

ECM 변환 및 양자화 기술

□ 김경용, 김동철, 손주형, 박진삼 / 월러스표준기술연구소

요약

본고에서는 ECM SW 버전 10.0에 포함된 기존 VVC 대비 확장된 변환 기법에 대해서 설명하고 새롭게 추가된 NSPT 및 sign prediction 기술들에 대하여 설명한다. 또한, 기존 VVC 대비 변경된 dependent quantization 기술 등 양자화 기술의 변화에 대해서도 설명한다.

1. 서론

비디오/오디오 코덱 국제 표준화 기구인 ISO/IEC SC29/WG11 MPEG (Moving Picture Experts Group) 과 ITU-T VCEG (Video Coding Experts Group)이 공동으로 설립한 JVET (Joint Video Experts Team)을 통해 2020년 7월에 HEVC 대비 2배 이상 부호화 성능을 개선한 VVC (Versatile Video Coding) 국제 표준[1]을 제정하였다. VVC 표준이 제정된 후 차세대 압축 기술을 탐색하는 단계가 시작되었다. 그러한 탐색의 한 축은 신경망

기반 비디오 압축 기술인 Neural Network-based Video Coding (NNVC)이다. NNVC는 AHG9을 통해 2년 동안 연구를 진행했지만 VVC 표준화를 위해 2019년에 중단되었다가, 여러 회사의 요청에 의해 2020년 10월 AHG11[2]이 설립되어 NNVC에 대한 연구가 다시 시작되었다. 차세대 압축 기술 탐색의 또 다른 축은 VVC보다 더욱 향상된 비디오 부호화 기술에 대한 것으로, 2021년 1월 회의에서 Qualcomm이 VVC 표준에 채택되지 못한 기술들과 템플릿 기반 움직임 보정 기술 등을 통해 random access 실험 환경에서 VVC 대비 -12% 정도의 부호화 성능을 개

선한 기고서[3-4]를 발표하면서 시작되었다. 이후 2021년 4월 AHG12[5]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)[6]을 기반으로 진행 중이다. 아래 <표 1>은 VVC 참조 소프트웨어인 VTM-11.0 대비 ECM-10.0[7]의 부호화 성능을 보여준다.

All Intra 실험 환경에서 -12.49%로 BD-rate 성능이 개선된 반면, 인코더/디코더 복잡도는 각각 867%/454% 증가되었다. 또한, 스트리밍 환경에서 주로 사용되는 Random Access 실험 환경에서는 -22.23%로 BD-rate 성능이 개선된 반면, 인코더/디코더 복잡도는 각각 744%/784% 증가되었다. 마지막으로 저지연 실험 환경에서는 -17.77%로 BD-rate 성능 개선에, 인코더/디코더 복잡도는 각각 621%/620% 증가되었다. 특히, ECM에서는 스크린 콘텐츠 영상에 대한 부호화 기술인 Intra Block Copy (IBC) 기술이 크게 개선되었고 Intra Template Matching Prediction (IntraTMP) 기술 등이 새롭게 포함되면서, 스크린 콘텐츠 영상에 대한 부호화 성능이 VVC 대비 크게 증가하였다. 본고에서는 ECM SW 버전 10.0에 포함된 기존 VVC 대비 확장된 transform 방법에 대해서 설명하고 새롭게 추가된 Non-Separable Primary Transform (NSPT) 및 부호 예측 (Sign Prediction) 기

술에 대하여 설명한다. 또한, 기존 VVC 대비 변경된 dependent quantization 기술에 대하여 설명한다.

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. 그리고, 성능 향상을 위해서, 1차 변환 과정에서 zero-out 과정이 삭제되었다. 단, 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, LFNST에서는 변환 커널의 개수를 증가시켰다. 또한, ECM에는 VVC 표준에 채택되지 못했던 부호 예측 기술이 포함되었으며, 새로운 변환 방법인 NSPT가 ECM에 추가되었다. 본 장에서는 ECM에 추가된 변환 기술에 대하여 설명한다.

