

ECM 변환 및 양자화 기술

□ 김경용, 김지영, 손주형, 곽진삼 / 월러스표준기술연구소

요약

본 고에서는 ECM 19.0에 포함된 기술 중에서 기존 VVC 대비 확장된 변환(Transform) 방법과 새롭게 추가된 NSPT, MTSS 및 Sign Prediction 기술에 대하여 살펴본다. 또한, VVC 대비 ECM에 새롭게 추가된 Shift Quantization Center 기술에 대해서 상세히 다루며, VVC 대비 변경된 8-State 기반의 Dependent Quantization 기술에 대해서 설명한다.

I. 서론

차세대 압축 기술에 대한 탐색은 2021년 1월 회의에서 Qualcomm이 VVC 표준에 채택되지 못한 기술들과 템플릿 기반 움직임 보정 기술 등을 통해 random access 실험 환경에서 VVC 대비 -12%의 부호화 성능을 개선한 기고서[1-2]를 발표하면서 시작되었다. 이후 2021년 4월 AHG12[3]가 설립되면서 본격적으로 Enhanced compression beyond VVC capability에 대한 탐색 실험이 시작되었다. Enhanced compression beyond VVC capability에 대한 탐색 실험은 참조 소프트웨어 ECM(Enhanced Compression Model)[4]을 기반으로 진행 중이다.

II. ECM에서 개선된 변환 기술

ECM에서는 UHD 해상도 이상의 영상에서 CTU 및 변환 블록의 최대 크기를 128에서 256으로 증가시켰다. 그 이외의 해상도 영상에서는 CTU 최대 크기가 128로 유지된다. CTU 및 변환 블록의 최대 크기가 256으로 증가되면서, ECM에서는 128, 256 포인트 변환 커널이 추가되었다. 또한 성능 향상을 위해서, 1차 변환 과정에서의 zero-out 과정이 삭제되었다. 반면, 2차 변환인 LFNST(Low-Frequency Non-Separable Transform)에서의 zero-out 과정은 삭제되지 않았다. 그리고, 인트라 예측 블록의 최대 크기도 128x128로 증가되었다. ECM의 부호화 성능을 향상시키기 위해서, MTS(Multiple Transform Selection)에서는 새로운 변환 커널들이 추가되었으며, LFNST에서

는 변환 커널의 개수를 증가시켰다. 또한, ECM에는 VVC 표준에 채택되지 못했던 부호 예측 기술이 포함되었으며, 새로운 변환 방법인 NSPT가 ECM에 추가되었다. 본 장에서는 ECM 변환 기술들을 채택된 순서에 따라 순차적으로 설명한다.

1. 개선된 MTS

ECM에서는 4가지의 1차 변환 커널 즉, DCT5, DST4, DST1, IDT(Identity Transform)가 추가되어 변환 커널의 개수가 총 7가지로 늘어났다. 여기서, IDT 변환 커널은 특정 조건에서만 활성화된다. 변환 커널의 개수가 늘어나게 되면, mts_idx에 대한 코드워드의 최대값도 함께 증가된다. mts_idx의 비트량 증가를 해결하기 위해서, 1차 변환 커널을 지시하는 mts_idx에 대한 코드워드의 최대값을 줄이기 위한 방안도 함께 제안되었다. 즉, mts_idx에 대한 코드워드의 최대값을 기존과 동일하게 유지될 수 있도록 변환 커널 세트를 변환 블록의 크기와 인트라 예측 모드를 사용하여 유도되도록 하였다. 변환 커널 세트를 유도하는 과정에서 인트라 예측 모드가 34보다 큰 경우에는 68에서 현재 인트라 예측 모드를 뺀 값을 사용한다. 그리고 인트라 예측 모드가 음수인 경우, 인트라 예측 모드 2를 사용하여 변환 커널 세트를 유도하며, 인트라 예측 모드가 67보다 같거나 큰 경우에는 인트라 예측 모드 66을 사용하여 변환 커널 세트를 유도한다. 또한, mts_idx 값이 ‘3’이고, 현재 변환 블록의 가로 및 세로의 크기가 모두 16 이하이면, 추가적인 조건에 따라 수직 또는 수평 방향의 변환 타입이 IDT 변환 타입으로 재설정된다. IDT 변환 타입으로 재설정되는 추가 조건으로서, 현재 블록의 화면 내 예측 모드 인덱스와 수평 방향 예측 모드 인덱스 (18) 간의 차이의 절대값이 미리 정해진 값보다 작다면, 수직 방향의 변환 타입은 IDT 변환 타입으로 재설정된다. 또한, 현재 블록의 화면 내 예측 모드 인덱스와 수직 방향 예측 모드 인덱스 (50) 간의 차이의 절대값이 미리 정해진 값

보다 작다면, 수평 방향의 변환 타입은 IDT 변환 타입으로 재설정된다.

IntraTMP를 통해 예측된 블록의 경우, 변환 선택에 사용할 인트라 예측 모드를 유도하기 위해 예측 블록에 DIMD(Decoder-side Intra Mode Derivation) 프로세스를 적용한다. 구체적으로, 각 예측 샘플에 대해 수평 및 수직 gradient를 계산하여 gradient의 히스토그램(HoG)을 구성한다. 그 후, 히스토그램 진폭(Amplitude) 값이 가장 큰 인트라 예측 모드를 MTS(Multiple Transform Selection) 변환 세트 결정에 사용한다.

