

H.264 부호화기의 빠른 인트라 예측 모드 결정

*정영미 **정봉수 ***전병우

성균관대학교, 정보통신 공학부

*ym1010.jung@samsung.com **bsjung@ece.skku.ac.kr ***bjeon@skku.edu

Fast Intra Prediction Mode Decision of H.264|AVC Encoder

*Jung, Young-Mi **Jung, Bong-Soo ***Jeon, Byeungwoo

Sungkyunkwan University

요약

H.264|AVC는 인트라 부호화 효율을 높이기 위해 공간 영역에서 주변 화소를 이용하여 다양한 방향에 대한 윗-왜곡 최적화 기법을 사용하여 최적의 인트라 예측 모드를 선택한다. 하지만 윗-왜곡 최적화 기법을 사용함에 따라 인트라 부호화에 높은 복잡도가 필요하게 되었다. 따라서 본 논문에서는 인트라 예측 모드 결정의 연산 복잡도를 감소시키고자 사전에 인트라 4x4 예측 모드들의 SATD(Sum of Absolute Transform Difference)를 계산하여 조기에 최우선 모드(Most Probable Mode)를 선택하는 방법을 제안하고, SATD의 값에 따라 제한된 후보 모드에 대해서만 윗-왜곡 최적화를 수행하여 연산 복잡도를 감소하는 방법을 제안한다. 또한 Vertical, Horizontal 그리고 DC모드는 인트라 4x4와 인트라 16x16의 공통적인 모드이므로 인트라 4x4에서 계산되어진 SATD값을 이용하여 인트라 16x16에서의 SAD 계산 복잡도를 줄이는 방법을 제안한다. 본 논문에서 제안하는 빠른 인트라 예측 모드 결정 기법은 연산 복잡도는 평균 61.4% 감소 시킨 반면 부호화 손실은 평균 3.09%에 불과하였다.

1. 서론

최근에는 유튜브, 판도라 TV 등 UCC(User Created Contents)에 대한 관심의 증가로 인해 사용자가 직접 콘텐츠를 제작하여 다른 사용자에게 배포하는 경우가 많아지고 있다. 이에 따라 영상통화와 같은 통신과 결합하여 송/수신이 가능한 양방향 영상 서비스로 빠르게 진화하고 있다. 따라서 이러한 양방향 영상 서비스를 위한 휴대단말상의 고속 비디오 부호화에 대한 요구가 중요한 이슈이다. 또한 사용자가 영상 콘텐츠를 자유로이 제작하고 편집하기 위해 인트라 부호화 기술이 중요하다. 특히, H.264|AVC 인트라 부호화 기술은 종래 JPEG2000과 비교하여 부호화 효율 대비 복잡도 성능이 우수하여 많은 관심을 받고 있다 [3].

H.264|AVC는 인트라 16x16 블록 예측과 인트라 4x4 블록 예측을 수행하여 공간 영역에서 다양한 방향으로 주변 블록의 화소 값을 이용하여 높은 부호화 효율을 달성한다. 그리고 최근에는 High Profile에 인트라 8x8 블록 예측도 적용이 되었다. H.264|AVC 참조 소프트웨어 모델인 JM (Joint Model)[6]은 부호화 효율을 극대화 하고 발생하는 비트량을 최소화 하면서 압축 성능을 높이기 위해 매크로블록의 부호화 모드를 결정할 때 윗-왜곡 최적화 기법(Rate-distortion optimization)을 적용 했다. 하지만 윗-왜곡 최적화 기법은 가장 좋은 예측 모드를 찾기 위해 발생할 수 있는 모든 예측 모드들에 대해 윗-왜곡 관계를 조사해야 하기 때문에 부호화 연산 복잡도가 매우 높아지는 단점이 있다. *

이러한 부호화기의 연산 복잡도를 줄이기 위한 고속 인트라 예측 모드 결정 기법의 연구가 활발히 진행되고 있다[2]-[6]. 빠른 인트라 예측 모드 결정을 위해 윗-왜곡 계산을 필요로 하는 예측 모드의 수를 제한

