

# H.264 표준을 위한 고속 모드 결정 방법

김희순, 김승환, 호요성  
광주과학기술원 정보통신공학과

## A Fast Mode Decision Method for H.264

Hee-Soon Kim, Seung-Hwan Kim, Yo-Sung Ho  
Gwangju Institute of Science and Technology (GIST)  
{[hskim03](mailto:hskim03@gist.ac.kr), [kshkim](mailto:kshkim@gist.ac.kr), [hoyo](mailto:hoyo@gist.ac.kr)}@gist.ac.kr

**Abstract:** H.264는 기존의 비디오 압축 표준들보다 높은 압축 효율을 제공하는데, 이는 증가된 모드 수와 복잡한 비트-왜곡 최적화 기법을 사용하는 모드 결정 방법에 기인한다. 본 논문에서는 각 모드의 발생 빈도와 복잡도를 고려하여 모드를 분류하고, 분류된 모드에 따라 선택적으로 비트-왜곡 최적화 과정을 적용한다. 제안한 알고리즘은 H.264 표준에 비해 부호화 시간을 평균 29.65%, 비트-왜곡 계산 복잡도를 89.12% 정도 감소시켰다.

**Keywords:** H.264, 고속모드결정, 비트-왜곡최적화

### I. 서론

비디오 신호의 부호화 효율을 높이기 위해, H.264 표준은 다양한 매크로블록 모드, 가변 블록 크기, 다수의 참조 영상 등과 같은 많은 부호화 매개 변수를 제공한다 [1-2]. 부호화 과정 중에, 이들 매개 변수를 포함하는 매크로블록 모드 중에 최적의 모드를 결정하는 것은 가장 중요하며 복잡하다.

H.264는 7가지 매크로블록 모드 중에 5가지의 SKIP, 16x16, 16x8, 8x16와 P8x8을 인터 모드에 할당한다. 또한, 이 중에 P8x8은 각각의 8x8 블록 안에 8x8, 8x4, 4x8와 4x4의 하위 매크로블록 모드 중에 하나를 선택할 수 있다. 그림 1에서 인터 모드와 하위 매크로블록 모드를 보여주고 있다.

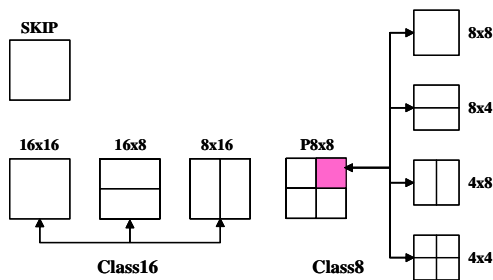


그림 1. 인터 모드

인터 모드와는 별도로, H.264는 두 가지 매크로블록 모드 Intra4x4와 Intra16x16을 인트라 모드로 제공한다. 인트라 모드는 이전에 부호화된 유효한 매크로블록의 경계 화소를 이용해 현재 매크로블록을 방향성 예측 한다. Intra4x4는 4x4 블록 단위로 현재 매크로블록을 예측하고, 각각의 4x4 블록당 9가지 예측 방향을 제공한다. Intra16x16은 매크로 블록 단위로 현재 매크로블록을 예측하고, 4가지의 예측 방향을 제공한다.

인터 모드에서 색차 신호는 휘도 신호의 움직임 벡터를 이용해 예측되지만, 인트라 모드에서 색차 신호는 휘도 신호와 별도로 예측된다. 인트라 모드에 대한 색차 신호 예측은 8x8 블록 단위로 이루어지며, 각 8x8 블록당 4가지 예측 방향을 제공한다.