2. 적응적 MTS

ECM에서는 Adaptive Intra MTS 기법이 추가되었으며, 이는 변환 커널 후보의 개수를 블록마다 적응적으로 변경함으로써, mts_idx를 부호화하기 위한 비트량을 감소시키는 방안이다. 변환 커널 후보의 개수는 변환 계수들의 절대값의 합에 따라 1개, 4개, 6개까지 적응적으로 결정될 수 있다. 변환 커널 후보의 개수에 따라 mts_idx를 시그널링하기 위한 최대 bin의 개수가 달라지게 된다. 만일 변환 계수들의 절대값의 합이 6보다 작거나 같으면, 변환 커널 후보의 개수가 하나이므로, mts_idx는 시그널링되지 않는다. 반면, 변환 계수들의 절대값의 합이 6보다 크고 32보다 작거나 같으면, 변환 커널 후보의 개수는 4개가 되고, 변환 계수들의 절대값의 합이 32보다 크면 변환 커널 후보의 개수는 6개가 된다.

3. Inter MTS

ECM에서는 현재 블록이 화면 간 부호화 모드인 경우, {(DST7, DST7), (DST7, DCT8), (DCT8, DST7), (DCT8, DCT8)}와 같이 4가지 조합 중에 하나의 변환 커널 세트가 사용되며, 어떤 변환 커널 세트가 사용되는지를 나타내는 인덱스가 CU 블록마다 시그널링된다. 그리고 복잡도 최적화를 위해서, 큰 해상도 영상의 경우에는 Inter

MTS가 적용될 수 있는 최대 CU 크기를 32x32로 설정하고, 그 이외의 해상도는 16x16으로 설정한다. 마지막으로, 4, 8, 16 포인트 변환에서 DST7과 DCT8은 separable KLT(Karhunen Loeve Transform)로 대체된다.

4. 개선된 Implicit MTS

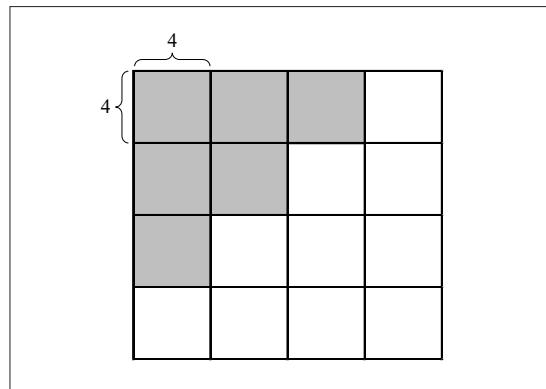
초기 Implicit MTS는 TU 크기와 인트라 예측 모드를 이용하여 테이블로부터 변환 커널을 유도하였다. 하지만 상기 방식은 테이블의 사용을 일반적인(또는 MPDIP) 인트라 모드로만 제한하고 있어, Non-regular 인트라 예측 모드들은 여전히 기본 DCT-II만을 사용했다. 이에 DIMD, TIMD, MIP, OBIC, Intra TMP, SGPM, EIP, IntraNN 부호화 블록 각각에 대해 새로운 추가 테이블이 정의되었다. 각 인트라 모드 타입에 대해 제1 예측 모드와 제2 예측 모드가 주어진다. 구체적으로, SGPM, MIP, EIP, TMP의 경우, 예측 블록으로부터 HoG를 유도한 후, HoG 값 중 첫 번째와 두 번째로 높은 모드로 제1 예측 모드와 제2 예측 모드가 각각 설정된다. DIMD, OBIC, TIMD, MPDIP의 경우, 예측에 사용된 가중치에 따라 제1 인트라 예측 모드와 제2 인트라 예측 모드가 설정되거나, 만약 제1, 제2 인트라 예측 모드가 같거나 제2 인트라 예측 모드가 없는 경우에는 주변 복원 픽셀을 이용한 HoG를 통해 유도된다. 최종적으로 최적의 변환 커널은 TU 크기, 제1 인트라 예측 모드, 그리고 인트라 예측 모드 간의 차이($diff = abs(ipm1 - ipm2)$)에 따라 테이블에서 선택된다.

5. 개선된 LFNST

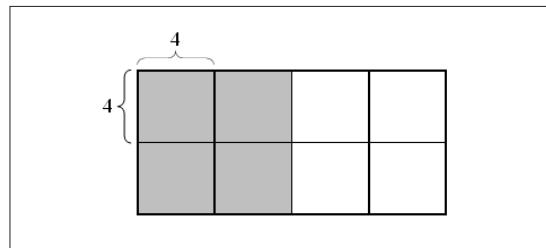
ECM에서 LFNST의 변환 세트는 35개로 늘어났으며, 각 세트당 3개의 후보가 존재한다. 현재 블록의 인트라 예측 모드에 따라 35개 LFNST 세트 중에서 하나의 세트가 선택된다. 이때, Planar와 DC 모드를 제외하고, 인트라 예측 모드가 2보다 작은 경우에는 모두 2번 LFNST 세

트가 적용되며, 인트라 예측 모드가 66보다 큰 경우에도 모두 2번 LFNST 세트가 적용된다. 각 LFNST 세트마다 3개의 커널 후보가 존재하며, 3개 후보 중에서 현재 블록에 대한 최적의 LFNST 커널 후보에 대한 인덱스 정보는 시그널링된다.

