2020년
ICT국제표준 마에스트로
주요이슈 분석서

[차세대 비디오 압축기술 표준화]

한국정보통신기술협회

표준 마에스트로 주요이슈 분석서
(차세대 비디오 압축기술 표준화)

1 개요

1.1 Overall 기술 및 표준화 Trend

○ MPEG-2 비디오 압축 표준이 개발되었을 때만 해도 MPEG 표준은 시장에서 독점적인 지위를 가지는 사실상 유일한 표준이었으나 20 년이 지난 오늘 날에는 다양한 표준화 단체에서 다양한 비디오 표준을 개발하고 있는 상황이다.

○ 그림 1 에 나타낸 것처럼 공식 표준화 단체인 ISO 의 MPEG (Moving Picture Expert Group)과 ITU-T 의 VCEG (Video Coding Expert Group) 및 이들의 연합 표준화 기구인 JVET (Joint Video Expert Team) 외에도 구글이 독자적으로 개발한 VP8 이나 VP9, 그리고 구글, 마이크로소프트 등이 주도하여 구성한 AOM (Alliance for Open Media)에서 개발한 AV1 까지 특정 표준이 시장에서 절대적인 우위를 점하지 못한 채 다수의 비디오 압축 표준이 시장에서 서로 경쟁하고 있는 춘추전국시대와 같은 상황이다.

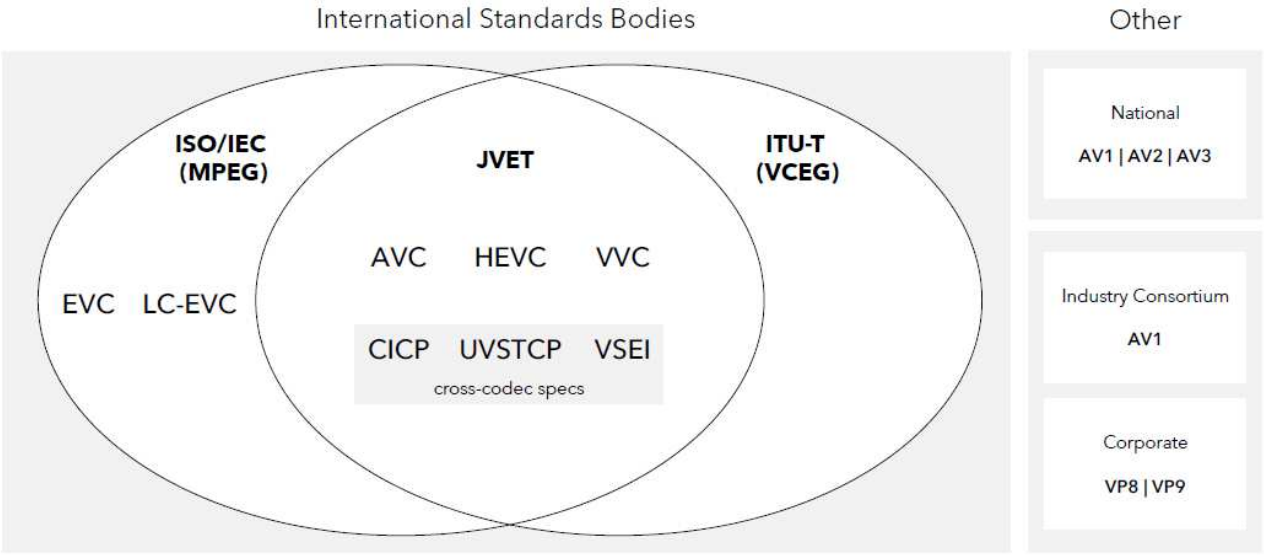


그림 1 비디오 코덱 표준화 현황

○ 시장에서 비디오 압축 표준을 결정하는 두 가지 큰 요소는 로열티 지급 필요 여부와 실시간 부호화 가능 여부이다. 스포츠 중계나 뉴스와 같은 실시간 서비스를 제공해야 하는 응용 분야에서는 실시간 부호화 지원 여부가 중요한 선택 기준이 되고, 응용 분야가 인터넷 브라우저가 주를 이루는 경우는 로열티 지원 여부가 또 다른 중요한 선택 기준이 되기도 한다. 이러한 선택 기준에 따라 응용 분야별로 표준의 채택 현황 및 향후 전망을 분류한 결과를 그림 2 에 나타냈어다. 검정색으로 표시한 코덱은 이미 활용되고 있

는 것들이고, 붉은 색으로 표시한 것들은 향후 활용이 예상되는 것들이다.

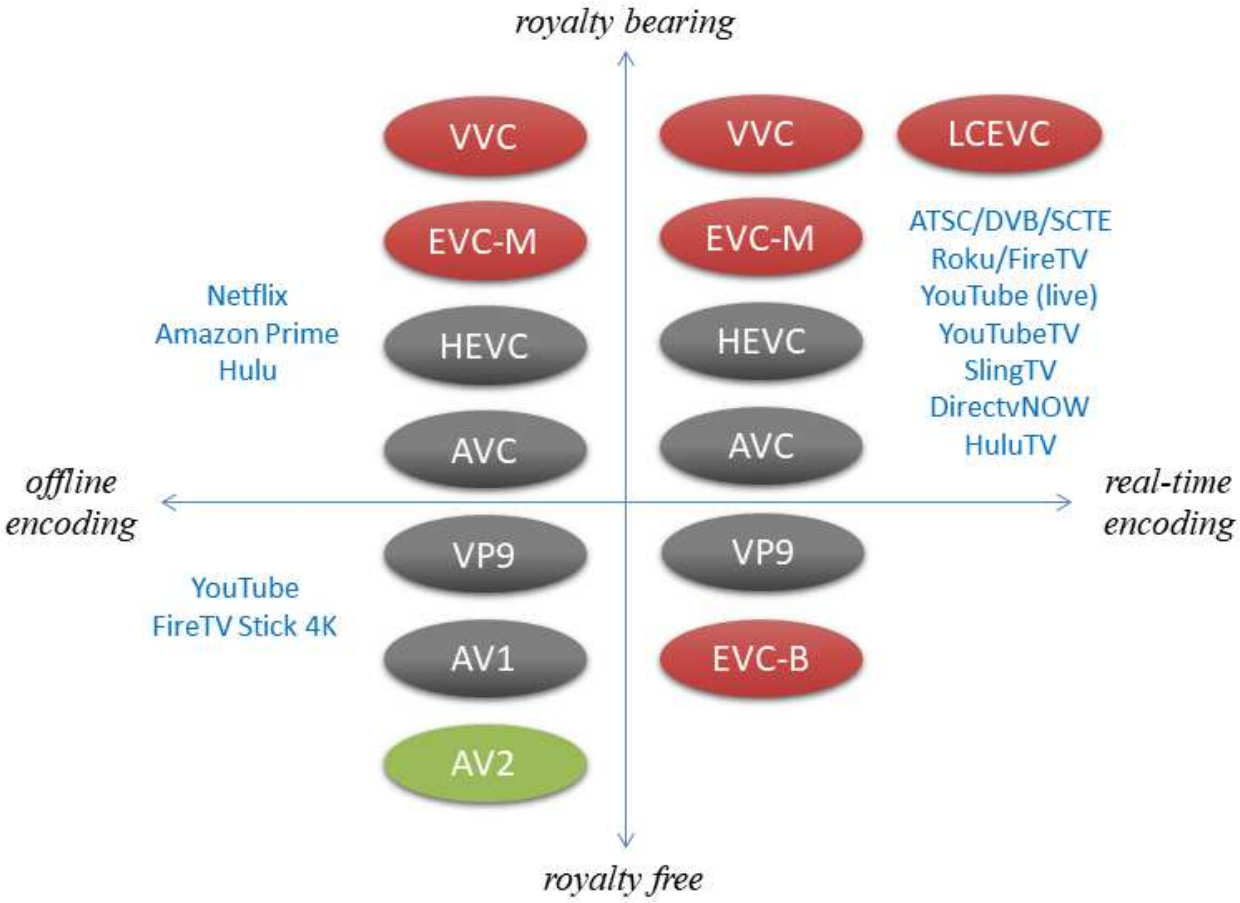


그림 2 비디오 코덱 상용화 동향

- 방송 표준인 ATSC 나 DVB 의 경우는 로열티가 존재하더라도 실시간 부호화 여부가 중요한 요소이므로 로열티 지급 필요 여부와 무관하게 MPEG 과 JVET 에서 개발하는 비디오 압축 표준들이 주로 쓰이고 있으며, 앞으로도 이런 상황이 지속될 것으로 예상된다.
- 실시간 부호화가 필요 없는 드라마 시리즈나 영화와 같은 컨텐츠를 제공하는 SVOD (Subscriptio-based Video on Demand) 서비스의 경우는 유료 서비스일 경우 로열티 지급이 요구되는 MPEG 과 JVET 의 비디오 압축 표준을 주로 활용하고, 무료 서비스일 경우 로열티 지급이 필요 없는 민간 표준화 단체의 비디오 압축 표준을 주로 활용하는 상황이다.
- 실시간 부호화가 필요하나 로열티 지급이 어려운 서비스, 예를 들어 YouTube 를 이용한 실시간 뉴스 공유 서비스 등을 위해서는 표준화된 지 오랜 시간이 지나 압축 성능이 상대적으로 낮지만 로열티 부담이 적은 AVC 가 여전히 활용되고 있는 상황이다.

1.2 JVET 코덱 개발 동향

ㅇ JVET 은 2020 년 7 월 제 19 차 회의(제 131 차 MPEG 회의)에서 VVC 표준의 Final Draft International Standard (FDIS) 를 발간함으로서 2018 년 4 월 Call for Proposal 을 시작한지 약 2 년 여 만에 1 차 버전의 표준화 작업을 마무리하였음. 개발된 FDIS 는 ISO 내 준비 작업과 3 개월 간의 투표 절차와 동시에 ITU 내 투표 절차를 거친 후 2021 년 초에 최종 ITU Recommendation H.266 와 ISO/IEC International Standards 23090-3 으로 동시에 발간될 것으로 예상된다.