<표 1> VTM-11.0 대비 ECM-10.0의 부호화 성능

		Over VTM-11.0ecm10				
		Y	U	V	EncT	DecT
All Intra Main 10	Overall	-12.49%	-23.52%	-24.63%	867.2%	454.4%
	Class F	-26.21%	-35.26%	-35.17%	541.1%	452.7%
	Class TGM	-39.28%	-47.12%	-46.45%	457.3%	496.2%
Random Access Main 10	Overall	-22.23%	-30.91%	-32.63%	744.0%	784.1%
	Class F	-28.30%	-36.71%	-37.10%	602.2%	482.5%
	Class TGM	-36.63%	-43.02%	-43.08%	570.5%	415.0%
Low delay B Main 10	Overall	-17.77%	-32.52%	-31.39%	621.7%	620.8%
	Class F	-24.60%	-37.20%	-36.64%	588.5%	498.3%
	Class TGM	-35.27%	-43.12%	-43.29%	560.3%	418.2%

1. 개선된 MTS

ECM에서는 4가지의 1차 변환 커널, 즉 DCT5, DST4, DST1, IDT (Identity Transform)가 추가되어 변환 커널

의 개수가 총 7가지로 늘어났다. 여기서, IDT 변환 커널은 특정 조건에서만 활성화된다. 변환 커널의 개수가 늘어나게 되면, mts_idx에 대한 코드워드의 최댓값도 함께 증가된다. mts_idx의 비트량 증가를 해결하기 위해서,

<표 2> 인트라 예측 모드에 대한 변환 커널 세트 매핑 테이블

		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	MIP
0	4x4	0	0	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4
1	4x8	5	5	6	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7	7	7	3	3	3	3	3	3	3	3	3	3	3	9
2	4x16	10	10	11	11	11	11	11	11	11	11	11	11	11	12	12	12	12	12	12	12	12	12	12	12	13	13	13	13	13	13	13	13	13	13	13	14
3	4x32	15	15	16	16	16	16	16	16	16	16	16	16	16	17	17	17	17	17	17	17	17	17	17	17	18	18	18	18	18	18	18	18	18	18	18	19
4	8x4	20	20	21	21	21	21	21	21	21	21	21	21	21	22	22	22	22	22	22	22	22	22	22	22	23	23	23	23	23	23	23	23	23	23	23	24
5	8x8	25	25	26	26	26	26	26	26	26	26	26	26	26	27	27	27	27	27	27	27	27	27	27	27	28	28	28	28	28	28	28	28	28	28	28	29
6	8x16	30	30	31	31	31	31	31	31	31	31	31	31	31	32	32	32	32	32	32	32	32	32	32	32	33	33	33	33	33	33	33	33	33	33	33	34
7	8x32	35	35	36	36	36	36	36	36	36	36	36	36	36	37	37	37	37	37	37	37	37	37	37	37	38	38	38	38	38	38	38	38	38	38	38	39
8	16x4	40	40	41	41	41	41	41	41	41	41	41	41	41	42	42	42	42	42	42	42	42	42	42	42	43	43	43	43	43	43	43	43	43	43	43	44
9	16x8	45	45	46	46	46	46	46	46	46	46	46	46	46	47	47	47	47	47	47	47	47	47	47	47	48	48	48	48	48	48	48	48	48	48	48	49
10	16x16	50	50	51	51	51	51	51	51	51	51	51	51	51	52	52	52	52	52	52	52	52	52	52	52	53	53	53	53	53	53	53	53	53	53	53	54
11	16x32	55	55	56	56	56	56	56	56	56	56	56	56	56	57	57	57	57	57	57	57	57	57	57	57	58	58	58	58	58	58	58	58	58	58	58	59
12	32x4	60	60	61	61	61	61	61	61	61	61	61	61	61	62	62	62	62	62	62	62	62	62	62	62	63	63	63	63	63	63	63	63	63	63	63	64
13	32x8	65	65	66	66	66	66	66	66	66	66	66	66	66	67	67	67	67	67	67	67	67	67	67	67	68	68	68	68	68	68	68	68	68	68	68	69
14	32x16	70	70	71	71	71	71	71	71	71	71	71	71	71	72	72	72	72	72	72	72	72	72	72	72	73	73	73	73	73	73	73	73	73	73	73	74
15	32x32	75	75	76	76	76	76	76	76	76	76	76	76	76	77	77	77	77	77	77	77	77	77	77	77	78	78	78	78	78	78	78	78	78	78	78	79