### II. H.264 매크로블록 모드 결정 방법

높은 압축 효율을 얻기 위해 다양한 매크로블록 모드 중에 가장 최적의 매크로블록 모드를 결정하는 것이 부호화의 핵심이다. 최적의 매크로블록 모드를 결정하기 위해 H.264는 비트-왜곡 최적화 기법을 사용한다 [3]. 이 비트-왜곡 최적화 기법은 각 매크로블록 모드에 대한 비트-왜곡 비용 (RDcost)을 구해야 한다. 인터 모드는 RDcost를 구하기 전에 움직임 벡터와 참조 영상을 결정해야 한다. 인터 모드에서 가변 블록에 대한 움직임 벡터와 참조 영상은 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  $\lambda_{motion}$ 는 Lagrangian 계수이고,  $0.85 \times 2^{Q/3}$ 의 제곱근 값을 갖는다. Q는 양자화 계수이다.  $R(s, r(MV, REF))$ 는 움직임 벡터와 참조 영상을 부호화하는데 필요한 비트량이고,  $SAD(s, r(MV, REF))$ 는 원 영상과 움직임 예측에 의해 복원된 영상과의 차에 대한 절대값의 합이고, 다음과 같이 계산된다.

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

여기서 H와 V는 각각 가변 블록의 너비와 높이이다.  $m_x$ 와  $m_y$ 는 움직임 벡터이고,  $m_t$ 는 참조 영상을 나타낸다.

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

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

여기서  $\lambda_{mode}$ 는  $\lambda_{motion}$ 의 제곱 값이다. M은 매크로블록 모드, 예측 방향 모드 또는 하위 매크로블록 모드이다.  $R(s, r, M)$ 은 M에 해당하는 모드를 부호화 했을 때 실질적

으로 발생하는 비트량이다.  $SSD(s,r;M)$ 는 원 영상과 복호된 영상과의 차이에 대한 제곱의 합이고, 다음과 같이 계산된다.

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

식 (3)과 식 (4)에서 매크로블록 모드와 다른 두 가지 모드의 차이점은 매크로블록 모드는 휘도와 색차 성분 모두를 고려하지만 다른 두 가지 모드는 휘도 성분만을 고려한다. 위와 같이 H.264는 모든 매크로블록 모드와 이와 연관된 매개 변수에 대한 RDcost를 구하고, 그 중에 최소값을 갖는 매크로블록 모드를 최적의 매크로블록 모드로 결정한다.

### III. 제안된 고속 모드결정 방법

표 1은 첫 번째 프레임을 제외한 100 프레임을 부호화 했을 때, 두 영상에 대한 각 매크로블록 모드(모드 번호는 순서대로 SKIP, 16x16, 16x8, 8x16, P8x8, Intra4x4, Intra16x16)의 발생빈도수이다.

표 1. 매크로블록 모드의 발생 빈도수

Mobile					Silent				
모드	양자화 파라미터				모드	양자화 파라미터			
	28	32	36	40		28	32	36	40
0	379	923	1960	3802	0	6320	6924	7603	8297
1	2354	3007	3941	4055	1	1183	1208	1073	825
2	1456	1648	1310	745	2	485	409	322	194
3	1399	1548	1353	737	3	607	532	414	282
8	4210	2671	1233	401	8	1103	628	268	78
9	0	0	0	0	9	92	73	63	36
10	3	4	4	1	10	11	27	58	89

표 1에서, 각 매크로블록 모드의 발생 빈도수는 영상의 특성과 양자화 계수에 따라 다름을 확인할 수 있다. 이러한 매크로블록 모드의 발생 빈도수를 고려하여, 모든 매크로블록 모드에 대해 비트-왜곡 최적화 과정을 수행하는 것 대신 매크로블록 모드를 4가지 {SKIP, Class16 (16x16, 16x8, 8x16), P8x8, 인트라 모드}로 분류하여 선택적으로 수행한다. 여기서, 매크로블록이 4개의 8x8블록으로 나뉘어진 매크로블록 모드를 기본 모드라고 정의한다.

#### 1. 조기 SKIP 모드 결정

SKIP 모드는 움직임 예측이나 잔여 영상에 대한 부호화가 필요하지 않기 때문에 7가지 매크로블록 모드 중에 가장 복잡도가 낮다. 이러한 중요성 때문에, SKIP 모드가 조기에 결정된다면 복잡도를 상당히 줄일 수 있다. 현재의 매크로블록 모드가 SKIP 모드인지 아닌지에 대한 결정은 식 (5)와 같이 SKIP 모드와 기본(Base) 모드에 대한 RDcost의 비교를 통해서 이루어진다.

$$J_{mode}(Base) > J_{mode}(SKIP) \quad (5)$$

만약 식 (5)가 만족된다면, 이것은 움직임 예측을 통한 부호화 이득이 없다는 것을 의미한다. 왜냐하면, 기본 모드는 4개의 8x8 블록으로 나뉘어짐으로써 Class16 모드들 보다 더 복잡한 움직임 예측을 할 수

있다. 하지만, 이러한 복잡한 움직임 예측에도 불구하고 식 (5)가 만족됐다는 것은 움직임 벡터를 위한 비트량만 늘어났을 뿐 SKIP 모드의 왜곡에는 큰 변화가 없기 때문이다.