ㅇ VVC 는 일반적인 비디오 압축 표준들과 마찬가지로 영상을 다양한 크기의 균일한 특성을 갖는 코딩 블록으로 분할한 후, 프레임 간의 움직임 예측 및 보상을 이용하여 중복성을 제거하고 이 과정에서 남은 잔여 신호 데이터를 변환과 양자화 및 엔트로피 부호화를 이용하여 부호화하는 기본적인 구조를 가지고 있으며, 그림 3 에 나타낸 것처럼 각 구성 요소별로 새로운 기술을 채택하여 압축 효율을 개선하였음.

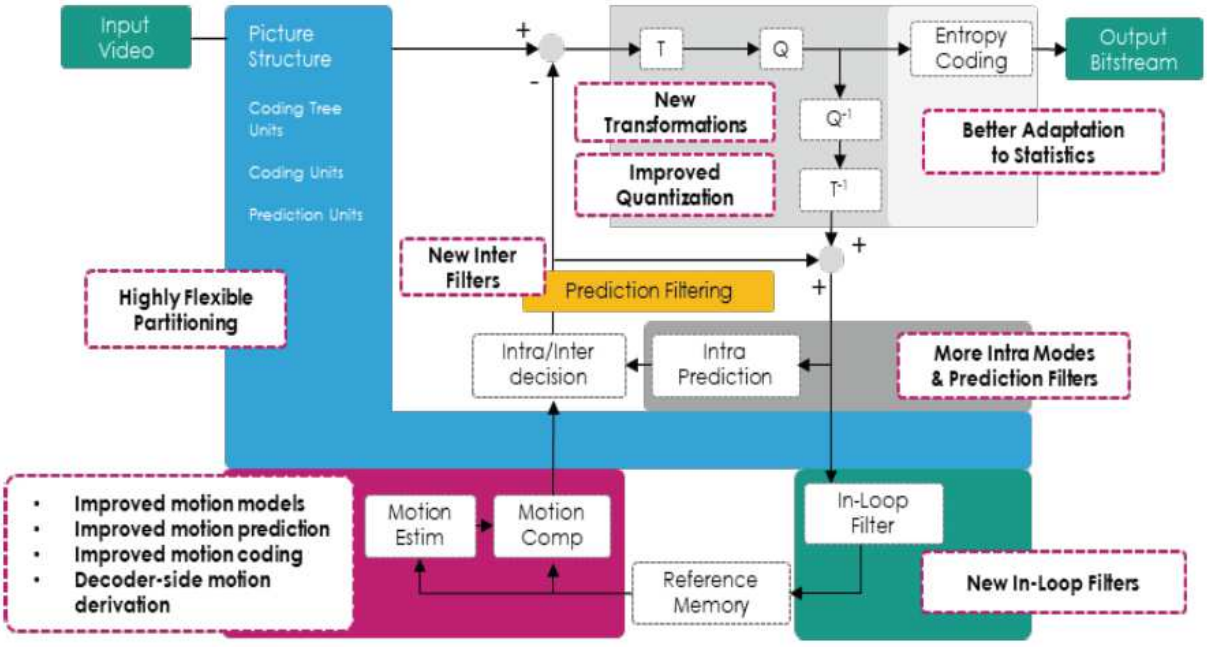


그림 3 VVC 코덱 구조 및 신규 채택 기술

- VVC 는 그림 4 에 나타낸 것처럼 기존의 HEVC 가 64x64 픽셀 크기의 블록에서 시작해서 이를 4 개의 블록으로 반복적으로 분할하여 8x8 픽셀 크기의 블록까지 분할해 나가는 구조와 달리 이보다 큰 128x128 픽셀 크기의 블록에서 시작하여 마찬가지로 4 개의 블록으로 반복적으로 분할해 나가되 4x4 픽셀 크기의 블록까지 분할해 나가는 구조에 경우에 따라서 블록을 세로나 가로의 2 개의 블록으로 분할하거나 3 개의 블록으로 분할하는 구조를 추가로 지원함으로써 좀 더 다양한 형태의 블록으로 나누어 영상을 부호화할 수 있도록 함.

- 이는 8K 와 같이 더 큰 크기의 영상 내에 존재할 수 있는 더 큰 크기의 균일한 특성을 갖는 블록을 단일 블록으로 부호화함으로써 압축 효율을 높이고, 좀 더 다양한 형태의 균일한 특성을 갖는 블록 형태에 대응하여 압축 효율을 높일 수 있는 장점이 있는 반면에 다양한 형태의 블록을 분할하여 이 중 최적의 분할 구조를 찾는 과정에서의 복잡도 증가라는 단점을 가지게 됨.

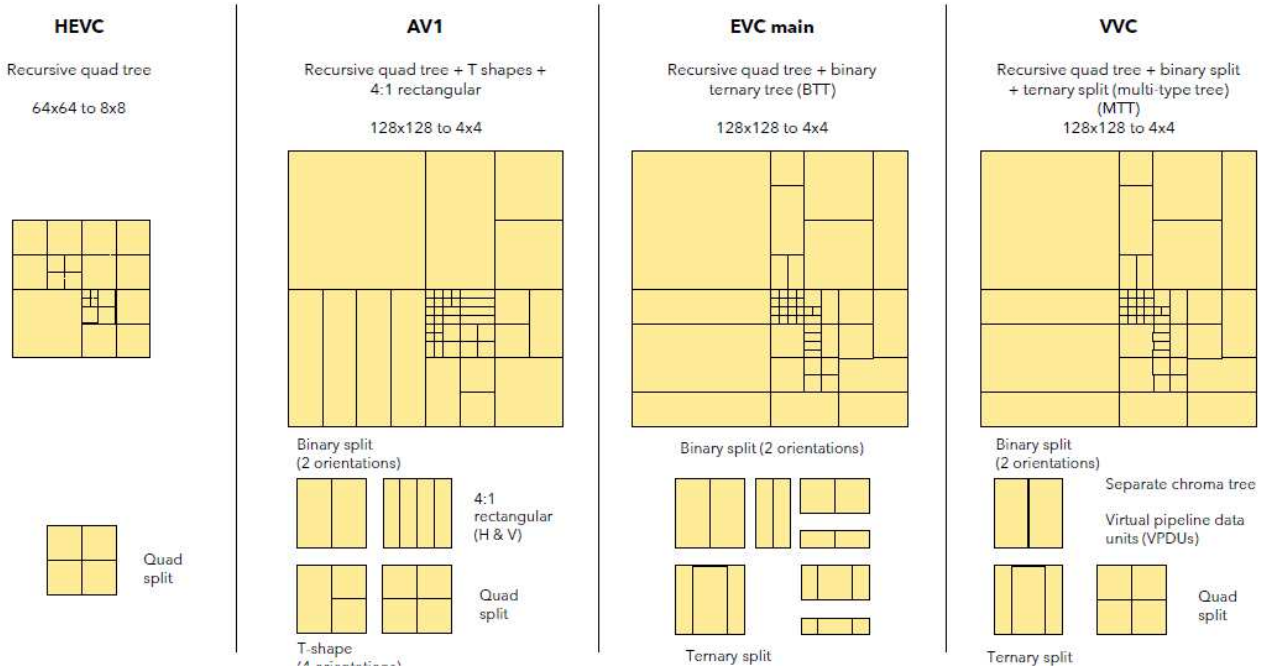
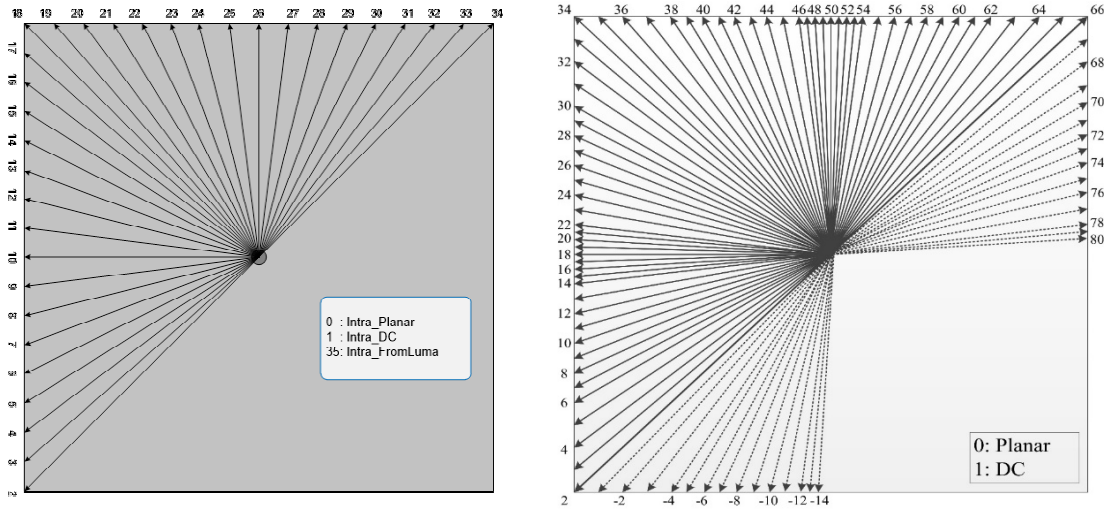


그림 4 코덱별 블록 분할 구조



HEVC V V C

그림 5 화면내 부호화 모드

- VVC 는 HEVC 가 총 36 개 방향의 화면 내 부호화 모드를 지원하였던 것에 비해 2 배 이상 많아진 총 87 개의 부호화 모드를 지원함. 그림 5 에 나타낸 것처럼 이들 중 20 개의 모드는 광각 모드이며, 65 개 모드는 일반 방향성 모드임.
- VVC 는 일반적인 Discreet Cosine Transform 외에 Multiple Transform Selection (MTS)를 화면 내 부호화 블록과 화면간 부호화 블록의 잔여 데이터 부호화에 이용함. 새로운 변환으로는 Low-Frequency Non-Separable Transform (LFNST)와 Matrix-Based Intra Predication (MIP) 및 Sub-block Transform (SBT) 등이 추가되었음.

