

HEVC 기반 스케일러블 영상 압축 표준 기술(SHVC) 소개

시작하기 앞서

최근의 멀티미디어 서비스 환경은 다양한 대역폭, 프로토콜 및 QoS(Quality of Service)를 갖는 이종망(3G/LTE, 와이파이, 유선랜)이 혼재 있으며 사용자 단말(텔레비전, 노트북, 태블릿, 스마트폰)들의 컴퓨팅 성능 및 접근성의 차이가 확대되고 있다. 이러한 환경에서 멀티미디어 콘텐츠를 접근/소비하기 위해서는 다양한 네트워크 및 사용자 환경에 대응하기 위한 비디오 부호화 방식이 요구되며 이러한 방식으로는 [그림 1]에서와 같이 스케일러블(scalable) 부호화 방식과 스케일러블을 지원하지 않는(non-scalable) 방식이 존재한다.

스케일러블을 지원하지 않는 방식은 콘텐츠를 소비하고자 하는 사용자 환경에 맞도록 다수의 원본 비디오 영상을 독립적으로 따로 부호화하거나 이미 부호화된 비트스트림을 각각의 소비환경에 적합하도록 트랜스코딩 하여야 하기 때문에 복잡한 연산량 및 대용량의 저장 장치를 필요로 한다는 문제가 있다.

한편, 스케일러블 부호화 방식은 하나의 비트스트림에 다양한 해상도/프레임율/화질 등을 지원하도록 계층적 방식으로 부호화함으로써 소비자 단말에서는 비트스트림의 일부만을 추출(extractor)하여 사용자 환경에 적합한 영상을 복호하는 방식이다. 이때, 추출 방법은 부호화된 비트스트림에 포함되어 있는 계층 식별자 정보를 분석하여 추출하고자 하는 계층 식별자를 가지는 기본단위들만을 출력한다. 스케일러블 부호화 방식은 한 번의 부호화와 단순한 추출을 통하여 다양한 소비 환경을 지원할 수 있다는 장점을 가지고 있으나, 부호화 및 복호화 복잡도가 높다는 단점을 가지고 있다.

