

Overview of VVC

□ 이종석, 박준택, 최한솔, 변주형, 심동규 / 광운대학교

요약

본 고는 ISO/IEC MPEG과 ITU-T VCEG이 참여하는 JVET (Joint Video Expert Team)에서 진행되고 있는 비디오 압축 표준 기술인 VVC(Versatile Video Coding)에 대하여 설명하고자 한다.

I. 서론

영상 데이터 압축에 관한 기술들은 데이터 압축 기술에 근거하여 발전해 왔다. 데이터에는 불필요하거나, 반복되거나, 혹은 인간이 인지할 수 없는 데이터가 포함되어 있는데 이러한 데이터를 제거하면 전체 데이터를 압축하여 더욱 적은 데이터로 표현할 수 있어 데이터량을 줄일 수 있을 것이다. 초기에는 이진 영상 압축 기술이 먼저 상용화되었고, 이후 정지영상 및 동영상으로도 확장되었다. 이진 영상이 아닌 다수의 밝기 값을 가지는 컬러 정지

영상의 압축을 위해서는 DCT(Discrete Cosine Transform), DST(Discrete Sine Transform)와 같은 주파수 변환을 적용하여 저주파 성분에 에너지를 모으고 이를 양자화하여 데이터를 줄인 후 엔트로피 압축하는 방법 등이 개발되었다. 이러한 변환/양자화 부호화 방법은 매우 많은 변형된 형태의 기술들로 발전되었으며, 실제 JPEG(Joint Picture Experts Group)[1]라는 표준 기술로 개발되었다. 정지 영상 이후에는, 디지털 동영상 서비스의 상용화를 목표로 정지 영상을 넘어 동영상을 압축하는 기술이 발전되었다. 동영상 압축 기술은 정지영상에 대한 모든 압축 기술을 포함하였고, 추가적으로 연속되는 영상들 간의 데이터의 중복성 및 유사성을 이용하여 데이터를 제거하는 방식으로 데이터량을 크게 줄이게 되어, 정지영상 압축 기술 대비 큰 압축 효율을 얻을 수 있게 되었다. 이러한 기술들을 바탕으로 MPEG(Moving Picture Experts Group)

에서는 MPEG-1, MPEG-2, MPEG-4라는 비디오 압축 기술을 개발하였고, VCEG(Video Coding Experts Group)에서는 H.261[2], H.262[3], H.263[4]이라는 표준화를 진행하여 표준기술로 사용되도록 하였다. 2000년 초 MPEG과 VCEG는 JVT(Joint Video Team)라는 임시 조직을 구성하여, 두 기관 공동의 표준인 H.264/AVC(Advanced Video Coding)[5] 표준 기술을 제정하였다. H.264/AVC 표준 기술은 모바일, 위성, 인터넷 기술 등이 발전함에 따라 단기간에 많은 상용화 상품으로 만들어졌다. 그러나, UHDTV(Ultra-High Definition Television)와 같은 고화질/고해상도 비디오 응용 제품의 필요성이 제기되고 특히 모바일 환경에서 VOD(Video On Demand)와 같은 서비스 수요가 증가함에 따라 H.264/AVC 표준 기술보다 더욱 높은 성능의 압축 기술을 필요로 하게 되었다. 이에 발맞춰 하드웨어 및 소프트웨어 기술의 발전으로 좀 더 복잡한 알고리즘의 구현이 가능해짐으로써 더 많은 계산량 및 더 높은 복잡성을 가지는 압축 기술 개발이 가능해졌다. MPEG과 VCEG이 공동으로 H.264/AVC보다 더 높은 압축률, 고화질/고해상도 비디오 지원 및 고속 병렬 처리를 목표로 HEVC(High Efficiency Video Coding)[6]라는 새로운 비디오 압축 기술을 개발하였다. HEVC는 2010년에 첫 미팅을 시작으로 2013년 1월에 표준화를 완료하였다. HEVC의 전체적인 기술의 구조는 Hybrid Video Coding 구조인 H.264/AVC와 유사하지만 각 모듈마다 기존의 방법보다 향상된 기술을 적용함으로써 압축 성능을 H.264/AVC 대비 주관적 화질 평가 기준 50%, 객관적 화질 평가 기준 약 35% 비트 감소시킬 수 있었다. HEVC 이후에 MPEG과 VCEG는 2015년 10월 JVET(Joint Video Exploration Team)을 구성하고 약 2년

6개월간의 표준화 미팅을 통하여 새로운 압축 기술 탐색을 거친 후 2018년 4월 JVET(Joint Video Expert Team)으로 이름을 변경하면서 새로운 비디오 압축 표준인 VVC(Versatile Video Coding)의 표준화를 시작하였다. VVC는 이름과 같이 UHD의 고화질 영상에 대한 처리뿐만 아니라 360도 비디오 및 HDR(High Dynamic Range) 비디오 서비스 지원을 목적으로하고 기존 HEVC보다 2배 이상의 압축 효율 향상을 목표로 2020년 표준화 완료를 위해 계속적으로 표준화를 진행하고 있다. 본 고에서는 현재까지 진행된 VVC에 채택된 표준 기술들과 HEVC 대비 달라진 기술들에 대하여 소개한다.

본 고의 구성은 다음과 같다. II장에서는 HEVC와 비교하여 VVC6.0에서 개선되거나 추가된 기술들에 대하여 설명한다. III장에서는 VTM6.0의 tool off test의 결과를 논의하고, IV장에서 결론을 맺는다.

II. VVC coding tools

1. 하이 레벨 신택스

VVC는 현재 32종류의 NAL 유닛을 사용함으로써 임의 접근, 시간 계층 접근, 비트스트림 분할 및 병합 등을 지원하고 있다. 임의 접근을 위한 픽처(picture)를 IRAP(Intra Random Access Picture)라고 부르고 IDR(Instantaneous Decoder Refresh), CRA(Clean Random Access), GRA(Gradual Random Access)와 같은 3가지 픽처 종류가 있다. IDR과 CLA 픽처는 픽처 전체가 인트라 모드로 코딩되어 급격히 비트레이트가 높아지는 현상이 발생한다. 이를 보완하고자 VVC에서는 저지연 응용을 위

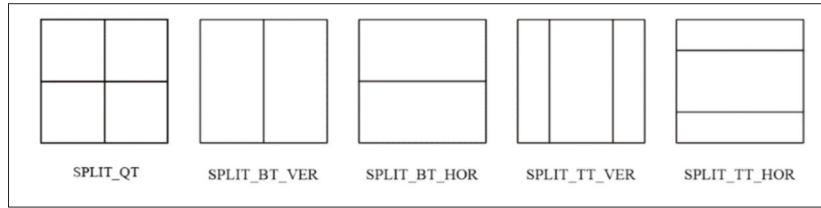
하여 HEVC에서 저지연 응용에 사용되던 Intra refresh와 유사한 기술이 적용된 GRA 픽처가 추가되었다. GRA 픽처 내 영역은 intra-coded 영역, clean 영역, dirty 영역과 같이 3가지 영역으로 구분되며, 여기서 clean 영역과 dirty 영역은 각각 에러 전과 현상이 절대 발생할 수 없는 영역과 발생할 수 있는 영역을 의미한다. GRA 픽처는 연속적으로 다수개가 나타나며, 현재 GRA 픽처의 clean 영역은 앞선 GRA 픽처들의 clean 영역을 참조하기 때문에 완벽하게 복원될 수 있다. dirty 영역은 clean 영역과는 달리 참조 영역의 제한이 없기 때문에 상황에 따라서 복호화가 되지 않을 수 있다. 이와 같은 복호화 과정을 통해서 첫 GRA 픽처 이후 특정 시점의 픽처부터는 완벽 복원이 보장되게 된다.