만약 식 (5)가 만족되지 않는다면, 이것은 움직임 예측이 SKIP 모드보다 효율적이라고 본다. 하지만 매크로블록에서 움직임의 복잡도는 예측하기 힘들다. 따라서 만약 식 (5)가 만족되지 않는다면, 기본 모드뿐만 아니라 다른 인트라 모드도 고려해야 한다.

#### 2. 선택적 인트라 매크로블록 모드 결정

P8x8 모드는 Class16 모드들 보다 더 복잡한 움직임 예측을 수행하기 때문에 더 높은 복잡도를 요구한다. 하지만 표 1이 보여주는 것처럼 P8x8 모드의 발생 빈도수는 Class16 모드의 발생 빈도수 보다 적다. P8x8 모드의 높은 복잡도와 적은 발생 빈도수에도 불구하고, 모든 매크로블록에 대해 P8x8의 RDcost를 계산한다는 것은 비효율적이다. 따라서, 본 논문은 P8x8 모드를 Class16 모드들과 식 (6)의 RDcost 비교를 통해 비트-왜곡 최적화 대상에서 제외한다.

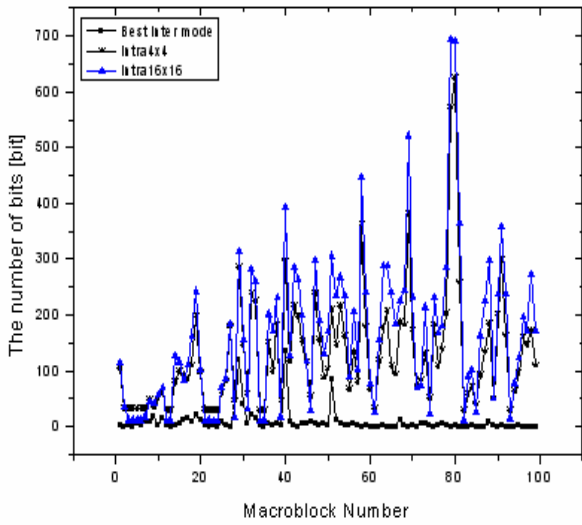
$$J_{mode}(Best16) > J_{mode}(Base) \quad (6)$$

Class16 모드 중에서 최적의 매크로블록 모드 (Best16)와 RDcost를 결정한다. 이 최적의 매크로블록 모드의 RDcost를 식 (5) 계산 중에 구한 기본 모드의 RDcost와 비교한다. 만약 식 (6)이 만족되지 않는다면, 이것은 Class16 보다 더 복잡한 움직임 예측이 비효율적임을 의미한다. 왜냐하면, 기본 모드는 Class16 모드 보다 더 복잡한 움직임 예측을 하지만 하위 매크로블록 모드의 어떤 조합으로 이루어진 P8x8 모드 보다는 같거나 더 단순한 움직임 예측을 하기 때문이다. 따라서 만약 식 (6)이 만족되지 않는다면, 비트-왜곡 최적화 과정에서 P8x8 매크로블록 모드를 제외시킨다. 만약 식 (6)이 만족되면, 현재의 가변 블록은 더 복잡한 움직임 예측을 요구한다는 것을 의미한다. 따라서 P8x8 모드를 비트-왜곡 최적화 과정에 포함시킨다.