ㅇ VVC 는 그림 6 에 나타낸 것처럼 HEVC 보다 확장된 다양한 화면간 예측 기술들을 채택하여 부호화 성능향상을 달성하였음.

- HEVC 의 스킵(skip) 및 머지(merge) 모드 기술에서는 Extended merge prediction, MMVD(Merge with Motion Vector Difference), SbTMVP(Subblock-based Temporal Motion Vector Prediction), TPM(Triangle Partitioning Mode)과 CIIP(Combined Inter and Intra Prediction) 모드가 채택되었으며, AMVP(Advanced MVP) 기술에서는 SMVD(Symmetric MVD), AMVR(Adaptive Motion Vector Refinement)와 BCW(Bi-prediction with CU-level Weights) 모드가 추가되었음.
- 디코더 기반 움직임 벡터 보정 기술들인 DMVR(Decoder-side Motion Vector Refinement) 및 BDOF(Bi-Directional Optical Flow) 모드를 추가하여 ₩ 부호화 효율을 개선하였음.

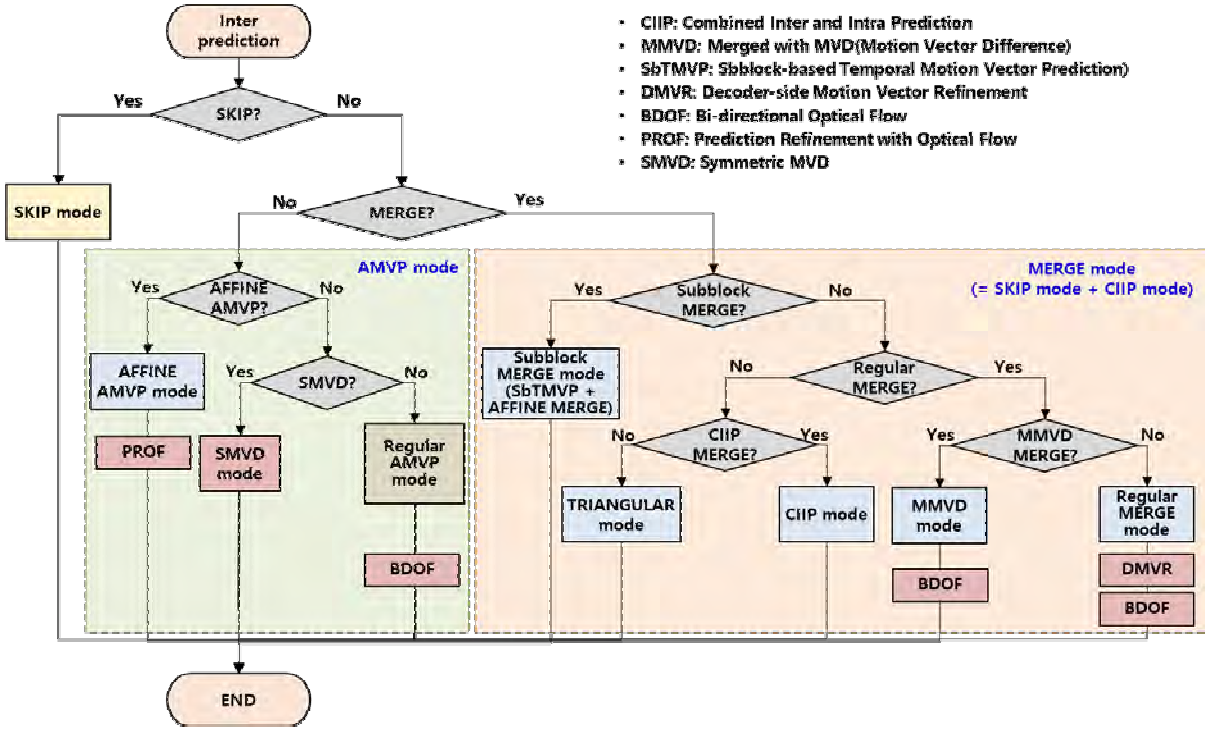


그림 6 VVC 화면간 예측 부호화 모드

ㅇ VVC 는 움직임 예측의 기준 영상으로 사용되는 복원 영상에 대해 루프내 필터를 적용하여 다양한 화질 열화를 제거하여 주관적인 화질을 향상시킬 뿐만 아니라 이후 화면간 예측 부호화시에 발생할 수 있는 오류를 최소화함.

ㅇ VVC 는 HEVC 와 동일한 CABAC 엔진을 이용하여 엔트로피 부호화를 수행하지만, 디코더에 저장되는 확률 테이블의 크기를 줄이고자 테이블 확률 업데이트 방식에서 간단한 수식을 통한 테이블 업데이트 방식을 이용하며, 좀 더 정확한 심볼의 확률 값을 얻기 위하여 서로 다른 2 개의 확률 모델을 사용함.

ㅇ VVC 는 여러 개의 레이어로 비트스트림을 구성할 수 있으며, 아울러 화면을 타일과 슬라이스로 나누어 비트스트림을 구성할 수 있음. 그림 7 는 이러한 타일과 슬라이스 구성의 예를 나타냄. 아울러 VVC 비트스트림은 독립적인 디코딩이 가능한 단위인 서브픽처로 구성할 수도 있음.

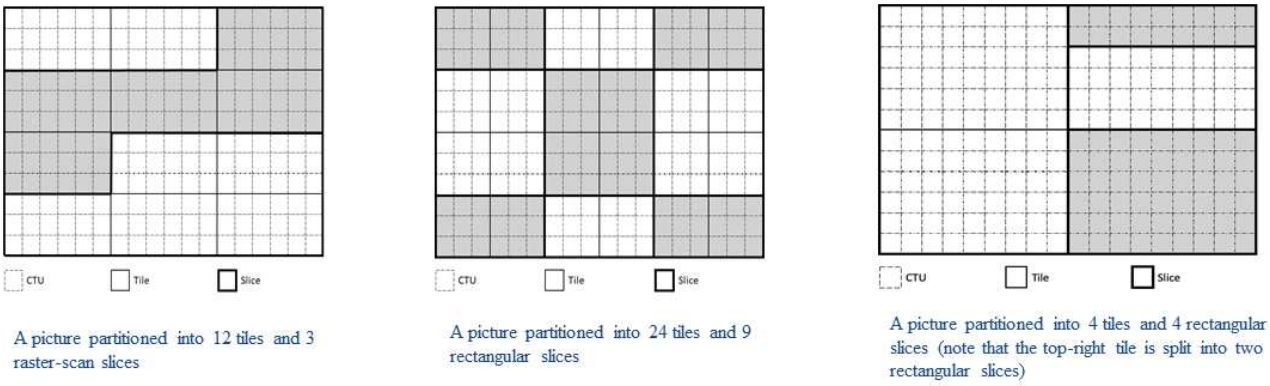


그림 7 VVC 슬라이스와 타일 구성의 예

ㅇ VVC 는 일반적인 자연 영상의 부호화외에 그래픽 데이터나 텍스트 데이터 위주로 구성되는 컴퓨터 화면 영상 부호화 효율을 개선하기 위해 다양한 Screen Content Coding 기술들을 추가하였음.

- VVC 에서는 그래픽 데이터가 유사한 픽셀이 반복되는 경우를 고려하여 다양한 방식의 화면 내 코딩 블럭 복사 (Intra Block Copy) 기술이 추가되었음.
- 팔레트 (Palette) 부호화 기술은 HEVC 에 표준화되었던 기술로 주변 픽셀과 동일하거나 유사한 픽셀이 반복되는 스크린 영상 부호화 효율 개선에 효과적임.

ㅇ JVET 에서 진행 중인 성능 시험 결과에 의하면 주관적 화질 평가 결과를 기준으로 VVC 는 HEVC 대비 최고 50% 이상의 압축 성능 개선을 보이는 것으로 알려지고 있음.

- 그림 8 는 HEVC 와 유사한 MOS 를 나타내는 VVC 의 비트율과 HEVC 비트율과의 차이를 나타내는 표로서 차이가 가장 적은 경우는 16.4%의 성능 개선만을 나타내었고, 차이가 가장 큰 경우는 51%의 성능 개선을 나타내었음. 흥미로운 점은 두가지 경우가 모두 동일한 비디오를 서로 다른 비트율로 압축하였을 때 나타났으며, HEVC 비트스트림과 VVC 비트스트림의 MOS 점수의 차이가 매우 크게 나나탐. 이는 비트율 변화에 따른 MOS 성능 변화가 매우 심하여 정확하고 일관된 MOS 점수를 갖도록 인

코더를 설정하기가 쉽지 않았음을 나타낸다고 볼 수 있음. 실제로 이런 경우는 실시간으로 부호화할 경우 주어진 비트율을 최대로 활용하면서 부호화하기 쉽지 않은 경우로 주어진 비트율보다 더 많은 비트율의 비트스트림을 생성할 수 없으므로 주어진 비트율보다 적은 비트율의 비트스트림을 생성하게 되어 실제 환경에서는 압축 성능 개선 효과가 기대에 미치지 못하게될 것으로 예상됨. MOS 점수의 차이가 신뢰 구간(CI : Confidence Interval) 내에 들어오는 경우들로 구성된 비디오의 경우 압축 성능 개선 비율이 35.5%에서 47.7%로 대략 이 정도 범위의 성능 개선을 보인다고 볼 수 있음.