2. 영상분할구조

동영상은 다수의 픽처로 구성되어 있고 하나의 픽처는 압축 효율, 병렬처리 등을 위하여 계층적으로 분할되어 처리된다. HEVC에서는 하나의 픽처를 하나 또는 다수개의 타일/슬라이스로 분할하고 다시 다수개의 CTU(Coding Tree Unit)로 분할했다. VVC는 먼저 하나의 픽처를 다수개의 서브픽처(sub-picture)로 분할할 수 있다. 서브픽처는 직사각형 모양의 슬라이스의 그룹으로 정의되며, 픽처를 부분적으로 독립 부/복호화 및 전송하는 기능을 지원하기 위하여 VVC에 추가되었다. 여러 개의 서브픽처는 각각 개별적으로 복원 가능하기 때문에 멀티 채널 입력을 하나의 픽처로 구성하는 응용에 있어서 편집이 용이하다는 장점이 있다. 하나의 서브픽처는 HEVC와 유사하게 타일/슬라이스로 분할될 수 있다. VVC에서는 HEVC와 유사한 타일과 슬라이스가 존재하고 직사각형 모드 슬라이스가 추가

되어 래스터스캔 모드 슬라이스와 함께 총 두가지의 슬라이스 모드를 지원하고 있다. 또한 VVC에는 새로운 픽처 분할 구조로 브릭이 추가되었다. 브릭은 타일을 횡방향으로 분할하여 생성되며 병렬처리의 기본 단위이다. VVC는 HEVC보다 높은 해상도의 영상을 처리하기 위하여 HEVC 대비 16배 커진 최대 256x256 크기의 CTU(Coding Tree Unit)를 부/복호화의 기본 단위로서 사용한다.

HEVC에서는 슬라이스는 다수개의 CTU로 분할되고 하나의 CTU는 쿼드 트리(QT : Quadtree)로 재귀적으로 분할되고 분할의 말단 노드를 CU(Coding Unit)라고 하였다. 이후에 CU는 예측 단위인 PU(Prediction Unit)와 변환 단위인 TU(Transform Unit)로 분할되어 예측과 분할이 수행되었다. VVC에서는 유연한 분할을 위하여 각 CTU는 쿼드 트리와 멀티타입 트리(MTT : Multi-Type Tree)로 재귀적으로 분할된다. CTU는 QT의 말단 노드에서 멀티타입 트리로 분할이 시작될 수 있으며 MTT는 BT와 TT로 구성된다. 분할의 말단 노드인 CU 단위로 예측 및 변환이 수행된다. MTT 구조에는 <그림 1>과 같이 수직이진분할모드(SPLIT_BT_VER), 수평이진분할모드(SPLIT_BT_HOR), 수직삼향분할모드(SPLIT_TT_VER), 수평삼향분할모드(SPLIT_TT_HOR)가 있다. 분할 시 휘도 블록의 쿼드 트리의 최소 블록 크기(MinQTSIZE)는 16x16이고 바이너리 트리의 최대 블록 크기(MaxBtSize)는 128x128, 트리플 트리의 최대 블록 크기(MaxTtSize)는 64x64이다. 바이너리 트리의 최소 블록 크기(MinBtSize)와 트리플 트리의 최소 블록 크기(MinTtSize)는 4x4, 멀티 타입 트리의 최대 깊이(MaxMttDepth)는 4로 지정되어 있다. 또한 VVC에서는 I 슬라이스의 부호화 효율을 높이기 위하여 휘도와 색차성분의 CTU 분할 구조



〈그림 1〉 쿼드 트리와 멀티타입 트리 분할 모드

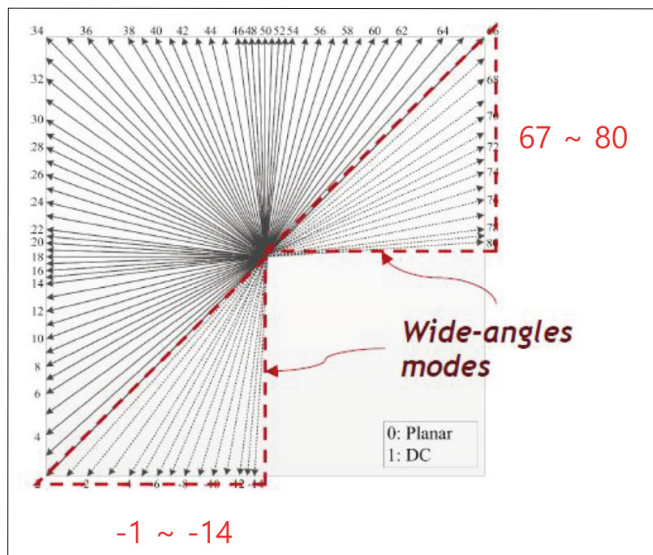
를 서로 다르게 사용하는 듀얼 트리를 채택하였다. 반대로 P와 B슬라이스에서 HEVC와 같이 한 CTU 내의 휘도와 색차 CTB들이 코딩 트리 구조를 공유하는 싱글트리로 분할된다.

3. 화면 내 예측

화면 내 예측은 현재 부호화하려는 블록의 주변에 존재하는 샘플들을 참조하여 현재 블록을 예측하는 방법을 말한다. 이때 참조 샘플들은 예측과 복원을 통해 재구성된 샘플을 의미한다. VVC는 DC,

Planar의 무방향성 예측 모드와 65개의 방향성 예측 모드 및 WAP(Wide-Angle intra Prediction), MRLP(Multiple Reference Line intra Prediction), CCLM(Cross-Component Linear Model), PDPC(Position Dependent intra Prediction Combination), ISP(Intra Sub-Partitions), MIP(Matrix-based Intra Prediction)의 화면내 예측 모드 및 기술이 적용되었다.

