

비트율-왜곡값 예측을 이용한 H.264 고속 모드 결정 방법

김 승 환^o, 호 요 성
 광주과학기술원 정보통신공학과
 {kshkim, hoyo}@gist.ac.kr

요 약

최근에 만들어진 비디오 압축 표준인 H.264 방식은 비트율-왜곡(rate-distortion) 최적화 기법을 통해 매크로블록 단위로 가능한 모든 모드들을 고려하여 최적의 부호화 모드를 결정하므로, 입력 동영상을 부호화하는데 많은 시간이 걸린다. 따라서 H.264의 부호화 시간을 단축하기 위해 고속으로 모드를 결정하는 방법이 필요하다. 본 논문에서는 이전 화면에서 얻은 비트율-왜곡값의 통계적 특성을 이용하여 현재 부호화할 매크로블록의 비트율-왜곡값을 예측하고, 불필요한 움직임 예측과 모드 결정 과정을 생략하는 새로운 고속 모드 결정 방법을 제안한다. 또한 이전 화면의 동일한 위치에 있는 매크로블록의 모드와 비트율-왜곡값 특성을 반영하여, 조기 SKIP 모드 결정 방법과 조기 16×16 모드 결정 방법을 제안한다. 본 논문에서 제안한 고속 인터 모드 결정 방법이 H.264의 참조 소프트웨어에 구현된 고속 모드 결정 방법에 비해 부호화 효율적인 면에서 거의 동일한 성능을 유지하면서 전체 부호화 시간을 평균적으로 79%정도 감소시키는 것을 컴퓨터 모의실험을 통해 확인했다.

1. 서론

최근에 완성된 비디오 압축 표준인 H.264/AVC는 비디오 압축 성능을 높이기 위해 비트율-왜곡 최적화(Rate Distortion Optimization) 기반의 모드 결정 방법을 사용했다 [1]. H.264는 기본(Baseline) 프로파일에서 총 7가지 매크로블록 모드를 지원하는데, 다섯 개의 인터 모드(SKIP, 16×16, 16×8, 8×16과 P8×8)와 두 개의 인트라 모드(I4×4, I16×16)가 있다. 이 중 P8×8 모드는 각각의 8×8 블록 안에서 8×8, 8×4, 4×8, 4×4 중의 하나로 나뉜다.

H.264는 높은 압축 효율을 얻기 위해 앞서 언급한 매크로블록 모드들 가운데 최적의 매크로블록 모드를 결정하기 위해 비트율-왜곡 최적화 기법을 사용한다. 하지만, 매크로블록마다 각 모드별로 비트율-왜곡값을 얻기 위해 부호 및 복호 과정을 수행해야 하므로 매우 많은 계산적인 복잡도가 요구된다.

본 논문에서는 이미 부호화된 매크로블록에서 얻은 최적 모드의 통계적 분포와 각 모드별 평균 비트율-왜곡값을 이용하여 현재 부호화할 매크로블록에서 발생할 비트율-왜곡값을 예측한다. 이러한 예측 방법을 이용하여 불필요한 움직임 예측 및 모드 결정 과정을 생략하는 새로운 고속 모드 결정 방법을 제안한다. 본 논문에서 제안한 고속 모드 결정 방법은 두 단계의 조기 SKIP 모드 결정 방법과 조기 16×16 모드 결정 방법을 포함하고 있다.

2. H.264 모드 결정

H.264 참조 소프트웨어에서 움직임 벡터와 참조 영상을 결정하기 위해 식 (1)과 같은 비트율-왜곡 최적화 기법을 사용한다.