하여 복잡도를 감소시키는 기법이 주된 연구내용이다. 이러한 예측 모드의 수를 제한하기 위해, Sobel 에지 연산자를 이용하여 에지 방향을 결정한 다음 선택한 후보 예측 모드에 대해서만 윗-왜곡 최적화를 수행하는 방법[4], 하나의 4x4블록을 2x2블록으로 만들고, 그 블록에 대한 수평, 수직, 대각선 방향에 대한 필터를 적용하여 DES (Dominant Edge Strength)를 계산한 다음 후보 예측 모드를 결정하고 윗-왜곡 최적화 계산을 통해 예측 모드를 결정하는 방법[5], Intensity gradient 기법을 적용하여 후보 예측 모드를 선택하는 방법[6] 등이 있다. 상기의 제안 고속 인트라 예측 기법은 제한된 후보 예측 모드만을 이용하여 윗-왜곡 최적화 기법을 수행하기 때문에 연산 복잡도의 감소는 많이 발생하지만, 부호화 효율이 크게 떨어지는 문제점이 있다.

본 논문은 인트라 4x4 예측과 인트라 16x16 예측에 각각 빠른 인트라 예측 모드 결정 방법을 적용한다. 인트라 4x4 예측 모드 결정 시 9가지 예측 방향에 대하여 사전에 SATD를 모두 계산한 후 가장 작은 SATD 값을 가지는 모드를 찾는다. 이때 찾은 모드가 최우선 모드와 같을 경우 윗-왜곡 계산을 수행하지 않고 조기에 최우선 모드를 선택하여 인트라 4x4 예측 모드의 복잡도를 줄이는 기법을 제안한다. 또한 vertical, horizontal, DC 모드는 인트라 4x4 예측과 인트라 16x16 예측에서 공통되는 모드이므로 인트라 4x4 예측 모드 시 계산된 SATD를 이용하여 인트라 16x16 예측 모드를 이용한다. 즉 인트라 4x4 예측 모드 시 하나의 매크로블록에서 발생하는 16개의 SATD 값을 모두 합하여 cost로 이용하고 인트라 16x16 예측 모드 시에는 이 값들을 비교하여 최적 모드를 결정하는 방법을 제안한다.

본 논문은 다음과 같은 순서로 구성된다. 2장에서 H.264|AVC의 인트라 예측 모드 결정 방법에 대해 살펴보고 3장에서는 제안하는 알고리

※ 논문은 과학기술부의 재원으로 한국과학재단 국가지정 연구실사업으로 수행된 연구임(R0A-2006-10286-0(2008))

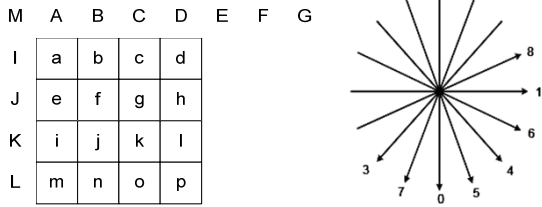


그림 1. 4x4 블록 인트라 예측 모드

즘에 대해 설명하고 4장에서 실험 및 결과를 보여주고 5장에서 결론을 짓도록 한다.

2. H.264|AVC에서 인트라 예측 방법

가. 인트라 4x4 예측 모드 결정

H.264|AVC는 공간 영역에서 주변 블록 화소의 에지 방향성 상관성을 이용하는 인트라 예측 부호화를 수행한다. 휘도 신호에 대한 인트라 4x4 예측은 그림 1에서처럼 4x4블록에 대해 8가지 방향에 따른 모드와 DC 모드(2번)로 총 9가지 모드의 예측 모드가 있다. 앞서 설명한 것과 같이 압축 성능 효율을 높이기 위해 H.264|AVC는 울-웨곡 최적화 기법을 적용하고 있으며 H.264 참조소프트웨어 부호화기는 매크로 블록 내의 16개의 4x4 블록에 대해서 9가지 예측 모드의 RD_Cost를 다음과 같이 계산하여 최적의 인트라 4x4 예측 모드를 결정한다.

$$J(s, c, IMODE | QP, \lambda_{MODE}) = SSD(s, c, IMODE | QP) + \lambda_{MODE} \cdot R(s, c, IMODE | QP) \quad (1)$$

$IMODE \in \{\text{all nine Intra4x4 modes}\}$

S는 원본 4x4블록이고 C는 복원된 4x4블록이다. $R(s, c, IMODE | QP)$ 는 인트라 4x4예측 모드 정보와 양자화 계수 값 그리고 차영상 계수들을 부호화하기 위해 필요한 비트량이다. SSD(Sum of squared distortion)는 원본 블록과 복원된 블록간의 왜곡을 나타낸다. 상기의 RD_Cost는 9가지 모든 방향에 대하여 계산되어 지고 그 중에서 가장 작은 값을 가지는 모드를 해당 블록의 최적 모드로 결정한다. 그러므로 RD_Cost 계산을 위해 변환 부호화, 양자화, 엔트로피 부호화가 반복적으로 이루어지므로 연산 복잡도가 매우 높게 나타난다.