VTM / HM		Rate Diff.	ΔMOS
DrivingPOV3	R1	**-51,0%**	0,5
DrivingPOV3	R2	-49,3%	1,0
DrivingPOV3	R3	-48,0%	0,7
DrivingPOV3	R4	-30,1%	1,7
DrivingPOV3	R5	-16,4%	0,9
Marathon2	R1	-47,7%	< CI
Marathon2	R2	-45,0%	< CI
Marathon2	R3	-43,6%	< CI
Marathon2	R4	-35,5%	< CI
Marathon2	R5	-36,7%	< CI
MountainBay2	R1	-48,8%	< CI
MountainBay2	R2	-47,4%	-0,6
MountainBay2	R3	-47,1%	< CI
MountainBay2	R4	-46,3%	-0,8
MountainBay2	R5	-45,7%	-0,5
NeptuneFountain3	R1	-39,7%	< CI
NeptuneFountain3	R2	-27,4%	0,5
NeptuneFountain3	R3	-27,1%	0,7
NeptuneFountain3	R4	-38,5%	< CI
NeptuneFountain3	R5	-39,6%	< CI
TallBuildings2	R1	-46,0%	-0,9
TallBuildings2	R2	-44,7%	< CI
TallBuildings2	R3	-43,3%	< CI
TallBuildings2	R4	-41,5%	< CI
TallBuildings2	R5	-32,6%	< CI

그림 8 VVC 압축 성능 시험 결과 표

- 그림 8 에서 HEVC 와 VVC 간의 MOS 차이가 큰 비디오의 경우를 그래프로 나타낸 그림 9 를 보면 파란색으로 나타낸 HEVC 의 성능은 낮은 비트율부터 높은 비트율까지 기울기가 거의 일정한 직선으로 증가함을 보여주는 반면에 오렌지색으로 나타낸 VVC 의 성능은 4,000kbsp 구간까지는 화질이 어느 정도 일정한 직선으로 증가하는 모습을 나타내는 반면그 이후 구간에서는 직선의 기울기가 낮아져서 거의 증가하지 않아 4,000kbps 의 MOS 점수와 이 보다 2 배의 비트을 할당하여 부호화한 8,000kbps 의 MOS 점수가 1 점도 차이 나지 않는 결과를 나타내고 있음. 이는 4,000kbps 에서 성능이 최고치에 이르러러 그보다 많은 비트를 할당하여도 화질이

거의 증가하지 않을 정도로 포화되었다는 것을 나타내는 것으로 이해할 수 있음. 이는 이러한 비디오의 경우 낮은 비트율로 압축할 경우 비트율 추가에 따른 성능 개선 효과가 두드러지게 나타타나 고 비트율에서는 비트율 추가 할당에 따른 화질 개선 효과가 거의 눈에 보이지 않게 되는 것을 의미함.

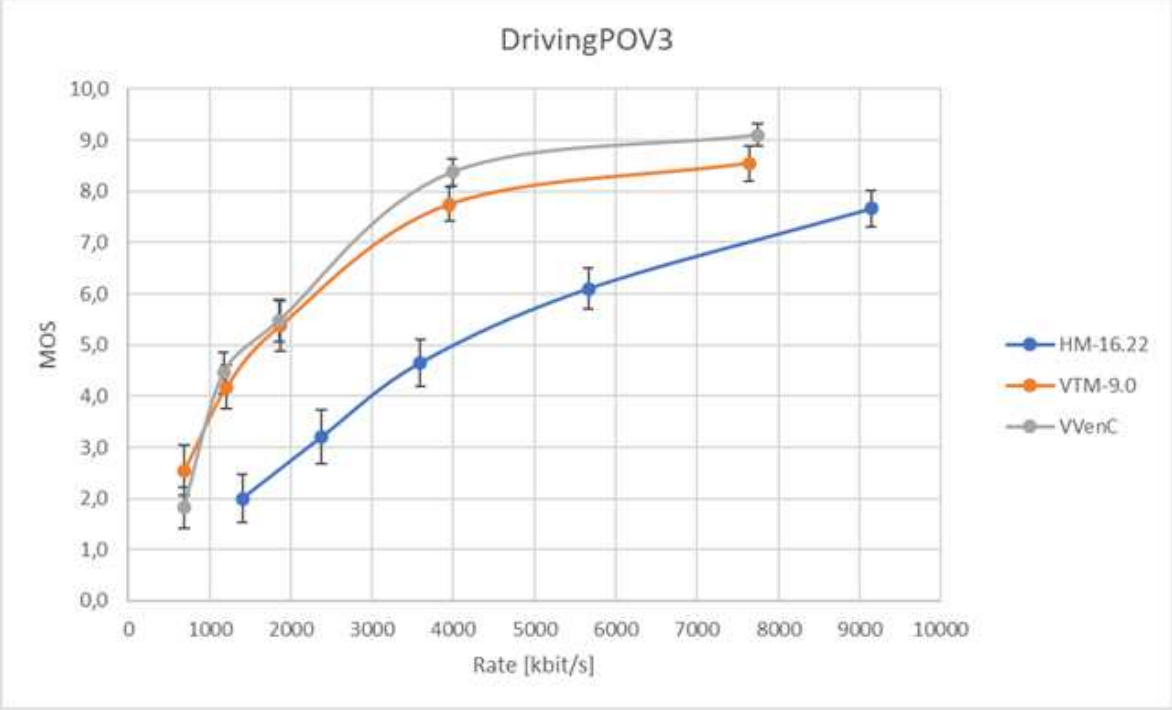


그림 9 VVC 압축 성능 시험 결과 그래프

1.3 MPEG 코덱 개발 동향

ㅇ VVC 를 개발한 JVET 은 ISO/IEC JTC 1 산하의 MPEG 과 ITU-T SG 16 으로 구성된 비디오 코덱 표준화 합동 팀인데, 이 중에서 ITU-T 산하의 Video Coding Expert Group (VCEG)은 VVC 외의 별도의 표준화 작업을 하고 있지 않은 반면, MPEG 은 비디오 그룹에서 별도로 코덱 표준화 작업을 진행중임.

ㅇ MPEG 에서 표준화하고 있는 비디오 코덱 중 하나인 ISO/IEC 23094-1 (MPEG-5) Essential Video Coding (EVC) 는 HEVC 상용화의 가장 큰 걸림돌로 지적된 복잡하고 불확실한 라이센싱 조건 문제를 해결하고자 표준화한 코덱으로 2019 년 1 월 기술 제안 요청서의 결과를 바탕으로 표준화를 시작하여 2020 년 4 월 표준화가 마무리되었음.

- EVC 는 화웨이, 퀄컴, 삼성 등 3 개사가 기술을 제안하여 베이스라인과 메인, 두 가지의 프로파일을 제정하였는데, 베이스라인의 경우 기술이 공개되었거나 특허가 출원된지 20 년이 지난 기술들로 구성되어 로열티 이슈가 전혀 없으며, 메인 프로파일의 경우 최초로 기술을 제안한 3 개사의 기술들로만 구성되어 3 개사만 합의하면 라이센싱이 가능하여 라이센싱 조건이 단순하고 명확하여 HEVC 가 가지고 있는 문제를 해결할 수 있을 것으로 예상됨. 아울러 베이스라인 프로파일의 경우 모든 툴을 선택

적으로 on/off 하여 사용할 수 있는 구조를 적용하고 있어 만약 특정 툴에 대해 라이센싱 이슈가 발생할 경우 해당 툴을 사용하지 않도록 하여 추후에라도 라이센싱 이슈에 대응할 수 있도록 대비하고 있음

- EVC 는 그림 10 에 나타낸 것처럼 통상적인 비디오 압축 표준처럼 움직임예측/보상, 잔여 신호에 대한 변환, 양자화 및 엔트로피 코딩 등으로 구성되며 매우 단순한 기술들로 구성된 베이스라인 프로파일과 달리 메인 프로파일의 경우 다양한 기술들을 제공하고 있음.

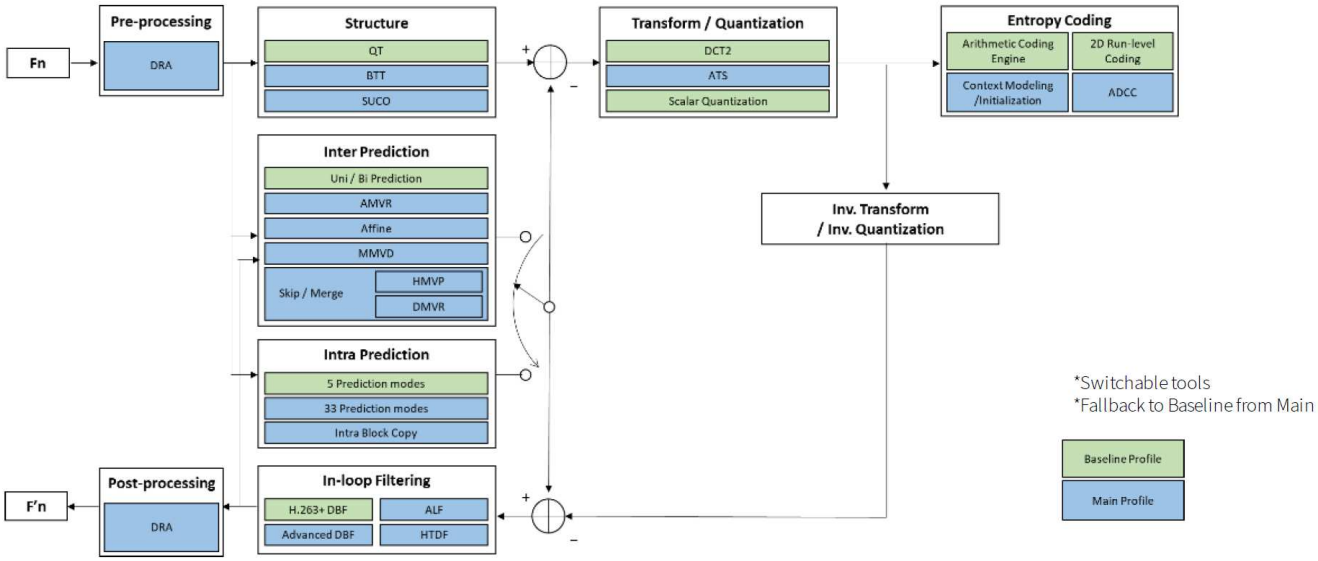


그림 10 EVC 구조 및 프로파일별 기술 구성

- 베이스라인 프로파일의 경우 최대 64x64 픽셀 크기의 블록을 4 개의 서브블록으로 지속적으로 분할해가는 쿼드트리 구조를 이용하며, 화면내 부호화시 픽셀간 예측은 DC, 가로, 세로, 두 방향의 대각선으로 예측하는 5 방향 예측을 이용함.
- 잔여 신호 부호화를 위해서는 2x2 크기부터 64x64 크기까지의 다양한 블록 크기의 DCT 를 이용하며, 지그재그 스캔과 이진 산술 부호화를 이용하여 변환 계수를 부호화함.
- 움직임 예측을 위해서는 여러 개의 기준 프레임을 이용할 수 있으며, 이웃하는블럭간의 움직인 벡터 예측 기술을 이용함.
- 메인프로파일의 경우 최대 128x128 크기의 블록을 이용하여, 쿼드 트리 외에 이진 및 삼진 트리 분할 (BTT : Binary and Triple Tree) 방법을 이용하여 블록을 분할할 수 있으며, 분할된 서브블럭들의 코딩 순서를 다양하게 변경하여 블록 간 예측을 수행할 수 있음,
- 화면 내 부호화시 픽셀 간 예측은 DC 와 이진 예측 외에 30 개의 각도에 의한 예측을 지원함.
- 움직임 벡터부호화는 다양한 해상도로 이웃하는 블록 간의 움직임 벡터를 예측 부호

화하는 AMVR (AdaptiveMotion Vector Resolution) 기술과 MMVD (Merge with Motion Vector Difference) 기술을 이용할 수 있으며, DMVR (Decoder-side MV Refinement)를 이용함.