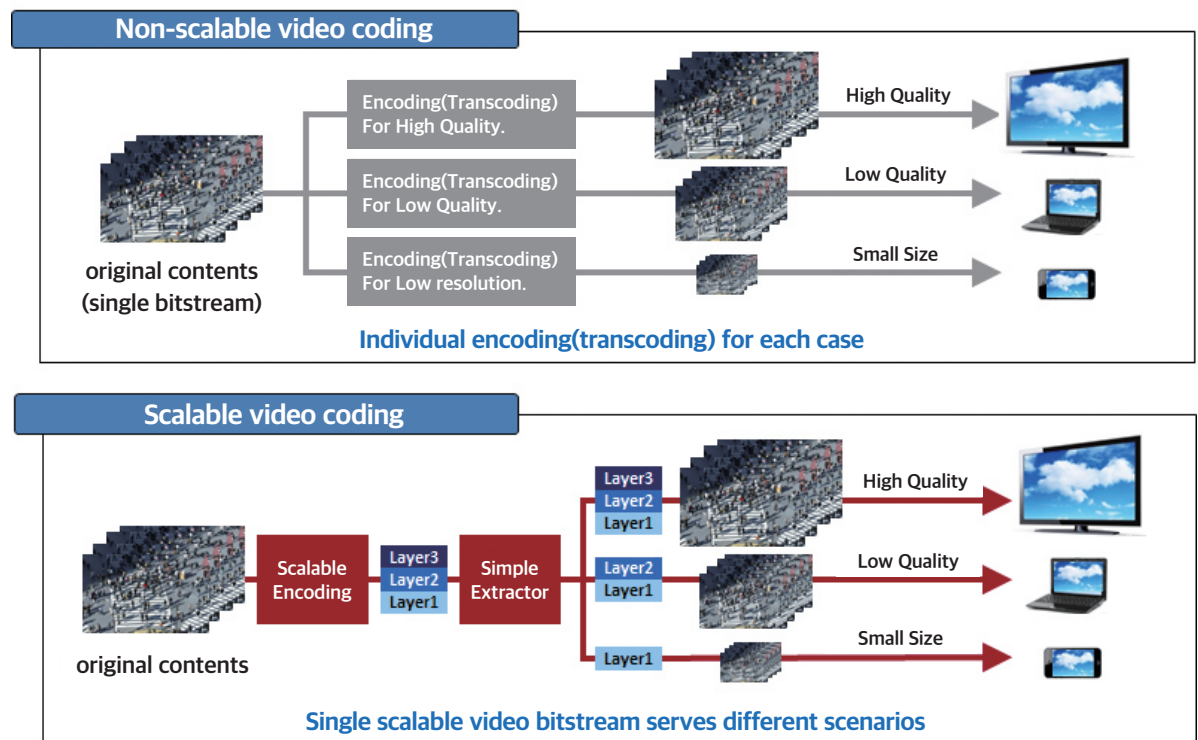


그림 1. (비)스케일러블 방식과 스케일러블 방식 개념도

스케일러빌리티(scalability) 기능을 지원하는 기존 부호화 표준은 MPEG-2, H.263, MPEG-4 Part 2 Visual 및 H.264/AVC(Advanced Video Coding) 등이 있으며, H.264/AVC 기반 스케일러블 부호화 방식인 SVC(Scalable Video Coding)[1]는 독립 부호화 방식(simulcast) 대비 높은 부호화 효율을 목표로 하여 시간적(temporal), 공간적(spatial) 및 화질적(SNR) 스케일러빌리티를 지원하도록 2007년에 개발이 완료되었다. SVC는 출력하고자 하는 최상위 계층의 픽처만을 디코딩하고 참조 관계가 있는 하위 계층의 픽처들에 대해서는 계층 간 예측에 필요한 최소한의 정보만을 디코딩하는 단일루프 디코딩(single-loop decoding) 방식을 채택하여 연산량 및 메모리 사용량을 줄일 수 있었다. 하지만 기본 계층의 복원된 차분(residual) 신호, 블록 파티션 정보 등을 얻기 위한 새로운 인터페이스 및 매크로블록 단위 '계층 간 예측 여부 지시자(base_mode_flag)' 정보로 인해 기존 H.264/AVC 코덱을 수정 없이 재사용을 하지 못함으로써 구현 복잡도가 증가하였다. 또한, H.264/AVC 표준화시 스케일러블 부호화 기술로의 확장성에 대한 고려가 이루어지지 않았고 H.264/AVC 표준화가 완료된 이후에 추가 표준안인 SVC 표준화를 시작함에 따라 스케일러빌리티 지원을 위한 최적의 구조설계가 이루어지지 못하였다. 결국, 이러한 문제들로 인하여 SVC는 화상회의 등의 시장에서 제한적으로 사용되었을 뿐 전반적으로 사용되지는 못하였다.

본 고에서는 기존 스케일러블 부호화 방식의 단점을 극복하기 위해 2014년 7월에 표준화가 완료된 HEVC 기반 스케일러블 비디오 부호화 방식인 SHVC[3] 기술을 소개한다. 세부적으로는 SHVC 코덱 구조 및 부호화 기술을 설명하며 표준화 작업에서 이루어진 부호화 성능 평가 방법과 성능에 대해 살펴본다.

SHVC 개요

ISO/IEC JTC1 MPEG(Moving Picture Experts Group)과 ITU-T VCEG(Video Coding Experts Group)의 영상 부호화 공동 협력팀인 JCT-VC(Joint Collaborative Team on Video Coding)는 차세대 영상 부호화 표준인 HEVC의 개정 표준으로 HEVC 기반의 스케일러빌리티 지원을 위한 SHVC 표준화를 진행하였다. SHVC는 2012년 10월 CfP(Call for Proposal)에 응답한 20개 기관들의 코덱들에 대한 기술 검토를 하는 것으로 본격적인 표준화가 시작되었으며, 2014년 8월 최종 개정 표준안(Final Draft Amendment, FDAM)을 발간하는 것으로 마무리가 되었다.