변환 블록의 크기에 따라 LFNST4, LFNST8, LFNST16 까지 총 3가지 종류의 커널이 존재한다. LFNST4/LFNST8/LFNST16은 각각 현재 변환 블록의 가로 혹은 세로의 크기가 4/8/16인 블록에 적용한다.

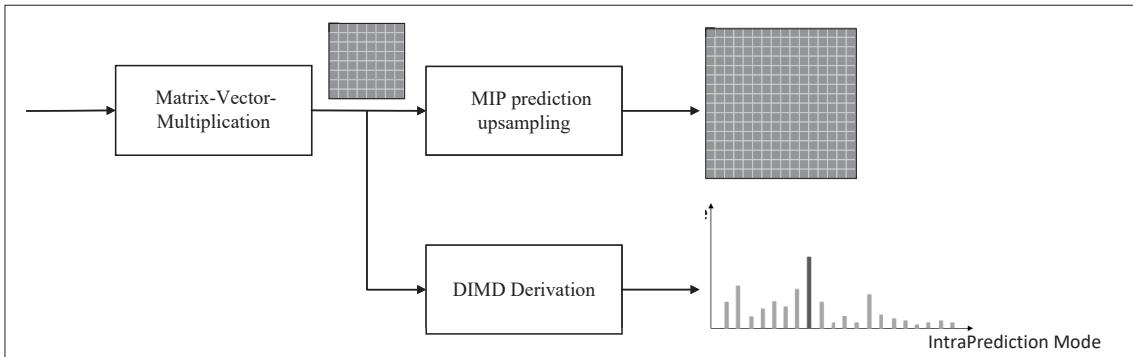


<그림 1> The ROI for LFNST16 (회색 서브 블록: ROI)



<그림 2> The ROI for LFNST8 (회색 서브 블록: ROI)

<그림 1>을 참고하면, LFNST 변환 후 2D 기준으로 왼쪽 위의 저주파 부분이 ROI (Region of Interest)이며, 미리 지정된 ROI 이외의 부분은 0으로 zero-out 처리된다. LFNST16의 경우, 96개 샘플만 필요하므로, <그림 1>과 같이 6개의 4x4 서브 블록을 제외한 나머지 흰색 서브 블록들은 zero-out 처리된다. LFNST8의 경우, 64개 샘플만



<그림 3> Derivation of LFNST set using MIP prediction block

필요하므로, <그림 2>와 같이 4개의 4×4 서브 블록을 제외하고 나머지 흰색 서브 블록들은 zero-out 처리된다. 마지막으로, LFNST4의 경우, zero-out 되는 영역이 없다.

인트라 예측 모드 중에는 기존 VVC에서의 MIP 부호화 모드와 ECM에서 새롭게 추가된 IntraTMP 모드가 존재한다. 해당 모드들에는 방향성 인트라 예측 모드가 없으므로, LFNST 세트를 유도하기 위해서 MIP 또는 IntraTMP를 사용하여 예측된 휘도 블록을 사용하여 인트라 예측 모드를 유도한다. MIP 모드에서는 <그림 3>과 같이 MIP를 사용하여 예측할 때, 업샘플링이 적용되기 전 블록에 DIMD(Decoder side Intra Mode Derivation)를 적용하여 방향성 인트라 예측 모드를 유도한 후, 유도된 인트라 예측 모드로 LFNST 세트를 유도한다. IntraTMP 부호화 모드에서도 IntraTMP 방법으로 예측된 블록에 DIMD(HoG)를 적용하여 인트라 예측 모드를 유도한 후, 유도된 인트라 예측 모드를 사용하여 MTS 혹은 LFNST 세트를 유도한다.

6. 새롭게 추가된 NSPT

NSPT는 non-separable 변환 방식으로 ECM에 새롭게 추가된 변환 방법이다. NSPT는 기존 separable 변환 방식인 DCT2와 non-separable 변환 방식인 LFNST를 대체하는 변환 방식으로, 16×16 변환 블록 크기보다

작은 블록에서 최초로 적용되었다. 즉, 변환 블록의 크기가 16×16 이하인 블록에 대하여, 변환 블록의 크기에 따라 NSPT 4×4 (16×16 kernel), NSPT 4×8 (32×20 kernel), NSPT 8×4 (32×20 kernel), NSPT 8×8 (64×32 kernel), NSPT 4×16 (64×24 kernel), NSPT 16×4 (64×24 kernel), NSPT 8×16 (128×40 kernel), NSPT 16×8 (128×40 kernel) 중에서 하나가 적용된다. 그 이후, 변환 블록의 가로 혹은 세로의 크기가 32인 경우에 대하여, NSPT 4×32 (128×20 kernel), NSPT 32×4 (128×20 kernel), NSPT 8×32 (256×24 kernel), NSPT 32×8 (256×24 kernel)가 추가로 적용되었다.