- 잔여 신호 변환은 블럭별로 코딩 블록의 크기보다 작은 크기의 변환 블록을 이용할 수 있으며, 변환 계수는 역방향 지그재그 스캔을 이용할 수 있음.
- HDTF 를 이용한 in-loop 필터를 적용하며, DRA(Dynamic Range Adjustment) 기능을 이용함.

ㅇ EVC 메인 프로파일은 주관적 화질 평가 결과 그림 11 에 나타낸 것처럼 HEVC 대비 약 50% 압축 성능 개선 효과를 나타내고 있으며, 베이스라인 프로파일은 주관적 화질 평가 결과 그림 12 에 나타낸 것처럼 AVC 대비 약 50% 압축 성능 개선효과를 나타내고 있음.

- 그림 13 에 나타낸 AVC, HEVC, EVC 베이스라인 프로파일, EVC 메인 프로파일 간의 비교 결과 그래프를 보면, 낮은 비트율에서는 EVC 베이스라인 프로파일의 성능이 HEVC 에 비해 20% 정도 낮은 것으로 나타나고 있으나, 높은 비트율에서는 EVC 베이스라인 프로파일의 성능이 HEVC 의 성능을 앞지르는 경우도 있는데, 이는 다른 비디오에 대해서도 일관적으로 나타나는 현상으로 EVC 가 HEVC 에 비해 고화질 비디오 서비스에는 더 적합한 것임을 알 수 있음.

SDR, Main, RA, 4K		HM16.20			ETM6.0			BD-rate (piecewise cubic)
		QPISlice	bits	MOS	QPISlice	bits	MOS	bits reduction
Class A	CatRobot	22	23098443	8.66	23	11693305	9.16	-46.2%
		27	6932117	8.44	29	4398668	8.47	
		32	3505155	7.56	34	2203761	7.91	
		37	1899506	5.75	39	1160625	5.97	
	BarScene	27	4732382	8.38	29	2603829	8.47	-47.9%
		32	1511056	6.88	35	896521	7.50	
		38	719676	4.38	41	434653	4.63	
		44	378746	1.44	47	244750	2.03	
	Dancers	22	26218368	8.41	24	11753572	8.72	-60.2%
		27	3085503	7.03	29	1280342	7.66	
		32	1036903	5.91	34	686214	6.56	
		37	560204	3.06	39	371672	3.47	
	DrivingPov	24	39860381	9.09	26	23448296	8.59	-45.5%
		28	17680603	8.28	30	10855738	8.47	
		32	9049443	7.47	34	5707010	7.66	
		38	3732064	4.19	40	2302969	4.97	
Average								-50.0%

그림 11 EVC 메인프로파일 압축 성능 시험 결과

SDR, Baseline, RA, 4K		JM19.0			ETM6.0			BD-rate (piecewise cubic)
		QPISlice	bits	MOS	QPISlice	bits	MOS	bits reduction
Class A	CatRobot	23	18249728	8.22	25	12647710	8.81	-44.0%
		27	9427642	7.81	29	6650028	8.25	
		32	5238860	5.81	35	3126513	5.94	
		39	2413263	2.38	43	1359661	2.66	
	BarScene	25	10638455	8.19	27	5376868	8.34	-47.9%
		27	5769534	7.66	29	3181024	7.69	
		32	2565515	5.72	34	1480618	6.06	
		39	1256031	1.91	41	687569	2.22	
	Dancers	24	10627865	7.34	26	5251676	8.28	-52.7%
		26	5292067	6.75	28	2775577	7.19	
		30	2485826	4.97	32	1358234	5.06	
		36	1301353	1.78	38	695513	3.44	
	DrivingPov	22	62531140	8.66	24	44390447	8.72	-55.9%
		26	32484781	8.09	28	20275946	8.47	
		31	15621178	5.75	34	8194473	6.28	
		37	7318298	3.09	40	3734092	4.00	
Average								**-50.1%**

그림 12 EVC 베이스라인프로파일 압축 성능 시험 결과

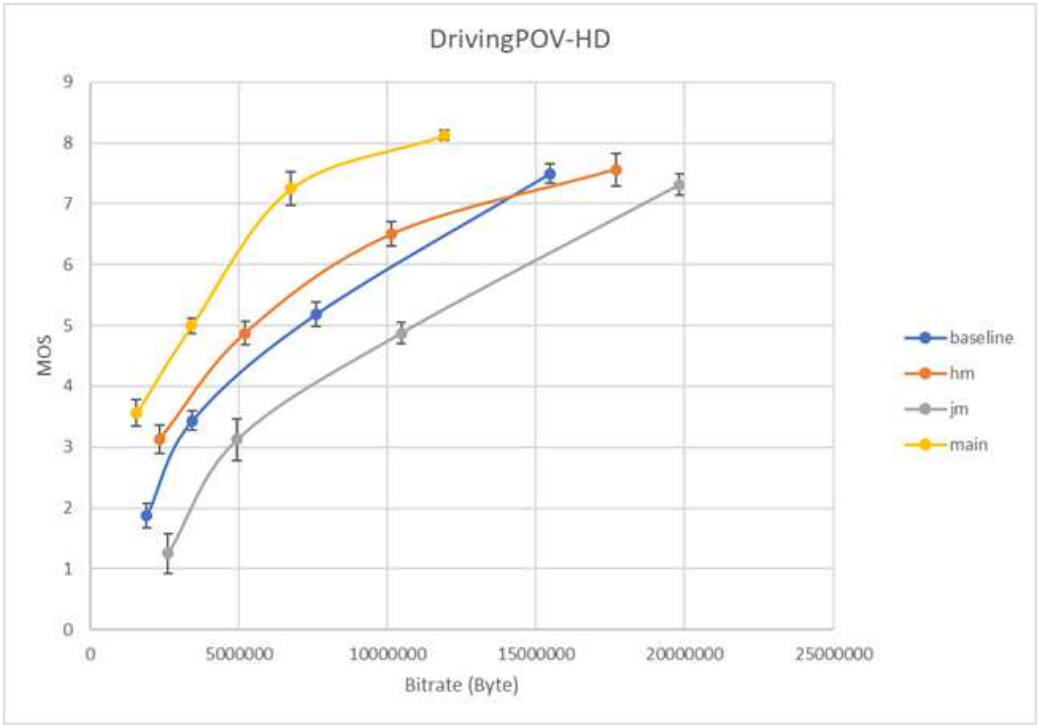


그림 13 EVC 성능 시험 결과 그래프

ㅇ MPEG 비디오 그룹에서 개발한 또 하나의 비디오 압축 기술인 LC-EVC 는 2018 년 10 월에 기술 제안 요청서가 발송되어 2019 년 4 월 기술 제안을 받아 표준 개발 작업을 시작하였으며, 2020 년 10 월 FDIS 를 발간함으로 표준화가 완료되었음.

- LCEVC 는 인코딩을 위해서 입력 영상을 1/2 해상도로 다운 샘플링한 후 이를 기존의 코덱으로 부호화하고, 이렇게 부호화된 결과 영상과 다운 샘플링된 원본 영상과의 차이를 1/2 해상도 서브스트림으로 부호화하고, 이 결과를 다시 원래 해상도의 원본 영상과 비교하여 전체 해상도 서브스트림으로 부호화하는 구조를 가짐 (그림 14 참조)

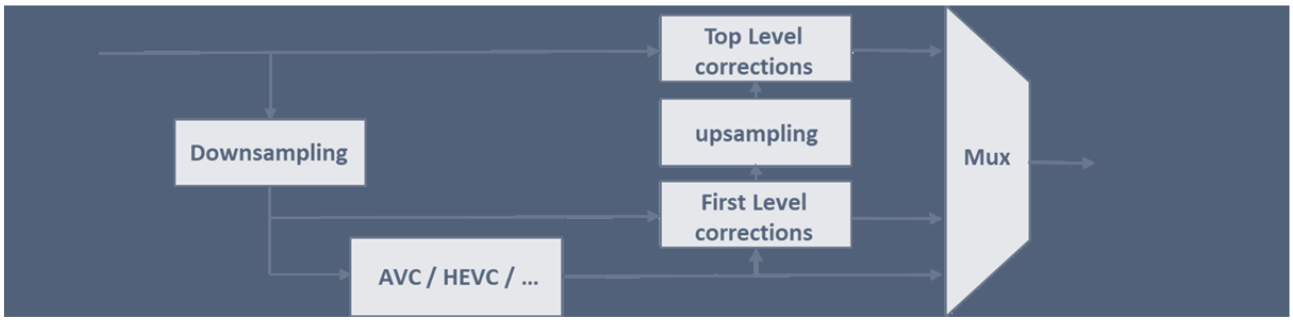


그림 14 LC-EVC 인코더 구조

- LC-EVC 비트스트림은 그림 15 에 나타낸 것처럼 원영상의 1/2 해상도의 기본 계층 비디오 비트스트림과, 1/2 해상도 향상 계층 비트스트림, 그리고 원본 해상도 향상 계층 비트스트림으로 구성되는데 통상적으로 기본 계층 비트스트림이 약 85% 이상을 차지하며, 기본 계층의 화질이 전체적인 영상의 화질에 결정적인 영향을 미치는 것으로 분석되고 있음.