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

여기서 λ_{motion} 는 Lagrangian 계수이고, $0.85 \times 2^{0.75}$ 의 제곱근 값을 갖는다. Q는 양자화 계수이다. $R(MV, REF)$ 는 움직임 벡터와 참조 영상을 부호화하는데 필요한 비트수이다. $SAD(s, r(MV, REF))$ 는 원 영상과 움직임 예측에 의해 복원된 영상과의 차에 대한 절대값의 합이다. 또한, 최적의 모드 결정은 식 (2)의 비트율-왜곡값을 비교하여 결정된다.

$$J_{mode}(s, r, M | \lambda_{mode}) = SSD(s, r, M) + \lambda_{mode} \cdot R(s, r, M) \quad (2)$$

여기서 λ_{mode} 는 λ_{motion} 의 제곱 값이다. M은 매크로블록 모드이다. $R(s, r, M)$ 은 M에 해당하는 모드를 부호화했을 때, 실제로 발생하는 비트수이며 $SSD(s, r, M)$ 는 원 영상과 복호된 영상과의 차이에 대한 제곱의 합이고, 이는 다음과 같이 계산된다.

$$SSD(s, r, M) = \sum_{x=0}^{H-1} \sum_{y=0}^{V-1} (s(x, y) - r(x - m_x, y - m_y))^2 \quad (3)$$

비트율-왜곡값 예측을 이용한 H.264 고속 모드 결정 방법

김 승 환^o, 호 요 성
 광주과학기술원 정보통신공학과
 {kshkim, hoyo}@gist.ac.kr

요 약

최근에 만들어진 비디오 압축 표준인 H.264 방식은 비트율-왜곡(rate-distortion) 최적화 기법을 통해 매크로블록 단위로 가능한 모든 모드들을 고려하여 최적의 부호화 모드를 결정하므로, 입력 동영상을 부호화하는데 많은 시간이 걸린다. 따라서 H.264의 부호화 시간을 단축하기 위해 고속으로 모드를 결정하는 방법이 필요하다. 본 논문에서는 이전 화면에서 얻은 비트율-왜곡값의 통계적 특성을 이용하여 현재 부호화할 매크로블록의 비트율-왜곡값을 예측하고, 불필요한 움직임 예측과 모드 결정 과정을 생략하는 새로운 고속 모드 결정 방법을 제안한다. 또한 이전 화면의 동일한 위치에 있는 매크로블록의 모드와 비트율-왜곡값 특성을 반영하여, 조기 SKIP 모드 결정 방법과 조기 16×16 모드 결정 방법을 제안한다. 본 논문에서 제안한 고속 인터 모드 결정 방법이 H.264의 참조 소프트웨어에 구현된 고속 모드 결정 방법에 비해 부호화 효율적인 면에서 거의 동일한 성능을 유지하면서 전체 부호화 시간을 평균적으로 79%정도 감소시키는 것을 컴퓨터 모의실험을 통해 확인했다.

1. 서론

최근에 완성된 비디오 압축 표준인 H.264/AVC는 비디오 압축 성능을 높이기 위해 비트율-왜곡 최적화 (Rate Distortion Optimization) 기반의 모드 결정 방법을 사용했다 [1]. H.264는 기본 (Baseline) 프로파일에서 총 7가지 매크로블록 모드를 지원하는데, 다섯 개의 인터 모드 (SKIP, 16×16, 16×8, 8×16과 P8×8)와 두 개의 인트라 모드 (I4×4, I16×16)가 있다. 이 중 P8×8 모드는 각각의 8×8 블록 안에서 8×8, 8×4, 4×8, 4×4 중의 하나로 나뉜다.

H.264는 높은 압축 효율을 얻기 위해 앞서 언급한 매크로블록 모드들 가운데 최적의 매크로블록 모드를 결정하기 위해 비트율-왜곡 최적화 기법을 사용한다. 하지만, 매크로블록마다 각 모드별로 비트율-왜곡값을 얻기 위해 부호 및 복호 과정을 수행해야 하므로 매우 많은 계산적인 복잡도가 요구된다.

본 논문에서는 이미 부호화된 매크로블록에서 얻은 최적 모드의 통계적 분포와 각 모드별 평균 비트율-왜곡값을 이용하여 현재 부호화할 매크로블록에서 발생할 비트율-왜곡값을 예측한다. 이러한 예측 방법을 이용하여 불필요한 움직임 예측 및 모드 결정 과정을 생략하는 새로운 고속 모드 결정 방법을 제안한다. 본 논문에서 제안한 고속 모드 결정 방법은 두 단계의 조기 SKIP 모드 결정 방법과 조기 16×16 모드 결정 방법을 포함하고 있다.