한편, 인트라 4x4 예측 모드는 한 매크로블록 내에 16개가 존재하여 이를 고정 길이 부호화를 통해 전송할 경우 많은 양의 오버헤드가 필요하여 효과적인 부호화 방법이 필요하다. 인트라 예측 모드는 인접 블록과의 공간적 상관도가 매우 높기 때문에, 이를 이용하

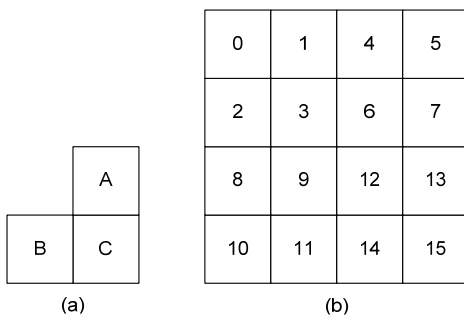


그림 2. 4x4 블록 인트라 예측 순서

여 16개의 인트라 예측 모드를 부호화 한다. 각 인트라 4x4 예측 모드의 정보는 그림 3(b)와 같은 순서로 전송되는데, 그림 3(a)에서처럼 좌측블록 B와 상위 블록 A의 모드 정보를 이용하여 부호화 효율을 높이고 있다. 주변 블록 A와 B의 예측 모드 가운데 값이 적은 예측모드 값을 최우선모드로 결정하여 현재 블록 C의 예측 모드가 최우선 모드와 같은 경우 1비트로 나타내어 부호화 효율을 높이고 있다. 현재 블록의 예측 모드가 최우선 모드와 같으면 prev_intra4x4_pred_mode_flag 라는 한 비트 플래그를 '1'로 하여 전송하고, 같지 않으면 '0'을 전송한 다음, 추가로 나머지 8가지 경우 중 한 가지를 가리키는 3비트의 부호어를 rem_intra4x4_pred_mode 라는 Syntax의 이름으로 고정장으로 부호화하여 전송한다. 추가로 전송될 3비트 고정장 부호어는 최우선 모드와 비교하여 현재 모드가 같거나 클 경우, 현재 모드에서 1을 뺀 값을 부호화한다. 표 1은 인트라 4x4 예측 모드에서 최우선 모드로 결정되는 확률분포를 나타내었다. 표 1에서 볼 수 있듯이 다양한 영상에 대해 최우선 모드로 결정될 확률이 0.6이상으로 매우 높음을 알 수 있다. 따라서 예측 모드에 대해 RD_Cost를 계산하기 전에 최우선 모드를 결정할 수 있다면, 부호화 효율 감소에 큰 효과가 있을 것이다.

나. 인트라 16x16 예측 모드 결정

휘도 신호에 대한 인트라 16x16 예측 모드에는 그림 3과 같이 vertical(0번), horizontal(1번), DC(2번) 그리고 plane(3번)의 총 4가지 모드가 있다. 부호화를 하려는 현재 매크로블록의 왼쪽과 위쪽에 이용할 수 있는 화소가 없다면 방향성을 가지지 않는 DC모드가 인트라 16x16의 최적의 예측 모드가 된다. 반대로 위쪽과 왼쪽에 주변 화소가 있다면 각 모드에 대하여 SAD(Sum of Absolute Different)값을 계산하여 그 값이 가장 작은 모드를 인트라 16x16의 최적 예측 모드로 결정한다. H.264의 참조 소프트웨어에서는 인트라 16x16 예측을 인트라 4x4 보다 먼저 수행한다. 하지만 인트라 16x16의 0,1,2 번 모드가 인트라 4x4와 공통된다는 점에 착안하여 빠른 인트라 4x4 예측을 먼저 수행하여 이때 계산되는 각 모드 별 SATD 값을 인트라 16x16 예측에 이용할 수 있을 것으로 예상된다.