그림 15 LC-EVC 비트스트림 구조

- LC-EVC 디코딩 과정은 앞서 설명한 기본 계층 코덱 비트스트림과 1/2 해상도 서브비스트스트림 및 전체 해상도 서브비스트림을 단계적으로 디코딩하고 블렌딩하는 과정으로 이뤄짐. 그림 16 에 나타낸 것처럼 우선 기본 계층 코덱 비트스트림을 디코딩하여 preliminary intermediate picture 를 생성한 후 이를 1/2 해상도 향상 계층 비트스트림을 디코딩한 결과와 합성하여 preliminary output picture 를 생성함. 이에 다시 전체 해상도 향상 계층 비트스트림을 복호화한 결과를 함성하여 combined output picture 를 생성함.

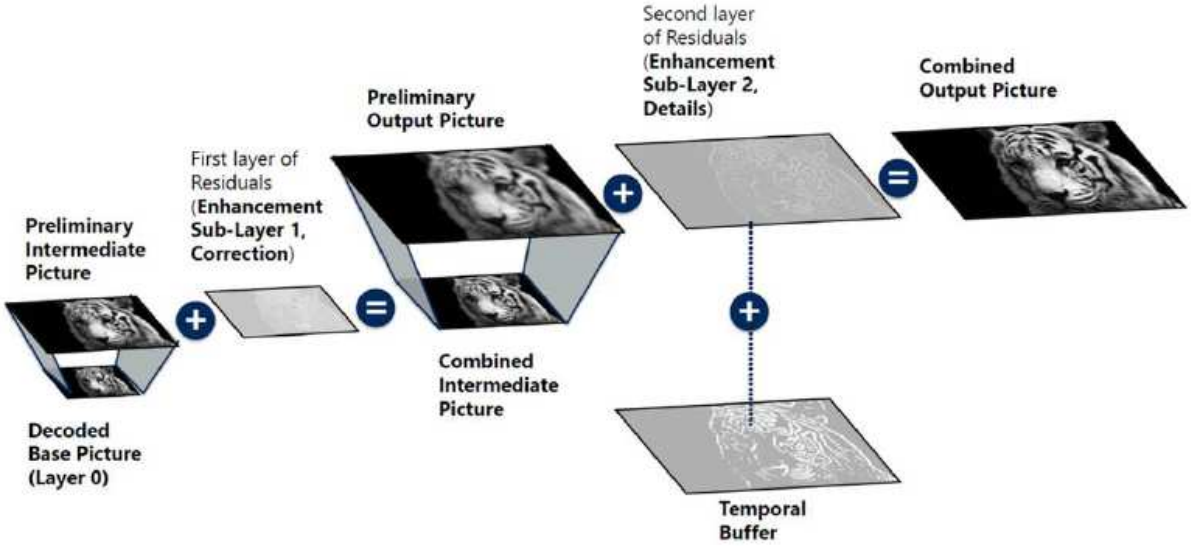


그림 16 LC-EVC 디코더 구조

- LC-EVC 의 성능은 기본 코덱의 영상을 특정 비트율로 부호화한 후 이를 복호화하여 업샘플링한 결과와 기본 코덱의 기본 계층과 LC-EVC 향상 계층의 조합으로 복호화한 결과를 비교하여 평가하는데, 현재까지 개발된 결과에 의하면 AVC 나 HEVC 와 비교하여 약 5% 정도 나은 성능을 보이는 것으로 알려지고 있음. (그림 17 참조)

AVC Base

Experiment 2 (JM)		
UHD LTM 4.1 (JM)	**vs. JM 1080p upsampled**	
	BD-rate-PSNR	BD-rate-MOS
Campfire_3840x2160_30	-4.28%	0.24%
CatRobot1_3840x2160_60	-4.07%	-17.43%
FoodMarket4_3840x2160_60_600	-2.66%	-4.79%
Fortnite_Part1_3840x2160p60	-12.35%	-8.70%
ParkRunning3_3840x2160_50	-3.64%	-10.68%
Average	**-5.40%**	**-8.27%**

HEVC Base

Experiment 2 (HM)		
UHD LTM 4.1 (HM)	**vs. HM 1080p upsampled**	
	BD-rate-PSNR	BD-rate-MOS
Campfire_3840x2160_30	-4.12%	-0.63%
CatRobot1_3840x2160_60	-4.57%	-5.25%
FoodMarket4_3840x2160_60_600	-3.57%	-0.09%
Fortnite_Part1_3840x2160p60	-14.34%	-7.68%
ParkRunning3_3840x2160_50	-2.13%	0.68%
Average	**-5.75%**	**-2.59%**

(negative value means better compression)

그림 17 LC-EVC 성능 비교

1.4 AOM (Alliance for Open Media) 코덱 개발 동향

ㅇ 2015 년 설립된 민간 표준화 단체인 AOM 은 웹 환경에서의 응용을 목표로 로열티가 없는 비디오 코덱 표준화를 목적으로 하고 있으며, 리눅스 파운데이션 산하의 프로젝트 중 하나로 운영되고 있음. AOM 은 13 개의 파운딩 멤버와 31 개의 프로모터 멤버로 구성되어 있음 (그림 18 참조)

ㅇ AOM 에서 2018 년 중순 개발 완료한 AV1 비디오 코덱은 구글이 개발 중이던 VP10, 모질라가 개발 중이던 Daala, 그리고 시스코가 개발 중이던 Thor 라는 로열티 프리 코덱 프로젝트의 기술들을 통합하여 개발되었으며, 그림 19 에 나타낸 것처럼 기본적으로는 통상적인 코덱들과 동일한 구조를 가지고 있음.

- AV1 이 가지는 특별한 기술 중 하나인 수퍼레졸루션은 VP9 에 포함되어 있던 기술로 그림 20 에 나타낸 것처럼 부호화시 입력 영상의 해상도를 1/2 로 줄여서 부호화하고 복호화 과정에서 인루프 필터를 이용하여 원래해상도로 업스케일링하는 방식으로 특정 프레임의 비트율을 급격하게 줄여야 하는 경우나 초기 지연 시간을 줄여야 하는 경우에 유용하게 활용할 수 있는 기술임.



그림 18 AOM 회원사

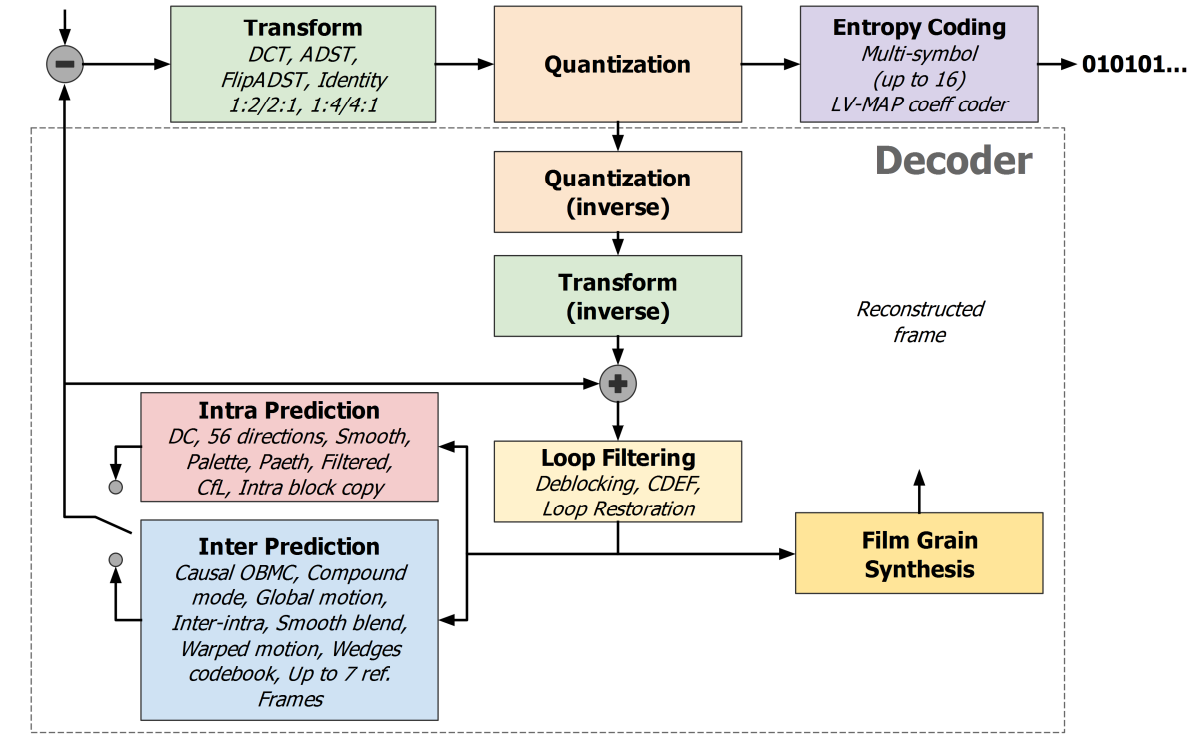


그림 19 AV1 코덱의 구조

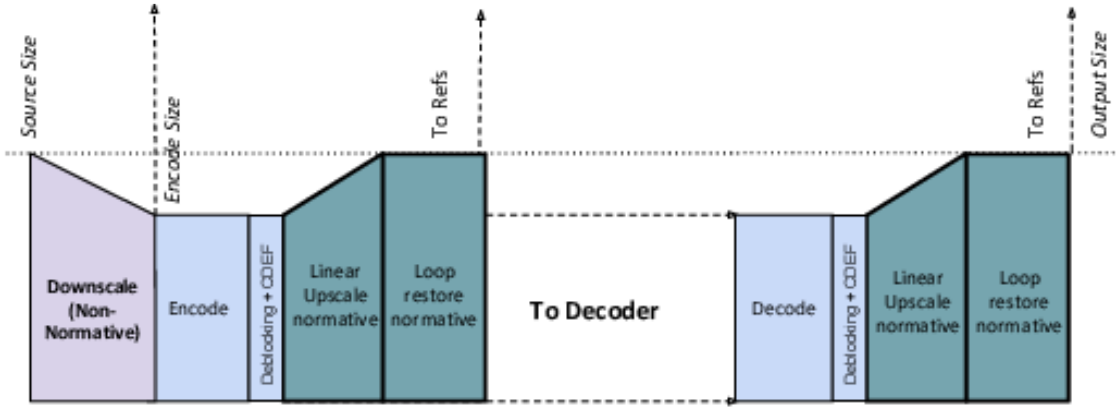


그림 20 AV1 수퍼 레졸루션 코딩 구조

- AV1 이 제공하는 특별한 기능 중 하나인 필름그레인 모드는 컨텐츠 제작자의 의도에 따라 비디오 자체에 노이즈가 있는 것처럼 제작된 영상을 부호화하기 위한 기법으로 노이즈를 모델링한 후 부호화시 노이즈를 제거한 후 부호화하고 복호화시 이를 다시 추가하는 기법으로 그림 21 에 나타낸 것처럼 매우 높은 비트율로도 부호화가 불가능한 결과를 낮은 비트율에서 보여주고 있음.