2. H.264 모드 결정

H.264 참조 소프트웨어에서 움직임 벡터와 참조 영상을 결정하기 위해 식 (1)과 같은 비트율-왜곡 최적화 기법을 사용한다.

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

여기서 λ_{motion} 는 Lagrangian 계수이고, 0.85×2^{Q3} 의 제곱근 값을 갖는다. Q는 양자화 계수이다. $R(MV, REF)$ 는 움직임 벡터와 참조 영상을 부호화하는데 필요한 비트수이다. $SAD(s, r(MV, REF))$ 는 원 영상과 움직임 예측에 의해 복원된 영상과의 차에 대한 절대값의 합이다. 또한, 최적의 모드 결정은 식 (2)의 비트율-왜곡값을 비교하여 결정된다.

$$J_{mode}(s, r, M | \lambda_{mode}) = SSD(s, r, M) + \lambda_{mode} \cdot R(s, r, M) \quad (2)$$

여기서 λ_{mode} 는 λ_{motion} 의 제곱근 값이다. M은 매크로블록 모드이다. $R(s, r, M)$ 은 M에 해당하는 모드를 부호화했을 때, 실제로 발생하는 비트수이며 $SSD(s, r, M)$ 는 원 영상과 복호된 영상과의 차이에 대한 제곱의 합이고, 이는 다음과 같이 계산된다.

$$SSD(s, r, M) = \sum_{x \in H, y \in V} (s(x, y) - r(x - m_x, y - m_y))^2 \quad (3)$$

3. 비트율-왜곡 예측 기반 모드 결정

본 논문에서는 비트율-왜곡 예측을 이용한 고속 모드 결정 방법을 제안한다. 그림 1은 비트율-왜곡 최적화 방법을 통해 각 매크로블록을 최적의 모드로 부호화했을 때 각 모드별로 양자화 계수와 모드에 따라 변하는 평균 비트율-왜곡값의 분포를 나타낸 것이다. 그림 1에서 알 수 있듯이, 최종 결정모드에 따라 비트율-왜곡값의 분포 특성이 다르며, 특히 SKIP 모드의 경우 가장 작은 비용값을 나타냈고, P8×8 모드와 인트라 4×4(I4×4) 모드의 경우 SKIP 모드에 비해 평균 5배 이상의 큰 비용값을 나타냈다.

그림 2는 양자화 계수에 따른 각 모드별 발생 빈도수를 나타낸 것인데, 일반적으로 SKIP 모드의 발생 빈도수가 가장 많으며, 그 다음으로 16×16 모드의 빈도수가 많음을 알 수 있다. 특히, SKIP 모드와 16×16 모드의 발생 빈도수가 대략 전체의 80%임을 고려하면, 초기 SKIP 모드 결정과 초기 16×16 모드의 결정이 고속 모드 결정에 매우 중요한 부분임을 알 수 있다. 따라서 본 논문에서는 이러한 특성들을 고려하여 각각 두 단계의 초기 SKIP 모드 결정 과정과 초기 16×16 모드 결정 방법을 제안하였다.

