

VVC 화면내 예측 및 부호화 주요 기술

□ 한희지, 최재륜, 권대혁, 최해철 / 한밭대학교

요약

VVC(Versatile Video Coding)는 국제 표준화 단체인 JVET(Joint Video Experts Team)에서 표준화가 진행되고 있는 새로운 국제 비디오 부호화 표준이다. 이 표준화에서는 기존 최신 비디오 부호화 표준인 HEVC(High Efficiency Video Coding)/H.265 대비 2배 이상의 부호화 성능을 목표로 다양한 부호화 방법들이 논의되고 있다. 본 고에서는 VVC의 새로운 부호화 모드 중 화면내 예측(intra prediction) 부호화 방법에 대해 소개한다. 화면내 예측은 현재 부호화를 진행하려는 블록의 주변에 이미 재구성된 샘플들을 참조하여 현재 블록을 예측하는 방법이다. 이 화면내 부호화 방법은 화면간 예측(inter prediction) 부호화 방법과 함께 부호화 효율 향상에 기여할 뿐만 아니라, 임의 접근(random access)을 가능하게 하고 부호화된 비트스트림의 에러 내성을 높인다. VVC는 화면내 부호화 예측 모드 종류를 최대 87개까지 확장하고 다양한 화면내 부호화 방법을 채택함으로써 기존 비디오 부호화 표준에 비해 높은 부호화 효율을 갖는다. 본 고에서는 VVC에 채택된 주요 화면내 부호화 방법들을 소개한다.

1. 서론

ISO/IEC MPEG(JTC1/SC 29/WG 11)과 ITU-T SG16은 새로운 비디오 부호화 표준을 개발하기 위해 JVET(Joint Video Experts Team)를 조직하였다. 이 표준은 2017년 10월 Call for proposal[1]을 받고 공식적인 표준화를 시작하였으며 VVC(Versatile Video Coding)로 명명되었다. VVC[2][3] 표준화는 2019년 7월까지 총 7차례 표준화 회의를 개최했으며, 국제적으로 다양한 기관 및 회사들의 경쟁을 통해 우수한 부호화 알고리즘을 VVC 표준안으로 채택하고 있다. 표준으로 채택된 부호화 방법은 공정한 검증과 용이한 부호화 툴 개발을 위해서 하나의 소프트웨어에 통합되는데, VVC의 경우 이를 VTM(VVC Test Model)이라고 한다. 현재까지 개발된 VTM은 임의 접근(random access) 실험 조건에서 HEVC/H.265보다 약 33.14% BD-rate의 부호화 이득을

갖는다고 보고된다[4]. VTM은 부호화 단위 및 구조, 화면간(inter) 예측, 화면내(intra) 예측, 보간(interpolation), 필터링(filtering), 변환(Transform), 양자화, 엔트로피(entropy) 부호화 방법 등에 대한 새롭고 다양한 알고리즘들을 포함하고 있다[5].

본 고에서 다루는 화면내 예측 부호화는 이미 복호된 공간적인 주변 정보를 참조 샘플로 이용하여 현재 부호화하려는 샘플을 효과적으로 예측하고 그 예측 값과 원래 샘플 사이의 예측 오차만을 부호화함으로써 부호화 효율을 높이는 방법이다. 참조 샘플로 이용할 수 있는 정보의 양이 시간적 주변 정보에 비해 적기 때문에 화면내 예측 방법은 화면간 예측 방법에 비해 일반적으로 낮은 부호화 성능을 갖지만, 새로운 물체의 출현 혹은 조명 변화가 심한 경우 등과 같이 시간적 주변 정보로부터 현재 화소를 예측하기 힘든 경우에는 화면내 예측이 더 우수한 성능을 가질 수 있다. 또한 부호화된 비트스트림에 대해 임의의 접근을 지원하거나 에러 내성을 높이기

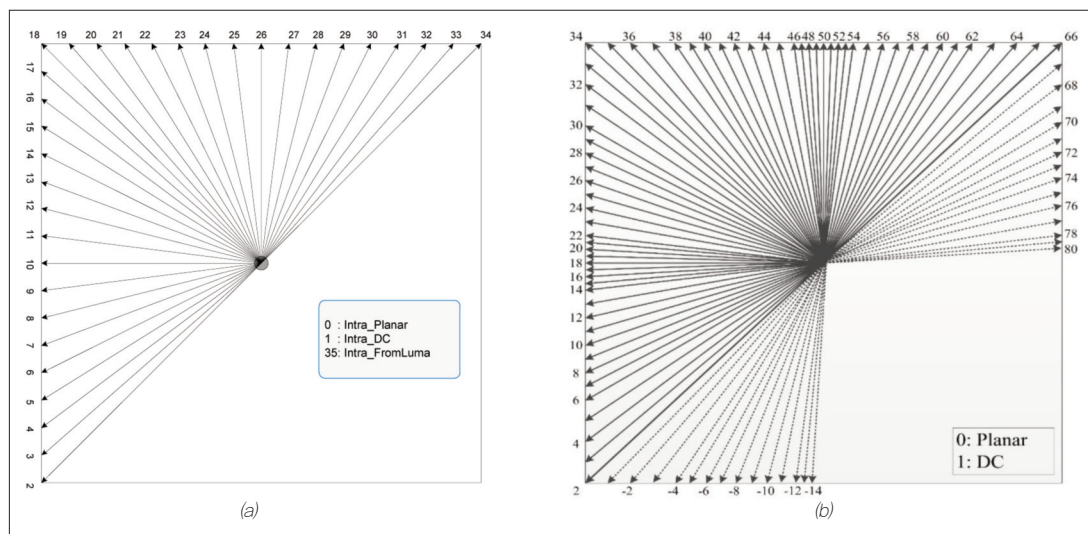
위해서 강제로 화면내 예측을 사용하기도 한다.

VVC에서는 화면내 부호화와 관련하여 Position Dependent Prediction Combination(PDPC)[6][7], Intra Sub-Partitions(ISP)[8]–[11], Multi-Reference Line(MRL)[12][13], Cross-Component Linear Model(CCLM)[14]–[22], Matrix-weighted Intra Prediction(MIP)[23]–[26], Mode Dependent Intra Smoothing(MDIS)[27] 등을 채택하였다. 본 고에서는 이 새로운 주요 화면내 예측 부호화 방법에 대해 소개한다.

II. VVC 화면내 부호화 방법

1. 화면내 예측 부호화 모드 및 모드 시그널링

VVC는 20개의 광각(wide) 방향성(angular) 모드



〈그림 1〉 화면 내 부호화 모드 (a) HEVC, (b) VVC

및 65개의 일반 방향성 모드를 포함하는 최대 87개의 부호화 모드를 지원한다[28][29]. 이는 이전 부호화 표준인 HEVC/H.265가 35개의 방향성 모드를 포함하여 37개의 부호화 모드를 지원한 것과 비교하여 2배 이상 부호화 모드가 증가한 것이다. <그림 1>(a)은 HEVC에서 지원하는 화면내 부호화 모드를 나타낸 것이며, <그림 1>(b)는 VVC에서 지원하는 화면내 부호화 모드를 나타낸 것이다. 이와 같이 VVC의 화면내 부호화 모드는 HEVC보다 더 세밀한 방향성에 대해 부호화할 수 있게 되었으며, 비정방형 부호화 블록(coding block)을 허용함에 따라 광각 방향성 모드를 사용하여 부호화할 수 있도록 설계되었다.