SHVC에서는 SVC 표준에서 지적되었던 문제점들을 극복하고자 HEVC의 표준화 기간에 스케일러빌리티 지원 구조 설계를 위한 AHG(ad-hoc group) 활동을 진행하였고 HEVC의 표준화가 완료되기 전에 SHVC 표준화를 시작하는 등 HEVC의 하이레벨 신택스(high-level syntax, HLS)를 설계하는데 있어 스케일러블 부호화 방식으로의 확장성을 고려하고자 노력하였다. 또한, SVC의 단점인 구현 복잡도를 줄이고 부호화 효율을 높이기 위해 디코딩하고자 하는 최상위 계층의 픽처가 참조하는 모든 하위 계층의 픽처들을 디코딩하는 다중 루프 디코딩(multi-loop decoding) 방식을 채택하였으며 HEVC의 블록 레벨 코딩 로직을 수정하지 않고 구현이 용이하도록 하위 참조 계층 픽처를 상위 계층의 복원 픽처 버퍼(Decoded Picture Buffer, DPB)에 추가하여 상위 계층의 시간 참조 픽처들과 함께 사용하는 참조 인덱스 프레임워크(reference index framework) 방식을 채택하였다.

[표 1]에서는 SHVC에서 지원하는 스케일러빌리티 종류를 SVC와 비교하여 보여주고 있다. SHVC에서는 기존 스케일러블 비디오 부호화 방식에서 사용된 시간, 공간, 화질 스케일러빌리티 뿐만 아니라, 기본 계층이 HEVC가 아닌 코덱을 지원하는 이종 코덱(hybrid codec) 스케일러빌리티, 계층 간 비트 심도 및 색영역이 다른 경우를 지원하기 위한 비트 심도 및 색영역 스케일러빌리티를 추가로 지원하고 있다.

프로파일 관련해서 SHVC는 Scalable Main Profile과 Scalable Main10 Profile을 지원하고 있으며, Scalable Main Profile을 지원하는 경우, HEVC 기본 계층은 Main Profile을 따라야 하며, Scalable Main10 Profile에 대해 HEVC 기본 계층은 Main10 Profile 또는 Main Profile에 일치하여야 한다.

스케일러빌리티 종류	SVC	SHVC	예	
			기본 계층	상위 계층
시간	○	○ (HEVC지원)	30fps	60fps
공간	○	○	HD(1920×1080)	UHD(3840×2160)
화질	○	○	34dB	36dB
이중 코덱	X	○	H.264/AVC	HEVC
비트 심도	X	○	8비트	10비트
색영역	X	○	BT.709	BT.2020