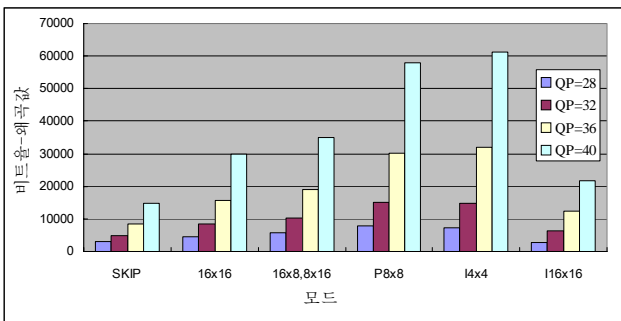


그림 1. 모드에 따른 평균 비트율-왜곡값 분포

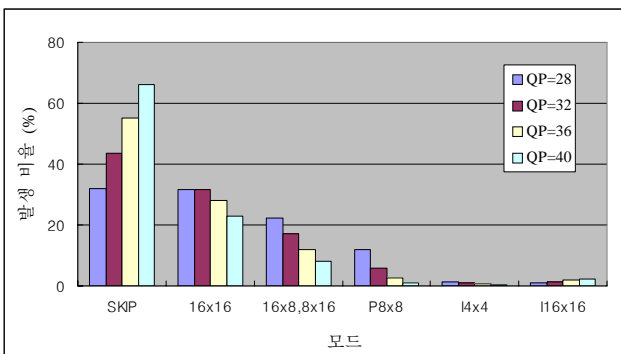


그림 2. 양자화 계수에 따른 모드별 발생 빈도 수

제안한 고속 모드 결정 알고리즘은 두 단계의 초기 SKIP 모드 결정 과정을 포함한다. 첫 번째 단계의 초기 SKIP 모드 결정에서는 이전 화면에서의 평균 SKIP 왜곡값과 이전 화면의 동일한 위치에 있는 매크로블록의 왜곡값을 이용하여 초기 SKIP 모드 여부를 결정한다.

그림 3에 나타난 것과 같이, SKIP 모드의 비트율 값이 왜곡값에 비해 매우 적으므로 SKIP 모드의 비트율-왜곡값이 아닌 왜곡값만을 이용하였다. 다음은 초기 SKIP 모드 결정을 위한 조건식이다.

$$D_c(SKIP|QP) < \delta \times \left\{ \frac{\alpha \times \overline{D_p(SKIP|QP)} + D_p(M|QP)}{\alpha + 1} \right\} \quad (4)$$

여기서 $D_c(SKIP|QP)$ 는 현재 부호화할 매크로블록의 SKIP 왜곡값, $\overline{D_p(SKIP|QP)}$ 은 이전 화면의 평균 SKIP 왜곡값을 나타낸다. α 는 이전 화면의 동일한 위치에 있는 매크로블록에서의 왜곡값과 이전 화면의 평균 SKIP 왜곡값 사이의 가중을 결정하는 상수 값을 나타내며, δ 는 제안한 고속 모드 결정 알고리즘에서 복잡도와 부호화 효율을 조절하기 위한 상수 값이다. 따라서 δ 값이 커짐에 따라 복잡도는 감소되며 부호화 효율은 증가한다.

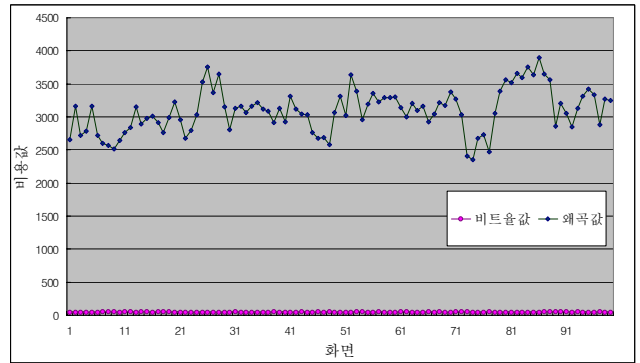


그림 3. 비트율값 및 비용값 (SKIP 모드)

두 번째 단계의 초기 SKIP 모드 결정은 16×16 모드의 움직임 벡터를 찾은 후 SKIP 모드의 움직임 벡터와 16×16 모드의 움직임 벡터가 동일하며 SKIP 모드의 비트율-왜곡값이 16×16 모드의 비트율-왜곡값보다 작을 경우 초기 SKIP 모드를 결정한다.