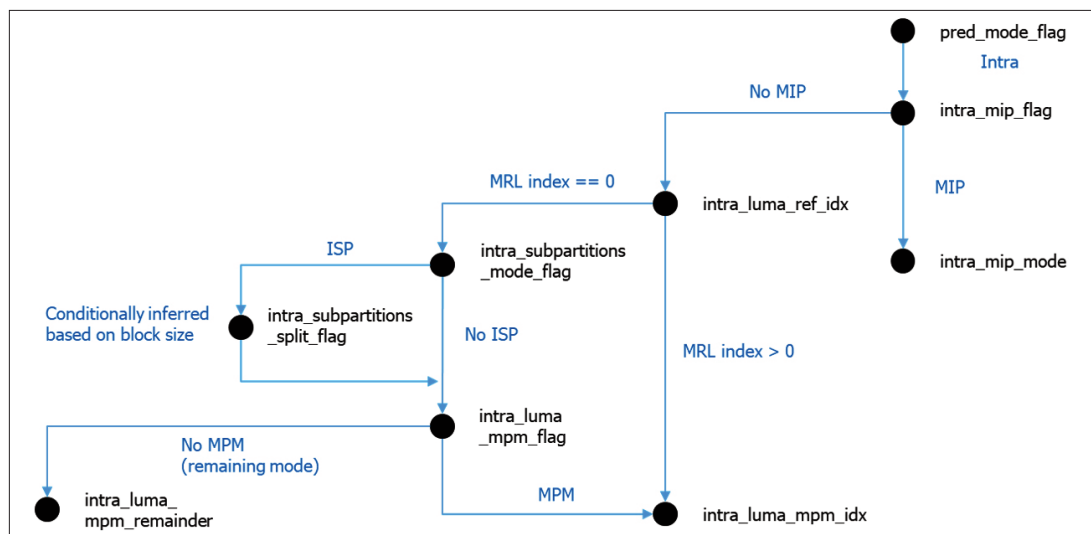
일반적으로 화면내 부호화에서는 현재 블록과 주변 블록이 서로 유사한 경우가 많으며, 이로 인하여 블록 간의 화면내 부호화 모드가 동일하거나 유사한 경우가 자주 발생한다. Most Probable Mode(MPM)은 이러한 유사성을 이용하여 확률이 높은 화면내 부호화 모드에 적은 비트를 할당하여 현재 블록의 화면내 부호화 모드의 전송 정보를 감소시킨다. 기존 표준인 HEVC/H.265의 화면내 부호화는 3개의 MPM을 지원하였으나, VVC의 화면내 부호화는 6개의 MPM을 지원한다[30]. 참조 소프트웨어인 VTM 6.0은 MPM 리스트를 현재 블록

의 좌측 블록 및 상단 블록의 화면내 부호화 모드에 따라 <표 1>과 같이 구성한다. 현재 블록의 좌측 블록의 모드(Left) 및 상단 블록의 모드(Above)의 화면내 부호화 모드 종류에 따라 크게 4가지로 구분하며, Left와 Above가 서로 다르며 두 모드가 모두 방향성 모드인 경우 Left와 Above의 차이에 따라 4가지로 더 구분하여 MPM 리스트를 생성한다. 좌측 블록 및 상단 블록이 화면내 부호화 모드가 아닌 경우 Planar 모드로 가정한다.

<그림 2>는 VVC에서 화면내 부호화 모드를 전송하는 방법을 나타낸 것이다. 현재 블록이 화면내 부호화인 경우 pred_mode_flag를 0으로 시그널링한다. 가장 먼저 intra_mip_flag를 통해 MIP 기술의 적용 여부를 시그널링한다. 만약 MIP 기술을 적용하는 경우, MIP의 부호화 모드를 추가 전송한다. MIP 기술이 적용되지 않는 경우, intra_luma_ref_idx를 통해 MRL 기술의 적용 여부를 시그널링하며, intra_luma_ref_idx가 1 혹은 2의 값으로 MRL 기술을 적용하는 경우에는 항상 MPM만을 사용하여 현재 블록의 화면내 부호화 모드를 알린다. MRL을 적용하지 않는 경우에는 intra_luma_ref_idx를 0으로 시그널링하며, 이 경우에는 ISP 기술의 적용 여부 및 ISP 모드를 결정하기 위한 구문요소를 시그널링하고 ISP 기술의 적용 여부와 관계없

<표 1> VVC의 화면내 부호화에서 지원하는 Most Probable Modes(MPMs)

조건	세부조건	6개의 MPM 모드
좌측 모드와 상단 모드가 방향성 모드이고 서로 동일한 경우		{Planar, Left, Left-1, Left+1, Left-2, Left+2}
좌측 모드와 상단 모드가 서로 다르며, 두 모드가 모두 방향성 모드인 경우	Max-Min=1	{Planar, Left, Above, Min-1, Max+1, Min-2}
	Max-Min>=61	{Planar, Left, Above, Min+1, Max-1, Min+2}
	Max-Min=2	{Planar, Left, Above, Min+1, Min-1, Max+1}
	Otherwise (2<Max-Min<61)	{Planar, Left, Above, Min-1, Min+1, Max-1}
좌측 모드와 상단 모드가 서로 다르며, 둘 중 하나만이 방향성 모드인 경우		{Planar, Max, Max-1, Max+1, Max-2, Max+2}
좌측 모드와 상단 모드가 모두 방향성 모드가 아닌 경우 (Planar, DC)		{Planar, DC, Angular50, Angular18, Angular46, Angular54}



〈그림 2〉VC 화면내 부호화 모드 전송 방법

이 MPM의 사용여부를 알리기 위해 `intra_luma_mpm_flag`를 전송한다. 만약 MPM을 적용하는 경우, `intra_luma_not_planar_flag` 및 `intra_luma_mpm_idx`를 통해 6개의 MPM 중 현재 블록이 사용하는 MPM을 지시한다. MPM을 사용하지 않는 경우에는 `intra_luma_mpm_remainder` 구문요소를 전송하여 MPM을 제외한 나머지 화면내 부호화 모드 중 하나를 지시한다.

2. Position Dependent Prediction Combination(PDPC)

기존 비디오 부호화 표준의 화면내 예측 방법은 블록 내 우측 하단영역과 같이 참조 화소에서 멀리 떨어진 픽셀일수록 예측 정확도가 떨어진다. 따라서 16x16 이하 크기를 가진 비교적 작은 크기의 블록에서 예측이 잘 동작되고, 그 이상의 큰 블록에서는 상대적으로 낮은 부호화 효율을 갖는다.

PDPC는 화소 위치, 예측 모드, 블록의 크기에 따

라 필터링된 참조 샘플과 필터링되지 않은 참조 샘플의 가중치 조합을 사용하여 참조 화소들 간의 차이 값으로 예측 값을 보정함으로써 위에서 언급한 문제를 해결한다. VVC에는 간소화된 PDPC (Simplified PDPC)가 채택되었다.

PDPC는 Planar, DC, 수평(horizontal), 수직(vertical), 그리고 좌-하단 각도 모드(bottom-left angular mode)와 그에 인접한 8개의 각도모드, 우-상단 각도모드(top-right angular mode)와 그에 인접한 8개의 각도모드에 대해 별도의 지시자 없이 각 화면내 부호화 모드에 적용된다.

DC, Planar, 수직, 수평 화면내 예측 모드인 경우, PDPC에 의한 최종 예측 샘플 $Pred(x, y)$ 은 다음 식 (1)에 따라 화면내 예측 모드 및 참조 샘플의 선형 조합으로 예측된다.