VVC에서는 HEVC와 다르게 정사각형이 아닌 직사각형의 CU 형태가 가능하다. 따라서 직사각형 모양의 블록에 대한 효율적인 방향성 예측을 위해



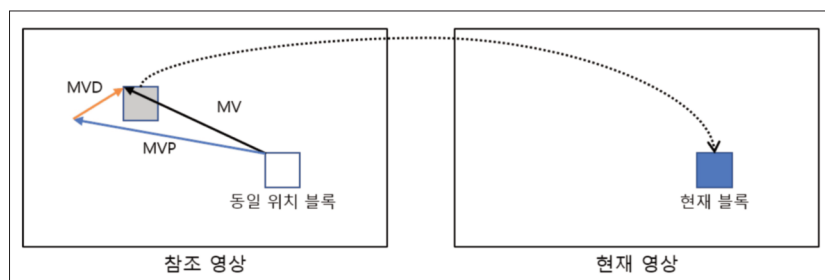
〈그림 2〉 VVC의 방향성 모드 및 WAP 모드[7-8]

WAP가 채택되었다. WAP는 현재 블록이 직사각형 일 때 추가적인 비트 전송 없이 특정 방향성 모드를 반대방향으로 예측을 수행하는 모드이다. 이때 WAP를 적용하는 모드는 현재 CU의 width와 height의 비율에 의해 결정되며, VVC의 방향성 모드 및 wide angle mode는 <그림 2>와 같다. VVC는 추가된 예측모드들을 효율적으로 전송하기 위해 기존 HEVC의 3-MPM(Most Probable Mode)에서 6-MPM으로 확장하여 사용한다. MRLP는 CU별로 `mrl_idx`를 통해 0(CU와 인접한 라인), 1, 3의 참조 샘플 라인 중 전송받은 참조 샘플 라인을 통해 화면내 예측을 수행하는 모드로, 현재 CU가 MPM으로 코딩되었을 때에만 MRLP가 사용된다. CCLM은 휘도 성분과 색차성분의 중복성을 이용하여 색차성분을 예측하는 모드이다. CCLM은 현재 CU의 이미 복원된 휘도 성분과 색차성분 간의 선형 모델을 구하고 해당 선형 모델을 통해 색차성분을 예측한다. PDPC모드는 현재 블록의 화면내 예측 모드 번호가 18(Vertical) 이하 또는 50(Horizontal) 이상 일시, 해당 모드로 예측한 예측 값과 픽셀 별 위치에 따라 왼쪽, 위쪽, 왼쪽 위 코너의 참조 샘플을 가중치 합을 통해 생성한 두 값을 가중치 합을 통해 최종 예측 값을 생성하는 모드이다. PDPC에서 사용하는 왼쪽 및 위쪽 참조 샘플의 위치와 가중

치 합의 가중치 값은 현재 CU의 모드 및 해당 픽셀의 위치에 따라 결정된다. ISP는 현재 블록을 HEVC의 PU 및 TU 분할과 유사하게 분할하는 것으로 블록 크기에 따라 각각 PU와 TU가 분할되지 않거나 2개 혹은 4개로 분할하여 각 서브 블록마다 예측을 수행하는 기술이다. 각 서브 블록은 모드를 동일하게 사용하고 하나의 서브 블록의 복호화가 끝나면 다음 서브 블록의 복호화가 진행되는 구조이다. 마지막으로 MIP는 사전에 트레이닝된 행렬을 기반으로 예측을 수행하는 모드이다. MIP는 각각의 블록크기별로 MIP모드의 수와 행렬의 크기가 정의되어 있으며 참조 샘플을 행렬의 크기에 맞게 다운 샘플링한 뒤, 모드 번호에 의해 결정된 행렬을 곱하고 예측 블록 크기에 맞도록 보간하여 예측 값을 생성한다. 이 때 MIP모드는 MPM을 사용하지 않고 truncated binary code로 모드 번호가 전송되며 MIP모드는 MPM 및 색차성분의 DM(Direct Mode)을 구성할 때 Planar모드로 매핑된다.

4. 화면 간 예측

화면간 예측은 DPB(Decoded Picture Buffer)에 저장되어 있는 이미 복호화된 픽처들을 이용하여 현재 픽처를 부호화하는 기술로 이전 픽처와 현재

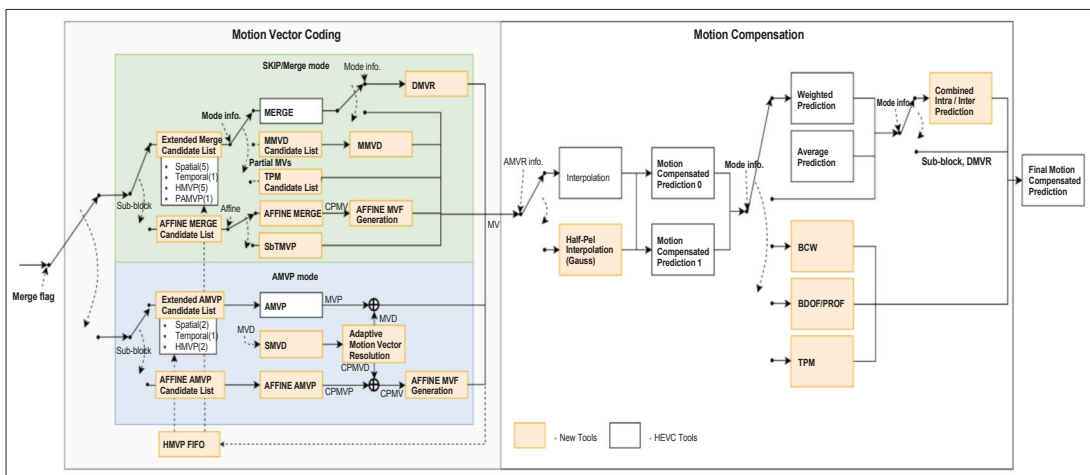


<그림 3> 화면 간 예측에서 MV 표현 예시

픽처 사이의 시간적 중복성을 제거하여 부호화 효율을 높이는 기술이다. 화면간 예측은 HEVC 이전부터 현재까지 지속적으로 사용되어왔으며 비디오 부호화에 있어서 매우 중요한 기술 중 하나이다. 일반적으로 화면간 예측은 참조 픽처와 MV(Motion Vector)를 이용하여 현재 PU(Prediction Unit)를 예측한다. 간단히 설명하자면 <그림 3>과 같이 현재 PU를 예측하기 위하여 참조영상의 어떤 위치에서 샘플들을 가져와 보상을 해주었는지 알려주는 방식이다. 참조영상에서 현재 PU의 왼쪽 상단과 동일한 위치를 기준으로 참조할 위치를 벡터로 표현한 것을 MV라고 한다. 일반적으로 MV 부호화는 MVP(Motion Vector Predictor)를 생성하고 부호화기에서 ME(Motion Estimation)로부터 예측된 MV와 차분하여 MVD(Motion Vector Difference)를 부호화하여 부호화 효율을 높이는 방법을 사용한다. HEVC에서는 MVP 생성을 위하여 AMVP(Advanced Motion Vector Predictor) 모드와 MERGE 모드를 사용하였다. AMVP모드는 주변으

로부터 MV 후보 리스트를 생성하고 리스트에서 사용된 MV의 인덱스와 MVD를 부호화하는 모드이며 MERGE 모드는 주변으로부터 AMVP 모드보다 더 많은 MV 후보들을 리스트에 저장하고 사용된 MV의 인덱스를 부호화한다.