NSPT는 LFNST와 비슷하게 35개의 커널 세트와 세트 당 3개의 후보로 되어 있으며, 인트라 예측 모드에 따라 세트가 정해지고, 시그널링된 정보에 따라 해당 세트 내의 3개 후보 중 하나가 결정된다. NSPT가 적용된 4×4 블록에는 zero-out이 수행되지 않고, 다른 블록에는 NSPT의 커널의 크기에 따라 zero-out되는 계수의 개수가 달라진다. 예를 들어, 4×8 블록 또는 8×4 블록에는 32×20 크기의 NSPT가 적용되므로, 32개 변환 계수 중 20개 변환 계수만 남고 나머지 12개 변환 계수는 zero-out된다.

7. 인터 모드를 위한 LFNST/NSPT

SBT(Sub-Block Transform) 부호화 블록을 포함한 인

터 부호화 블록의 경우, 현재 블록의 예측 신호를 기반으로 DIMD 방식의 HoG를 생성하여, 가장 높은 진폭에 해당하는 첫 번째 DIMD 인트라 예측 모드를 인트라 예측 모드로 사용한다. 그리고, 유도된 인트라 예측 모드는 LFNST/NSPT 변환 세트를 선택하는 데 사용되며, 인트라 부호화 과정과 마찬가지로 선택된 커널을 사용하여 변환 이 수행된다.

GPM(Geometric Partitioning Mode) 부호화 블록의 경우, DIMD 방식의 HoG로부터 유도된 첫 번째 DIMD 인트라 예측 모드 외에도 두 번째로 높은 HoG 진폭에 해당하는 DIMD 인트라 예측 모드가 추가적으로 사용된다. 그리고 LFNST/NSPT 변환 세트를 선택하는 데 사용되는 인트라 예측 모드는 첫 번째 인트라 예측 모드와 두 번째 인트라 예측 모드 중 최적인 것이 선택되며, CU 레벨 플래그로 최적 정보가 시그널링된다. 그리고, Neural network 기반으로 인트라 예측 블록을 생성하는 모드인 Intra NN에서도 GPM 모드에서 사용하는 변환 커널 세트를 선택하는 방식을 사용한다.

SBT 부호화 블록에서 0이 아닌 변환 계수가 존재하는 TU에 대해 LFNST/NSPT의 수행 여부를 지시하는 정보가 시그널링된다. 이때 인트라 예측 모드는 TU 영역의 예측 신호를 이용한 DIMD(HoG) 방식으로 유도되며, 이렇게 유도된 모드가 LFNST/NSPT 변환 커널 선택에 사용된다.

8. Intra LFNST/NSPT를 위한 MTSS (Multiple transform set selection)

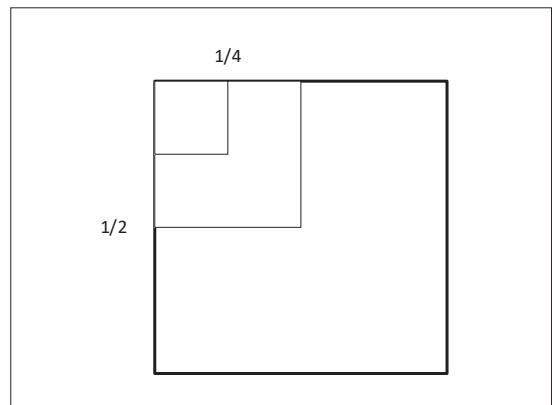
Intra NN 및 GPM 모드에 적용되던 인트라 LFNST/NSPT를 위한 MTSS 기술이 DIMD, OBIC, TIMD, SGPM, MIP, EIP 및 IntraTMP 모드로 확장되었다. LFNST/NSPT 변환 커널 선택에 사용되는 제2 인트라 예측 모드는 DIMD 방식의 HoG 값 중 두 번째로 높은 모드로부터 유도되며, 이때 HoG 유도 과정의 연산 복잡도를 줄이기 위해 서브 샘플링(Subsampling) 기법이 적용된다.

NSPT가 적용된 블록은 예측 모드 타입에 따라 NSPT 커널 세트를 선택하는 방식이 적용된다. 일반적인 인트라 모드는 Set 0을 선택하고, 예측 모드 타입이 TIMD, DIMD, OBIC, MIP, EIP인 경우는 Set 1이 선택되고, 예측 모드 타입이 IntraTMP, Inter인 경우는 Set 2가 선택된다. 만약 제1 인트라 예측 모드와 제2 인트라 예측 모드 간의 차이가 2 이하인 경우, 제1 인트라 예측 모드를 사용하여 일반 인트라 모드 커널 세트에서 커널이 선택된다.