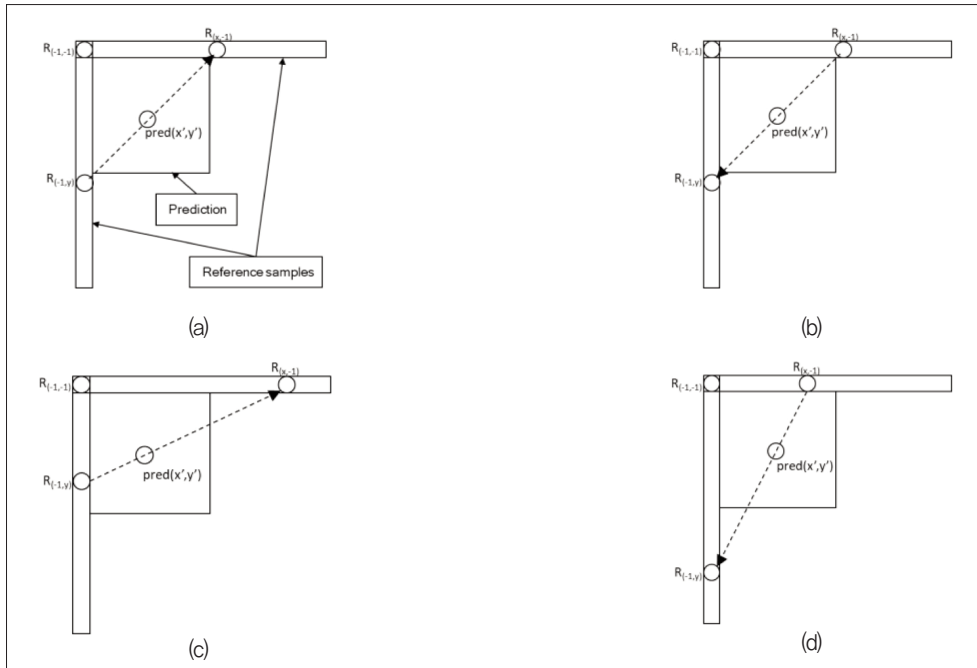
$$\begin{aligned}
 &Pred(x, y) \\
 &= (wL \times R_{-1,y} + wL \times R_{x,-1} - wL \times R_{-1,-1} + \\
 &\quad (64 - wL - wT + wTL) \times Pred(x, y) + 32) \gg 6
 \end{aligned} \tag{1}$$

여기서 $R_{x,i}$, $R_{-1,y}$ 는 각각 현재 샘플 (x, y) 의 상단(top)과 좌측(left)에 위치한 참조 샘플을 나타내고, $R_{-1,-1}$ 은 현재 블록의 좌 상단(top-left corner)에 위치한 참조 샘플을 나타낸다. $Pred(x, y)$ 는 기존 화면내 예측 모드에 의해 획득된다. 가중치 값인 wL , wT , wTL 은 현재 화소의 위치 (x, y) 에 따라 결정된다. PDPC가 DC, Planar, 수평 및 수직 화면내 예측 모드에 적용되는 경우, HEVC에 있는 DC 모드 경계 필터 또는 수평/수직 모드 에지 필터가 필요하지 않으며 PDPC가 이를 대체한다. 우-상단(top-right) 대각선 방향, 좌-하단(bottom-left) 대각선 방향, 그리고 이들의 근접 화면내 예측 모드들의 경우, PDPC에서 사용하는 참조 샘플의 위치가 상기의 식 (1)과 달리 식 (2)와 같이 변경된다.

$$\begin{aligned}
 &Pred(x', y') \\
 &= (wL \times R_{1,y} + wT \times R_{x,-1} - wTL \times R_{-1,-1} + \\
 &\quad (64 - wL - wT + wTL) \times Pred(x', y') + 32) \gg 6
 \end{aligned} \tag{2}$$

〈그림 3〉은 상기 식 (2)에서의 참조 샘플 $R_{x,-1}$, $R_{1,y}$ 및 $R_{-1,-1}$ 를 보여준다. 현재 화소의 위치가 (x', y') 일 때, 참조 샘플 $R_{x,-1}$ 의 좌표 x 는 $x = x' + y' + 1$ 이며, 참조 샘플 $R_{1,y}$ 의 좌표 y 는 $y = x' + y' + 1$ 이다.

PDPC를 상기 대각선 방향에 인접한 angular mode에 적용하는 경우, 참조 샘플 $R_{x,-1}$ 과 $R_{1,y}$ 는 분수 샘플 위치(fractional sample position)에 위치될 수 있다. 이 경우, 간소화를 위하여 분수 샘플 위치의 값을 구하지 않고 가장 가까운 정수 샘플 위치의 샘플 값을 $R_{x,-1}$ 과 $R_{1,y}$ 로 사용한다. PDPC의 가중치는 예측 모드에 따라 다르며 〈표 2〉와 같다. 표에



〈그림 3〉 대각선 및 대각선에 인접한 방향성 모드에 적용되는 PDPC 참조 화소. (a) 우-상단 대각선 모드, (b) 좌-하단 대각선 모드, (c) 우-상단 대각선 모드와 인접한 모드, (d) 좌-하단 대각선 모드와 인접한 모드

〈표 2〉 예측 모드에 따른 PDPC 가중치의 예

Prediction modes	wT	wL	wTL
DC and Planar	$32 \gg ((y \ll 1) \gg \text{shift})$	$32 \gg ((x \ll 1) \gg \text{shift})$	0
Horizontal	$32 \gg ((y \ll 1) \gg \text{shift})$	0	wT
Vertical	0	$32 \gg ((x \ll 1) \gg \text{shift})$	wL
Diagonal top-right	$16 \gg ((y' \ll 1) \gg \text{shift})$	$16 \gg ((x' \ll 1) \gg \text{shift})$	0
Diagonal bottom-left	$16 \gg ((y' \ll 1) \gg \text{shift})$	$16 \gg ((x' \ll 1) \gg \text{shift})$	0
Adjacent diagonal top-right	$32 \gg ((y' \ll 1) \gg \text{shift})$	0	0
Adjacent diagonal bottom-left	0	$32 \gg ((x' \ll 1) \gg \text{shift})$	0

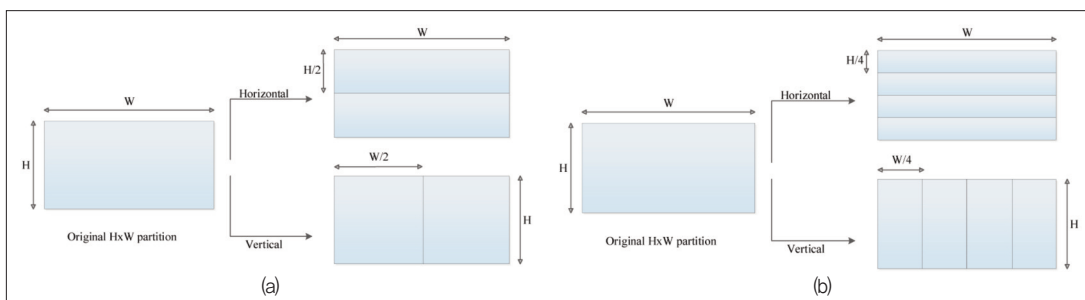
서 shift는 식 (3)이며, 여기서 width와 height는 CU의 가로와 세로 크기이다.

$$\text{shift} = (\log_2(\text{width}) - 2 + \log_2(\text{height}) - 2 + 2) \gg 2 \quad (3)$$