<표 3> 변환 커널 세트 테이블

	0	1	2	3	4	5		0	1	2	3	4	5		0	1	2	3	4	5		0	1	2	3	4	5
0	18	24	17	23	8	12	20	18	24	23	19	12	17	40	8	12	19	24	11	17	60	18	12	14	21	6	21
1	18	3	7	22	0	16	21	18	24	2	17	0	23	41	18	13	1	22	2	24	61	16	11	1	22	2	17
2	18	2	17	22	3	23	22	17	3	4	22	2	13	42	6	2	17	21	19	22	62	16	11	1	22	2	17
3	18	3	15	17	12	23	23	18	12	19	23	3	15	43	16	12	11	19	8	15	63	16	13	15	7	14	19
4	18	12	3	19	10	13	24	18	12	19	23	3	10	44	8	12	11	24	8	15	64	8	12	1	19	16	23
5	18	12	19	23	13	24	25	6	12	18	24	13	19	45	6	12	17	21	13	18	65	6	12	7	9	13	21
6	18	12	17	2	3	23	26	6	12	2	21	13	18	46	6	12	17	21	2	18	66	6	12	13	2	7	18
7	18	2	17	22	12	23	27	17	11	1	22	2	18	47	16	2	17	21	1	11	67	16	12	1	21	11	17
8	18	2	11	17	22	23	28	16	17	3	11	12	23	48	6	17	19	23	12	16	68	16	11	7	19	12	15
9	18	12	19	23	3	10	29	8	12	19	23	11	24	49	6	19	23	12	16	22	69	8	12	7	11	14	21
10	16	12	13	24	7	8	30	16	13	7	12	12	19	50	6	7	11	21	9	12	70	6	12	7	11	8	9
11	16	2	11	23	12	18	31	8	12	1	18	18	22	51	16	12	11	1	7	21	71	6	12	7	11	2	21
12	13	17	2	22	12	18	32	17	11	1	12	12	18	52	6	12	11	1	17	21	72	6	12	1	11	21	22
13	17	11	2	21	12	18	33	6	11	17	12	12	18	53	6	12	11	21	1	16	73	6	7	11	16	9	12
14	16	13	19	22	3	10	34	8	11	14	12	12	22	54	8	7	9	11	12	21	74	6	12	7	11	9	21
15	18	12	13	7	14	22	35	6	12	11	14	14	16	55	6	12	7	11	14	21	75	6	12	7	11	13	17
16	16	12	11	1	18	22	36	6	12	11	1	17	21	56	6	12	7	11	1	21	76	6	12	11	21	2	7
17	17	13	3	22	12	18	37	6	12	11	2	17	21	57	16	12	11	1	2	21	77	6	12	1	11	2	7
18	6	12	1	22	13	17	38	6	11	21	1	12	17	58	6	11	17	21	1	12	78	6	12	7	11	16	21
19	16	12	13	15	2	23	39	16	12	11	7	1	5	59	6	12	7	11	9	21	79	6	12	7	11	9	16

<표 4> 변환 커널 테이블

	0	1	2	3	4	5	6	7	8	9	10	11	12
0	DCT8	DCT8	DCT8	DCT8	DCT8	DST7	DST7	DST7	DST7	DST7	DCT5	DCT5	DCT5
1	DCT8	DST7	DCT5	DST4	DST1	DCT8	DST7	DCT5	DST4	DST1	DCT8	DCT7	DCT5

	13	14	15	16	17	18	19	20	21	22	23	24
0	DCT5	DCT5	DST4	DST4	DST4	DST4	DST4	DST1	DST1	DST1	DST1	DST1
1	DST4	DST1	DCT8	DST7	DCT5	DST4	DST1	DCT8	DST7	DCT5	DST4	DST1

1차 변환 커널을 지시하는 mts_idx에 대한 코드워드의 최댓값을 줄이기 위한 방안도 함께 제안되었다. 즉, mts_idx에 대한 코드워드의 최댓값을 기존과 동일하게 유지될 수 있도록 변환 커널 세트를 변환 블록의 크기와 인트라 예측 모드를 사용하여 유도되도록 하였다. 이를 위해 ECM에서는 변환 블록의 크기와 인트라 예측 모드를 기반으로 <표 2>의 테이블을 구성하였다. 만일, 현재 변환 블록의 크기가 4x8이고, 인트라 예측 방향성 모드가 5일 경우, <표 2>에서 6이 선택된다.

