왜곡도와 함께 발생비트를 고려한다. 이를 위해 라그랑지안 함수에 기반한 비용함수를 사용한다^[3]. 각 블록에 대한 움직임 벡터 결정과 매크로블록의 최적모드를 결정하기 위해 사용되는 이 함수는 왜곡도와 발생 비트량을 나타내는 항과 발생 비트량의 가중치 값인 라그랑지안 계수 (Lagrangian multiplier)로 이루어져 있다. 이 비용함수는 JM 참조 소프트웨어에 구현되어 사용되고 있다.

1. 비트율-왜곡 최적화 기법

비트율-왜곡 최적화 기법에서는 각 매크로블록 모드에 대한 비트율-왜곡 비용을 계산한다. 인터모드에서는 이 값을 계산하기 전에 움직임 벡터와 참조영상을 결정해야 한다. 가변 블록에 대한 움직임 벡터와 참조영상은 식 (1)의 최소값으로 결정된다.

$$J_{motion}(MV, REF | \lambda_{motion}) = SAD(s, r(MV, REF)) + \lambda_{motion} \cdot R(MV, REF) \quad (1)$$

여기서 λ_{motion} 은 라그랑지안 계수로 양자화 계수에 의존하는 값이며, 인터모드의 경우에는 $0.85 \times 2^{QP/3}$ 의 값을 가진다. $R(s, r(MV, REF))$ 는 움직임 벡터와 참조영상을 부호화하는데 필요한 비트량으로 표 참조 연산에 의해 얻어진다. MV 와 REF 는 각각 움직임 벡터와 참조영상을 나타낸다. $SAD(s, r(MV, REF))$ 는 원 영상과 움직임 예측에 의해 복원된 영상과의 차에 대한 절대값의 합이고, 다음과 같이 정의된다.

$$SAD(s, r(MV, REF)) = \sum_{x \in H, y \in V} |s(x, y) - r(x - m_x, y - m_y)| \quad (2)$$

여기서 s 와 r 은 각각 원 영상과 복호된 영상의 화소값

을 나타낸다. H 와 V 는 각각 가변 블록의 넓이와 높이이며, m_x 와 m_y 는 움직임 벡터이다.

P8x8 모드에서 각 8x8 블록에 대한 최적의 하위 매크로블록 모드의 결정, 인트라모드에서 최적의 예측 방향 결정과 매크로블록에 대한 최적모드 결정은 식 (3)의 최소값에 의해 결정된다.

$$J_{mode}(s, r, M, \lambda_{mode}) = SSD(s, r, M) + \lambda_{mode} \cdot R(s, r, M) \quad (3)$$

여기서 λ_{mode} 는 λ_{motion} 의 제곱값이다. M 은 매크로블록 모드, 예측방향 모드 또는 하위 매크로블록 모드이다. $R(s, r, M)$ 은 M 에 해당하는 모드로 부호화했을 때, 실제로 발생하는 비트량이며, H.264 참조 소프트웨어에서는 이 값을 계산하기 위해 모든 부호화 과정이 수행된다. s 와 r 은 각각 원 영상과 복호된 영상의 화소값을 나타낸다. $SSD(s, r, M)$ 는 원 영상과 복호된 영상의 차이에 대한 제곱의 합이고, 다음과 같이 계산된다.

$$SSD(s, r, M) = \sum_{x \in H, y \in V}^{H, V} (s(x, y) - r(x - m_x, y - m_y))^2 \quad (4)$$

P8x8 모드의 하위 모드결정과 인트라모드에서 예측 방향 결정의 경우에 식 (3)과 식 (4)는 휘도 성분만을 고려하여 계산하고, 최종 모드결정을 위한 경우는 휘도와 색차 성분 모두에 대해서 계산을 수행한다.

2. H.264 부호기의 복잡도

그림 2는 H.264 부호기에서 주요 모듈별 수행시간의 비율을 나타낸다^[4]. 그림 2에서 보듯이, 비용 생성을 위한 변환과 모드 결정 과정이 많은 부분을 차지하는 것을 알 수 있다. 모드 결정 과정에서 사용되는 식 (3)의 발생 비트량을 계산하기 위해서는 각 모드에 대해 DCT/Hadamard 변환 및 양자화/역양자화, 역 DCT/

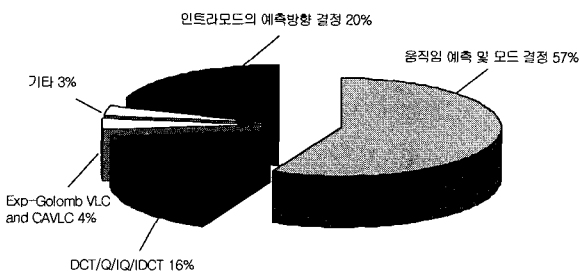


그림 2. 모듈별 수행시간 비율
Fig. 2. Run-Time Percentage of H.264 Modules.