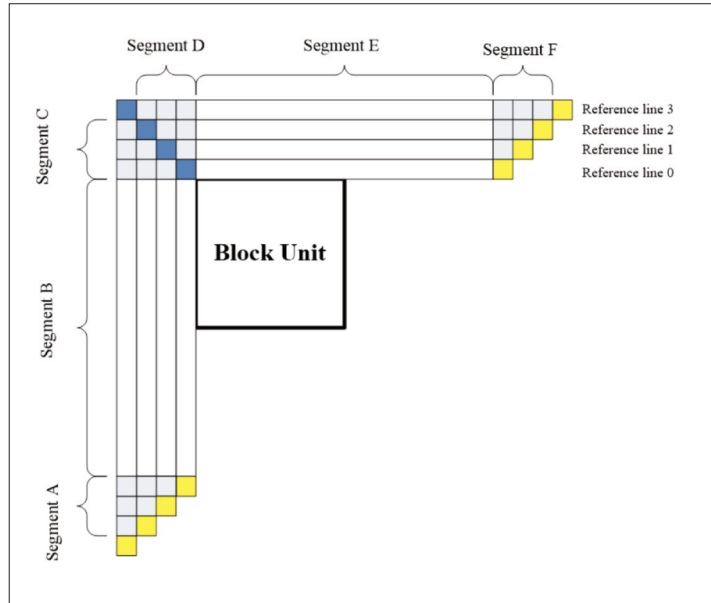
3. Intra Sub-Partitions(ISP)

ISP(Intra Sub-Partitions)는 블록의 크기에 따라 수직 또는 수평 방향으로 2개 또는 4개의 분할 영역(sub-partition)으로 나누고 각 분할 영역에 대해 동일한 화면내 예측 모드로 예측을 수행하는 방법이다. ISP가 적용 가능한 최소 블록 크기는 4×8 또는 8×4 이다. 〈그림 4〉와 같이 블록의 크기가 4×8 또는 8×4 이면 2개로 분할하며, 이보다 크

면 4개로 분할한다. ISP 블록이 $M \times 128$ ($M \leq 64$) 및 $128 \times N$ ($N \leq 64$)인 경우, VVC의 처리 단위인 VPDU(Virtual Pipeline Data Unit)로 인해 문제가 발생할 수 있다. 예를 들어, 단일 트리(single-tree) 구조에서 $M \times 128$ CU는 $M \times 128$ 휘도 TB(Transform Block) 및 2개의 대응하는 $M/2 \times 64$ 색차 TB를 갖는다. CU가 ISP를 사용하는 경우, 휘도 TB는 4개의 $M \times 32$ TB(수평 분할만 가능)로 나뉘며, 각각은 64×64 블록보다 작다. 그러나, 현재의 ISP 설계에서 색차 블록은 분할되지 않는다. 따라서 두 색차 성분의 크기는 32×32 블록보다 크다. 마찬가지로 ISP를 사용하여 $128 \times N$ CU로 비슷한 상황을 만들 수 있다. 따라서 이 두 경우에는 64×64 복호기 파이프라인에 문제가



〈그림 4〉 ISP의 블록 분할 예. (a) 4×8 , 8×4 CU인 경우, (b) 4×8 혹은 8×4 보다 큰 CU의 경우



〈그림 5〉 현재 블록에 인접한 4개의 참조 라인들의 예

발생한다. 이러한 이유로 ISP를 사용할 수 있는 CU 크기는 최대 64×64 로 제한된다. 〈그림 5〉는 ISP의 블록 분할 예를 보여준다. 모든 분할 영역은 최소 16개의 샘플이 있는 조건을 충족해야 한다.

ISP에서는 각 분할 영역의 재구성된 샘플 값을 다음 분할 영역의 예측을 위해 이용할 수 있고, 이것이 각 분할 영역에 대해 반복적으로 처리된다.

모든 분할 영역은 동일한 화면내 예측 모드를 공유한다. ISP와 다른 부호화 툴 간의 상호작용은 다음과 같다. 블록의 MRL 인덱스가 0이 아닌 경우 ISP를 사용하지 않는다. 분할 영역에 대한 엔트로피 코딩 시의 계수 그룹 크기는 〈표 3〉과 같이 16개의 샘플을 갖는다. 표에 나타난 상위 4가지 블록 크기는 ISP에 의해 분할되어 생성될 수 있는 것들이며, 나머지 블록 크기는 ISP가 적용되지 않은 경우와 같이 4×4 의 계수 그룹 크기를 유지한

다. Coded Block Flag(CBF)의 부호화와 관련하여, 분할 영역 중 하나 이상에 0이 아닌 CBF가 있다고 가정할 수 있으므로 이전 모든 분할 영역의 CBF가 0이라면 마지막 분할 영역의 CBF는 전송하지 않고 1로 추론한다. 변환 크기의 경우 16 포인트보다 긴 길이의 모든 ISP 변환에 대해 DCT-II를 사용하도록 제한한다. 또한, CU가 ISP 부호화 모드를 사용하는 경우 PDPC는 적용하지 않으며 Multiple Transform Selection 플래그를 0으로 추론한다.

〈표 3〉 엔트로피 코딩 계수 그룹의 크기

Block Size	Coefficient group Size
$1 \times N, N \geq 16$	1×16
$N \times 1, N \geq 16$	16×1
$2 \times N, N \geq 8$	2×8
$N \times 2, N \geq 8$	8×2
All other possible $M \times N$ cases	4×4

4. Multi-Reference Line(MRL)

MRL(Multiple Reference Line)은 화면내 예측을 수행할 때 더 많은 수의 참조 라인(0, 1, 3 lines)을 사용하는 기술이다. <그림 5>에서 세그먼트 A 및 F의 샘플은 재구성된 샘플로부터 가져오는 것이 아니라, 각각 세그먼트 B 및 E로부터 가장 가까운 샘플로 채워진다. HEVC의 화면내 예측은 <그림 6>에서의 참조 라인 0, 즉 가장 가까운 참조 라인만을 지원한다. MRL에는 이뿐만 아니라 <그림 5>의 참조 라인 1과 참조 라인 3의 2개 참조 라인을 추가적으로 사용하여 3개 참조 라인 중 한 개 참조 라인을 선택할 수 있다. 선택된 참조 라인의 인덱스(mrl_idx)를 별도로 시그널링하여 화면내 예측 과정에 필요한 참조 화소를 생성한다. 참조 라인 1 혹은 참조 라인 3이 선택된 경우, 즉 0보다 큰 참조 라인 인덱스인 경우, 항상 MPM을 사용하며 비 MPM 모드(remaining mode)는 사용할 수 없으며 PDPC를 적용하지 않는 제약이 있다. 현재 CTU 라인의 바깥 쪽에 대해서는 MRL을 사용할 수 없으며, 이

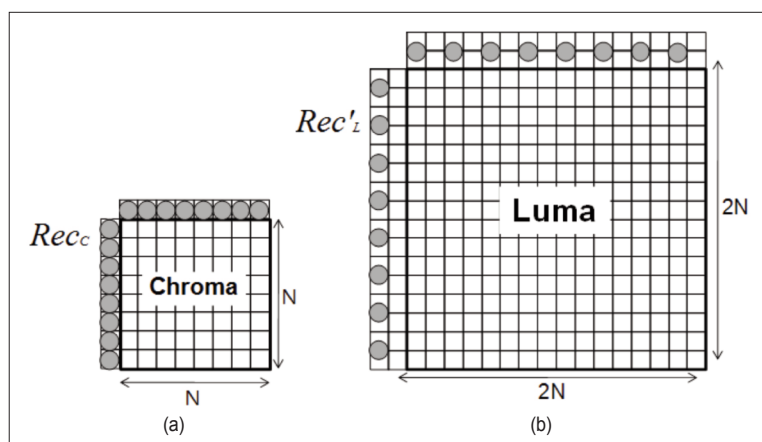
에 따라 CTU 내부 블록의 첫 번째 라인에 해당하는 CU에 대해서는 MRL을 사용할 수 없다.