다음으로, 이전 단계에서 선택된 6이 가리키는 항목을 <표 3>에서 찾으면 {18, 12, 17, 2, 3, 23}을 얻을 수 있다. 그리고, 시그널링된 현재 블록의 mts_idx가 1인 경우, {18, 12, 17, 2, 3, 23} 중에서 12가 선택된다. 마지막으로 <표 4>에서 12가 가리키는 항목의 변환 커널인 {DCT5, DCT5}가 현재 변환 블록을 위한 수평 및 수직 방향의 변환 커널이 된다.

<표 2>를 사용하여 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 조건은 현

재 블록의 화면 내 예측 모드 인덱스와 수평 방향 예측 모드 인덱스 (18) 간의 차이의 절댓값이 미리 정해진 값보다 적다면, 수직 방향의 변환 타입은 IDT 변환 타입으로 재설정된다. 또한, 현재 블록의 화면 내 예측 모드 인덱스와 수직 방향 예측 모드 인덱스 (50) 간의 차이의 절댓값이 미리 정해진 값보다 적다면, 수평 방향의 변환 타입은 IDT 변환 타입으로 재설정된다.

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 절댓값의 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 절댓값의 합이 6보다 작거나 같으면, 변환 커널 후보의 개수가 하나이므로, mts_idx는 시그널링되지 않는다. 그렇지 않고 만일 변환 계수들의 절댓값의 합이 6보다 크고 32보다 작거나 같으면, 변환 커널 후보의 개수는 4개가 되고, 변환 계수들의 절댓값의 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, 어떠한 변환 커널 세트가 사용되는지를 나타내는 인덱스가 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 KLT로 대체된다.

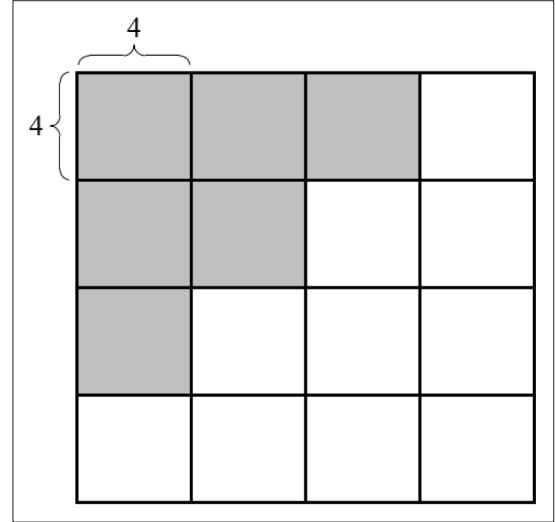
4. 개선된 LFNST

ECM에서 LFNST의 변환 세트는 35개로 늘어났으며, 각 세트당 3개의 후보가 존재한다. <표 5>는 인트라 예측 모드와 LFNST set index 간의 매핑을 나타낸다. <표 5>를 참고하면, 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 세트가 적용되며, 인트라 예측 모드가 66보다 큰 경우에도 모두 2번 LFNST 세트가 적용된다.

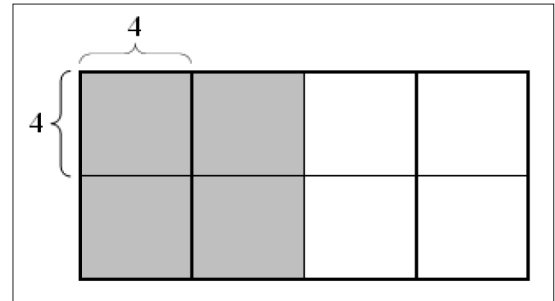
각 LFNST 세트마다 3개의 커널 후보가 존재하며, 3개 후보 중에서 현재 블록에 대한 최적의 LFNST 커널 후보에 대한 인덱스 정보는 시그널링된다.