Hadamard 변환과정, 그리고 엔트로피 부호화 과정을 거치게 된다. 엔트로피 부호화 과정에서 사용되는 Exp-Golomb VLC와 CAVLC, CABAC는 다소 복잡한 계산 과정을 필요로 하지만, 실제 부호기에서 차지하는 복잡도는 그렇게 크지 않다.

따라서 부호기의 복잡도를 줄이기 위해서는 고속 모드결정 방법의 개발이 필수적이다. 특히, 중요한 것은 같은 매크로블록을 부호화하더라도 양자화 계수에 따라 최적모드의 발생 비율이 변한다는 것을 고려하여, 최적모드가 될 확률이 매우 적은 모드에 대한 움직임 예측 단계를 효과적으로 생략하는 것이다.

3. H.264의 모드결정 과정

그림 3은 H.264 참조 소프트웨어 JM 9.5^[5]의 모드결정 과정을 나타낸다. 그림 3에서 과정 (A)는 인트라모드를 위한 움직임 벡터와 참조영상을 결정하는 과정이다. 16x16과 16x8, 8x16의 각 모드에 대해 식 (1)의 비용함수가 최소값을 가지는 움직임 벡터와 그 때의 참조영상이 저장된다. P8x8 모드의 하위 최적모드를 결정하는 과정은 8x8과 8x4, 4x8, 4x4 모드에 대해 움직임 벡터와 참조영상을 결정하는 과정과 식 (3)을 통해 4개의 모드 중에서 최적모드를 결정하는 두 과정으로 이루어져 있다. 마지막으로, SKIP 모드의 움직임 벡터를 계산한 후에 과정 (B)에서는 식 (3)을 이용하여 최적모드를

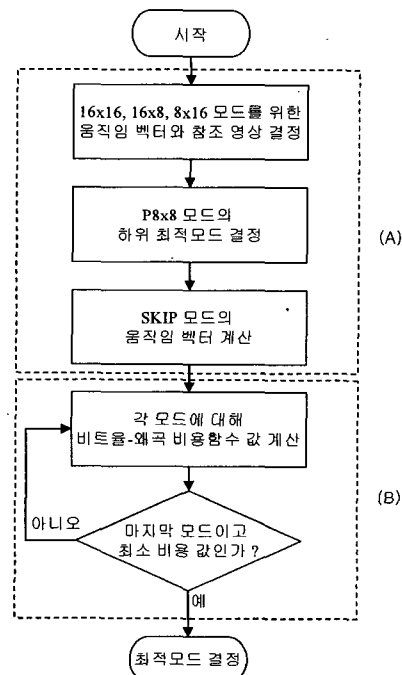


그림 3. 최적모드 결정과정
Fig. 3. Procedure for Choosing Best Mode.

결정하게 된다. 이 과정에서는 인터모드들을 포함하게 되며 모든 모드들에 대해서 비용함수 값이 최소가 되는 모드가 최적모드가 된다.

H.264 참조 소프트웨어 JM 9.5 에서는 식 (3)을 사용하여 최적모드를 결정하는 고복잡도 모드 (High Complexity Mode) 외에 식 (1)의 계산 결과가 최소값을 가지는 모드를 최적모드를 선택하는 부최적화 (sub-optimal) 모드를 지원하는데, 이것은 참조 소프트웨어에서 비트율-왜곡 최적화 옵션을 저복잡도 모드 (Low Complexity Mode)로 설정했을 때 동작한다. 이 모드는 고복잡도 모드에 비해 적은 부호화 시간을 필요로 하지만, 실제 발생하는 비트량을 고려하지 않기 때문에 고복잡도 모드로 부호화 했을 때보다 발생 비트율 측면에서 좋지 않은 성능을 보인다.

4. 기존의 고속 모드결정 방법들

H.264 부호기의 복잡도를 증가시키는 주요 요인인 모드결정 단계의 시간을 줄이고자 많은 방법들이 제안되고 있다. 그 중에서 초기에 연구된 대표적인 방법은 모드 분류를 이용한 고속 모드결정 방법이다^[6]. 이 방법의 핵심은 초기 SKIP 모드 결정과 인터모드에 대한 모드 분류라고 할 수 있다. 이 방법에서 SKIP 모드는 식 (5)를 만족할 때 결정된다.

$$J_{mode}(SKIP|\lambda_{mode}) < J_{mode}(16 \times 16|\lambda_{mode}) \quad (5)$$

식 (5)에서 보는 바와 같이, 모드결정 과정 초기에 SKIP과 16×16 모드에 대한 라그랑지안 함수값을 계산한다. 만약, SKIP 모드에 대한 값이 작으면 매크로블록이 더 이상 나뉘어지지 않아도 된다고 판단하고 최적모드를 SKIP으로 결정하는 것이다.