5. Cross-Component Linear Model (CCLM)

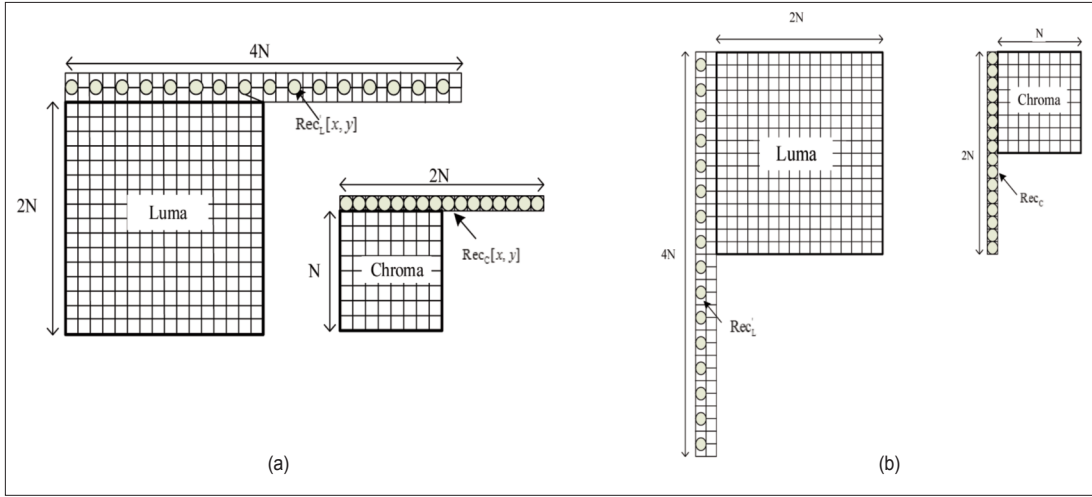
CCLM(Cross-Component Linear Model)은 휘도 신호와 색차 신호 간에 존재하는 중복성을 제거하여 부호화 예측 성능을 향상하기 위한 기술이다. 이를 위하여, CCLM은 색차 신호의 샘플과 동일한 위치의 재구성된 휘도 성분 샘플간의 연관성을 계산한 선형 모델(linear model)을 통하여 예측한다.

$$pred_c(i, j) = \alpha \cdot rec'_l(i, j) + \beta \quad (4)$$

상기 식에서 $pred_c(i, j)$ 는 CU에서 예측된 색차 성분 샘플을 나타내고 $rec'_l(i, j)$ 는 동일한 CU에서 재구성된 휘도 성분의 다운-샘플링된 샘플이다. 선형 모델 계수 α 및 β 는 별도로 시그널링하지 않고 주변 샘플로부터 유도한다.



<그림 6> LM_CCLM 모드에서 사용되는 주변 샘플



〈그림 7〉 (a) T_CCLM에서 사용되는 주변 샘플 (b) L_CCLM에서 사용되는 주변 샘플

다중 방향성 선형 모델(MDLM, Multi-Directional Linear Model)은 초기 CCLM 방법에서 발전하여 VVC에 채택된 기술이다. 〈그림 6〉과 〈그림 7〉에서와 같이, MDLM은 선형 모델 계수를 유도할 때 상단 및 좌측 주변 샘플을 이용하는 LM_CCLM 모드, 상단 주변 샘플만 이용하는 T_CCLM 모드, 좌측 주변 샘플만 이용하는 L_CCLM 모드를 지원한다. 현재 색차 성분 블록의 크기가 $W \times H$ 일 때, 선형 모델 계수의 유도를 위해 이용되는 주변 샘플의 W' 와 H' 는 다음과 같이 설정된다. LM_CCLM인 경우 $W'=W$, $H'=H$ 이며, T_CCLM 모드인 경우 $W'=W+H$ 이며, L_CCLM 모드인 경우 $H'=W+H$ 이다.

CCLM은 계산 복잡도를 감소시키기 위해서 선형 모델의 계수 α 및 β 를 최대 4개의 주변 색차 화소 및 같은 위치에 대응되는 4개의 다운-샘플링된 휘도 화소로부터 유도한다. LM_CCLM 모드이며 상단 및 좌측 주변 샘플이 사용 가능한 경우, $(W'/4, -1)$, $(3W'/4, -1)$, $(-1, H'/4)$, $(-1, 3H'/4)$ 4개의 샘플들을

선택한다. T_CCLM 모드이거나 상단 주변 샘플들만 사용할 수 있는 경우, $(W'/8, -1)$, $(3W'/8, -1)$, $(5W'/8, -1)$, $(7W'/8, -1)$ 4개의 샘플들을 선택한다. L_CCLM 모드이거나 좌측 주변 샘플들만 사용할 수 있는 경우 $(-1, H'/8)$, $S(-1, 3H'/8)$, $(-1, 5H'/8)$, $(-1, 7H'/8)$ 4개의 샘플들을 선택한다. 선택된 위치의 주변 휘도 샘플 4개를 다운-샘플링하고 비교하여 두 개의 작은 값(X_A^o 과 X_A^i)과 두 개의 큰 값(X_B^o 과 X_B^i)을 찾아낸다. X_A^o , X_A^i , X_B^o 과 X_B^i 에 대응하는 색차 성분 샘플을 Y_A^o , Y_A^i , Y_B^o , Y_B^i 로 표시하여, 식 (5)와 같이 작은 값들의 평균과 큰 값들의 평균인 X_A , X_B , Y_A , Y_B 를 유도한다.

$$\begin{aligned} X_A &= (X_A^o + X_A^i + 1) \gg 1; X_B = (X_B^o + X_B^i + 1) \gg 1; \\ Y_A &= (Y_A^o + Y_A^i + 1) \gg 1; Y_B = (Y_B^o + Y_B^i + 1) \gg 1; \end{aligned} \quad (5)$$

최종적으로 선형 모델 파라미터 α, β 는 식 (6)과 같이 구한다.

$$\alpha = \frac{Y_a - Y_b}{X_a - X_b}, \quad \beta = Y_b - \alpha \cdot X_b \quad (6)$$

상기에서 휘도 성분의 다운 샘플링은 4:2:0 YCbCr 규격일 때 서로 크기가 다른 휘도 블록과 색차 블록의 대응 위치를 맞추기 위해서 적용된다. VVC에서는 두 가지 유형의 다운-샘플링 필터를 재구성된 휘도 성분 샘플에 선택적으로 적용할 수 있으며 수직 방향과 수평 방향으로 2:1 다운-샘플링 비율을 가진다. 선택된 다운-샘플링 필터를 `sps_cclm_colocated_chroma_flag` 구문요소로 지정하며, SPS 레벨로 시그널링한다.

`sps_cclm_colocated_chroma_flag`가 0 및 1인 경우 각각 식 (7) 및 식 (8)과 같은 다운 샘플링을 수행한다.

현재 블록이 CTU 경계에 존재하는 경우, 라인

버퍼 제한을 위해 근접 1개의 휘도 라인만을 사용하여 다운-샘플링을 수행한다.