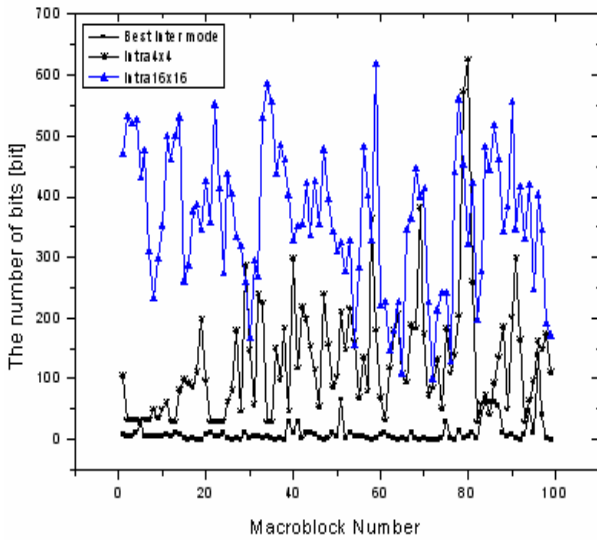
식 (6)의 조건에 따라 P8x8을 선택적으로 비트-왜곡 최적화 과정에서 제외시킴으로써, P8x8에 대한 움직임 예측 과정 또한 생략될 수 있고 부호화 시간을 줄일 수 있다. 식 (6)의 조건에 따라 선택적으로 인트라 모드에 대해 비트-왜곡 최적화 과정을 수행하고 최적의 인트라 모드를 결정한 후에, 선택적 인트라 매크로블록 모드 결정 과정으로 진행된다.

#### 3. 선택적 인트라 매크로블록 모드 결정

표 1에서 보인 것처럼, 인트라 모드의 발생 빈도수는 다른 매크로블록 모드에 비해 매우 미비하다. 인트라 모드의 발생 빈도수가 미비하다고 해서 인트라 프레임에서 인트라 모드를 생략한다면 공간 상관도가 높은 특성을 지닌 영상에서 영상 전체의 비트-왜곡 특성은 나빠질 것이다. 따라서 선택적으로 인트라 모드를 비트-왜곡 최적화 과정에서 제외하는 것이 바람직하다. 선택적 인트라 모드 결정을 위해, 본 논문은 최적의 인트라 모드와 인트라 모드들 사이의 왜곡-비트 특성을 이용한다. 그림 2는 최적의 인트라 모드와 인트라 모드들 사이의 발생 비트량 관계를 하나의 인트라 프레임에 대해 보여주고 있다.



(a) News 영상



(b) Mobile 영상

그림 2. 발생 비트량 비교

그림 2에 보인 것처럼, 최적의 인터 모드에 대한 발생 비트량은 인트라 모드의 발생 비트량 보다 작거나 비슷하다. 따라서 왜곡-비트 곡선에서 만약 인트라 모드의 왜곡이 최적의 인터 모드에 대한 왜곡보다 작다면 인트라 모드의 RDcost가 최적의 인터 모드에 대한 RDcost보다 더 작은 경우가 있을 수 있다. 하지만 반대의 경우에는 최적의 인터 모드에 대한 RDcost가 인트라 모드의 RDcost 보다 항상 작다.

인트라 모드의 실질적인 발생 비트량은 인트라 모드를 부호화하지 않는 이상 알 수 없다. 하지만 위의 최적의 인터 모드와 인트라 모드의 발생 비트량 관계를 고려하여 더 이상 RDcost에서 발생 비트량은 고려하지 않는다. 최적의 인터 모드에 대한 왜곡은 이미 선택적 인터 매크로블록 모드 결정 과정에서 구했다. 인트라 모드의 왜곡 ( $MSBE$ )은, 그림 3에 보인 것처럼 부호화

될 매크로블록과 이미 부호화된 왼쪽과 오른쪽 매크로블록의 유효한 경계 화소를 이용하여 간략화 한다.