표 1. SVC와 SHVC에서 지원하는 스케일러빌리티 종류 비교

SHVC 코덱 구조

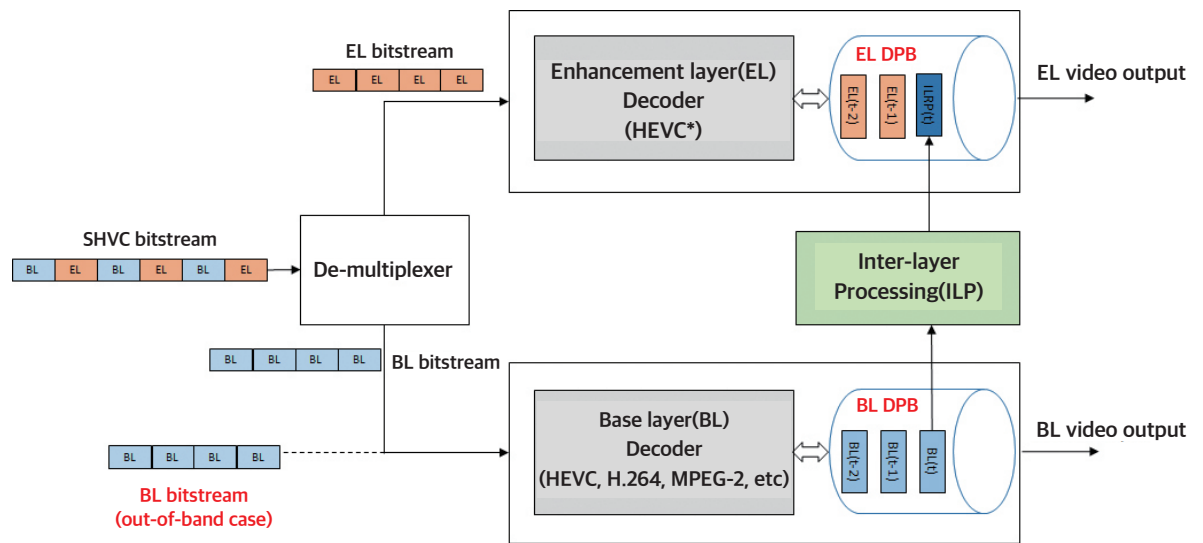


그림 2. 두 계층을 지원하는 SHVC 복호화기 구조

[그림 2]는 참조 인덱스 프레임워크 기반의 두 계층을 지원하는 SHVC 코덱 구조를 디코더 관점에서 간략히 도식화한 그림이다. 예를 들어 기본 계층(base layer, BL)은 HD(1920x1080)이며 상위 계층(enhancement layer, EL)은 UHD(3840x2160)를 지원할 수 있다. BL 비트스트림은 그림에서와 같이 in-band로 SHVC 비트스트림에 포함되어 전송되거나, 별도 수단을 통해 out-of-band 내에서 얻을 수 있다. 이미 널리 사용되고 있는 부호화 방식과의 역호환성을 제공하기 위해 BL 코덱은 HEVC뿐만 아니라 H.264/AVC, MPEG-2 또는 Google VP9과 같은 비표준 코덱을 지원할 수 있다. 한편, EL 코덱은 계층 간 참조(inter-layer dependency) 정보, 스케일러빌리티 타입 정보 등이 [그림 3]에서와 같이 하이레벨 신택스(HLS)로 지칭이 되는 비디오 파라미터 세트(video parameter sets), 시퀀스 파라미터 세트(sequence parameter sets), 픽처 파라미터 세트(picture parameter sets), 슬라이스 세그먼트 헤더 레벨에서만 추가되었을 뿐 코딩 유닛 레벨은 HEVC와 동일하다. 계층 간 예측을 위해 기본 계층에서 디코딩되어 DPB에 저장된 픽처에 대해 상위 계층과의 공간 해상도(spatial resolution), 비트심도(bit depth) 및 색 영역(color gamut) 등의 차이를 보정하기 위한 계층 간 처리(inter-layer processing, ILP)를 통해 계층 간 참조(inter-layer reference, ILR) 픽처를 생성하게 된다. ILP는 텍스처 리샘플링(texture resampling), 움직임 필드 리샘플링(motion field resampling)과 색 영역 매핑(color mapping) 모듈을 포함하고 있다. 그 후 ILR 픽처를 상위 계층 디코더의 DPB에 추가함으로써, 상위 계층의 시간 참조 픽처들과 함께 화면 간 예측에 사용한다는 것을 제외하고는 기본적으로 HEVC와 동일한 코덱 구조를 가진다. HEVC와 동일한 코딩 유닛 레벨 구조를 가짐으로써, 기존 HEVC 코덱을 재사용하여 SHVC 코덱을 구현할 수 있게 되었다.

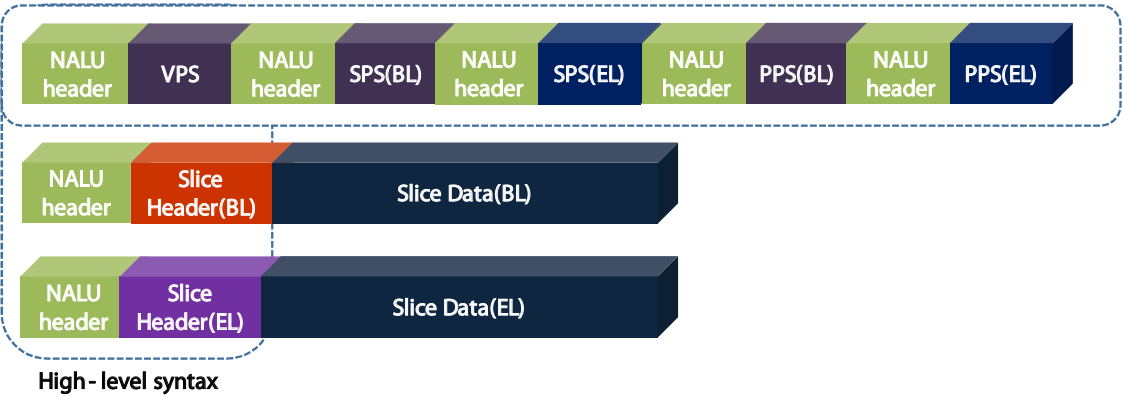


그림 3. SHVC 비트스트림 구조

텍스처 리샘플링(texture re-sampling)