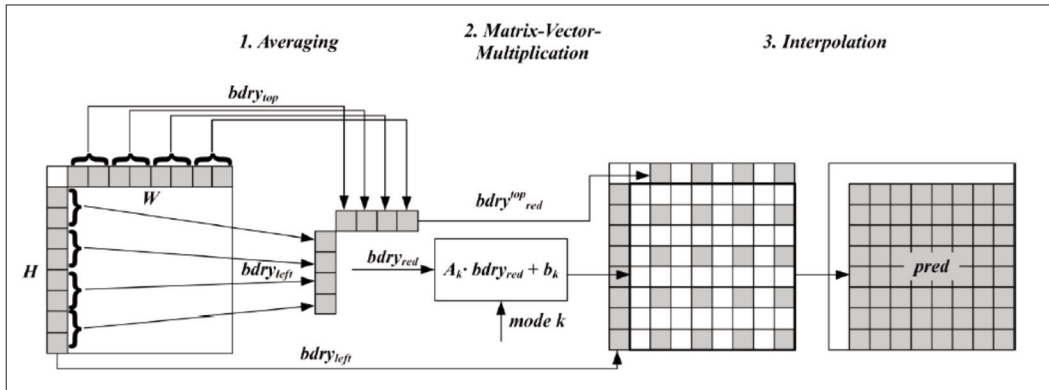
6. Matrix-weighted Intra Prediction(MIP)

휘도 샘플을 예측하기 위한 MIP는 블록 좌측의 재구성된 주변 화소와 블록 상단의 재구성된 주변 화소를 이용한다. 재구성된 주변 화소를 사용할 수 없는 경우, 기존의 화면내 예측과 동일하게 주변 화소를 생성한다. 예측 신호의 생성은 <그림 8>과 같이 경계 평균화, 행렬 벡터 곱, 선형 보간 과정 등 세 단계로 이루어진다.

첫번째로 경계 평균화 과정은 주변 샘플들을 평균화한다. <그림 9>와 같이, 경계 샘플들 중에서 현재 블록의 폭(W)과 높이(H) 크기가 모두 4인 경우,

$$\text{rec}_L'(i, j) = \left[\frac{\text{rec}_L(2i-1, 2j-1) + 2 \cdot \text{rec}_L(2i-1, 2j) + \text{rec}_L(2i+1, 2j-1) + \text{rec}_L(2i-1, 2j) + 2 \cdot \text{rec}_L(2i, 2j) + \text{rec}_L(2i+1, 2j) + 4}{4} \right] \gg 3 \quad (7)$$

$$\text{rec}_L'(i, j) = \left[\frac{\text{rec}_L(2i, 2j-1) + \text{rec}_L(2i-1, 2j) + 4 \cdot \text{rec}_L(2i, 2j) + \text{rec}_L(2i+1, 2j) + \text{rec}_L(2i, 2j+1) + 4}{4} \right] \gg 3 \quad (8)$$



<그림 8> MIP의 예측 수행 과정

$$bdry_{red} = \begin{cases} [bdry_{red}^{top}, bdry_{red}^{left}] & \text{for } W = H = 4 \text{ and } mode < 18 \\ [bdry_{red}^{left}, bdry_{red}^{top}] & \text{for } W = H = 4 \text{ and } mode \geq 18 \\ [bdry_{red}^{top}, bdry_{red}^{left}] & \text{for } \max(W, H) = 8 \text{ and } mode < 10 \\ [bdry_{red}^{left}, bdry_{red}^{top}] & \text{for } \max(W, H) = 8 \text{ and } mode \geq 10 \\ [bdry_{red}^{top}, bdry_{red}^{left}] & \text{for } \max(W, H) > 8 \text{ and } mode < 6 \\ [bdry_{red}^{left}, bdry_{red}^{top}] & \text{for } \max(W, H) > 8 \text{ and } mode \geq 6 \end{cases} \quad (9)$$

각 경계면은 평균을 취해 2개씩의 샘플로 줄인다. 그 외 다른 모든 경우는 각 경계면의 샘플에 평균을 취해 4개씩의 샘플로 줄인다. 즉, 입력 경계 $bdry_{red}^{top}$ 와 $bdry_{red}^{left}$ 는 블록 크기에 따라 주변 경계 샘플들을 평균화하여 작은 경계 $bdry_{red}^{top}$ 과 $bdry_{red}^{left}$ 로 줄인다. 다음으로, 식 (9)와 같이 두 개의 줄어든 경계 $bdry_{red}^{top}$ 와 $bdry_{red}^{left}$ 로 W, H, 및 MIP 모드에 따라 줄어든 경계 벡터 $bdry_{red}$ 를 구성한다.

두 번째 행렬 벡터 곱 과정은 평균화된 샘플에 대해 행렬 벡터 곱셈 후 오프셋을 추가하여 원래 블록의 서브 샘플링된(subsampled) 화소 세트에 대한 예측 신호를 생성한다. 즉, 입력 벡터 $bdry_{red}$ 로부터 획득하는 예측 신호 $pred_{red}$ 는 $W_{red} \times H_{red}$ 크기의 다운 샘플링된 블록에 대한 예측 신호이며, W_{red} 와 H_{red} 는 각각 식 (10), 식 (11)과 같이 정의된다.

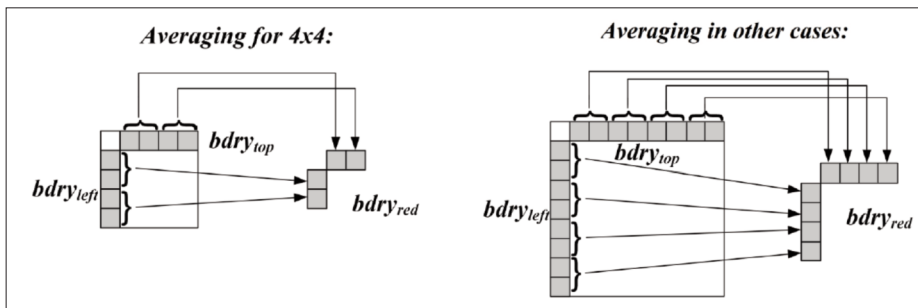
$$W_{red} = \begin{cases} 4 & \text{for } \max(W, H) \leq 8 \\ \min(W, 8) & \text{for } \max(W, H) > 8 \end{cases} \quad (10)$$

$$H_{red} = \begin{cases} 4 & \text{for } \max(W, H) \leq 8 \\ \min(H, 8) & \text{for } \max(W, H) > 8 \end{cases} \quad (11)$$

줄어든 예측 신호 $pred_{red}$ 는 식 (12)와 같이 행렬 벡터 곱셈과 오프셋을 더하는 과정을 통해 획득한다.

$$pred_{red} = A \cdot bdry_{red} + b \quad (12)$$

여기서 행렬 A는 행의 개수가 $W_{red} \times H_{red}$ 이고 열의 개수가 4개(W_{red} 와 H_{red} 가 모두 4인 경우) 또는 8개(W_{red} 와 H_{red} 가 모두 4인 경우를 제외한 나머지 모든 경우)인 행렬이다. 오프셋 b는 $W_{red} \times H_{red}$ 크기의 벡터이다. 예측 신호를 생성하는데 사용되는 행



(그림 9) 블록 크기 4x4 경우(좌)와 그 외 모든 경우(우) MIP의 평균화 과정

렬 A 와 오프셋 벡터 b 는 S_0, S_1, S_2 세트 중 하나의 세트에서 선택하며, 여기서 하나의 세트는 식 (13)과 같이 $idx = idx(W, H)$ 에 의해 결정된다.

$$idx(W, H) = \begin{cases} 0 & \text{for } W = H = 4 \\ 1 & \text{for } \max(W, H) = 8 \\ 2 & \text{for } \max(W, H) > 8 \end{cases} \quad (13)$$

세트 S_0 는 각각 16행과 4열을 가진 18개의 행렬과 16행을 가진 18개의 오프셋 벡터로 구성되며, 크기가 4×4 인 블록에 사용된다. 세트 S_1 는 16행과 8열을 가진 10개의 행렬과 16 행을 가진 10개의 오프셋 벡터로 구성되며, 크기가 $4 \times 8, 8 \times 4$ 및 8×8 인 블록에 사용된다. 마지막으로 세트 S_2 는 각각 64행과 8열의 6개의 행렬과 64행의 6개 오프셋 벡터로 구성되며, 상기 이외의 블록들에 사용된다. 상기 행렬은 전치 가능하기 때문에 S_0, S_1, S_2 세트로 각각 35, 19, 11개의 MIP 모드를 지원할 수 있다.