(c) Coded at 13 112 kbps (d) With noise synthesis at 5 799 kbps

그림 21 AV1 필림그레인 코딩 결과

- AV1 의 성능은 그림 22 에 나타낸 것처럼 HEVC 와 비교하여 UHD 컨텐츠의 경우 평균 7.3%, HD 의 경우 3.8% 더 높은 압축 효율을 나타내는 것을 보여주고 있음. 특히 컨텐츠에 따라서는 HEVC 보다 압축율이 낮은 경우도 보여주고 있음.

Resolution Group	A (UHD)				B (HD)			
Codec	PSNR		VMAF		PSNR		VMAF	
Sequence\BD-rate	AV1	VTM	AV1	VTM	AV1	VTM	AV1	VTM
AirAcrobatic	-12.1%	-25.5%	-12.0%	-28.6%	-2.6%	-21.7%	4.2%	-20.0%
CatRobot	-6.2%	-38.0%	-12.8%	-39.6%	-4.0%	-37.7%	-10.4%	-41.2%
Myanmar	4.3%	-17.2%	1.3%	-21.3%	6.5%	-15.5%	3.5%	-18.6%
CalmingWater	-15.5%	-21.5%	-9.6%	-18.9%	-15.7%	-22.6%	-10.2%	-19.6%
ToddlerFountain	-6.6%	-18.7%	-2.0%	-17.4%	-8.1%	-18.2%	-3.7%	-16.4%
LampLeaves	-6.8%	-26.2%	-6.2%	-26.1%	-2.8%	-23.7%	-0.4%	-24.8%
DaylightRoad	-3.8%	-38.0%	-12.4%	-40.3%	-0.9%	-37.6%	-10.4%	-42.4%
RedRock	-3.5%	-32.5%	-9.0%	-37.9%	0.8%	-31.5%	-6.4%	-37.7%
RollerCoaster	-15.3%	-39.9%	-14.5%	-41.7%	-7.9%	-38.9%	-11.5%	-39.8%
Average	-7.3%	-28.5%	-8.6%	-30.2%	-3.8%	-27.5%	-5.0%	-28.9%

그림 22 AV1 성능 비교

- AOM 은 현재 AV1 의 뒤를 이을 새로운 비디오 코덱인 AV2 표준화에 착수한 것으로 알려지고 있는데, AV2 는 2023 년 완료를 목표로 하고 있으며, AV1 대비 25 ~ 30% 정도의 성능 개선을 목표로 하고 있는 것으로 알려지고 있음.

1.5 MP-AI (Moving Picture, Audio and Data Coding by Artificial Intelligence) 코덱 개발 동향

ㅇ MP-AI 는 MPEG 의 의장이었던 Leonardo Chariglione 가 MPEG 의 구조조정 이후에 MPEG 의장직을 사임하고 2020 년 9 월 새롭게 설립한 민간 표준화단체로 AI 기술을 이용한 비디오, 오디오 등의 멀티미디어 및 일반 데이터 압축 표준 개발을 그 목적으로 하고 있음.

ㅇ MP-AI 는 기존 표준화 단체와 달리 표준 개발에 착수하기 전에 표준의 라이센싱 구조를 정하는 Framework License 을 결정하고 이에 동의하는 기관만 해당 표준화에 참여하도록 하는 특징을 가지고 있음. Framework License 는 일반적인 표준의 라이센싱 조건에서 구체적인 라이센싱 금액을 제외한 것으로 표준 완료 후 빠른 기간 내에 구체적인 라이센싱 조건이 결정되도록 하는데 기여할 수 있을 것으로 예상됨.

ㅇ MP-AI 는 현재 분야별로 표준화 아이템을 선정하고 각 아이템별로 표준화 목표와 요구사항등을 구체화 하는 과정에 있음. 비디오 압축 표준으로는 MPEG 에서 개발한 EVC 를 기반으로 AI 기술을 도입하여 EVC 의 성능을 뛰어 넘는 새로운 표준을 개발하고자 하는 MPAI-EVC (MPAI AI-Enhanced Video Coding) 를 선정하였음.

- MPAI-EVC 는 그림 23 에 나타낸 것처럼 10 가지의 비디오 부호화 요소 기술 각각에 AI 기술을 적용하여 기존 EVC 대비 적어도 25% 더 높은 압축 성능을 얻는 것을 목표로 하고 있음.

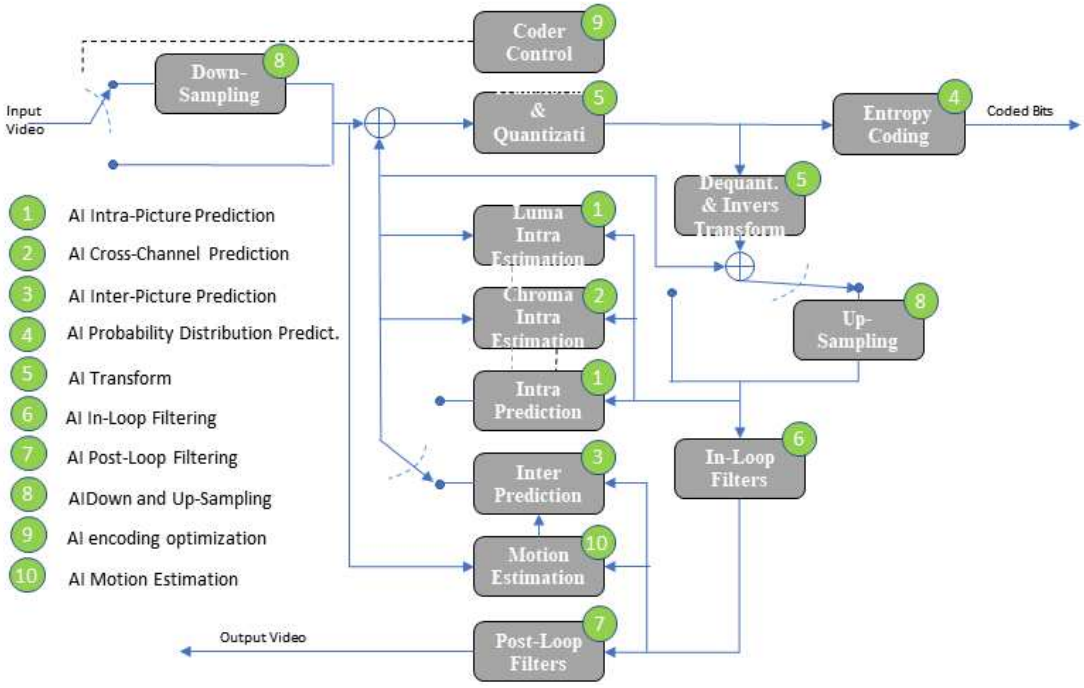


그림 23 MPAI-EVC 개념도

2 비디오 압축 국제 표준화 영향력 확대 방향 및 전략

2.1 국제표준화에서 한국의 취약점

ㅇ 비디오 압축 국제 표준화와 관련하여 한국 전문가들은 JVET 과 MPEG 비디오 그룹 내에서 질적, 양적으로 큰 성장을 이루고 있다. 표준화에 참여하는 전문가의 수도 지속적으로 증가하고 있을 뿐만 아니라, 제안되는 기고서의 숫자와 표준에 채택되는 기고서의 숫자도 지속적으로 증가하고 있다. 그러나, 여전히 표준 관련 리더쉽 면에서는 타국 전문가들에 비해 열세를 보이고 있다. 아래 표는 VVC 의 표준화가 한창 진행 중이던 시점에서 AHG 목록을 보여주고 있는데 한국 전문가는 극히 일부분인 5 명에 그치고 있어 상당히 열세에 있음을 보여주고 있다.