VVC는 <그림 4>의 블록도와 같이 HEVC의 화면간 예측 방법을 기초로 sub-PU 기반 예측의 AFFINE 모드, ATMVP(Advanced Temporal Motion Vector Predictor) 모드가 있고 PU 기반 예측의 MMVD(Merge with MVD), TRIANGLE 모드가 있다. 추가적으로 각 모드의 성능 향상을 위하여 HMVP(History based MVP), PAMVP(Pairwise Average MVP), CIIP(Combined Intra/Inter Prediction), AMVR(Adaptive Motion Vector Resolution), BDOPF(Bi-Directional Optical-Flow), BCW(Bi-predictive with CU Weights) 등이 있다. <표 1>은 많은 화면간 예측 기술들을 간단하게 분류하여 정리한 표이다.



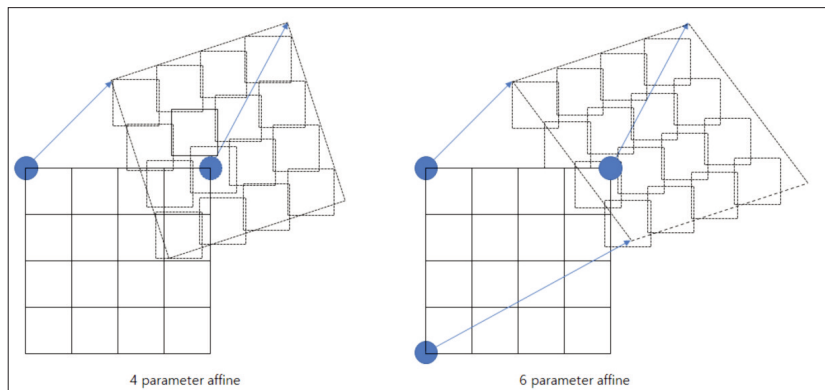
<그림 4> VVC의 화면 간 예측 블록도

〈표 1〉 VVC와 HEVC의 화면 간 예측에 포함된 기술 비교

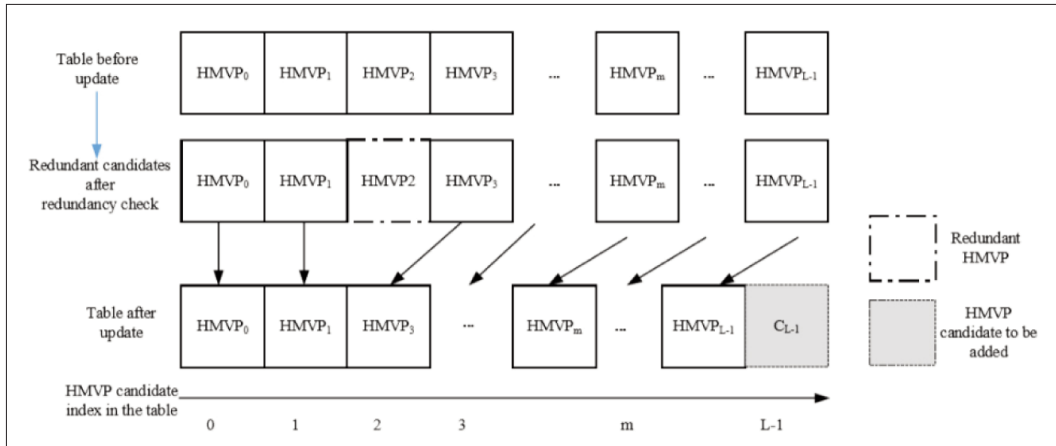
		HEVC	VVC
Motion Vector Coding	Index+MVD coding	AMVP	AMVP AFFINE INTER
	Index coding	MERGE	AFFINE MERGE SbTMVP Spatial MERGE Temporal MERGE HMVP Pairwise MVP TRIANGLE MMVD
Motion Compensation		Weighted Bi-Prediction	BCW CIIP
Additional Tools			AMVR SMVD Wrap-around MC

VVC에 새로 추가된 많은 기술들 중 AFFINE 모드는 HEVC 대비 매우 큰 변화 중 하나이다. AMVP와 MERGE 모드에서 모두 사용되고 부호화 효율 또한 높은 기술이다. 기존의 비디오 부호화 표준에서는 블록의 평행이동만을 고려하여 MC (Motion Compensation)를 수행하기 때문에 zoom in/out, rotation과 같이 현실에서 발생하는 움직임을 제대로 보상하지 못하는 단점이 있었다. 이를 보

완하여 VVC에서는 4개 또는 6개의 파라미터를 이용하는 affine 움직임 모델을 화면 간 예측에 적용하였다. AFFINE 모드는 CPMV(Control Point Motion Vector) 부호화에 대하여 AMVP 또는 MERGE 모드로 나뉜다. CPMV는 〈그림 5〉의 화살표와 같이 현재 블록의 왼쪽 상단 및 오른쪽 상단의 affine 움직임 모델을 표현하는 벡터이다. VVC는 계산 복잡도를 고려하여 픽셀단위 affine 움직임 보



〈그림 5〉 4, 6-파라미터 AFFINE 모드의 CPMV와 서브블록 단위 움직임 보상 예시



〈그림 6〉 HMVP FIFO 테이블 갱신 과정 예시

상을 수행하지 않고 4x4 블록 단위로 움직임 보상을 수행하였다. 즉, 4x4 블록 단위로 보면 기존 움직임 보상과 동일하지만, 전체 PU 관점에서는 affine 움직임 보상이 되었다고 볼 수 있다.

또 다른 큰 변화로는 TRIANGLE 모드가 있다. TRIANGLE 모드는 영상 내부의 물체의 경계 부분에서 높은 빈도수로 선택되는 모드로 현재 예측 블록을 대각선으로 나누어 각각의 MV를 이용하여 현재 예측 블록을 생성한다. 대각선으로 나누어 각각 예측할 경우에 대각선 경계에서 불연속이 심하게 발생하게 된다. VVC에서는 이를 해결하기 위하여 마스크를 이용하여 대각선 양쪽의 예측 신호를 블렌딩하여 해결하였다. 일반 MERGE 모드에서 HEVC 대비 큰 변화는 HMVP 모드의 추가일 것이다. HMVP는 현재 부호화된 MV를 HMVP 테이블에 저장하고 이후 부호화되는 블록에서 저장된 MV들을 사용하는 방법이다. HMVP 테이블은 〈그림 6〉과 같이 FIFO(First-In-First-Out) 형식으로 관리되며 블록단위 부호화 및 복호화 과정에서 계속적으로 업데이트되기 때문에 현재 블록보다 더 넓은 범위의

주변 블록들로부터 MVP를 얻을 수 있다. 즉, HMVP는 MV 부호화에 대한 기존보다 더 강한 공간적 중복성 제거를 위한 기술이라고 할 수 있다.