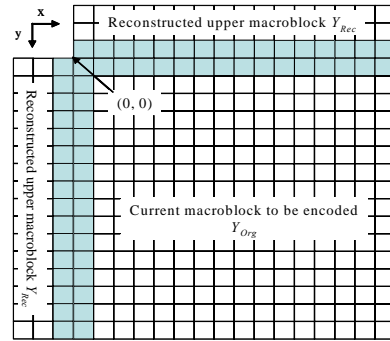


그림 3. 경계 화소들

인트라 모드에 대한 왜곡은 식 (7)에 의해 구해진다.

$$MSBE = \frac{W_{up}}{16} \cdot \sum_{i=0}^{15} [Y_{Org}(x+i, y) - Y_{Rec}(x+i, y-1)]^2 + \frac{W_{left}}{16} \cdot \sum_{i=0}^{15} [Y_{Org}(x, y+i) - Y_{Rec}(x-1, y+i)]^2 \quad (7)$$

여기서  $W_{up}$ 와  $W_{left}$  은 0 또는 1이 될 수 있고, 이들 값들은 경계 화소의 유효성 여부에 달려있다. 두 값이 0이면 인트라 모드는 고려하지 않는다.

만약  $MSBE$ 가 최적의 인터 모드에 대한 평균 왜곡 (ME)보다 작으면 인트라 모드에 대한 비트-왜곡 최적화 과정을 수행하고, 그렇지 않으면 최적의 인터 모드가 최종 매크로블록 모드가 된다. 그림 4에 고속 모드 결정에 관한 전체 흐름도를 요약한다.

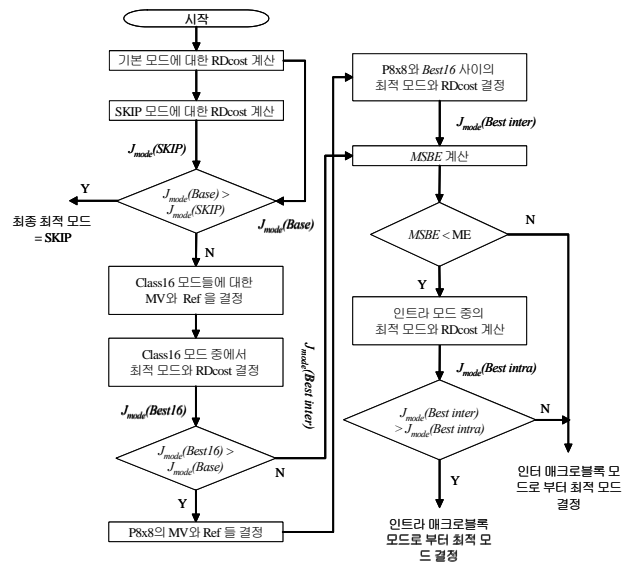


그림 4. 고속 모드 결정에 대한 전체 흐름도

#### IV. 실험 결과

본 논문에서는 JM 6.1에 제안한 알고리즘을 구현하였다. 제안한 알고리즘의 효율성을 보이교자, QCIF 해상도 영상 100 프레임에 갖는 6개의 각기 다른 비디오 시퀀스를 이용 하였다. 프레임 율은 30fps이고 IPPP 구

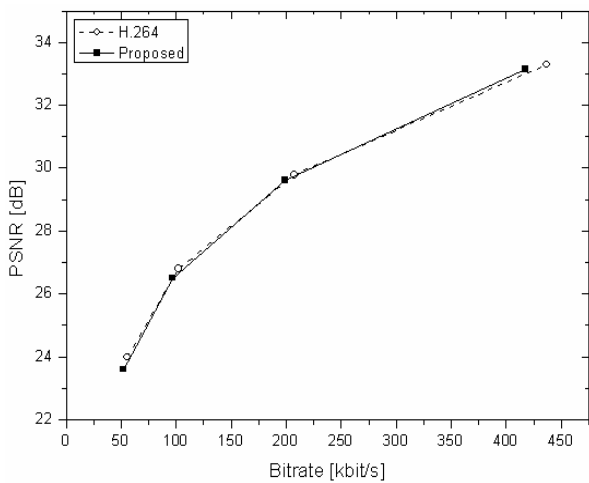
조의 Baseline 프로파일을 이용했다. 움직임 예측에서 참조 영상은 4개를 사용했으며, 움직임 예측의 탐색 범위는  $\pm 32$ 이다. 비트-왜곡 곡선을 유도하기 위해 실험은 4개의 양자화 계수에 대해 수행되었다.

H.264와 제안된 알고리즘 사이의 성능 비교 호화 시간차( $\Delta T$ )와 비트-왜곡 계산복잡도의 차( $\Delta N$ )를 이용하였다.