초기 SKIP 모드 조건을 만족하지 않아 인터모드 결정 과정이 필요한 경우는 P8×8 모드 중에서 8×8 모드를 기본 모드 (Base Mode)라 하고 16×16, 16×8, 8×16 모드의 J_{mode} 값 중에서 가장 작은값 (Best16)과 비교한다. 이를 식으로 나타내면 다음과 같다.

$$J_{mode}(Best16|\lambda_{mode}) > J_{mode}(Base|\lambda_{mode}) \quad (6)$$

식 (6)의 조건을 만족하게 되면 P8×8 모드는 최적모드 결정 과정에 포함되며 그렇지 않으면 이하 모드결정 과정에서 생략된다.

이 방법에서 간과하고 있는 것은 대부분의 영상에서는 SKIP 모드와 16×16 모드의 발생빈도가 가장 높다는 것이다. 즉, 모드결정 과정의 초기에 P8×8에 대한 고려

는 효율적으로 모드결정 과정의 복잡도를 줄일 수 없다. 설령, P8×8 모드가 최적모드가 나오는 경우라 할지라도 초기에 SKIP, 16×16, 16×8, 8×16에 대한 결과로 P8×8 모드를 최적모드 결정과정에서 고려해야 하는지를 예측할 수 있어야 한다. 왜냐하면, P8×8 모드는 II장 3절에서 설명한 바와 같이 매크로블록 내의 각 8×8 블록에 대한 움직임 예측과 모드결정 과정이 모두 포함되어 16×16이나 16×8, 8×16의 경우보다 복잡도가 높기 때문이다.

고속 모드결정을 위한 다른 방법은 J_{motion} 값과 J_{mode} 값의 상관도를 이용하는 방법이다^[7]. 움직임 벡터를 예측하기 위해 식 (1)이 사용되는데 일반적으로 이 값이 작을수록 식 (3)의 값이 작을 확률이 높다고 예측할 수 있다. 왜냐하면, 발생 비트의 대부분은 움직임 보상 과정에서 생기는 잔여 데이터가 차지하기 때문이다.

이 방법에서는 식 (5)와 같은 초기 SKIP 모드 결정 방법이 사용되며, 또한 위에서 설명한 바와 같이 J_{motion} 값과 J_{mode} 값의 상관도를 이용한다. 즉, 16×16, 16×8, 8×16의 J_{motion} 값 중에서 16×16에 해당하는 값이 가장 작다면 P8×8에 대한 J_{motion} 및 J_{mode} 계산 과정을 생략한다. 이런 경우는 현재 매크로블록을 더 작은 블록으로 나눈다 하더라도 왜곡도나 발생 비트량 측면에서 효율이 좋지 않을 것이라고 예측되기 때문이다. 이를 식으로 나타내면 다음과 같다.

$$J_{motion}(16 \times 16) < J_{motion}(16 \times 8, 8 \times 16) \quad (7)$$

본 논문에서도 이와 같이 J_{motion} 과 J_{mode} 값의 상관성을 이용한다.

5. 참조 소프트웨어의 고속 모드결정 방법

H.264 참조 소프트웨어 JM 9.4 버전부터 구현되어 있는 고속 모드결정 방법은 초기 SKIP 모드결정을 통해 불필요한 계산을 생략한다^{[8][9]}. 초기에 16×16 모드에 대한 움직임 예측을 수행하여 움직임 벡터와 참조영상을 결정한 후, 비트율-왜곡치를 계산한다. 이 계산과정을 통해 나온 결과가 다음과 같은 4 가지의 조건을 모두 만족할 때 SKIP 모드로 결정되며, 최적모드가 SKIP 모드로 결정되면 이하 모든 모드결정 과정은 생략된다.

- (1) 움직임 보상 블록이 16×16
- (2) 참조영상은 바로 이전 영상
- (3) 16×16 모드의 움직임 벡터와 SKIP 모드의 움직임 벡터가 같음

(4) 잔여 데이터의 비트량이 0 (CBP = 0)

SKIP 모드는 움직임 벡터가 주변의 이웃한 3 개의 매크로블록으로부터 예측된 움직임 벡터와 같은 16×16 모드의 한 경우로 생각할 수 있다. SKIP 모드의 비트율-왜곡치는 움직임 예측 과정, 참조영상 결정, 잔여 데이터의 계산과정 없이 움직임 벡터를 주변의 블록으로부터 예측된 움직임 벡터로 가정하고 비트율-왜곡치를 계산하여 다른 모드와 비교를 수행한다.

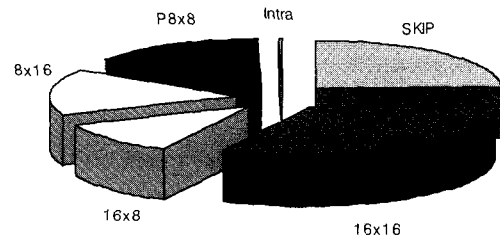
H.264 표준에서 SKIP 모드는 잔여 데이터의 비트량이 0 이어야 하지만, 실제 SKIP 모드의 비트율-왜곡치 계산 결과 비트량이 0이 아니어도 다른 모드에 비해 그 값이 작다면, SKIP 모드로 결정될 수 있다. 따라서, 16×16 모드의 비트율-왜곡치 계산 결과로부터 얻어진 네번째 조건은 초기에 SKIP 모드를 결정하는데 충분한 조건을 제공한다. CBP (coded block pattern)는 8×8 크기의 휘도블록 4개와 색차블록 2개로 이루어진 매크로블록에서 어떤 블록이 0이 아닌 DCT 변환값들을 가지고 있는지를 나타낸다. 따라서 CBP 값은 잔여 데이터의 DCT 계수들의 양자화된 값들이 모두 0일 때 0의 값을 갖는다. 본 논문에서는 위 4가지의 조건을 모두 만족하지 않아서 초기에 SKIP 모드로 결정되지 않더라도, 모드결정 과정에서 이 값을 이용한다.