참조 계층 픽처와 현재 계층 픽처가 서로 다른 공간 해상도 또는 다른 비트 심도를 가지는 경우에는 현재 계층 픽처와 동일한 공간 해상도 및 비트 심도를 갖는 계층 간 참조 픽처를 생성하기 위해 참조 계층의 디코딩된 픽처의 화소값들에 대해 리샘플링 과정을 적용하게 된다. 참조 계층의 공간 해상도를 현재 계층의 공간 해상도로 리샘플링하기 위해 참조 계층 복원 픽처의 화소값들에 적용되는 보간 필터는 [표 2]와 [표 3]에서처럼 휘도 신호에 대하여 8탭, 색차 신호에 대하여 4탭 필터 계수를 사용한다. 필터 계수는 6비트의 정밀도를 갖는데, 기본적으로 HEVC의 움직임 보상 보간에서 사용되는 DCT-IF(Discrete Cosine Transform-based Interpolation Filter)와 동일하며, 계층 간 해상도의 차이가 2배, 1.5배인 경우뿐만 아니라 다양한 배율을 지원하도록 1/16 픽셀 위치의 위상들에 대한 필터 계수를 추가하여 설계하였다. SHVC에서는 기본 계층 픽처가 인터레이스-필드(interlaced-field)이고 상위 계층 픽처가 프로그레시브-프레임(progressive-frame)인 경우, Interlaced-field to progressive-frame 스케일러빌리티를 지원하기 위한 위상 조정(phase adjustment) 정보가 PPS에 전송되어 진다. 또한, ROI(region-of-interest) 스케일러빌리티를 위한 참조 계층 및 향상 계층에 대한 cropping 정보들이 PPS에 전송되어 진다.

위상 p	보간 필터 계수							
	fL[p, 0]	fL[p, 1]	fL[p, 2]	fL[p, 3]	fL[p, 4]	fL[p, 5]	fL[p, 6]	fL[p, 7]
0	0	0	0	64	0	0	0	0
1	0	1	-3	63	4	-2	1	0
2	-1	2	-5	62	8	-3	1	0
3	-1	3	-8	60	13	-4	1	0
4	-1	4	-10	58	17	-5	1	0
5	-1	4	-11	52	26	-8	3	-1
6	-1	3	-9	47	31	-10	4	-1
7	-1	4	-11	45	34	-10	4	-1
8	-1	4	-11	40	40	-11	4	-1
9	-1	4	-10	34	45	-11	4	-1
10	-1	4	-10	31	47	-9	3	-1
11	-1	3	-8	26	52	-11	4	-1
12	0	1	-5	17	58	-10	4	-1
13	0	1	-4	13	60	-8	3	-1
14	0	1	-3	8	62	-5	2	-1
15	0	1	-2	4	63	-3	1	0

표 2. 휘도 신호 리샘플링 필터 계수

위상 p	보간 필터 계수			
	fc[p, 0]	fc[p, 1]	fc[p, 2]	fc[p, 3]
0	0	64	0	0
1	-2	62	4	0
2	-2	58	10	-2
3	-4	56	14	-2
4	-4	54	16	-2
5	-6	52	20	-2
6	-6	46	28	-4
7	-4	42	30	-4
8	-4	36	36	-4
9	-4	30	42	-4
10	-4	28	46	-6
11	-2	20	52	-6
12	-2	16	54	-4
13	-2	14	56	-4
14	-2	10	58	-2
15	0	4	62	-2

표 3. 색차 신호 리샘플링 필터 계수

움직임 필드 리샘플링(motion field resampling)

HEVC는 움직임 정보 부호화를 위해 현재 블록 주변 블록의 움직임 정보와 시간상으로 대응되는 참조 픽처에 속한 블록의 움직임 정보를 예측 후보로 하여 현재 블록의 움직임 정보를 부호화하는 AMVP(Advanced MV Prediction)와 Merge 모드를 지원하고 있다. SHVC에서는 계층 간 참조 픽처를 시간 움직임 예측 후보(temporal motion vector predictor, TMVP)를 구하기 위한 대응 픽처(collocated picture)로 지정하여 참조 계층 픽처의 움직임 정보를 TMVP로 사용하도록 하였다. 이로 인해 HEVC 코딩 유닛 레벨 알고리즘을 수정하지 않고 계층 간 움직임 예측이 가능해졌다.