HMVP 이외에 MERGE 모드의 후보로 PAMVP가 있다. PAMVP는 HEVC의 Combined bi-predictive MERGE 모드에서 수정된 모드로 각각 L0/L1 리스트의 움직임 정보를 평균을 취하여 후보를 생성하여 리스트에 추가한다. 마지막으로 MMVD 모드는 MERGE 후보의 인덱스와 추가적인 MVD를 간소화된 시그널링을 통해 MV를 부호화하는 기술이다. MMVD는 AMVP와 같이 MVD를 전송하지만 AMVP와의 차이점은 MVD를 표현하는 방법이 x 또는 y축의 방향과 크기를 테이블로 만들고 그에 해당하는 인덱스만 전송한다. 또한 MMVD와 다르게 복호화기에서 MVD를 시그널링하는 것이 아니라 복호화기에서 양방향 매칭 기술을 통해 재탐색하여 MV를 보정해주는 기술인 DMVR (Decoder-side Motion Vector Refinement)도 채택되었다.

MC 과정은 단 방향 예측, 양방향 예측에 따라 추

가 과정들이 발생하게 된다. 참조 블록을 보간하여 소수 픽셀 단위로 정밀하게 예측 블록을 가져올 수 있다. VVC는 HEVC와 동일한 보간 필터를 사용한다. 예외적으로 MV의 정확도가 1/2-pel 정확도인 경우 가우시안 필터를 사용할 수 있다. 양방향 예측에 대하여 HEVC는 부호화기에서 정해놓은 가중치를 이용하여 양방향으로 생성된 예측신호를 가중 합하여 최종 예측 블록을 만들었다. 하지만, 일반적으로 부호화기의 복잡도를 감소시키기 위해서 가중치를 탐색하지 않고 두 신호의 평균으로 사용하였다. VVC에서는 부호화기 복잡도를 낮추면서 가중치를 사용하기 위하여 가중치의 조합을 테이블로 만들고 가중치 조합의 인덱스를 전송하는 방식을 채택하였고 이를 BCW라고 부른다. VVC는 BCW 이외에 예측신호를 생성한 이후에 보다 움직임 보상을 잘하기 위하여 픽셀 단위 움직임을 Optical flow를 이용하여 예측하여 부호화 효율을 높이는 BDOF와 PROF 기술도 포함되어 있다. 또는 화면내 예측신호를 생성하여 가중 합산하여 최종 예측 신호를 생성하는 CIIP 기술도 사용하고 있다.

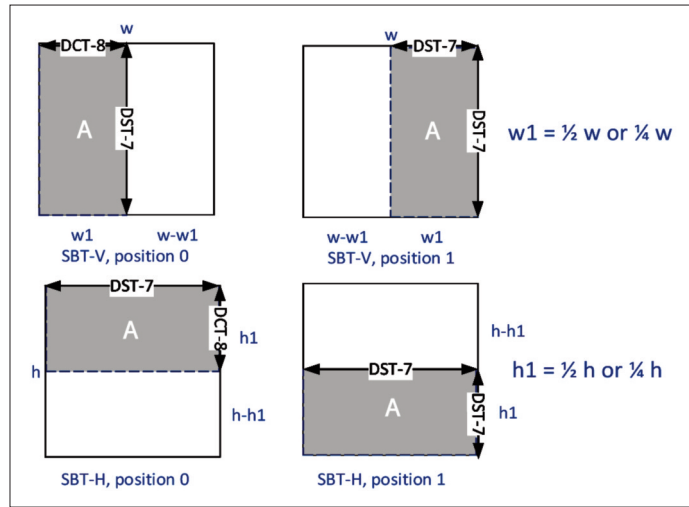
5. 변환 및 양자화

비디오 압축에서의 변환은 공간 영역의 차분 신

호를 주파수 영역으로 변환할 시 저주파로 계수가 집중되는 특징과 사람이 주관적으로 고주파 성분의 양자화 오차에 둔감하다는 특징을 이용한 것으로, 변환 이후 양자화 및 엔트로피 부호화를 통해 효율적으로 데이터 압축을 수행할 수 있다. 일반적으로 공간 영역의 입력 영상에서 주파수 영역으로의 변환은 입력 영상을 구성하는 주파수 성분을 표현하는 기저 벡터와 입력 데이터 간의 내적을 통해 수행될 수 있다. HEVC에서는 화면내 예측을 통해 생성된 4x4 휘도 잔차 블록은 DST(Discrete Sine Transform) 기반 기저 벡터를 통해 변환하며, 나머지 잔차 블록에 대해서는 DCT(Discrete Cosine Transform) 기반의 기저 벡터를 사용하여 변환을 수행한다. 또한 HEVC에서는 RQT(Residual Quad Tree) 기술을 통하여 하나의 블록에 대하여 변환 블록을 쿼드 트리 형태로 분할하며 RQT를 통해 분할된 각 변환 블록에 대하여 변환과 양자화를 수행하고 난 후 모든 계수가 0이 되는 경우의 부호화 효율을 높이기 위해 cbf(coded block flag)를 전송한다. HEVC와 달리 VVC에서는 여러 변환 기저를 선택적으로 사용하여 변환을 수행하는 MTS(Multiple Transform Selection) 기술이 채택되었다. <표 2>는 MTS 플래그와 예측 모드에 따른 변환 기저를 나타낸다.

<표 2> MTS 플래그와 예측 모드에 따른 변환 기저

MTS_CU_flag	MTS_Hor_flag	MTS_Ver_flag	Intra		Inter	
			Hor	Ver	Hor	Ver
0			DCT2			
1	0	0	DST7	DST7	DCT8	DCT8
	0	1	DST8	DST7	DCT7	DCT8
	1	0	DST7	DST8	DCT8	DCT7
	1	1	DST8	DST8	DCT7	DCT7



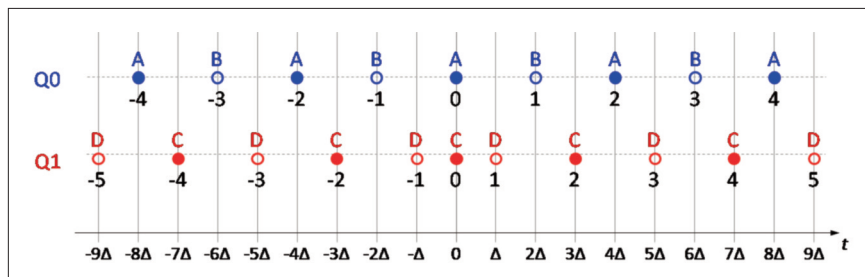
〈그림 7〉 SbT 분할의 예시

VVC에서는 CU를 RQT를 통해 TU로 분할하지 않으며 SbT(Sub-block Transform) 기술을 통해 TU분할과 유사한 기능을 수행한다. SbT는 화면간 예측 블록에만 적용되며 RQT와 달리 현재 블록을 〈그림 7〉과 같이 수직 혹은 수평 방향으로 $\frac{1}{2}$ 혹은 $\frac{1}{4}$ 크기로 분할한 뒤 그 중 하나의 블록에 대해서만 변환을 수행하는 기술이다. 수직으로 분할된 경우 맨 왼쪽 혹은 맨 오른쪽 블록에 대해 변환을 수행하고, 수평으로 분할된 경우 맨 위쪽 혹은 맨 아래쪽 블록에 대하여 변환을 수행한다.