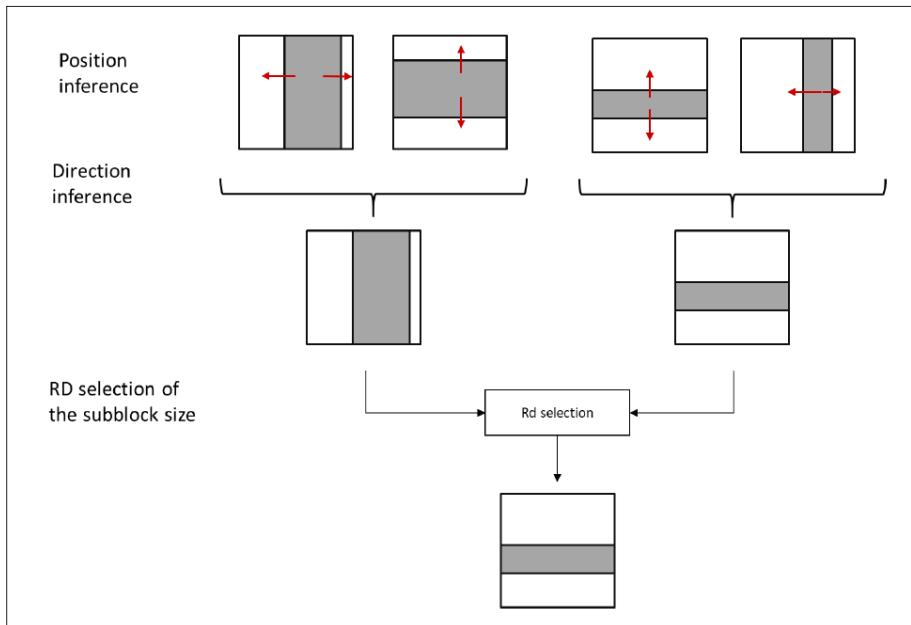
SGPM 모드의 경우, LFNST/NSPT 변환 커널 세트를 결정하기 위해 기존처럼 분할 각도를 사용하는 대신, SGPM 예측 블록에 DIMD(HoG) 방식을 사용하여 두 개의 인트라 예측 모드를 유도하는 방식을 사용한다. 이렇게 유도된 두 개의 인트라 예측 모드를 SGPM의 기존 인트라 예측 모드들과 비교하여, 만약 유도된 모드가 특정 임계값 내에서 SGPM의 기존 인트라 예측 모드와 유사하다면 유도된 모드를 사용하여 커널을 결정하고, 그렇지 않다면 기존 ECM 방식대로 분할 각도 기반의 인트라 예측 모드를 사용한다.

9. SBT(Sub Block Transform)의 개선

SBT에 4개의 코너(모서리) 방식이 새롭게 도입되었다. <그림 4>와 같이, 서브 블록의 크기는 ($\text{너비}/2$) x ($\text{높이}/2$)



<그림 4> Corner subblock



<그림 5> SBT with direction and position inference

이/2) 또는 (너비/4) x (높이/4)가 될 수 있다. 변환 커널은 암묵적으로 유도되며, DCT-8과 DST-7의 조합이 사용된다. 기존 SBT와 마찬가지로, 분할 모드(split mode)와 0이 아닌 변환 계수를 가진 서브 블록의 위치 정보가 시그널링된다. 또한 LFNST 변환도 적용되며, 해당 서브 블록들에 대해 어떤 변환을 선택할지가 시그널링된다.

〈그림 5〉를 참고하면, SBT를 위한 서브 블록 분할의 위치와 방향이 시그널링 없이 암묵적으로 유도되는 암묵적 SBT 모드가 추가되었다. 별도의 플래그(Flag)를 사용하여 시그널링되며, 직사각형 형태로 분할된 경우에만 적용된다. 암묵적 SBT의 TU 경계는 디블록킹 필터 과정에서 고려된다. 각 서브블록의 크기와 방향에 대해, 인코더는 〈그림 5〉의 변환 부분(회색 영역)의 최적 위치를 평가한다. 그런 다음, 최적의 위치를 기반으로 각 세분화에 대한 방향이 결정되며, 최적의 세분화 방식을 결정하게 된다. 선택되는 위치는 인터 예측으로부터 얻은 그래디언트의 합을 최대화하는 서브 블록의 위치이고, 방향은 선택

된 각 최적 위치에서 그래디언트의 합이 최대가 되는 방향이 선택된다.

10. 변환 계수 코딩의 개선

변환 계수는 정규 모드와 바이패스(bypass) 모드를 사용하여 코딩된다. ECM은 VVC 대비 정규 모드로 코딩되는 gtN의 범위를 식 (1)과 같이 대폭 늘렸다. N이 7인 경우, 최대 8개의 선택스가 정규 모드로 코딩되고 나머지 선택스는 바이패스 모드로 코딩된다.

$$\begin{aligned} \text{abs_coeff_level} = & \text{sig} + \text{gt1} + \text{gt2} + \text{gt3} + \dots \\ & + \text{gtN} + \text{par} + 2^{\text{remLevel}} \end{aligned} \quad (1)$$

변환 계수에 대한 예측 코딩도 적용된다. 특정 위치의 계수가 0이 아닐 확률이 미리 지정된 값보다 큰 경우, sig_coeff_flag와 abs_level_gt1_flag을 1로 간주할지 여

부를 지시하는 플래그가 시그널링된다. 만일 해당 플래그가 1이면, sig_coeff_flag와 abs_level_gt1_flag는 시그널링되지 않고, 1로 간주된다. 그렇지 않고 해당 플래그가 0이면, sig_coeff_flag와 abs_level_gt1_flag은 기존 방식대로 시그널링된다.