참조 계층 픽처의 움직임 정보를 현재 계층의 TMVP로 사용하기 위해서는 현재 계층의 해상도에 맞게 매핑(mapping)하는 절차가 필요하다. [그림 4]는 참조 계층과 현재 계층이 2배 공간 해상도 차이인 경우의 움직임 정보 매핑 방법을 나타내며 (a)는 참조 계층의 복원된 블록, (b)는 ILR 픽처의 16×16 블록들을 나타낸다. HEVC에서 16×16 블록 단위로 움직임을 저장하기 때문에 SHVC에서도 16×16 단위로 매핑을 진행한다. 즉, 현재 계층의 픽처를 16×16 단위의 블록으로 분할했을 경우에 매핑하고자 하는 대상 16×16 블록의 중앙(xPCtr, yPCtr)에 위치한 샘플에 대응하는 참조 계층 블록의 샘플 위치(xRef, yRef)에서 (4, 4)만큼 이동한 샘플 위치(xRL, yRL)를 찾고 해당하는 샘플을 포함하고 있는 16×16 블록의 움직임을 현재 블록의 움직임 정보로 매핑한다. 이때, 움직임 벡터의 크기는 2배 해상도 차이를 고려하여 현재 계층의 해상도에 맞게 스케일링 한다. 기본 계층이 HEVC가 아닌 이종 코덱의 경우에는 움직임 필드 정보에 대한 인터페이스가 존재하지 않기 때문에 움직임 필드 리샘플링 과정을 수행하지 않는다.

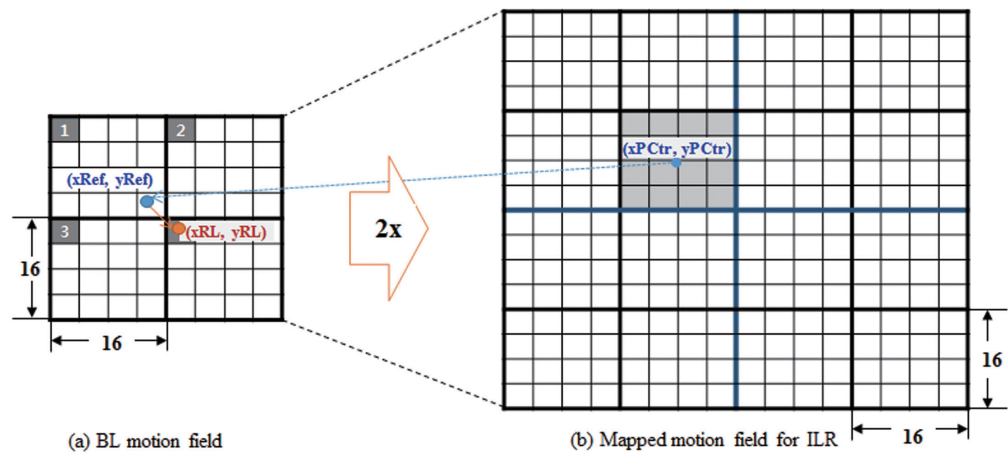


그림 4. 움직임 필드 리샘플링 예

색영역 매핑(color mapping)

색영역 스케일러빌리티(CGS, color gamut scalability)는 한 가지 종류 이상의 색영역을 지원하는 기능으로, HD 방송의 경우는 Rec. BT. 709, UHD 방송의 경우는 Rec. BT. 2020의 색영역을 사용하도록 권고됨에 따라 중요하게 부각되었다. [그림 5]에서 보는 바와 같이 Rec. BT. 2020의 색영역이 Rec. BT. 709보다 넓은 색영역을 포함하고 있기 때문에 더 다양한 색을 표현할 수 있다.

이때, 더 넓은 색영역을 효과적으로 표현하기 위해서 더 많은 비트심도가 요구된다. 따라서, 방송에서 지원하는 색영역을 고려하는 경우에, HD 영상을 채널당 8비트의 비트심도로 표현하고 UHD 영상을 채널당 10비트의 비트심도로 표현하는 것이 일반적이며, 이에 따른 영향으로 색영역 매핑은 비트심도 스케일러빌리티도 함께 제공

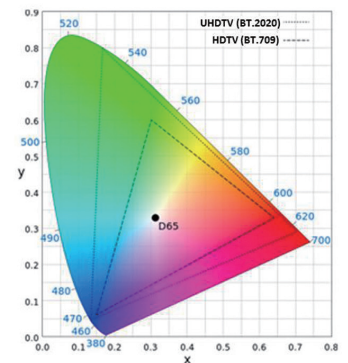


그림 5. HDTV(BT.709)와 UHD TV(BT.2020)를 위한 색영역 비교

하도록 설계되었다. 예를 들어, 기본 계층은 HD 방송을, 향상 계층은 UHD 방송을 제공하는 경우에, 기본 계층의 Rec. BT. 709의 HD 픽처를 부/복호화하고, 디코딩된 해당 픽처를 Rec. BT.2020의 색영역으로 변환(비트심도 변환 포함)한 후 업샘플링하여 계층 간 참조 픽처로 사용한다. 색영역 변환 시 3D Look-Up Table 방식이 적용되며, 사용되는 3D LUT은 향상 계층의 입력 픽처와 기본계층의 복원 픽처를 이용한 학습을 통해 결정된다.