III. 제안하는 고속 모드결정 방법

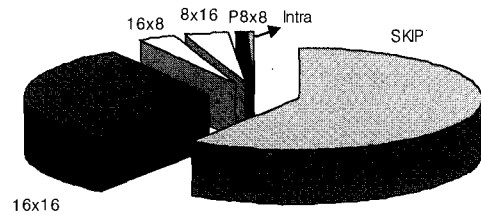
본 장에서는 제안하는 고속 인터모드 결정방법에 대해 설명한다. 본 논문에서 제안하는 고속 인터모드 결정방법의 가장 큰 특징은 같은 매크로블록을 부호화 하더라도 양자화 계수에 따라 최적모드의 발생 분포가 변한다는 점을 고려했다는 점이다. 즉, 양자화 계수의 값이 증가함에 따라 SKIP이나 16×16모드의 발생빈도가 많아진다는 점에 기반하여 CBP의 값을 이용하여 초기에 모드결정 과정을 종료시킨다. 또한, 부최적화모드 결정방법 (sub-optimal mode decision)의 결과를 이용하여 16×16 모드로 초기에 최적모드를 결정하는 방법을 제안한다.

1. 양자화 계수에 따른 모드 분포의 변화

같은 영상에 대해서도 양자화 계수값이 변하면 모드 분포의 변화가 심하며, 일반적으로 양자화 계수값이 큰 경우에는 SKIP 모드와 16×16 모드가 최적모드가 되는 빈도가 증가한다. 그림 4는 양자화 계수가 각각 28과 40



(a) 양자화 계수가 28인 경우



(b) 양자화 계수가 40인 경우

그림 4. 양자화 계수에 따른 최적모드 분포

Fig. 4. Distribution of Best Mode According to QP.

인 경우에 QCIF 크기의 FOREMAN 영상에 대해 결정된 최적모드의 분포를 보여준다.

그림 4에서 보는 바와 같이, 양자화 계수가 크면 SKIP 모드와 16×16 모드가 많이 발생하는 것을 볼 수 있는데, 이는 DCT 연산 후 큰 양자화 계수로 인한 발생 비트량의 변화로 최적모드가 변했기 때문이다. 즉 식 (3)에서 $\lambda_{mode} \cdot R(s,r,M)$ 값이 변하는 것이다. 매크로블록이 복잡한 화면의 부분이라 할지라도, 양자화 계수가 커질수록 DCT 과정을 거치면 결과 계수들이 0을 가지는 경우가 많아지며, 따라서 최적모드가 SKIP이나 16×16으로 결정되는 빈도수가 증가하게 되는 것이다.

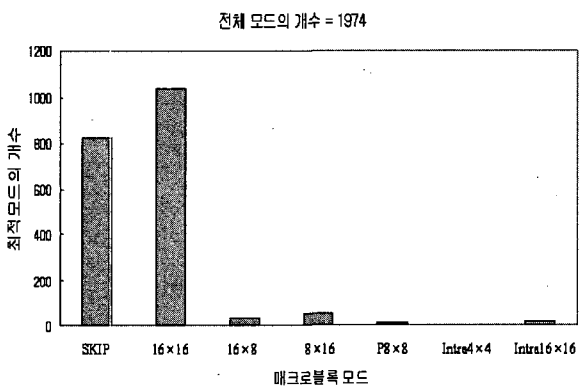
기존의 고속 모드결정 방법에서는 이 부분을 고려하지 않았기 때문에 저비트율로 부호화된 경우에 최적모드 결정을 위해 고려할 필요가 없는 모드에 대한 움직임 예측 및 비용함수 계산을 생략하지 못했다. 본 논문에서는 이를 고려하기 위해 양자화 계수의 변화를 반영하는 CBP 값을 이용한다.