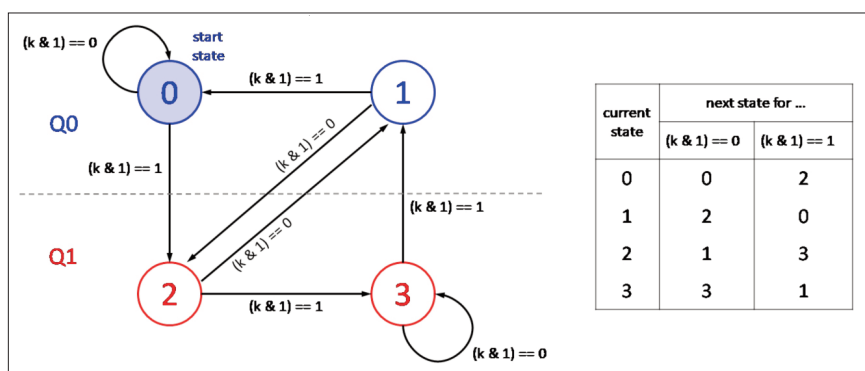
또한 VVC에서는 DCT 또는 DST를 통해 주파수 영역으로 변환된 잔차 신호를 추가 변환하는 secondary transform 기술인 LFNST(Low Frequency Non-Separable Transform)가 추가되었다. LFNST는 좌측 상단의 4×4 또는 8×8 의 저주파수 영역에 대해서 변환을 추가적으로 수행하여 잔차 계수를 좌측 상단으로 집중시킨다.

양자화 과정은 변환계수에 대하여 현재 블록의 QP(Quantization Parameter)에 의해서 계산된 특

정값으로 나누어 주는 과정이다. QP 값이 크다면 대부분의 변환 계수가 0이 되기 때문에 이후 엔트로피부호화 과정에서 부호화 효율이 올라간다. 복호화 과정에서는 이미 양자화된 값은 다시 나누어 준 값을 곱한다고 해도 원래의 값으로 복원할 수 없다. 이로 인해서 복원된 영상에서 블록 단위로 오차가 생기고 blocking, ringing 현상이 발생한다. 다시 말하면, 양자화 과정은 변환 계수에 대하여 엔트로피 부호화를 위하여 심볼화해주는 과정이라고 생각하면 된다. HEVC에서는 0~51 QP 값을 사용하는 양자화기를 사용했다. VVC에서는 영상의 크기가 더욱 커지고 부호화 효율이 높아졌기 때문에 0~63 QP를 사용하도록 확장되었다. 또한 하나의 양자화기를 사용하는 것이 아닌 두 개의 양자화기를 사용하는 DQ(Dependent Quantization) 방법을 채택하였다. DQ는 〈그림 8〉과 같이 두개의 양자화기 Q0과 Q1을 사용하여 양자화를 수행하며 이는 시그널링 없이 〈그림 9〉와 같이 상태 전이 모델을 통해 현재 상태에 따라서 다음 변환 계수에 사용할



〈그림 8〉 DQ의 양자화기[9]



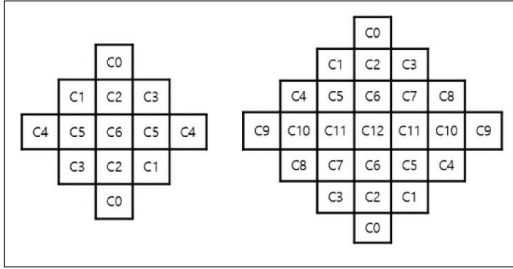
〈그림 9〉 DQ를 위한 상태전이 모델[9]

양자화기가 선택된다.

6. 인루프 필터

VVC의 Loop Filter는 HEVC부터 존재했던 DF(Deblocking Filter)와 SAO(Sample Adaptive Offset)에 ALF(Adaptive Loop Filter)와 LMCS(Luma Mapping with Chroma Scaling)가 추가되었으며 복호화기를 기준으로 inverse-LMCS, DF, SAO, ALF의 순서로 수행된다. DF는 CU 경계에서 발생하는 블록킹 열화 현상을 개선하기 위한 기술이다. 블록들이 서로 다른 QP로 부호화되고 복원될 경우 복원된 영상의 블록 경계에서 화소 값의 불연속성이 발생할 수 있으며 이는 양쪽 CU의 QP 값의

차가 커질수록 불연속성은 더욱 커진다. 따라서 VVC에서는 8x8 그리드와 블록의 경계가 겹치는 경우 DF를 수행하고 필터링의 방식은 HEVC와 거의 동일하지만 몇 가지 차이점이 있다. 첫 번째로 HEVC에서는 디블록킹 필터의 강도는 평균 양자화 파라미터 qPL에 의해 유도된 변수 β 와 tC에 의해 결정되었지만 VVC에서는, 복원된 luma 샘플의 레벨에 따라서 qPL에 offset을 더하여 필터의 강도를 결정한다. 두 번째는 VVC에서 최대 QP 값이 51에서 63으로 증가됨에 따라, QP 값에 기반하여 디블록킹 파라미터 tC를 유도하는 테이블이 수정되었다. 또한 QP 값이 큰 경우 블록킹 열화 현상과 더불어 영상의 에지 영역에서 링잉 현상(Ringing artifacts)이 발생할 수 있고 이를 개선하기 위하여



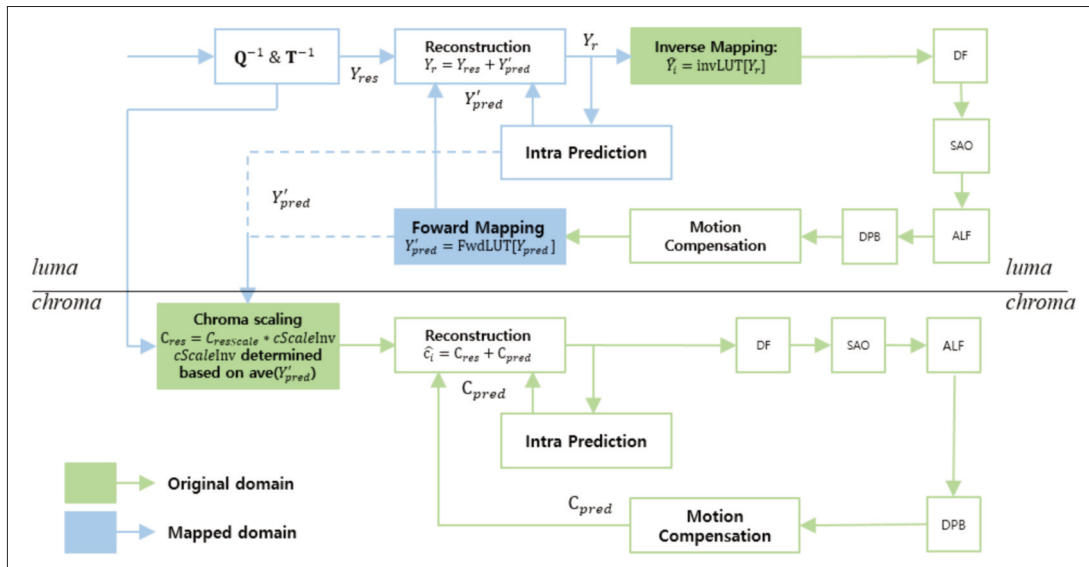
〈그림 10〉 ALF 필터 모양

SAO를 수행한다. SAO는 DF가 수행된 후의 복원 샘플에 픽셀 단위로 오프셋을 더해 원본 영상과 복원 영상과의 에러를 최소화한다. 에지에 대하여 에러를 보정하는 에지 오프셋(Edge offset)과 특정 밝기 값 대역에 대하여 에러를 보정하는 밴드 오프셋(Band offset)으로 세분화된다.