표 1. 인트라 4x4 예측 최우선 모드 확률

Sequence	QP	P(MPM)	P(others)
Foreman (30fps, 100frames)	28	0.60	0.40
	32	0.68	0.32
	36	0.74	0.26
	40	0.79	0.21
Paris (30fps, 100frames)	28	0.61	0.39
	32	0.66	0.34
	36	0.72	0.28
	40	0.76	0.24
Silent (30fps, 100frames)	28	0.51	0.49
	32	0.60	0.40
	36	0.69	0.31
	40	0.79	0.21
Stefan (30fps, 100frames)	28	0.57	0.43
	32	0.59	0.41
	36	0.62	0.38
	40	0.65	0.35

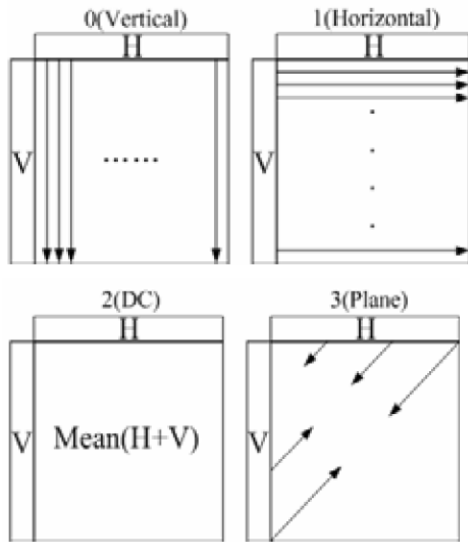


그림 3. 16x16 블록 인트라 예측 모드

3. 제안하는 인트라 예측 모드 결정 방법

가. 인트라 4x4 예측 모드 결정 방법

앞에서 살펴본 바와 같이 인트라 4x4 예측에서 최우선 모드로 선택될 확률이 매우 높음을 확인하였다. 따라서 9가지 예측 모드에 대해 반복적인 울-왜곡 비용 계산을 수행하기 전 최우선 모드를 조기에 결정한다면, 울-왜곡 비용 계산 수행 횟수를 줄여 H.264/AVC 부호화기의 인트라 4x4 예측 모드 탐색을 빠르게 수행할 수 있을 것이다.

그림 4에 제안하는 조기 최우선 모드를 결정하는 순서도를 나타내었다. 인트라 4x4 예측 방향을 결정하기 위해 부호화 블록이 입력되면 9가지 예측방향에 대해 각각 SATD 값을 계산한다. 그림 4의 cand_imode2는 cand_imode1 다음으로 작은 SATD 값을 가지는 모드이다. 9가지 예측 방향에 대해 모든 SATD 값 계산이 완료되면 결정된 cand_imode1과 MPM(most probable mode: 최우선 모드)과 비교한다. 이 때 cand_imode1과 최우선 모드가 같은 경우 현재 입력된 부호

화 블록의 예측 모드를 최우선 모드로 조기에 결정한다. 그러므로 조기에 최우선 모드가 결정되면 기존과 비교했을 시 입력되는 모든 부호화 블록에 대해 9번 반복적으로 수행되던 울-왜곡 계산 과정이 한 번으로 줄어들어 연산 복잡도도 같이 줄어들게 된다.

반면에 cand_imode1과 최우선 모드가 다른 경우에는 SATD 계산 시 미리 결정되었던 cand_imode1과 cand_imode2의 두 모드에 대해서만 울-왜곡 계산을 수행하여 최소값을 가지는 모드를 현재 입력된 부호화 블록의 최적 예측 모드로 결정한다. 이는 최우선 모드가 조기에 결정되지 않더라도 9번의 반복적인 울-왜곡 비용 계산을 수행하는 것이 아니라 가장 작은 SATD 값을 가지는 두 가지 모드에 대해서만 울-왜곡 비용을 계산함으로써 부호화 실행 시간을 더 줄일 수 있다.

나. 인트라 16x16 예측 모드 결정 방법

그림 4의 SAD배열은 인트라 4x4 예측과 인트라 16x16 예측 모두에서 접근이 되면 인트라 4x4 예측 시 계산되는 SATD 값을 저장하는 변수이다. 앞서 설명한 것과 같이 공통되는 모드인 0,1,2 번 모드에 대해 인트라 4x4 예측을 수행할 시 계산되는 SATD 값을 각 모드 별로 모두 더해 SAD의 해당 인덱스에 저장되며 하나의 매크로블록은 총 16개의 4x4 블록으로 이루어지므로 총 16개의 SATD 값이 저장되게 된다. 예를 들어 현재 Imode가 0(vertical)이라면 SAD[0]에 0번 모드에서 발생하는 모든 SATD 값을 다 더한 후 저장한다. 이 과정은 인트라 4x4 예측에서 수행이 되며 SAD에 최종적으로 저장된 값은 인트라 16x16에서 사용된다. 인트라 16x16 예측에서는 별도의 SAD를 계산하는 과정을 수행하지 않는 대신에 미리 저장된 각 모드에 해당하는 SAD 배열 값을 읽어 최소값을 갖는 모드를 최적의 모드로 결정한다.