- (1) 움직임 보상 블록의 크기가 16×16
- (2) SKIP 움직임 벡터와 16×16 움직임 벡터가 동일
- (3) 참조 영상은 바로 이전 영상
- (4) $J_c(SKIP) < J_c(16 \times 16)$

위에서 언급한 두 단계의 초기 SKIP 모드 조건을 만족하지 않는 경우에는 두 단계의 초기 16×16 모드 결정 여부를 체크한다. 첫 번째 초기 16×16 모드 결정은 현재 16×16 모드의 비트율-왜곡값과 이전 화면에서의 평균 16×16 모드의 비트율-왜곡값을 고려하여 식 (5)와 같이 결정하며, 두 번째 초기 16×16 모드 결정은 식 (6)을 이용하여 결정한다.

$$J_c(16 \times 16) < \delta \cdot \overline{J_p(16 \times 16)} \quad (5)$$

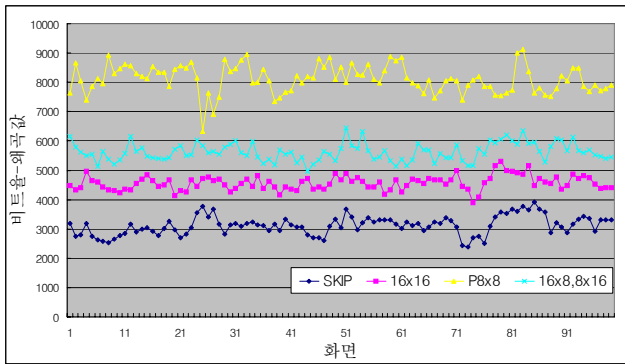
$$J_c(16 \times 16) < J_c(16 \times 8) \ \&\& \ J_c(16 \times 16) < J_c(8 \times 16) \quad (6)$$

조기 16×16 모드 결정 단계 이후에는 현재 얻어진 최소 비트율-왜곡값과 이전 화면에서 통계적 관찰을 통해 예측된 P8×8 모드에 대한 평균 비트율-왜곡값을 이용하여 P8×8 모드 결정 과정의 생략 여부를 결정한다. 이는, 그림 4 (a)에 나타난 바와 같이, P8×8 모드의 비트율-왜곡값이 다른 모드들에 비해 매우 큰 값을 갖고, P8×8 모드의 발생 확률 또한 매우 낮아 식 (7)을 이용해 효율적으로 해당 모드의 유효성을 검사한다.

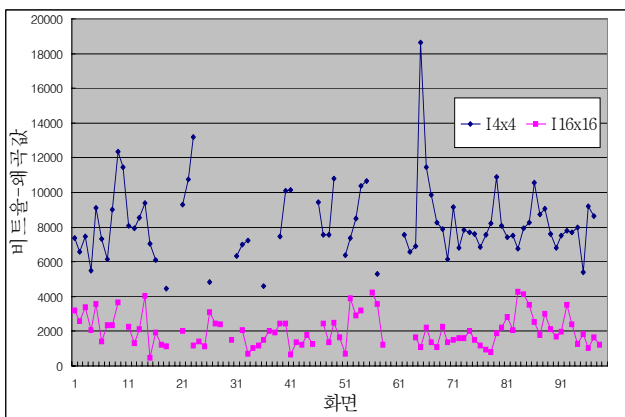
$$Min_RDCost \{16 \times 8, 8 \times 16\} < \delta \cdot \overline{J_p(P8 \times 8)} \quad (7)$$

인트라 4×4 모드(I4×4)의 경우도, P8×8 모드와 같이, 다른 모드들에 비해 더 많은 헤더 비트를 요구하므로 상대적으로 다른 모드의 비트율-왜곡값보다 큰 값을 갖는 동시에 발생 확률도 매우 낮다. 이러한 점을 고려하여 현재 얻어진 최적의 비트율-왜곡값이 식 (8)에 설정된 임계값보다 작은 경우 I4×4 모드에 대한 비트율-왜곡 예측 과정을 생략하는 방법을 제안했다. 그림 4(b)에는 인트라 모드의 비트율-왜곡값을 나타냈다.

$$Min_RDCost \{16 \times 8, 8 \times 16, P8 \times 8\} < \delta \cdot \overline{J_p(I4 \times 4)} \quad (8)$$



(a) 평균 비트율-왜곡값 (인터모드)



(b) 평균 비트율-왜곡값 (인트라 모드)

그림 4. 모드별 비트율-왜곡값의 확률 분포

그림 5에는 본 논문에서 제안한 고속 모드 결정 방법의 전체적인 구조를 도식적으로 나타냈다.