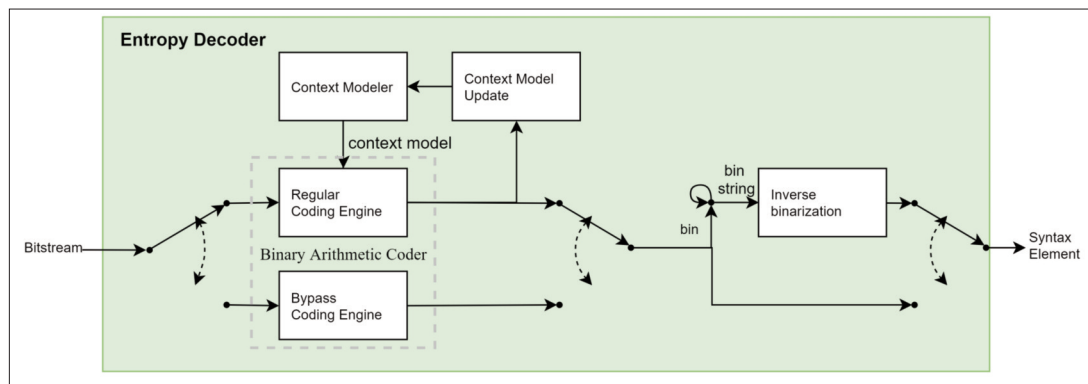
ALF는 VVC 복호화 과정에서 가장 마지막에 수행되는 기술로 최종적으로 부/복호화 손실을 보상해주는 기술로 부호화 복잡도는 높지만 부호화 효

율은 매우 뛰어난 기술이다. 부호화기에서 4x4샘플 단위로 25개의 클래스로 구별하고 각 클래스 별로 부/복호화 손실을 최소화하는 필터를 생성하여 필터셋을 부호화해준다. 추가로 클래스와 필터에 대한 맵핑 정보를 전송해 복호화기에서 부호화기와 동일하게 4x4샘플 단위로 클래스를 분류하여 전송된 필터셋 중에서 클래스에 해당하는 필터를 사용한다.

LMCS는 HDR 영상의 특성을 반영하기 위한 기술로 전처리 및 후처리로 수행되던 HDR 보정 기술을 VVC 코덱 내부에 적용한 기술이라고 할 수 있다. Luma-mapping은 piece-wise 선형 모델을 통해 휘도 값을 재맵핑하고 예측 신호의 평균 휘도 값에 따라 색차성분의 잔차 값을 스케일링해주는 기술이다. piece-wise 선형 모델은 부호화기에서 최적화되어 APS를 통해 복호화기로 전송된다. 〈그림 11〉은 LMCS가 포함된 복호화기의 블록도이다. HEVC와 같은 구조에 Inverse Mapping 과정과 Forward Mapping 과정이 추가되었고 위에서 설명한 것과



〈그림 11〉 LMCS의 블록도



〈그림 12〉 VVC의 엔트로피 복호화 블록도

같이 색차 성분의 잔차 신호를 스케일링해주는 과정이 추가되었다. 〈그림 11〉에서 파란색 영역은 매핑된 도메인을 의미하고 녹색 부분은 원래 신호 도메인을 의미하며 화면 내 예측 같은 경우에는 항상 매핑된 도메인에서 모든 처리가 이루어진다. 하지만, 화면 간 예측의 경우는 참조 픽처를 사용하기 때문에 Forward Mapping을 하여 매핑 도메인으로 변환하고 예측을 수행하고 잔차 신호와 합하여 복원된 신호를 다시 역 매핑하여 원래 신호 도메인으로 변환하여 DPB에 저장한다. Chroma scaling은 휘도신호와 대응하는 색차신호의 연관성을 이용하여 색차성분의 잔차신호의 압축 효율을 높이는 기술이다. 휘도신호의 예측신호를 이용하여 유도된 스케일계수로 스케일링된 색차신호의 잔차신호가 전송된다. 복호화기에서는 부호화기와 같은 방법으로 스케일계수를 유도하고 전송받은 색차신호의 잔차신호와 곱하여 최종 잔차신호를 생성한다.

7. 엔트로피 코딩

VVC의 엔트로피 부호화 및 복호화는 CABAC (Context-based Adaptive Binary Arithmetic

Coding)의 기술로 동작한다. 〈그림 12〉는 VVC의 엔트로피 디코더의 블록도이다. VVC CABAC는 HEVC와 동일한 알고리즘을 사용하지만, CABAC 엔진 내부에 구현적인 부분에서 차이가 있다. 복호화기에서 저장되는 확률 테이블의 크기를 줄이고자 테이블 확률 업데이트 방식에서 간단한 수식을 통한 테이블 업데이트 방법을 채택하였다. 또한 더 정확한 심볼의 확률 값을 얻기 위하여 2개의 서로 다른 확률 모델을 사용한다.

III. VTM-6.0 tool 별 성능

JVET에서는 매 회의마다 VVC에 포함된 많은 기술들의 부호화 효율 검증을 수행하고 있다. 〈표 3〉, 〈표 4〉는 각각 VTM6.0에 포함된 주요 기술 별 AI (All Intra)환경 및 RA(Random Access)환경에서의 실험 결과이다. 각각의 실험은 성능을 검증하고자 하는 기술과 다른 기술들과의 조화를 고려하여 성능을 측정하기 위하여 tool off test가 진행되었다[10]. Tool off test는 검증하고자 하는 기술을 끄고 실험하였을 때 감소되는 성능을 조사하는 방법

〈표 3〉 VTM tool “off” test (AI configuration, VTM6.0 anchor)[10]

Tools		BDR-Y	BDR-U	BDR-V	EncTime	DecTime
Block structure	CST	0.43%	9.85%	9.79%	154%	104%
Prediction	ISP	0.41%	0.29%	0.32%	86%	97%
	MIP	0.54%	0.16%	0.16%	91%	101%
	MRLP	0.39%	0.18%	0.22%	99%	101%
	CCLM	1.60%	14.79%	15.90%	99%	100%
	IBC	0.64%	0.62%	0.69%	62%	100%
Residual Coding	BDPCM	0.00%	-0.02%	0.03%	107%	100%
	MTS	1.25%	0.73%	0.77%	85%	98%
	LFNST	0.96%	0.19%	0.44%	105%	101%
	DQ	2.03%	-0.84%	-0.91%	97%	106%
	JCCR	0.56%	0.35%	0.46%	99%	101%
In-loop	SAO	0.01%	0.14%	0.17%	100%	99%
	ALF	2.38%	2.94%	3.65%	97%	88%
	LMCS	1.09%	-1.03%	-0.67%	99%	99%