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)

<표 5> 인트라 예측 모드에 따른 LFNST 세트 매핑 테이블

Intra pred. mode	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
LFNST set index	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Intra pred. mode	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49
LFNST set index	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20	19
Intra pred. mode	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	
LFNST set index	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	

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 필요하므로, <그림 2>와 같이 4개의 4x4 서브 블록을 제외하고 나머지 흰색 서브 블록들은 zero-out 처리된다. 마지막으로, LFNST4의 경우, zero-out되는 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 방향성 인트라

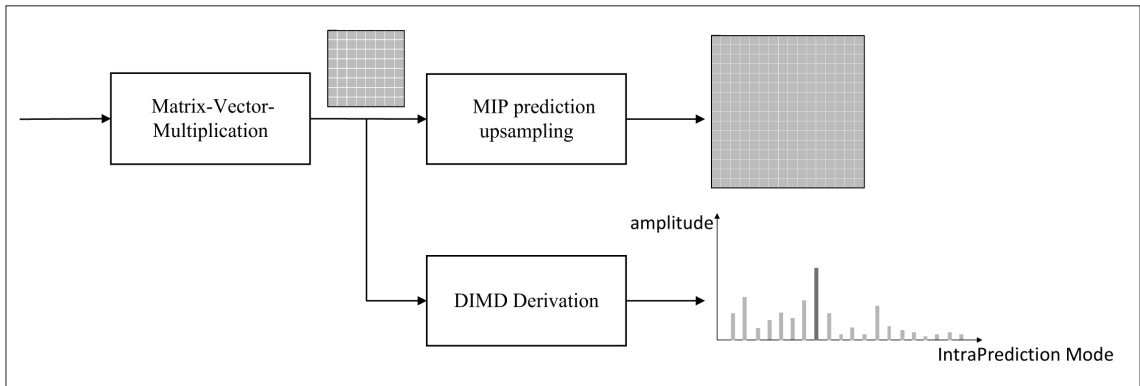
예측 모드를 유도한다.

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.

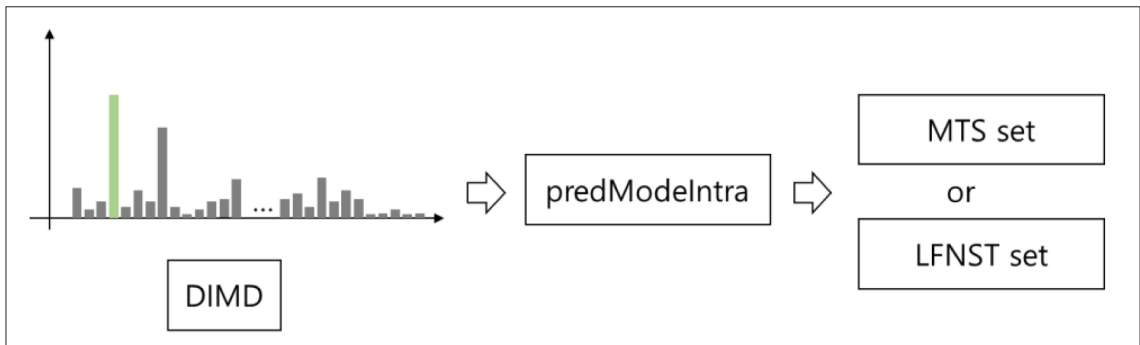
IntraTMP 부호화 모드에서도 <그림 4>와 같이 IntraTMP 방법으로 예측된 블록에 DIMD를 적용하여 방향성 인트라 예측 모드를 유도한 후, 유도된 인트라 예측 모드를 사용하여 MTS 혹은 LFNST 세트를 유도한다.

5. 새롭게 추가된 NSPT

NSPT는 non-separable 변환 방식으로 ECM에 새롭게



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



<그림 4> Derivation of MTS / LFNST set using IntraTMP prediction block

추가된 변환 방법이다. NSPT는 기존 separable 변환 방식인 DCT2와 non-separable 변환 방식인 LFNST를 대체하는 변환 방식으로, 16x16 변환 블록 크기보다 작은 블록에서 최초로 적용되었다. 즉, 변환 블록의 크기가 16x16보다 같거나 작은 블록에 대하여, 변환 블록의 크기에 따라 NSPT4x4(16x16 kernel), NSPT4x8(32x20 kernel), NSPT8x4(32x20 kernel), NSPT8x8(64x32 kernel), NSPT4x16(64x24 kernel), NSPT16x4(64x24 kernel), NSPT8x16(128x40 kernel), NSPT16x8(128x40 kernel) 중에서 하나가 적용된다. 그 이후, 변환 블록의 가로 혹은 세로의 크기가 32인 경우에 대하여, NSPT4x32(128x20 kernel), NSPT32x4(128x20 kernel), NSPT8x32(256x24 kernel), NSPT32x8(256x24 kernel)가 추가로 적용되었다.