2. 부최적모드 결정방법의 이용

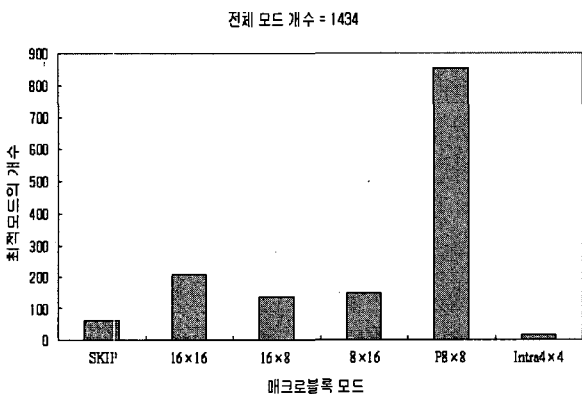
참조 소프트웨어에서 비트율-왜곡 최적화 옵션을 저 복잡도 모드로 설정하면 식 (1)의 J_{motion} 가 최소값을 가지는 모드가 최적모드로 결정된다. 이것은 그림 3의 (A) 부분에 해당되며, 식 (3)의 J_{mode} 를 이용한 (B) 부분

은 수행되지 않는다. 실제로 움직임 벡터를 결정할 때에도 식 (3)을 이용하면 비트율과 왜곡도 측면에서 더 정확한 계산이 가능하지만, 복잡도가 증가하기 때문에 참조 소프트웨어에서는 식 (1)과 같이 근사화된 계산을 이용한다. II장 3절에서 언급한 바와 같이, 저복잡도 모드로 설정하면 고복잡도 모드로 설정했을 때보다 비트율에서 좋지 않은 성능을 보이며, 이는 실제 발생 비트를 계산하지 않고 움직임 벡터와 참조화면만을 고려하기 때문이다. 그러나, 이 결과는 실제 고복잡도 모드로 최적모드를 결정했을 때와 유사한 경우가 많으며, 우리는 이 값을 최적모드 결정을 위한 후보로 생각할 수 있다.

그림 5는 부최적화 모드결정 방법으로 결정된 부최적모드와 최적모드 사이의 관계를 보여준다^[7]. 그림 5(a)는 FOREMAN 영상에서 부최적모드가 16×16일 경우를, 그림 5(b)는 MOBILE 영상에서 부최적모드가 P8×8인 경우를 나타낸다. 실험은 QCIF 영상의 50 프레임 을 이용하여 수행되었다.



(a) FOREMAN 영상에서 부최적모드가 16×16일 때



(b) MOBILE 영상에서 부최적모드가 8×8일 때

그림 5. 최적모드와 부최적모드의 관계

Fig. 5. Relation between Best Mode and Sub-Optimal Best Mode.

영상에서 매크로블록이 배경이나 움직임이 느린 부분에 속해 있을 경우에는 최적모드가 SKIP이나 16×16 모드 결정되는 경우가 많다. 이는 큰 블록을 사용하여 매크로블록을 부호화해도 왜곡도가 충분히 작기 때문에 작은 블록으로 부호화할 필요가 없기 때문이다. 블록을 작게 나누어 부호화하여 왜곡도가 감소하더라도, 각 블록에 대한 정보를 복호기에 알려주어야 하기 때문에 비트량이 증가하게 된다.

본 논문에서는 이런 점을 고려하기 위해서 16×16, 16×8, 8×16 모드들 중 부최적모드가 16×16으로 결정되면, P8×8 모드가 최적모드가 되는 경우가 적다고 판단하여 그림 3의 (A)와 (B) 단계에서 P8×8 모드를 제외시킨다. 이것은 매크로블록을 작게 나누어도 부호화 효율이 증가되지 않는 경우를 조기에 발견함으로써, 부호기의 복잡도를 줄이는 효과를 가져온다.

3. 양자화 계수를 고려한 고속 모드결정 방법

그림 6은 본 논문에서 제안하는 가변 부호화 비트율을 위한 고속 인터모드 결정 방법의 흐름도를 보여준다. 우선 16×16 모드에 대해서 참조영상 및 움직임 벡터를 결정한 후, 비트율-왜곡치를 계산하고 조기 SKIP 모드 결정 조건을 만족하는지 검사한다. 만약, 조기 SKIP 모드 조건이 만족되면 16×8과 8×16, P8×8 모드에 대한 참조영상 결정 및 움직임 벡터 계산을 생략하기 때문에 효율적으로 부호화 시간을 단축할 수 있다.

조기 SKIP 모드 조건을 만족하지 않더라도 16×16, 16×8, 8×16 모드 중 부최적모드가 16×16 이고 CBP가 0의 값을 가지면, 현재 부호화되는 매크로블록은 영상의 복잡한 부분이 아니며 움직임이 일정하거나 거의 없는 부분이라고 생각할 수 있다. 또한, 양자화 계수값이 크다면 DCT 계수들의 값이 작아 CBP 값이 0이 나오는 경우가 증가하므로 매크로블록을 작은 단위의 블록으로 나눌 필요가 없다. 따라서 P8×8 모드에 대한 세부블록 모드결정 과정에 필요한 참조영상 결정 및 움직임 벡터 계산을 생략한다. 이 과정을 통해 양자화 계수값이 큰 경우에 SKIP이나 16×16 모드의 발생빈도가 증가하는 경우를 고속 모드결정 과정에 반영할 수 있다.