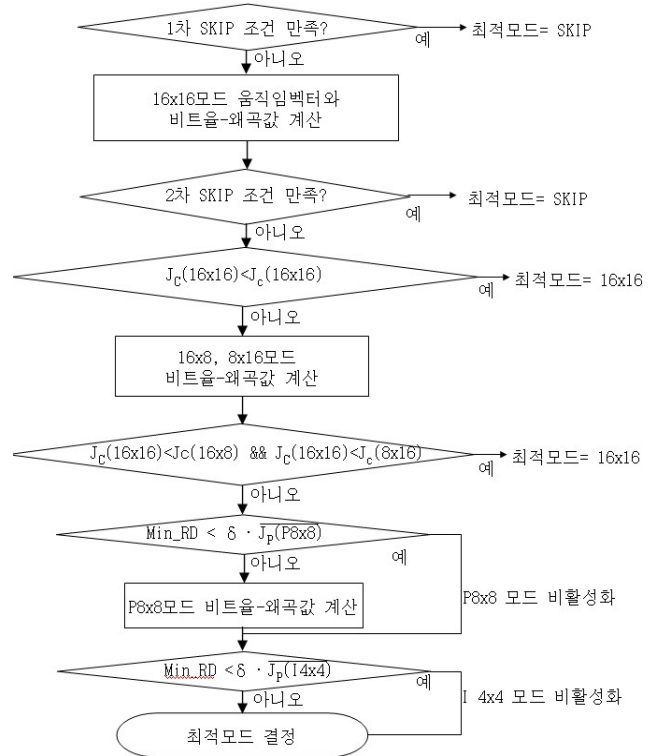


그림 5. 제안한 고속 모드 결정 방법

4. 실험 결과 및 분석

본 논문에서는 JM 12.4에서 제안한 고속 모드 결정 방법을 구현했다. 이 방법의 효율을 보이기 위해, 100 프레임을 갖는 해상도 CIF(352×288) 크기의 각기 다른 9개의 비디오 시퀀스를 이용하여, IPPP 구조의 베이스라인(Baseline) 프로파일로 부호화했다. 움직임 예측에 2개의 참조영상을 사용했으며, 움직임 예측을 위한 탐색 범위는 16이다. 비트율-왜곡 곡선을 유도하기 위해 4개의 양자화 계수 (28, 32, 36, 40)를 사용했다.

제안한 고속 모드 결정 방법과 H.264 참조 소프트웨어의 고속 모드 결정 방법의 성능 비교를 위해 평균 PSNR과 비트율을 사용했다 [3]. 실험은 고속 복잡도 모드(Fast High Complexity Mode)에서 수행되었으며, 수행시간은 각 양자화 계수(28, 32, 36, 40)별 부호화 시간의 평균으로 계산했다 [4] [5].

<표 1>부터 <표 4>까지는 양자화 계수에 따른 성능을 비교하기 위해 실험한 결과를 나타낸다. <표 1>부터 <표 4>에서 알 수 있듯이, 제안한 고속 인터 모드 결정 방법은 JM 참조 소프트웨어에 구현된 고속 모드 결정 방법에 비해 평균 79%정도 부호화 시간을 단축시켰다. 또한, PSNR과 발생 비트율에 대해서도 거의 동일한 성능을 나타냈다. 특히, 각 실험 영상마다 모드별 발생 빈도와 비트율-왜곡값에 있어서 확률 분포의 차이 때문에 부호화 시간 감소의 비율도 각기 다르게 나타남을 확인했다.