NSPT는 LFNST와 비슷하게 35개의 커널 세트와 세트당 3개의 후보로 되어 있으며, 인트라 예측 모드에 따라 세트가 정해지고, 시그널링된 정보에 따라 해당 세트 내의 3개 후보 중 하나가 결정된다.

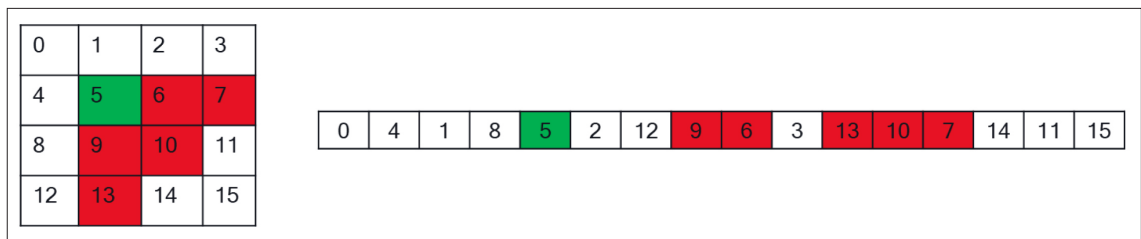
NSPT가 적용된 4x4 블록에는 zero-out이 수행되지 않고, 다른 블록에는 NSPT의 커널의 크기에 따라 zero-out

되는 계수의 개수가 달라진다. 예를 들어, 4x8 블록 또는 8x4 블록에는 32x20 크기의 NSPT가 적용되므로, 32개 변환 계수 중 20개 변환 계수만 남고 나머지 12개 변환 계수는 zero-out된다.

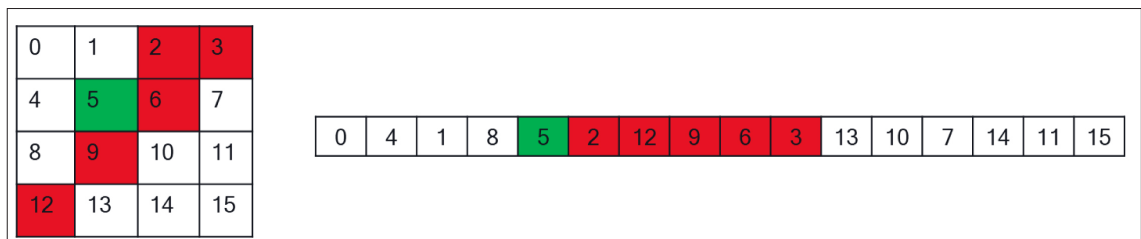
6. Context modeling for LFNST/NSPT

ECM에서는 LFNST 및 NSPT 변환 계수에 대한 context modeling 방법이 변경되었다. 예를 들어, <그림 5>에서 5번 스캔 위치의 변환 계수를 디코딩할 때, 기존 VVC에서는 2D 기준으로 인접한 변환 계수들(<그림 5>에서 9, 6, 13, 10, 7)을 기반으로 context가 modeling되었다. 하지만, ECM에서는 좀 더 심플한 구조로 변경되었으며, <그림 6>과 같이, 스캔 순서상 이전에 복원된 5개의 변환 계수들(<그림 5>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



<그림 5> Context modeling for LFNST/NSPT coefficients in VVC



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

변환 계수를 파싱한 후, lfnstIdx가 파싱되었다면, ECM에서는 모든 last_sig_coeff_pos 선택스가 파싱된 후 lfnstIdx가 파싱되며, 이후 남아있는 변환 계수 부분이 파싱된다.