P8×8 모드를 비활성화하고 SKIP 모드의 움직임 벡터가 16×16 모드의 움직임 벡터와 같다면, 16×8과 8×16의 모드 또한 최적모드 결정을 위한 비트율-왜곡치 계산에서 배제한다. CBP의 값이 0 이면서 이 조건을 만족하면 조기에 16×16 모드를 최적 인터모드로 결정하고 인트라모드와 비교하는 것이다. 이러한 조건들을 모

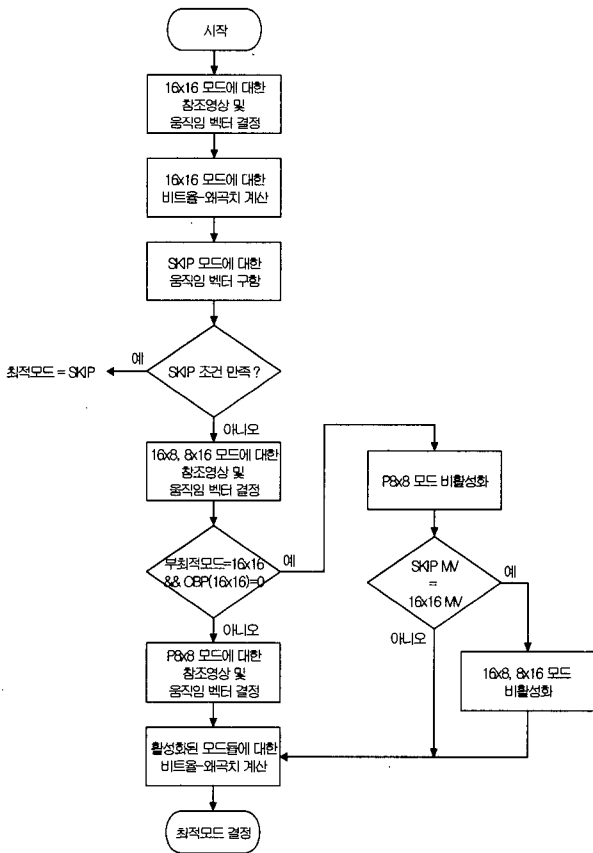


그림 6. 제안하는 고속 모드결정 방법의 흐름도
Fig. 6. Flowchart of Proposed Fast Mode Decision.

두 검사한 후, 마지막으로 인트라모드를 포함한 활성화된 인터모드들에 대해서 비트율-왜곡치를 계산하여 최적모드를 결정하게 된다.

IV. 실험결과

본 논문에서는 JM 9.5에 제안한 고속 모드결정 방법을 구현하였다. 제안한 방법의 효율성을 보이교자, 100 개의 화면을 가지는 QCIF 크기의 각기 다른 8개의 비디오 영상을 이용하여 IPPP 구조의 Baseline 프로파일과 Main 프로파일로 각각 부호화했다. Baseline 프로파일의 경우 가변길이 부호화 방법으로 CAVLC를 사용했으며, Main 프로파일의 경우 CABAC를 사용했다.

움직임 예측을 위해 참조영상은 5개를 사용했으며, 탐색 범위는 ±16 이다. 비트율-왜곡 곡선을 유도하기 위해 4개의 양자화 계수 (28, 32, 36, 40)를 사용했으며 Hadamard 변환을 사용했다^[10].

제안한 고속 모드결정 방법과 H.264 참조 소프트웨어의 고속 모드결정 방법의 성능 비교를 위해 평균 PSNR과 비트율^[11]을 사용했다. 실험은 고속 고복잡도

표 1. 실험결과 (Baseline 프로파일).

Table 1. Experimental Results (Baseline Profile).

영상	ΔPSNR (dB)	ΔBit Rate (%)	Avg. ΔT (%)
AKIYO	+ 0.02	- 0.44	22.0
NEWS	0	- 0.17	15.7
COASTGUARD	0	- 0.12	19.7
CONTAINER	+ 0.02	- 0.43	19.4
SILENT	0	+ 0.07	18.1
FOREMAN	+ 0.02	- 0.35	24.7
CARPHONE	+ 0.02	- 0.35	25.6
MOBILE	0	+ 0.04	11.9
평균	+ 0.01	- 0.21	19.6

표 2. 실험결과 (Main 프로파일)

Table 2. Experimental Results (Main Profile).

영상	ΔPSNR (dB)	ΔBit Rate (%)	Avg. ΔT (%)
AKIYO	- 0.03	+ 0.47	20.0
NEWS	+ 0.01	- 0.21	14.5
COASTGUARD	0	+ 0.16	18.4
CONTAINER	0	- 0.05	20.5
SILENT	- 0.01	+ 0.14	18.5
FOREMAN	0	- 0.02	22.1
CARPHONE	- 0.03	+ 0.47	24.3
MOBILE	0	- 0.01	12.1
평균	- 0.01	+ 0.12	18.8

표 3. 양자화 계수에 따른 성능 변화

Table 3. Performance Change According to QP.