세 번째 선형 보간 과정은 두 번째 과정에서 얻은 다운 샘플링된 블록의 예측 신호로부터 원래블록 크기의 예측 신호를 생성하는 과정이다. <그림 11>과 같이 두 번째 과정에서 얻은 $Pred_{red}$ 신호와 주변 샘플 값을 이용하여 수직 및 수평 방향으로 보간함으로써 남은 화소들의 예측 신호를 생성한다. <그림 10>과 같이 예측 신호를 생성하는데 필요한

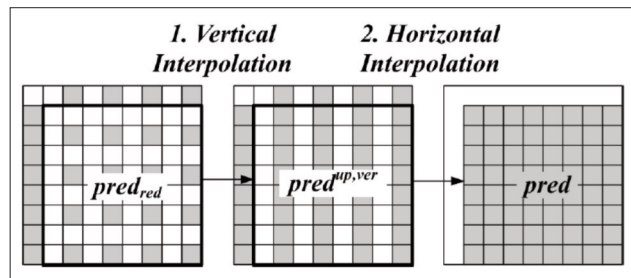
배열과 오프셋 벡터는 배열의 S_0, S_1, S_2 세트에서 구한다.

$intra_mip_flag$ 로 휘도 성분 블록에 대해 MIP의 적용 여부를 알리며, $Intra_mip_mpm_flag$, $intra_mip_mpm_idx$, $intra_mip_mpm_remainder$ 는 휘도 성분 샘플의 MIP 모드이다. MIP를 위해 MPM 목록이 사용된다. MIP 모드는 컨텍스트 부호화 bin 없이 truncated binary code로 부호화된다.

7. Reference Sample Smoothing Filter 및 Interpolation Filters

VVC는 방향성 화면내 예측 모드의 참조 샘플을 얻기 위해 [1,2,1] smoothing 필터, Gaussian 보간 필터, Cubic 보간 필터를 이용한다. HEVC에서는 방향성 예측 모드에서 화면내 예측 블록을 생성하기 위해 2-탭 선형 보간 필터를 적용하였으나, VVC는 방향성 화면내 예측 모드를 위하여 간략화된 6비트의 4-탭 Gaussian 보간 필터 혹은 Cubic 보간 필터를 선택적으로 사용한다.

Smoothing 필터, Gaussian 보간 필터, Cubic 보간 필터는 화면내 부호화 모드에 따라 다음과 같이 적용한다. 우선, 화면내 부호화 모드를 3가지 그룹으로 분류한다. 그룹 A는 수평과 수직 예측 모드



<그림 10> MIP의 보간 과정

〈표 4〉 변환 블록 크기($nTbS$)에 따른 $intraHorVerDistThres$

	$nTbS=2$	$nTbS=3$	$nTbS=4$	$nTbS=5$	$nTbS=6$	$nTbS=7$
$intraHorVerDistThres[nTbS]$	24	14	2	0	0	0

이며, 그룹 B는 45도의 배수인 대각(diagonal) 모드들을 포함하며, 그룹 C는 그 외 방향성 모드들이다. 그룹 A에 대해서는 아무 필터도 적용하지 않는다. 그룹 B에 대해서는 45도의 배수이므로 정수 위치의 화소를 참조하기 때문에 보간 필터를 적용할 필요가 없으며 참조 샘플에 대해 [1, 2, 1] smoothing 필터만 적용한다. 그룹 C에 대해서는 [1, 2, 1] smoothing 필터를 적용하지 않고 조건에 따라 4-탭 Gaussian 혹은 4-탭 Cubic 필터를 적용하여 소수 위치의 참조 샘플을 얻는다. 방향성 모드가 수직 혹은 수평 모드 중 가까운 것과의 차이 값이 Table X의 임계치($intraHorVerDistThres$)보다 큰 경우 소수 위치의 참조 샘플 값을 얻기 위해 Gaussian 필터를 선택하며, 그 외의 경우 Cubic 필터를 이용한다.

III. 결론

본 고에서는 VVC의 화면내 예측 모드 및 새롭게 추가된 화면내 부호화 기술에 대해 살펴보았다. VVC는 총 87개의 부호화 모드를 지원함으로써 37개의 HEVC/H.265보다 훨씬 더 다양한 방향으로

예측이 가능하다. 이러한 다양한 예측 모드는 예측 오차를 감소시키고 결과적으로 부호화 성능을 얻을 수 있다. 부호화 효율을 위한 새로운 부호화 기술로서, 주변 샘플 간의 차이로 위치 별 예측 신호를 보정하는 PDPC, 블록을 수평 혹은 수직으로 2 혹은 4등분하고 동일한 예측 모드를 적용하는 ISP, 더 많은 수의 참조 라인을 선택적으로 사용하는 MRL, 색차 화소와 대응하는 재구성된 휘도 화소의 선형 모델로 색차 화소를 예측하여 휘도 신호와 색차 신호의 중복성을 줄이는 CCLM, 평균값으로 다운 샘플링 된 주변화소에 학습된 행렬을 곱하고 오프셋을 더하여 블록을 예측하는 MIP, 방향성 모드에서 소수 위치의 화소를 더 정확히 예측하기 위한 Gaussian 및 Cubic 보간 필터 등이 있다. VVC는 상기 부호화 기술의 조합으로써 HEVC/H.265보다 약 23% BD-rate의 부호화 효율을 갖는다. VVC의 표준화 완료까지 새로운 툴이 더 추가되고 복잡도 감소를 위한 최적화가 진행될 예정이다.

현재 VVC 화면내 부호화를 위한 인코더 복잡도는 HEVC/H.265 대비 3000%가 넘는다. 따라서 VVC 표준화가 완료되어도 VVC 표준의 활용성을 높이기 위해서는 인코더 복잡도를 낮추기 위한 고속화 연구가 절실히 필요하다.