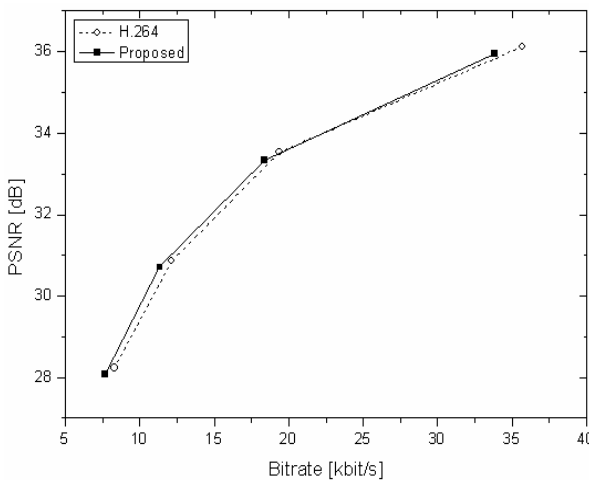
$$\Delta N = \frac{\#N[H.264] - \#N[proposed]}{\#N[H.264]} \times 100[\%] \quad (8)$$

$$\Delta T = \frac{T[H.264] - T[proposed]}{T[H.264]} \times 100[\%]$$

비트-왜곡 비용계산은 4x4 블록의 정수 DCT를 수행하는 순간마다 요구된다. 따라서 비트-왜곡 계산 복잡도는 4x4 블록의 정수 DCT를 기본단위로 한다.



(a) Mobile 영상



(b) Container 영상

그림 5. 비트-왜곡 곡선

H.264와 제안된 알고리즘 사이의 PSNR와 비트율 비교는 그림 5와 같은 두 비트-왜곡 곡선 사이의 차이에 근거하여 계산된다 [4]. 나머지 4개의 비트-왜곡 곡선은 지면 관계상 생략한다. 표 2에 제안한 알고리즘과 H.264 표준 사이의 성능 비교를 요약했다.

표 2. 실험 결과와 비교

Sequence	$\Delta BR$ [%]	$\Delta PSNR$ [dB]	$\Delta N$ [%]	$\Delta T$ [%]
Mobile	-0.092	-0.003	85.6	30.24
Silent	1.353	-0.068	92.4	32.52
Foreman	-2.599	-0.142	85.1	27.41
Container	2.983	-0.163	94.4	31.51
News	-2.099	0.122	90.4	28.50
coastguard	-2.428	0.078	86.8	27.70
Average	-0.480	-0.029	89.12	29.65

### V. 결론

본 논문에서 매크로블록 모드의 발생 빈도와 복잡도를 고려하여 매크로블록 모드를 분류하고 이것을 통해 비트-왜곡 최적화 과정에서 조사될 매크로블록 모드 수를 제한하였다. 기본 모드를 이용하여 SKIP 모드와 P8x8 모드를 다른 인터 모드와 분류하였으며, 인트라 모드를 왜곡-비트 특성을 이용하여 다른 매크로블록 모드와 분류하여 선택적으로 비트-왜곡 최적화 과정에 적용했다. 실험 결과를 통해 보인 것처럼 제안한 알고리즘은 심각한 화질저하나 비트량 증가 없이 부호기 복잡도를 H.264 표준과 비교하여 부호화 시간은 평균 29.65%, 비트-왜곡 계산 복잡도는 평균 89.12% 감소시켰다.

### 감사의 글

본 연구는 광주과학기술원(GIST)과 광주과학기술원 실감방송 연구센터를 통한 대학IT연구센터(ITRC), 그리고 교육부 두뇌한국21(BK21) 정보기술사업단의 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 Doc. JVT-G050, 2003.
- [2] I. Richardson, H.264 and MPEG-4 video compression: video coding for next-generation multi media, Wiley, 2003.
- [3] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini and G. J. Sullivan, "Rate-Constrained Coder Control and Comparison of Video Coding Standards", *IEEE Trans. Circuit and System for Video Technology*, vol. 13, no. 7, pp. 688-703, July 2003.
- [4] G. Bjontegaard, "Calculation of Average PSNR Difference between RD-curve", ITU-T Q.6/16, Doc. VCEG-M33, Apr. 2001.