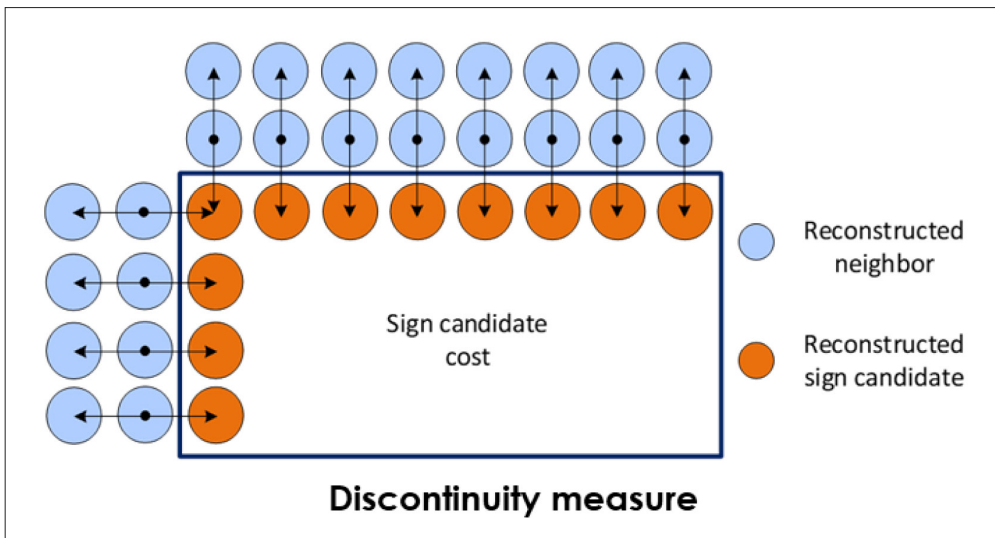
7. Coefficient sign prediction

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, 이를 사용하여 복원을 진행한 후, 아래 식 (1)을 사용하여 최소 코스트를 갖는 변환 계수 조합 후보를 변환 계수 부호에 대한 예측 부호 조합으로 하며, 예측 부호 조합이 실제 부호와 같은지를 시그널링하는 방법이다. 예를 들어, 변환 계수가 2개 존재할 경우, 부호 조합 후보 개수는 (+, +), (+, -), (-, +), (-, -) 로 4가지가 된다. 이때 코스트

를 계산하는 수식은 <그림 7>과 같이 변환 블록의 경계 부분 샘플값을 통해 계산되며, 수식식에서 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 예측값에 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 확장되었

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



<그림 7> 코스트 계산에 사용되는 샘플 위치

으며, 변환 계수의 절댓값 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

ECM에서 양자화 기술은 VVC 대비 크게 변화되지 않았다. Dependent quantization에서 8-State 기반 전이 절차가 추가되었지만, 이는 VVC 표준화 당시 제안된 기술로 복잡도 문제로 인해 채택되지 못했던 기술이다. 그 이외에는 shift quantization center 기술이 VVC 대비 ECM에 새롭게 추가된 기술이다.

1. 8-State 기반의 전이 절차

VVC에서 dependent quantization 방법은 다른 복원 레벨을 갖는 2개의 양자화기 Q0, Q1과 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}”으로 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_coff_flag에 대한 컨텍스트를 선택하는 부분, 그리고 bypass 모드로 부호화된 변환 계수 레벨의 ZeroPos 파라미터를 매핑하는 부분들이 수정되었으며, 자세한 사항은 ECM Algorithm description[7] 문서의 pseudocode를 참고하기 바란다.

2. 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 기존의 역양자화된 변환 계수 레벨의 값이 조정된다. 수정된 역양자화된 변환 계수 레벨값(\hat{x}_i)은 식 (2)를 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 (3)을 통해 계산된다. 그리고 α 의 최적화된 값은 63이고, 이를 $|T|$ 에 적용하면, T 는 식 (4)와 같이 정수 단위로 변경되어 적용된다.