〈표 4〉 VTM tool “off” test (RA configuration, VTM6.0 anchor)[10]

Tools		BDR-Y	BDR-U	BDR-V	EncTime	DecTime
Block structure	CST	0.10%	4.27%	4.81%	104%	100%
Prediction	ISP	0.20%	0.25%	0.19%	97%	100%
	MIP	0.30%	0.39%	0.41%	97%	101%
	MRLP	0.20%	0.13%	0.09%	100%	100%
	CCLM	0.94%	11.95%	13.87%	99%	100%
	IBC	-0.03%	0.02%	0.06%	93%	100%
	AFF	3.02%	2.15%	2.01%	86%	98%
	SbTMC	0.50%	0.40%	0.42%	101%	101%
	AMVR	1.39%	2.36%	2.41%	81%	101%
	TPM	0.34%	0.63%	0.51%	99%	101%
	BDOF	0.77%	0.35%	0.29%	98%	97%
	CIIP	0.29%	0.12%	0.02%	99%	101%
	MMVD	0.52%	0.55%	0.55%	93%	103%
	BCW	0.41%	0.49%	0.52%	93%	101%
	DMVR	0.81%	1.11%	1.11%	100%	97%
	PROF	0.48%	0.16%	0.09%	99%	99%
	SIF	0.28%	0.72%	0.69%	96%	100%
	SMVD	0.27%	0.26%	0.28%	88%	98%
Residual Coding	MTS	0.70%	0.62%	0.56%	93%	100%
	SBT	0.39%	0.06%	0.01%	95%	100%
	LFNST	0.67%	0.20%	0.56%	94%	100%
	DQ	1.76%	-0.24%	-0.61%	102%	103%
	JCCR	0.49%	0.46%	-0.15%	99%	101%
	RDPCM on	-0.03%	-0.07%	-0.02%	99%	100%
In-loop filtering	LMCS	1.63%	-1.30%	-1.42%	94%	100%
	SAO	0.08%	0.28%	0.35%	100%	98%
	ALF	4.53%	5.14%	5.10%	97%	87%

으로 양수 값이 클수록 좋은 부호화 효율이 높은 기술을 뜻한다.

IV. 결 론

본 고에서는 현재 진행되고 있는 비디오 압축 표준 기술인 VVC 비디오 부호화 표준을 소개하고 VVC에 포함된 기술들을 간략히 소개하였다. VVC에서는 영상분할구조가 QT와 MIT로 확장되었고, 휘도와 색차성분 CTB의 듀얼 트리 분할이 가능해졌다. 화면 내 예측에서는 방향성 예측 모드가 65가지로 확장되었고 WAP, MRLP, CCLM, PDPC 등과 같은 성능이 높은 기술들이 추가되었다. 화면 간

예측에서는 AFFINE, ATMVP, CIIP, HMVP 등의 기술이 추가되어 부호화 효율이 크게 향상되었다. 변환 단계에서는 서브블록 단위로 다양한 기저를 사용하여 변환을 수행하고 추가로 secondary transform을 수행한다. 양자화 단계에서는 두 개의 양자화기를 사용하는 DQ 기술이 추가되었고, 인루프 필터 단계에서는 ALF와 LMCS 기술이 추가되어 부호화 효율과 주관적 화질이 향상되었다. 엔트로피 코딩은 2개의 확률 모델을 사용하는 CABAC으로 변경되었다. 2020년 표준화 완료로 목표로 하여 VVC의 부호화 효율 향상과 복잡도 감소를 위한 최적화 연구가 앞으로 계속적으로 진행될 예정이다.

참 고 문 헌

- [1] Joint Photographic Experts Group, ISO/AEC JTC 1/SC29/WG 10, Digital compression and coding of continuous-tone still images - part 1: Requirements and guidelines, 1994.
- [2] Video Codec for Audiovisual Services at px64 kbit/s, ITU-T Rec. H.261, version 1: Nov. 1990, version 2: Mar. 1993.
- [3] Generic Coding of Moving Pictures and Associated Audio Information Part 2: Video, ITU-T Rec. H.262 and ISO/IEC 13818-2 (MPEG 2 Video), ITU-T and ISO/IEC JTC 1, Nov. 1994.
- [4] Video Coding for Low Bit Rate Communication, ITU-T Rec. H.263, Nov. 1995 (and subsequent editions).
- [5] Advanced Video Coding for Generic Audio-Visual Services, ITU-T Rec. H.264 and ISO/IEC 14496-10 (AVC), ITU-T and ISO/IEC JTC 1, May. 2003 (and subsequent editions).
- [6] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," IEEE Trans. Circuits Syst. Video Technol., vol. 22, no.12, pp. 1649-1668, 2012.
- [7] J. Chen, Y. Ye, S. H. Kim, "Algorithm description for Versatile Video Coding and Test Model 6 (VTM 6)", JVET-O2002, Jul. 2019.
- [8] L. Zhao, X. Zhao, S. Liu and X. Li, "CE3-related: Unification of angular intra prediction for square and non-square blocks," JVET-L0279, 2018.
- [9] H. Schwarz, T. Nguyen, D. Marpe and T. Wiegand, "CE7: Transform Coefficient Coding and Dependent Quantization (Tests 7.1.2, 7.2.1)", JVET-K0071, 2018.
- [10] W.-J. Chien, J. Boyce, W. Chen, Y.-W. Chen, R. Chernyak, K. Choi, R. Hashimoto, Y.-W. Huang, H. Jang, R.-L. Liao, S. Liu, "JVET AHG report: Tool reporting procedure (AHG13)", JVET-P0013, Oct. 2019.

필자 소개



이종석

- 2016년 2월 : 광운대학교 전자공학과 학사
- 2018년 2월 : 광운대학교 전자공학과 석사
- 2018년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- 주관심분야 : 영상압축, 컴퓨터비전, 고해상도 위성영상 처리



박준택

- 2017년 2월 : 광운대학교 컴퓨터공학과 학사
- 2019년 2월 : 광운대학교 컴퓨터공학과 석사
- 2019년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전



최한솔

- 2018년 2월 : 광운대학교 컴퓨터공학과 학사
- 2018년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 석사과정
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전



변주형

- 2019년 2월 : 광운대학교 컴퓨터공학과 학사
- 2019년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 석사과정
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전



심동규

- 1993년 2월 : 서강대학교 전자공학과 공학사
- 1995년 2월 : 서강대학교 전자공학과 공학석사
- 1999년 2월 : 서강대학교 전자공학과 공학박사
- 1999년 3월 ~ 2000년 8월 : 현대전자 선임연구원
- 2000년 9월 ~ 2002년 3월 : 바로비전 선임연구원
- 2002년 4월 ~ 2005년 2월 : University of Washington Senior research engineer
- 2005년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 교수
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전