그리고, 0이 아닌 변환 계수의 마지막 위치 정보를 나타내는 신택스에 대한 코딩 방안도 변경되었다. 0이 아닌 변환 계수의 마지막 위치 정보가 유효 영역(zero-out)을 고려한 영역)의 마지막 위치인지를 알려주는 플래그가 시그널링된다. 만일 해당 플래그가 1인 경우, suffix에 대한 코딩은 생략된다. 그렇지 않고 해당 플래그가 0인 경우, 기존 suffix에 대한 코딩이 적용된다.

11. Context modeling for LFNST/NSPT

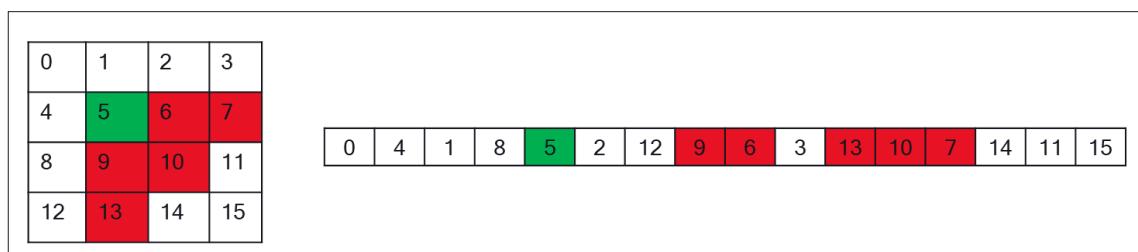
ECM에서는 LFNST 및 NSPT 변환 계수에 대한 context modeling 방법이 변경되었다. 예를 들어, <그림 6>에서 5번 스캔 위치의 변환 계수를 디코딩할 때, 기존 VVC에서는 2D 기준으로 인접한 변환 계수들(<그림 6>에서 9, 6,

13, 10, 7)을 기반으로 context가 modeling되었다. 하지만, ECM에서는 좀 더 심플한 구조로 변경되었으며, <그림 7>과 같이, 스캔 순서상 이전에 복원된 5개의 변환 계수들(<그림 6>에서 2, 12, 9, 6, 3)을 사용하여 context가 modeling된다.

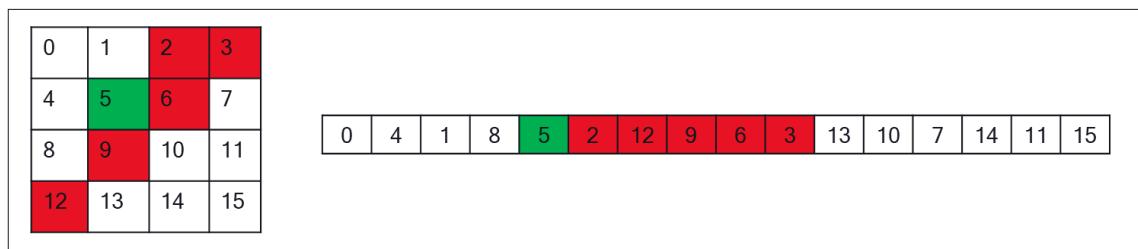
추가적으로, LFNST 및 NSPT가 적용되지 않은 블록에는 기존의 방법을 유지시키기 위해서, lfnstIdx가 변환 계수를 파싱하는 과정으로 입력된다. 즉, VVC에서는 모든 변환 계수를 파싱한 후, lfnstIdx가 파싱되었다면, ECM에서는 모든 last_sig_coeff_pos 신택스가 파싱된 후 lfnstIdx가 파싱되며, 이후 남아있는 변환 계수 부분이 파싱된다.

12. Coefficient sign prediction

Coefficient sign prediction은 ECM에서 변환 계수에 대한 부호를 예측하는 방법이다. 실제 구현은 조금 다르지만, 기본 아이디어는 변환 계수들의 부호에 대하여 양수와 음수의 조합을 통해 변환 계수 조합 후보를 구성하고,



<그림 6> Context modeling for LFNST/NSPT coefficients in VVC



<그림 7> Context modeling for LFNST/NSPT coefficients in ECM

$$\text{cost} = \sum_{x=0}^w |(-R_{x,-1} + 2R_{x,0} - P_{x,1}) - r_{x,1}| + \sum_{y=0}^h |(-R_{-1,y} + 2R_{0,y} - P_{1,y}) - r_{1,y}| \quad (2)$$