4. 실험 결과

본 논문에서 제안하는 빠른 인트라 예측 모드 결정 기법의 성능을 검증하기 위해 이를 JM12.4 참조소프트웨어[7]에 구현하였다. 표 2는 실험에 사용한 시퀀스와 부호화기의 설정 파라미터를 나타낸다. QP는 총 4가지 (28, 32, 36, 40)에 대해서 실험 하였다. 제안한 기법의 부호화

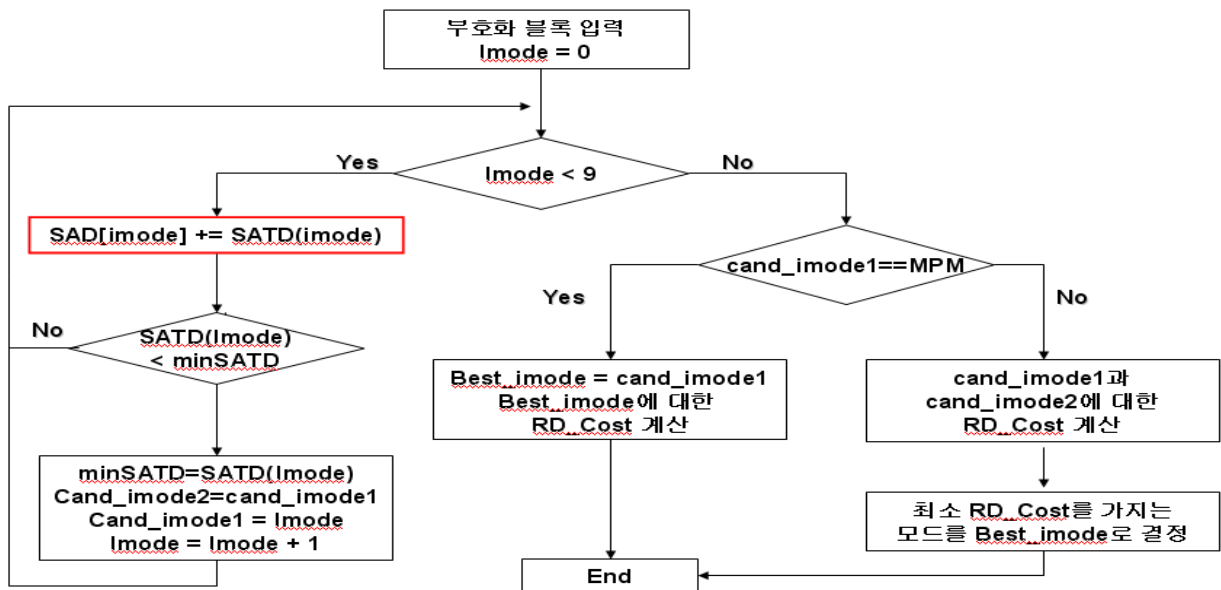


그림 4. 제안 알고리즘 순서도

표 3. 절약된 부호화 시간 비교

Sequence Information	
Sequence Name	Container, Foreman, Paris, Silent
Frame Size	CIF 352 * 288
Video Format	30fps Total Encoding Frames : 150
H.264 Encoder - Jm12.4	
Block Mode	Intra 4x4 & Intra 16x16
RDOptimization	High Complexity Mode Enable
Intra 16x16 Prediction	Disable
Entropy Coding	CAVLC
Coding Structure	Intra Only Coding

성능 평가에는 BDBR(%)와 BDPSNR(dB)를 사용하였다[8]. 여기서 BDBR(%)의 (-)와 BDPSNR(dB)의 (+) 부호는 각각 제안 부호화방법의 성능향상을 나타낸다. 또한 부호화 속도는 다음과 같이 AST(Average saving time)로 평가하였다.

$$AST(\%) = \frac{Time(full_search) - Time(proposed)}{Time(full_search)} \times 100 \quad (2)$$

여기서 (+)부호는 부호화 속도 향상을 (-) 부호화 속도 감소를 나타낸다. 상기의 성능평가는 인트라 4x4 예측 모드결정시 전체 검색 방법을 기준으로 Tsai방법[6]과 본 논문에서 제안하는 방법의 결과를 비교하였다.