참고문헌

- [1] A Segall, V Baroncini, J Boyce, J Chen, and T Suzuki, Joint Call for Proposals on Video Compression with Capability beyond HEVC, ITU-T/ISO/IEC JVET, JVET-H1002, Oct. 2007
- [2] B. Bross, J. Chen, and S. Liu, Versatile Video Coding (Draft 6), ITU-T/ISO/IEC JVET, JVET-O2001, Jul. 2019
- [3] Versatile Video Coding Test Model Software 6.0
https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/tree/VTM-6.0
- [4] F. Bossen, X. Li, A. Norkin, and K. Sühring, JVET AHG report: Test model software development (AHG3), ITU-T/ISO/IEC JVET, JVET-O0003, Jul. 2019
- [5] J. Chen, Y. Ye, and S. Kim, Algorithm description for Versatile Video Coding and Test Model 6 (VTM 6), ITU-T/ISO/IEC JVET, JVET-O2002, Jul. 2019
- [6] G. Van der Auwera, V. Seregin, A. Said, A. K. Ramasubramonian, and M. Karczewicz, CE3: Simplified PDPC (Test 2.4.1), ITU-T/ISO/IEC JVET, JVET-K0063, Jul. 2018
- [7] J. Lee, H. Lee, S.-C. Lim, J. Kang, and H. Y. Kim, Non-CE3: Modification of PDPC, ITU-T/ISO/IEC JVET, JVET-M0238, Jan. 2019
- [8] S. De-Luxán-Hernández, V. George, J. Ma, T. Nguyen, H. Schwarz, D. Marpe, and T. Wiegand, CE3: Intra Sub-Partitions Coding Mode (Tests 1.1.1 and 1.1.2), ITU-T/ISO/IEC JVET, JVET-M0102, Jan. 2019
- [9] S. De-Luxán-Hernández, B. Bross, T. Nguyen, V. George, B. Stabernack, H. Schwarz, D. Marpe, and T. Wiegand, AHG16/Non-CE3: Restriction of the maximum CU size for ISP to 64x64, ITU-T/ISO/IEC JVET, JVET-N0308, Mar. 2019
- [10] K. Misra, A. Segall, M. Ikeda, T. Suzuki, K. Andersson, and J. Enhorn, Non-CE11: On ISP transform boundary deblocking, ITU-T/ISO/IEC JVET, JVET-N0473, Mar. 2019
- [11] S. De-Luxán-Hernández, V. George, G. Venugopal, J. Brandenburg, B. Bross, T. Nguyen, H. Schwarz, D. Marpe, and T. Wiegand, Non-CE3: Proposed ISP cleanup, ITU-T/ISO/IEC JVET, JVET-O0502, Jul. 2019
- [12] A. K. Ramasubramonian, G. Van der Auwera, T. Hsieh, V. Seregin, L. Pham Van, M. Karczewicz (Qualcomm), S. De-Luxan Hernandez, B. Bross, T. Nguyen, V. George, B. Stabernack, H. Schwarz, D. Marpe, and T. Wiegand, CE3-1.6: On 1xN and 2xN subblocks of ISP, ITU-T/ISO/IEC JVET, JVET-O0106, Jul. 2019
- [13] A. K. Ramasubramonian, G. Van der Auwera, L. Pham Van, and M. Karczewicz, Non-CE3: DC intra prediction mode alignment, ITU-T/ISO/IEC JVET, JVET-O0426, Jul. 2019
- [14] X. Ma, H. Yang, and J. Chen, CE3: Tests of cross-component linear model in BMS1.0 (Test 4.1.8, 4.1.9, 4.1.10, 4.1.11), ITU-T/ISO/IEC JVET, JVET-K0190, Jul. 2018
- [15] C.-M. Tsai, C.-W. Hsu, C.-Y. Chen, T.-D. Chuang, Y.-W. Huang, and S.-M. Lei, CE3.5.8: Line buffer reduction for LM chroma, ITU-T/ISO/IEC JVET, JVET-L0085, Oct. 2018
- [16] G. Laroche, J. Taquet, C. Gisquet, and P. Onno, CE3: Cross-component linear model simplification (Test 5.1), ITU-T/ISO/IEC JVET, JVET-L0191, Oct. 2018
- [17] J. Choi, J. Heo, S. Yoo, L. Li, J. Choi, J. Lim, and S. Kim, CE3 : CCLM with line buffer restriction (Test 5.2.7), ITU-T/ISO/IEC JVET, JVET-L0136, Oct. 2018
- [18] Y. Yasugi, F. Bossen, and E. Sasaki, Non-CE3: CCLM table reduction and bit range control, ITU-T/ISO/IEC JVET, JVET-M0064, Jan. 2019
- [19] P. Hanhart and Y. He, CE3: Modified CCLM downsampling filter for "type-2" content (Test 2.4), ITU-T/ISO/IEC JVET, JVET-M0142, Jan. 2019
- [20] J. Huo, Y. Ma, S. Wan, Y. Yu, M. Wang, K. Zhang, L. Zhang, H. Liu, J. Xu, Y. Wang, J. Li, S. Wang, and W. Gao, CE3-1.5: CCLM derived from four neighbouring samples, ITU-T/ISO/IEC JVET, JVET-N0271, Mar. 2019
- [21] Z. Zhang, R. Sjöberg, and R. Yu, Non-CE6: On LFNST transform set selection for a CCLM coded block, ITU-T/ISO/IEC JVET, JVET-O0219, Jul. 2019
- [22] Y. Zhao, H. Yang, C.-M. Tsai, T.-D. Chuang, C.-W. Hsu, Y.-W. Huang, and S.-M. Lei, Draft text for CCLM restriction to reduce luma-chroma latency for chroma separate tree, ITU-T/ISO/IEC JVET, JVET-O1124, Jul. 2019
- [23] J. Pfaff, B. Stallenberger, M. Schäfer, P. Merkle, P. Helle, T. Hinz, H. Schwarz, D. Marpe, and T. Wiegand, CE3: Affine linear weighted intra prediction (CE3-4.1, CE3-4.2), ITU-T/ISO/IEC JVET, JVET-N0217, Mar. 2019

- [24] J. Pfaff, P. Merkle, P. Helle, H. Schwarz, D. Marpe, T. Wiegand, A. K. Ramasubramonian, T. Biatek, G. Van der Auwera, L. Pham Van, M. Karczewicz, J. Choi, J. Heo, J. Lim, M. Salehifar, S. Kim, K. Kondo, M. Ikeda, T. Suzuki, Z. Zhang, K. Andersson, R. Sjöberg, J. Ström, P. Wennersten, T. C. Ma, X. Xiu, Y. Wen, C. H. J. Jhu, X. Wang, J. Huo, Y. Ma, S. Wan, and Y. Yu, Non-CE3: Simplifications of MIP, ITU-T/ISO/IEC JVET, JVET-O0925, Jul. 2019
- [25] Z. Zhang, R. Sjöberg, and R. Yu, Non-CE6: On LFNST transform set selection for a CCLM coded block, ITU-T/ISO/IEC JVET, JVET-O0219, Jul. 2019
- [26] K. Naser, F. Le Léanec, and T. Poirier, Non-CE6 Interaction between Implicit MTS and LFNST and MIP, ITU-T/ISO/IEC JVET, JVET-O0529, Jul. 2019
- [27] A. Filippov, V. Ruffitskiy, J. Chen, G. Van der Auwera, A. K. Ramasubramonian, V. Seregin, T. Hsieh, and M. Karczewicz, CE3: A combination of tests 3.1.2 and 3.1.4 for intra reference sample interpolation filter, ITU-T/ISO/IEC JVET, JVET-L0628, Oct. 2018
- [28] N. Choi, Y. Piao, K. Choi, and C. Kim, CE3.3 related: Intra 67 modes coding with 3MPM, ITU-T/ISO/IEC JVET, JVET-K0529, Jul. 2018
- [29] F. Racape, G. Rath, F. Urban, L. Zhao, S. Liu, X. Zhao, X. Li, A. Filippov, V. Ruffitskiy, and J. Chen, CE3-related: Wide-angle intra prediction for non-square blocks, ITU-T/ISO/IEC JVET, JVET-K0500, Jul. 2018
- [30] L. Li, J. Heo, J. Choi, J. Choi, S. Yoo, and J. Lim, CE3-6.2.1: Extended MPM list, ITU-T/ISO/IEC JVET, JVET-L0165, Oct. 2018

필자소개



한희지

- 2018년 : 한밭대학교 정보통신공학과 학사
- 주관심분야 : 영상처리, 영상 압축, 비디오 코딩, 신호 처리



최재륜

- 2018년 : 한밭대학교 정보통신공학과 학사
- 주관심분야 : 영상처리, 영상 압축, 비디오 코딩, 신호 처리



권대혁

- 2012년 : 한밭대학교 멀티미디어공학과 학사
- 2014년 : 한밭대학교 멀티미디어공학과 석사
- 주관심분야 : 영상처리, 영상 압축, 비디오 코딩, 신호 처리

필자 소개



최해철

- 1997년 : 경북대학교 전자공학과 학사
- 1999년 : 한국과학기술원 전기및전자공학과 석사
- 2004년 : 한국과학기술원 전기및전자공학과 박사
- 2004년 ~ 2010년 : 한국전자통신연구원 방송미디어연구부 선임연구원
- 2010년 ~ 현재 : 한밭대학교 정보통신공학과 교수
- 주관심분야 : 영상처리, 영상 압축, 비디오 코딩, 신호 처리, 컴퓨터 비전, 딥 러닝