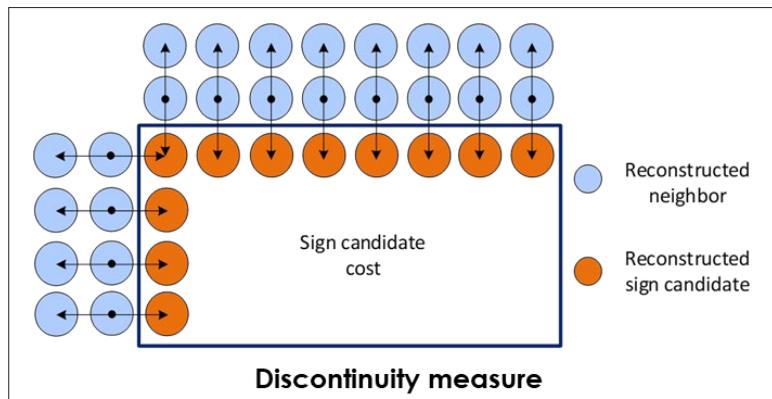
이를 사용하여 복원을 진행한 후, 식 (2)를 사용하여 최소 코스트를 갖는 변환 계수 조합 후보를 변환 계수 부호에 대한 예측 부호 조합으로 결정하고, 예측 부호 조합이 실제 부호와 같은지를 시그널링하는 방법이다. 예를 들어, 변환 계수가 2개 존재할 경우, 부호 조합 후보 개수는 (+, +), (+, -), (-, +), (-, -)로 4가지가 된다. 이때 코스트를 계산하는 수학식은 <그림 8>과 같이, 변환 블록의 경계 부분 샘플 값을 통해 계산되며, 수학식에서 대문자 R은 복원된 주변 샘플이고, 대문자 P는 현재 예측 샘플이며, 소문자 r은 현재 블록의 오차 샘플이다. 기존에는 부호에 대한 엔트로피 코딩 시, context model 없이 bypass 형태(0.5의 확률)로 부호화되었다면, 이 방법에서는 변환 계수 기반의 context model을 사용하여 부호 예측 값에 대한 엔트로피 코딩을 수행함으로써, 부호화 효율을 높인 기술이라고 볼 수 있다.

이러한 coefficient sign prediction 방법에서 예측할 Sign의 최대 개수는 8이고, 복잡도 감소를 위해서 QP 22에서 예측할 Sign의 최대 개수를 4로 줄였다. 이때, 어떤 변환 계수의 sign을 예측할지가 중요하다. ECM 초기 버전에서는 top-left 4x4 블록 내 변환 계수 중에서 raster

scan 순서로 최초 8개의 변환 계수가 선택되었다. 여기서 변환 계수의 값은 q index 즉, 양자화 변환 계수를 의미한다. 현재 ECM 버전에서는 4x4 블록이 32x32로 확장되었으며, 변환 계수의 절대값 크기가 큰 값 순서부터 작은 값 순서로 8개가 선택된다. 이때, 변환 블록의 크기가 32x32가 아닌 경우에는 가로 세로 각각에서 최소 크기가 사용된다. 또한, coefficient sign prediction 방법은 1차 변환 계수뿐만 아니라 LFNST 계수에도 적용되며, LFNST에서는 top-left 4x4 블록 내의 최대 4개 변환 계수에 대한 부호가 예측된다.

III. ECM에서 개선된 양자화 기술

VVC 대비 ECM에는 Shift Quantization Center 기술이 새롭게 추가되었으나, 이를 제외한 전반적인 양자화 기술의 큰 틀은 VVC와 유사하다. VVC 표준화 당시 높은 복잡도 문제로 채택되지 못했던 8-State 기반의 Dependent Quantization이 추가되었다. 또한, 상태(State) 개수가 기존 4개(VVC)에서 8개로 늘어남에 따라 Trellis 기반의 경



<그림 8> 코스트 계산에 사용되는 샘플 위치

$$\hat{x}_i \leftarrow ((1024 - T[|y_i|]) * Q^{-1}(y_i) + T[|y_i|] * Q^{-1}(y'_i)) >> 10 \quad if \quad |T| > |y_i| > 0 \quad (3)$$

$$y'_i = y_i + (y_i > 0? 1:-1) \quad (4)$$

$$T = [0,63,31,21,15,12,10,9,7,7,6,5,5,4,4,4,3,3,3,3,3,2,2,2,2,2,2,2,2,2,1,\dots,1] \quad (5)$$

로 탐색이 더욱 정교해졌으며, 결과적으로 부호화 효율이 대폭 개선되었다.

1. Shifting quantization center

Shifting quantization center는 역양자화된 변환 계수 레벨의 조정값이 quantization index에 따라 반비례하도록 설계함으로써 부호화 효율을 증가시킨 기술이다. 다시 말해, quantization index의 크기가 커질수록 역양자화된 변환 계수 레벨의 조정값의 크기가 점점 작아지도록 기준의 역양자화된 변환 계수 레벨의 값이 조정된다. 수정된 역양자화된 변환 계수 레벨 \hat{x}_i 은 식 (3)을 통해 계산된다. 여기서 T 는 $[0, \alpha, \frac{\alpha}{2}, \frac{\alpha}{3}, \dots, \frac{\alpha}{\alpha}]$ 이고, y_i 는 quantization index이고, y'_i 는 식 (4)를 통해 계산된다. 그리고 α 의 최적화된 값은 63이고, 이를 $|T|$ 에 적용하면, T 는 식 (5)와 같이 정수단위로 변경되어 적용된다. 추가적으로 quantization index 값이 0이거나 혹은 63보다 클 경우, Shifting quantization center 기술은 적용되지 않는다.