표 3과 표 4는 제안 방법의 부호화 속도와 부호화 효율을 나타낸다. 여기서 “제안 알고리즘1”은 빠른 인트라 4x4 예측 모드 결정 방법만 적용하였고, “제안 알고리즘2”는 추가로 빠른 인트라 16x16도 적용한 것이다. 표 3과 표 4에서 볼 수 있듯이 제안 알고리즘1의 경우 약 54% 연산 시간을 줄이는 효과가 있다. 또한 제안 알고리즘2와 Tsai 방법은 연산 시간 감소는 비슷하지만, 제안 알고리즘의 부호화 효율이 Tsai 방법보다 BDPSNR은 약 0.05dB, BDBR은 약 0.4%정도 향상되어, 부호화 속도 및 압축 성능 측면에서 우수함을 알 수 있다.

5. 결론

본 논문은 H.264/AVC의 인트라 부호화에서 인트라 4x4 예측의 연산 복잡도 감소를 위한 빠른 인트라 4x4 예측 모드 선택 방법을 제안하였다. 인트라 4x4 예측에서 제안하는 방법은 율-왜곡 값을 계산하기 전에 9가지 방향에 대한 SATD값을 이용하여 조기에 최우선 모드를 선택하는 방법과 최소 SATD값을 가지는 2개의 후보 예측 모드에 대해서만 율-왜곡 값을 계산하여 최적의 예측 모드를 선택하는 방법이다. 인트라 16x16 예측에서 제안하는 방법은 SAD를 구하는 과정을 수행하지 않고 인트라 4x4 예측에서 계산된 SATD 값을 모두 더한 후 최소값을 가지는 모드를 최적의 예측 모드로 선택하는 방법이다. 제안 방법은 연산 복잡도를 평균 61.4%감소 시켰고, BDPSNR은 평균 0.19dB 감소, BDBR은 평균 3.09% 증가되었다.

표 3. 절약된 부호화 시간 비교

Sequence	Tsai 논문[6]	제안 알고리즘1	제안 알고리즘2
	AST	AST	AST
Container	53.78	38.85	62.46
Foreman	55.93	34.60	61.17
Paris	52.69	35.14	60.56
Silent	51.83	31.02	61.59
Average	53.56	34.90	61.44

표 4. BDPSNR과 BDBR 비교

Sequence	Tsai 논문[6]		제안 알고리즘1		제안 알고리즘2	
	BDPSNR [dB]	BDBR [%]	BDPSNR [dB]	BDBR [%]	BDPSNR [dB]	BDBR [%]
Container	-0.27	3.69	-0.09	1.16	-0.23	3.81
Foreman	-0.33	4.68	-0.09	1.25	-0.17	3.04
Paris	-0.16	1.87	-0.08	0.92	-0.19	2.44
Silent	-0.20	3.58	-0.05	0.93	-0.15	3.09
Average	-0.24	3.45	-0.08	1.07	-0.19	3.09

6. 참고 문헌

- [1] 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.
- [2] 이제운, 전병우, “H.264 동영상 압축을 위한 고속 부호화 모드 결정 방법,” 대한전자공학회 논문지 SP편 제41권 제6호, 65-173쪽, 2004년 11월.
- [3] Y. W. Huang, B. Y. Hsieh, T. C. Chen, and L. G. Chen, “Analysis, fast algorithm, and VLSI architecture design for H.264/AVC intra-frame coder,” IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 3, pp.378-401, Mar. 2005.
- [4] 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 Syst. Video Technol., vol. 15, no. 7, pp.813-822, Jul. 2005.
- [5] J.-C. Wang, J.-F. Wang, J.-F. Yang, and J.-T. Chen, “A Fast Mode Decision Algorithm and Its VLSI Design for H.264/AVC intra-Prediction,” IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 10, pp.1414-1422, Oct. 2007.
- [6] A.-C. Tsai, A. Paul, J.-C. Wang, “Intensity Gradient Technique for Efficient Intra-Prediction in H.264/AVC,” IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 5, pp.694-698, MAY 2008.
- [7] <http://bs.hhi.de/~suehring/tml/download/jm12.4.zip>
- [8] G. Bjontegaard, “Calculation of average PSNR differences between RD-curves,” VCEG Contribution VCEG-M33, Austin, Apr. 2001.