추가적으로 quantization index 값이 0이거나 혹은 63보다 클 경우, Shifting quantization center 기술은 적용되지 않는다.

$$\hat{x}_i \leftarrow ((1024 - T[|y_i|]) * Q^{-1}(y_i) + T[|y_i|] * Q^{-1}(y'_i)) >> 10 \quad \text{if } |T| > |y_i| > 0 \quad (2)$$

$$y'_i = y_i + (y_i > 0 ? 1 : -1) \quad (3)$$

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

IV. 결론

본고에서는 ECM에서 기존 VVC 대비 확장된 transform 방법과 새롭게 추가된 NSPT 및 Sign prediction 기술에 대하여 살펴보았다. 또한, 기존 VVC 대비 변경된 dependent quantization 기술에 대해서도 살펴보았다. ECM에서는 변환 기술 대비 양자화 기술은 VVC 대비 변화가 크지 않다. Dependent quantization이 VVC의 부호화 효율을 크게 증가시킨 점을 고려하면, ECM에서도 새로운 양자화 기술에 대한 적극적인 연구가 필요하다. 본고에서 다룬 ECM 기술들이 제안된 기고서 리스트는 다음과 같다.

- JVET-W0119: LFNST extension with large kernel
- JVET-Y0141: Sign prediction improvement
- JVET-W0103: Extended Intra MTS
- JVET-Y0142: Adaptive Intra MTS
- JVET-Y0159: Inter MTS uses fixed 4 candidates
- JVET-AA0133: Inter MTS optimization
- JVET-AB0067: Modification of LFNST for MIP coded block

- JVET-AC0130: NSPT replacing DCT-II + LFNST for certain block shapes
- JVET-AE0086: Large NSPT kernels (4x32, 8x32)
- JVET-AC0115: Modifications of MTS/LFNST for Intra TMP coded block
- JVET-AE0125: Shifting quantization center
- JVET-AE0102: Context modelling for transform coefficients for LFNST/NSPT

ECM 10.0은 Random Access 실험 환경에서 -22.23%로 BD-rate 성능이 개선되어, 차세대 비디오 표준을 제정하기 위한 CfE (Call for Evidence)를 요청하기 충분한 정도가 되었다. 하지만 VVC 표준이 제정된지 3년 이상 지났는데도 방송 및 스트리밍 시장에서는 VVC 기술을 적극적으로 채택하고 있지 않은 상황이다. 따라서 JVET 표준화 단체에서는 차세대 비디오 표준화를 시작하는 것보다는 탐색 실험 (Exploration Experiment) 단계를 계속 진행하고 있다. 브라질에서 VVC 기반의 “TV 3.0” 지상파 방송이 2025년 초부터 시작되고, 다양한 시장에서 VVC 기술에 대한 채택이 활발해지게 되는 시점에, 차세대 비디오 코덱 표준화도 본격적으로 진행될 것으로 예상된다.

참 고 문 헌

- [1] B. Bross, Y.-K. Wang, Y. Ye, S. Liu, J. Chen, G. J. Sullivan, and J.-R. Ohm, “Overview of the versatile video coding (vvc) standard and its applications,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 10, pp. 3736-3764, 2021.
- [2] A. Alshina, S. Liu, J. Pfaff, M. Wien, P. Wu, Y. Ye, “JVET AHG report: Neural-network-based video coding (AHG11),” *JVET-T0011*, Oct. 2020.
- [3] Y. Chang et al., EE2: Tests of compression efficiency methods beyond VVC, *JVET Document JVET-U0100*, December 2020.
- [4] Y. Chang et al., EE2: Tests of compression efficiency methods beyond VVC, *JVET Document JVET-V0120*, April 2021.
- [5] M. Karczewicz et al., JVET AHG report: Enhanced compression beyond VVC capability (AHG12), *JVET Document JVET-V0012*, April 2021.
- [6] ECM reference software, <https://vcgit.hhi.fraunhofer.de/ecm/ECM>
- [7] M. Coban, and et al., “Algorithm description of Enhanced Compression Model 10 (ECM 10),” *JVET Document JVET-AE2025*, July 2023.

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 : 윌러스표준기술연구소 수석연구원
- 주관심분야 : 동영상 압축 및 영상신호처리



김 동 철

- 2004년 : 경희대학교 전자정보공학부 학사
- 2006년 : 경희대학교 전파공학과 석사
- 2006년 ~ 2016년 : LG전자 차세대표준연구소 책임연구원
- 2016년 ~ 현재 : 윌러스표준기술연구소 수석연구원
- 주관심분야 : 동영상 압축 및 영상신호처리



손 주 형

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



곽 진 삼

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