Dependent Quantizer에 적용되던 이 Shifting Quantization Center 방식을, Uniform Quantizer를 사용하는 일반적인 Residual Coding과 변환 스킵(Transform Skip) 블록을 위한 Residual Coding에도 동일하게 확장 적용하여 성능을 향상시켰다.

2. 8-State 기반의 Dependent Quantizer

VVC에서 dependent quantization 방법은 다른 복원 레벨을 갖는 2개의 양자화기 Q_0, Q_1 과 2개의 양자화기

간의 전이 절차(switching procedure)로 구성된다. 2개의 양자화기를 변환 계수에 따라 선택적으로 사용하게 함으로써, 원본 영상과 복원된 영상 간의 평균적인 화질 저하를 최소화하여 부호화 효율을 높이는 효과가 있다. ECM에서는 2개의 양자화기 간의 4-State 기반의 전이 절차가 “{[0, 2], {2, 0}, {1, 3}, {3, 1}}”에서 “[{0, 1}, {2, 3}, {1, 0}, {3, 2}]”로 변경되었으며, 2개의 양자화기 간의 8-State 기반의 전이 절차(“{[0, 2], {5, 7}, {1, 3}, {6, 4}, {2, 0}, {4, 6}, {3, 1}, {7, 5}}”)가 추가되었다. 4-State 기반의 전이 절차를 사용할지 혹은 8-State 기반의 전이 절차를 사용할지는 슬라이스 레벨 단위로 선택할 수 있다.

또한, 8-State 기반의 전이 절차가 적용되면서, quantization index를 변환 계수 레벨로 매핑하는 부분, sig_coeff_flag에 대한 컨텍스트를 선택하는 부분, 그리고 bypass 모드로 부호화된 변환 계수 레벨의 ZeroPos 파라미터를 매핑하는 부분들이 수정되었으며, 자세한 사항은 ECM Algorithm description[5] 문서의 pseudo code를 참고하기 바란다.

IV. 결 론

본 고에서는 ECM 19.0에 포함된 기술 중에서 기존 VVC 대비 확장된 변환(Transform) 방법과 새롭게 추가된 NSPT, MTSS 및 Sign Prediction 기술에 대하여 살펴보았다. 또한, VVC 대비 ECM에 새롭게 추가된 Shift Quantization Center 기술에 대해서 상세히 다루었으며, VVC 대비 변경된 8-State 기반의 Dependent Quantization 기술에 대해서 다루었다. 최근 트렌드는 무

조건적인 성능 개선을 넘어 실용적인 효율성을 중시하고 있다. 특히 변환과 양자화는 예측 과정만큼의 높은 복잡도를 요구하는 만큼, 복잡한 조합을 단순화하고 암묵적 유도

기술을 활용해 성능과 복잡도의 균형을 맞추는 연구가 집중적으로 수행되어야 할 것이다.

참 고 문 헌

- [1] Y. Chang et al., EE2: Tests of compression efficiency methods beyond VVC, JVET Document JVET-U0100, December 2020.
- [2] Y. Chang et al., EE2: Tests of compression efficiency methods beyond VVC, JVET Document JVET-V0120, April 2021.
- [3] M. Karczewicz et al., JVET AHG report: Enhanced compression beyond VVC capability (AHG12), JVET Document JVET-V0012, April 2021.
- [4] ECM reference software, <https://vcgit.hhi.fraunhofer.de/ecm/ECM>
- [5] M. Coban, and et al., “Algorithm description of Enhanced Compression Model 19 (ECM 19),” JVET Document JVET-AN2025, Jan, 2026.

저 자 소 개

김 경 용



- 2007년 : 경희대학교 컴퓨터공학과 학사
- 2009년 : 경희대학교 컴퓨터공학과 석사
- 2013년 : 경희대학교 컴퓨터공학과 박사
- 2013년 ~ 2021년 : LG전자 HE사업본부 책임연구원
- 2021년 ~ 현재 : 월러스표준기술연구소 수석연구원
- 주관심분야 : 동영상 압축 및 영상신호처리

김 지 영



- 2023년 : 국립한밭대학교 정보통신공학과 학사
- 2025년 : 국립한밭대학교 지능미디어공학과 석사
- 2025년 ~ 현재 : 월러스표준기술연구소 주임연구원
- 주관심분야 : 동영상 압축 및 영상신호처리

저자 소개

손주형



- 2001년 : 연세대학교 전자공학과 학사
- 2003년 : 서울대학교 전기공학부 석사
- 2008년 : 서울대학교 전기공학부 박사
- 2008년 ~ 2012년 : LG전자 미래IT융합연구소 책임연구원
- 2012년 ~ 현재 : 월러스표준기술연구소 상무
- 주관심분야 : 차세대 통신/방송 표준, 동영상 압축 표준

곽진삼



- 1998년 : 서울대학교 전기공학부 학사
- 2000년 : 서울대학교 전기공학부 석사
- 2004년 : 서울대학교 전기공학부 박사
- 2007년 ~ 2012년 : LG전자 차세대통신연구소 책임연구원
- 2012년 ~ 현재 : 월러스표준기술연구소 대표
- 주관심분야 : 차세대 통신/방송 표준, 동영상 압축 표준