Title	Chairs
Project management (AHG1)	J.-R. Ohm, G. J. Sullivan (co-chairs)
Draft text and test model algorithm description editing (AHG2)	B. Bross, J. Chen (co-chairs), J. Boyce, **S. Kim**, S. Liu, Y.-K. Wang, Y. Ye (vice-chairs)
Test model software development (AHG3)	F. Bossen, X. Li, K. Sühring (co-chairs)
Test material and visual assessment (AHG4)	T. Suzuki, M. Wien (co-chairs), V. Baroncini, R. Chernyak, A. Norkin, J. Ye (vice-chairs)
Memory bandwidth consumption of coding tools (AHG5)	R. Hashimoto (chair), T. Ikai, X. Li, D. Luo, H. Yang, M. Zhou (vice-chairs)
360° video coding tools, software and test conditions (AHG6)	J. Boyce and Y. He (co-chairs), **K. Choi**, J.-L. Lin, Y. Ye (vice-chairs)
Coding of HDR/WCG material (AHG7)	A. Segall (chair), E. François, W. Husak, S. Iwamura, D. Rusanovskyy (vice-chairs)
Layered coding and resolution adaptivity (AHG8)	S. Wenger and A. Segall (co-chairs), M. M. Hannuksela, Hendry, S. McCarthy, Y.-C. Sun, P. Topiwala, M. Zhou (vice-chairs)
Neural networks in video coding (AHG9)	S. Liu and Y. M. Li (co-chairs), B. Choi, K. Kawamura, Y. Li, L. Wang, P. Wu, H. Yang (vice-chairs)
Encoding algorithm optimization (AHG10)	A. Duenas, A. Tourapis (co-chairs), S. Ikonin, A. Norkin, R. Sjöberg, J. Le Tanou, J.-M. Thiesse (vice-chairs)
Screen content coding (AHG11)	S. Liu (chair), J. Boyce, A. Filippov, Y.-C. Sun, J. Xu, H. Yang (vice-chairs)
High-level parallelism and coded picture regions (AHG12)	S. Deshpande (chair), M. M. Hannuksela, R. Sjöberg, R. Skupin, W. Wan, Y.-K. Wang S. Wenger (vice-chairs)
Tool reporting procedure (AHG13)	W.-J. Chien, J. Boyce (co-chairs), W. Chen, Y.-W. Chen, R. Chernyak, **K. Choi**, R. Hashimoto, Y.-W. Huang, H. Jang, R.-L. Liao, S. Liu (vice-chairs)
Operation modes for low latency support (AHG14)	J.-M. Thiesse (chair), S. Deshpande, A. Duenas, Hendry, K. Kazui, R. Sjöberg, A. Tourapis (vice-chairs)
Quantization control (AHG15)	R. Chernyak (chair), E. François, C. Helmrich, S. McCarthy, A. Segall (vice-chairs)
Implementation studies (AHG16)	M. Zhou (chair), **J. An**, E. Chai, K. Choi, S. Sethuraman, T. Hsieh, X. Xiu (vice-chairs)
High-level syntax (AHG17)	R. Sjöberg, J. Boyce (co-chairs), S. Deshpande, M. M. Hannuksela, R. Skupin, A. Tourapis, Y.-K. Wang, W. Wan, S. Wenger (vice-chairs)
Lossless and near-lossless coding tools (AHG18)	T. Nguyen, T.-C. Ma (co-chairs), M. Ikeda, S. Iwamura, **H. Jang**, X. Zhao (vice-chairs)

2.2 취약점 개선을 위한 전략 (접근방법 등)

ㅇ 국내 전문가들이 기술적으로 부족하지 않음에도 불구하고 리더쉽 포지션을 차지하지 못하는 가장 큰 이유는 표준화 경력이 상대적으로 짧기 때문으로 분석된다. 앞서 제시된 표에서 의장직을 수행하고 있는 전문가들은 적어도 10 년 이상 표준화 활동에 참여한 전문가들로서 안정적으로 표준화를 리드할 수 있을 것이라는 신뢰를 주고 있으나 국내 전문가들은 이에 비해 상대적으로 표준화 경력이 짧다는 단점을 가지고 있다. 따라서, 국내 전문가들의 취약점을 개선하기 위해서는 표준화에 지속적으로 참여할 수 있는 환경과 지원이 필요할 것으로 판단된다.

2.3 우리나라의 리더쉽 확대 방안

ㅇ 비디오 압축 분야 표준화에 있어서 특정 기술의 채택 여부는 대부분 객관적인 실험 결과 비교를 통해서 이뤄지게 되므로 상대적으로 의장단이 특정 기술의 채택 여부와 관련하여 행사할 수 있는 영향력이 제한적이라고 볼 수도 있으나, 의제 설정 및 진행은 전적으로 의장단들의 결정에 따르게 되므로 특정 기관이나 기술에 유리하거나 불리한 환경을 조성할 수 있으므로 리더쉽을 확보하는 것은 매우 중요하다.

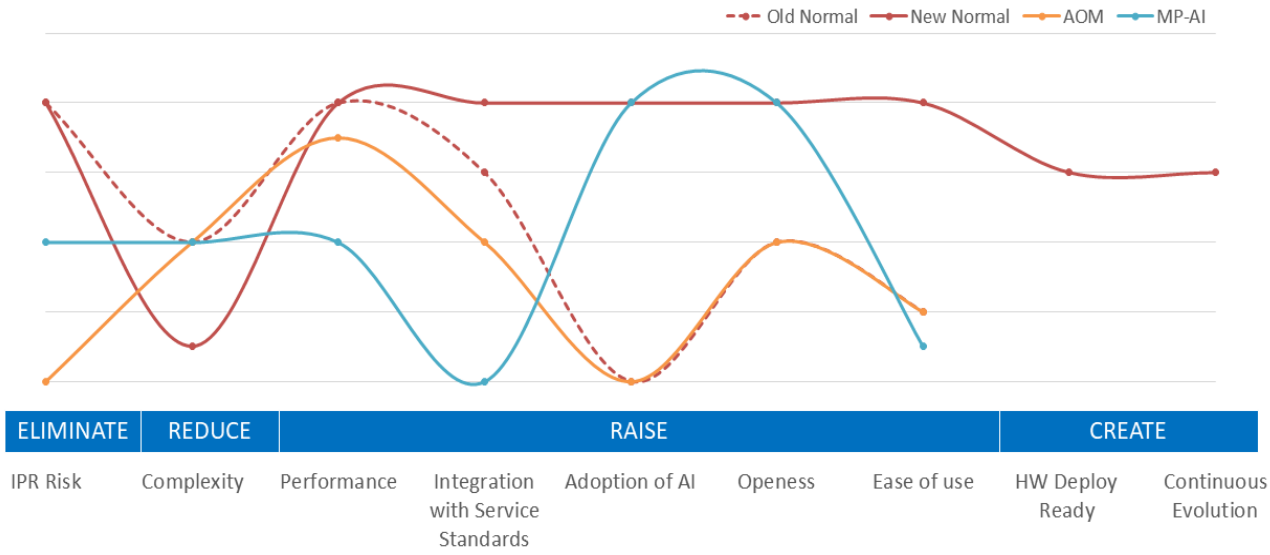
ㅇ 국제표준화 리더쉽 확대 전략

- 국내 전문가들의 경우 해외 전문가들에 비해 뒤떨어지지 않는 기술적인 능력을 보유하고 있음. 그러나, 리더쉽 확보를 위해 지속적이고 안정적인 표준화 참여에 대한 보장과 의장단 업무 수행을 위해 필요한 다른 표준화 역량, 예를 들어 영어 구사 능력, 등이 부족한 경우가 많음. 아울러, 의장단의 경우 단순히 기술을 제안하고 반영하는 휠동 외에 개인이나 기업에 직접적인 이득이 없으나 표준화 단체의 유지와 운영을 위해 필요한 업무를 수행해야 하는 경우가 많은데 이러한 업무 수행에 대한 전문가가 속한 기관에서의 이해와 지원이 부족하여 충분한 역략을 보유하고도 의장단 업무를 수행하지 못하는 경우가 많음.
- 국내 전문가들의 국제 표준화 리더쉽 확대를 위해서는 표준화에 참여하는 개별 기업들의 표준화에 대한 인식의 제고가 절실함. 표준화를 단순히 지적재산권을 확보하는 행위로만 보지 않고 새로이 제정되는 표준을 통해 새로운 시장을 형성하여 새로운 수익을 창출하는 과정으로 인식하고 표준의 성공과 빠른 보급이 표준화에 참여한 모든 기업의 이익에 직결된다는 점을 인식하고 표준화 활동에 참여하는 전문가들이 지적재산권을 확보하는 것 외에 표준 자체의 성공을 위해서 필요한 활동에도 적극적으로 참여할 수 있는 분위기를 조성하는 것이 리더쉽 확대에 큰 도움이 될것임.

3. 시사점 및 결언

○ 앞에서 살펴본 것처럼 최근 비디오 압축 표준을 개발하는 공식 국제 표준화 단체 외에 업계 표준을 개발하는 민간 표준화 단체들이 등장하면서 다수의 비디오 압축 표준이 시장에서 경쟁하는 양상을 띠게 되었다. 공식 표준화 기구에 비해 민간 표준화 단체들은 해당 단체에 참여하는 개별 기업의 시장 영향력에 따라 표준화 과정에서의 영향력이 좌우되는 특징을 가지고 있어 중소기업과 대학 전문가들의 비중이 매우 높은 국내 전문가들에게는 상대적으로 불리한 환경일 수 밖에 없다. 따라서, 공식 표준화 기구인 MPEG이 지속적으로 시장 우위를 점하는 것이 유리하다고 판단할 수 있다.

○ 아래 그래프는 표준의 시장 우위를 결정하는 다양한 요소를 상호 비교한 그래프이다. 그래프의 위쪽 구간에 위치할수록 각 요소에 대해서 시장 우위를 점할 수 있음을 나타낸다. 붉은 점선으로 나타낸 그래프는 현재 MPEG의 상황을 나타낸 것이고, 붉은 실선으로 나타낸 그래프는 MPEG이 시장 우위를 점하기 위해 변화되어야 할 특징들을 나타낸다.



- MPEG의 경우 최근 IPR 라이센싱 관련 불확실성 리스크가 커지고 있는 상황이므로 이를 해소하는 것이 필요하며, 비디오 압축 표준에 대한 기본 요구 사항인 높은 압축률과 낮은 복잡도를 유지하며, 서비스 표준과의 통합을 강화할 필요가 있다. 아울러, AI 기술 활용도를 높이고, 업계에서 쉽게 표준에 접근할 수 있도록 노력할 필요가 있다.
- 비디오 압축 표준의 경우 상용화를 위해서 하드웨어 구현이 필수적인데 빠른 하드웨어 구현을 위해 표준 개발 단계에서 하드웨어 구현을 위한 소프트웨어 개발과 현재의 7년 ~ 10년보다 짧은 주기로 새로운 표준을 개발하여 지속적으로 시장의 필요에 대응하는 등의 새로운 접근도 고려해 볼 필요가 있다.