항목	양자화 계수 28		양자화 계수 40	
	H.264 고속모드	제안한 방법	H.264 고속모드	제안한 방법
PSNR (dB)	35.67	35.64	28.15	28.15
Bit Rate (kbit/s)	108.34	108.03	24.34	24.19
부호화 시간 (초)	(60.1)	(45.2)	(50.8)	(27.5)

모드 (Fast High Complexity Mode)에서 수행되었으며, 수행시간은 각 양자화 계수별 부호화 시간의 평균으로 계산하였다. 고속 움직임 예측 방법은 참조 소프트웨어에 구현되어있는 UMHexagonS와 CBFPS를 사용했다^{[12][13]}. H.264의 고속 모드결정 방법과 제안한 고속 모드 결정 방법의 수행시간 비교를 위해 ΔT를 다음과 같이 정의한다.

$$\Delta T = \frac{Time(H.264 \text{ 고속모드}) - Time(\text{제안한 방법})}{Time(H.264 \text{ 고속모드})} \times 100\% \quad (8)$$

또한, 각 양자화 계수에 대한 ΔT 의 평균을 계산하기 위해 $Avg. \Delta T$ 를 다음과 같이 정의한다.

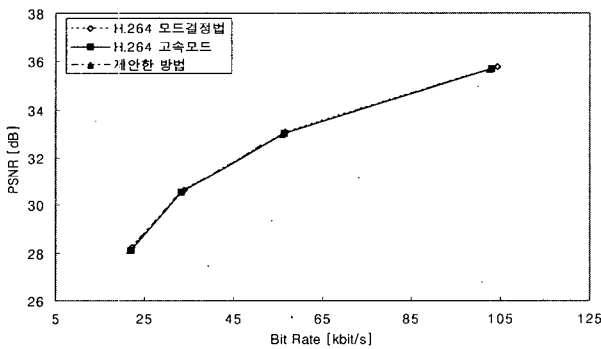
$$Avg. \Delta T = \frac{\Delta T(28) + \Delta T(32) + \Delta T(36) + \Delta T(40)}{4} [\%] \quad (9)$$

괄호안의 숫자는 양자화 계수를 의미한다.

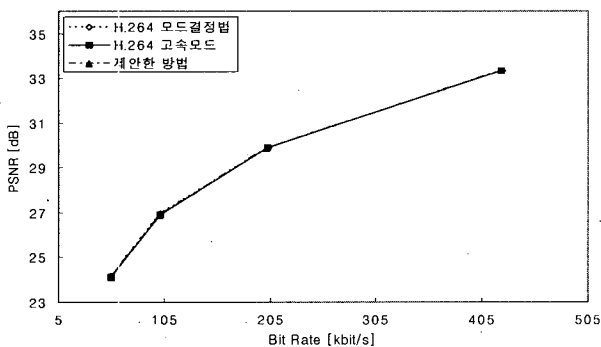
표 1과 표 2는 성능 비교를 위해 실험한 결과를 나타낸다. 표의 값들은 제안한 방법과 H.264의 고속 모드결정 방법을 비교한 PSNR과 비트율의 변화 및 부호화 시간을 나타낸다. 표 1은 Baseline 프로파일에 대한 결과이며, 표 2는 Main 프로파일로 부호화 했을 때의 결과를 나타낸다.

표 1과 표 2에서 보는 바와 같이, 제안한 고속 인터모드 결정방법은 H.264의 고속 모드결정 방법에 비해 Baseline 프로파일의 경우 평균 19.6%, Main 프로파일의 경우는 18.8% 부호화 시간을 단축시켰다. PSNR과 발생 비트량에 대해서는 프로파일에 따라 다소 차이가 있지만, 그 변화가 미비하여 부호화 효율의 변화없이 부호기의 복잡도를 감소시켰다는 것을 알 수 있다.

표 3은 Baseline 프로파일로 FOREMAN 영상을 부



(a) FOREMAN 영상



(b) MOBILE 영상

그림 7. 비트율-왜곡 곡선

Fig. 7. Rate-Distortion Curves.

호화하기 위해 H.264의 고속 모드결정 방법과 제안한 고속모드 방법을 사용했을 때, 양자화 계수가 28과 40인 경우에 대한 PSNR 값과 발생 비트량 및 부호화 시간을 나타낸다. 부호화 시간에서 괄호 안의 값은 움직임 예측 과정을 위해 소요되는 시간을 나타낸다.

표 3에 보인 것처럼, 제안한 방법은 양자화 계수가 커짐에 따라 부호화 시간을 더 단축시키는데, 이것은 저비트율에서 큰 QP로 인해 CBP가 0이 나오는 경우가 증가하기 때문이다. 또한, 부호화 시간 단축이 움직임 예측 단계에서 일어난 것을 알 수 있는데, 이것은 제안한 방법이 양자화 계수의 증가에 따라 최적모드 결정과정에 불필요한 P8×8 모드에 대한 움직임 예측 단계를 효율적으로 생략했기 때문이다.

그림 7은 FOREMAN과 MOBILE 영상에 대한 비트율-왜곡 곡선을 나타낸다.