<표 1> 영상에 따른 성능변화 (QP=28)

영상	△PSNR(dB)	△비트율(%)	△T(%)
FOREMAN	-0.11	1.52	-72.72
NEWS	0.02	0.71	-86.96
AKIYO	0.06	-1.60	-90.51
FOOTBALL	-0.01	3.01	-59.27
COASTGUARD	0.02	0.71	-79.55
SILENT	-0.09	-0.73	-81.13
CONTAINER	-0.02	-3.86	-90.57
PARIS	0.11	1.17	-80.35
MOBILE	-0.07	-2.15	-78.08
평균	-0.01	-0.13	-79.91

<표 2> 영상에 따른 성능변화 (QP=32)

영상	△PSNR(dB)	△비트율(%)	△T(%)
FOREMAN	0.02	0.34	-74.13
NEWS	-0.07	-1.92	-84.40
AKIYO	0.05	0.11	-89.60
FOOTBALL	-0.01	2.47	-59.48
COASTGUARD	0.06	1.01	-77.30
SILENT	-0.02	1.54	-83.01
CONTAINER	0.06	0.12	-89.13
PARIS	0.00	1.10	-78.84
MOBILE	-0.09	0.52	-78.40
평균	0.00	0.59	-79.36

<표 3> 영상에 따른 성능변화 (QP=36)

영상	△PSNR(dB)	△비트율(%)	△T(%)
FOREMAN	0.06	0.22	-75.55
NEWS	-0.08	-0.81	-83.50
AKIYO	-0.01	-1.10	-87.60
FOOTBALL	-0.04	1.30	-59.67
COASTGUARD	-0.06	-1.58	-78.27
SILENT	-0.03	0.36	-82.24
CONTAINER	0.02	0.39	-88.34
PARIS	-0.07	-1.71	-76.38
MOBILE	-0.02	1.52	-77.36
평균	-0.02	-0.16	-78.99

<표 4> 영상에 따른 성능변화 (QP=40)

영상	△PSNR(dB)	△비트율(%)	△T(%)
FOREMAN	-0.01	-0.12	-75.03
NEWS	-0.17	-2.30	-82.00
AKIYO	-0.01	0.81	-86.95
FOOTBALL	-0.07	-0.01	-53.84
COASTGUARD	-0.10	-1.34	-79.15
SILENT	-0.04	1.46	-81.36
CONTAINER	-0.06	0.22	-87.45
PARIS	-0.06	-1.88	-75.24
MOBILE	0.02	1.09	-77.91
평균	-0.06	-0.23	-77.66

5. 결론

본 논문에서는 주어진 영상을 부호화할 때 이전 화면의 매크로블록들의 모드 발생 빈도, 비트율-왜곡값 및 이전 화면에서 동일한 위치에 있는 매크로블록의 비트율-왜곡비용 값을 고려하여, 불필요한 움직임 예측 및 모드 결정 과정을 생략하는 새로운 고속 모드 결정 방법을 제안했다. 제안한 고속 모드 결정 방법에서는 비트율-왜곡값 예측을 통해 두 단계의 조기 SKIP 모드 결정과 두 단계의 조기 16x16 모드 결정 과정을 제안했다. 또한, P8x8 모드와 인트라 4x4 모드와 같이 많은 계산적 복잡도를 요구하지만 발생 빈도가 낮은 모드들을 위해 통계적 비트율-왜곡값 예측 방법을 통한 모드 생략 방법을 제안하였다. 컴퓨터 모의실험 결과를 통해 보인 것처럼, 제안한 고속 모드 결정 방법은 H.264 고속 모드 결정 방법과 비교하여, PSNR과 발생 비트율에서 거의 동일한 성능을 유지하면서 평균 79%의 부호화 시간을 감소시켰다.

감사의 글

본 연구는 광주과학기술원(GIST) 실감방송연구센터(RBRC)를 통한 정보통신대학 IT연구센터(ITRC)의 지원에 의한 것입니다.

참고문헌

- [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-G050, 2003.
- [2] JVT reference software version 12.4, available online at: http://iphome.hhi.de/suehring/tml/download/old_jm/
- [3] G. Bjontegaard, "Calculation of Average PSNR Difference between RD-curve," ITU-T Q.6/16, Doc. VCEG-M33, April 2001.
- [4] Z. Chen, P. Zhou, and Y. He, "Fast Motion Estimation for JVT," JVT-G016, March 2003.
- [5] G. Y. Kim, B. Y. Yoon, and Y. S. Ho, "A Fast Inter Mode Decision Algorithm in H.264/AVC for IPTV Broadcasting Services," Proc. VCIP 2007, Jan. 2007.