V. 결 론

본 논문에서는 같은 영상을 부호화하더라도 양자화 계수에 따라 최적모드의 발생 빈도가 다르게 나타나는 점을 고려하여, 불필요한 움직임 예측 및 모드결정 과정을 생략하는 새로운 고속 모드결정 방법을 제안했다. 즉, 16×16 모드의 CBP 값이 0일 경우 P8×8 모드를 최적모드 결정과정에서 제외하고, SKIP 모드의 움직임 벡터와 16×16 모드의 움직임 벡터까지 같은 경우에는 조기에 인터모드를 16×16 모드로 결정하여 부호화 시간을 효율적으로 단축했다. 제안한 고속 인터모드 결정 방법은 H.264 고속 모드결정 방법과 비교하여, PSNR 값과 발생 비트율의 변화가 거의 없으면서 Baseline 프로파일의 경우 19.6%, Main 프로파일의 경우는 18.8%의 부호화 시간을 감소시켰다. 특히, 양자화 계수가 커질수록 부호화 시간이 더욱 감소함을 확인했다.

참 고 문 헌

- [1] ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC, "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification," JVT Doc. JVT-G050, 2003.
- [2] 최용일, 전병우, "H.264 표준의 가변 움직임 블록을 위한 고속 움직임 탐색 기법," 전자공학회 논문지 제41권 SP편 제6호, 209-220쪽, 2004년 11월.
- [3] G. J. Sullivan and T. Wiegand, "Rate-Distortion Optimization for Video Compression," *Signal*

Processing Magazine, IEEE, Vol. 15, pp. 74-90, Nov. 1998.

[4] Y. U. Huang, B. Y. Hsieh, T. C. Chen, and L. G. Chen, "Analysis, Fast Aglorithm, and VLSI Architecture Design for H.264/AVC Intra Frame Coder," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 15, No. 3, pp. 378-401, March 2005.

[5] JVT Reference Software Version 9.5, Available Online at: http://iphome.hhi.de/suehring/tml/download/old_jm/jm95.zip

[6] H. S. Kim, S. H. Kim, and Y. S. Ho, "Fast Mode Decision Algorithm Using Mode Classification for H.264," *ITC, Thiland*, pp. 51-58, Nov. 2004.

[7] G. Y. Kim, S. H. Kim, H. S. Kim, and Y. S. Ho, "Fast Mode Decision Algorithm for H.264 Based on Motion Cost," *European Signal Processing Conference*, Sept. 2005.

[8] 이제운, 전병우, "H.264 동영상 압축을 위한 고속 부호화 모드 결정 방법," 전자공학회 논문지 제41 권 SP편 제6호, 165-173쪽, 2004년 11월.

[9] B. Jeon and J. Lee, "The Fast Mode Decision with Fast Motion Estimation," *JVT Doc. JVT-N013*, Jan. 2005.

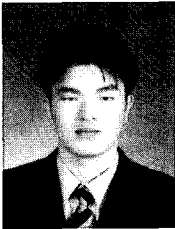
[10] G. Sullivan and G. Bjontegaard, "Recommended Simulation Common Conditions for H.26L Coding Efficiency Experiments on Low-resolution Progressive-Scan Source Material," *ITU-T Q.6/16, Doc. VCEG-N81*, Sept. 2001.

[11] G. Bjontegaard, "Calculation of Average PSNR Difference between RD-Curve," *ITU-T Q.6/16, Doc. VCEG-M33*, April 2001.

[12] Z. Chen, P. Zhou, and Y. He, "Fast Integer Pel and Frantional Pel Motion Estimation for JVT," *JVT Doc. JVT-F017*, Dec. 2002.

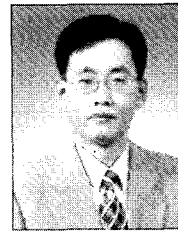
[13] Z. Chen, P. Zhou, and Y. He, "Fast Motion Estimation for JVT," *JVT Doc. JVT-G016*, March 2003.

저 자 소 개



김 근 용(학생회원)
 2004년 2월 광운대학교
 전자공학부 졸업(학사).
 2006년 2월 광주과학기술원 정보
 통신공학과 졸업(석사).
 2006년 2월~현재 한국전자통신
 연구원 광통신연구센터
 연구원

<주관심분야: H.264, IPTV, 영상신호 처리 및 압축, 디지털 신호처리>



호 요 성(평생회원)
 1981년 2월 서울대학교
 전자공학과 졸업(학사).
 1983년 2월 서울대학교 대학원
 전자공학과 졸업(석사).
 1983년 3월~1995년 9월 한국전자
 통신연구원 선임연구원.

1989년 12월 미국 University of California, Santa Barbara Department of Electrical and Computer Engineering(박사),
 1990년 1월~1993년 5월 미국 Philips 연구소 Senior Research Member.
 1995년 9월~현재 광주과학기술원 정보통신 공학과 교수.

<주관심분야: 디지털 신호처리, 영상신호 처리 및 압축, 초저속 영상통신, 디지털 TV와 고선명 TV, MPEG 표준, 다시점 비디오 부호화, 실감방송>