SHVC 부호화 성능

JCT-VC에서는 공통 실험 조건[4]을 정의하고, 이를 기반으로 SHVC 참조 소프트웨어인 SHM(SHVC test model)[5]을 이용하여 제안한 기술을 실험하고 평가하였다. 공통 실험 조건에서 기본 계층과 향상 계층의 해상도 비율이 2배 및 1.5배인 공간 해상도 스케일러빌리티 경우와 기본 계층과 향상 계층의 해상도가 동일하나 화질이 다른 화질(SNR) 스케일러빌리티 경우에 대한 실험을 수행하였다. 또한, 기본 계층과 향상 계층의 색 영역이 다른 색 영역(CSG) 스케일러빌리티 경우는 [표 4]와 같은 조건으로 실험을 수행하였다.

	BL/EL 스케일러빌리티	BL 입력영상	EL 입력영상
Case 1	공간해상도+비트심도+색영역	8bit / BT.709 / 1080p	10bit / BT.2020 / 2160p
Case 2	색영역	10bit / BT.2020 / 1080p	10bit / BT.2020 / 1080p

표 4. CGS 스케일러빌리티 실험 조건

조건	spatial		SNR	CGS	
	2x	1.5x		case 1	case 2
QP(BL/EL)	BL _{QP} = {22,26,30,34} EL _{QP} =BL _{QP} +Δ{0,2}		BL _{QP} = {26,30,34,38} EL _{QP} =BL _{QP} +Δ{-4,-6}	BL _{QP} = {26,30,34,38} EL _{QP} =BL _{QP} +Δ{0,2} EL _{QP} =BL _{QP} +Δ{0,-4}	
HEVC(BL+EL) simulcast(%)	-16.5	-27.0	-21.0	-18.8	-29.6
HEVC highest single-layer(%)	19.0	16.0	14.3	20.8	27.7

표 5. RA(Random Access) 실험 조건에서의 부호화 성능

[표 5]의 실험 결과를 통해, 두 계층을 SHVC를 이용하여 계층 간 예측 부호화를 수행하는 경우는 HEVC를 이용하여 단일 계층 부호화 방식으로 향상 계층만을 부호화하는 경우에 비해 평균 19% 정도의 비트량이 증가하지만, 기본 계층과 향상 계층을 HEVC를 이용하여 독립적으로 동시 부호화하는 경우에 비해 평균 23% 정도의 비트량이 감소함을 확인할 수 있다.

또한, JCT-VC에서는 SHVC 표준화 완료 이후, SHVC verification test[6] 일환으로 주관적 화질 평가를 통한 MOS(mean opinion score) 기반 HEVC simulcast 부호화 대비 SHVC의 비트율 감소량을 측정하였다. 주관적 화질 평가는 540p부터 4K/UHD까지 다양한 시퀀스를 대상으로 공간(2배, 1.5배), 화질 및 CGS 스케일러빌리티 경우에 대해 수행이 되었다. [그림 6]의 평가 결과는 SHVC로 두 계층을 계층 간 예측 부호화하였을 경우 평균적으로 40%~60% 정도의 적은 비트량으로 HEVC를 이용하여 독립 부호화한 경우와 동일한 주관적 화질을 얻을 수 있다는 것을 보여주고 있다.

따라서 다양한 소비환경에 비디오 콘텐츠를 서비스하는 경우에, 소비환경에 맞추어 독립적으로 부호화하는 것보다 SHVC를 사용하는 것이 네트워크 자원 및 콘텐츠 저장/관리에 효율적이라 할 수 있을 것이다. 그러나 SHVC는 여러 계층을 동시에 부호화 및 복호화하므로 인해 부호화 및 복호화시 연산량이 증가하고, 메모리 사용량 또한 50% 이상 증가한다는 단점도 존재한다.

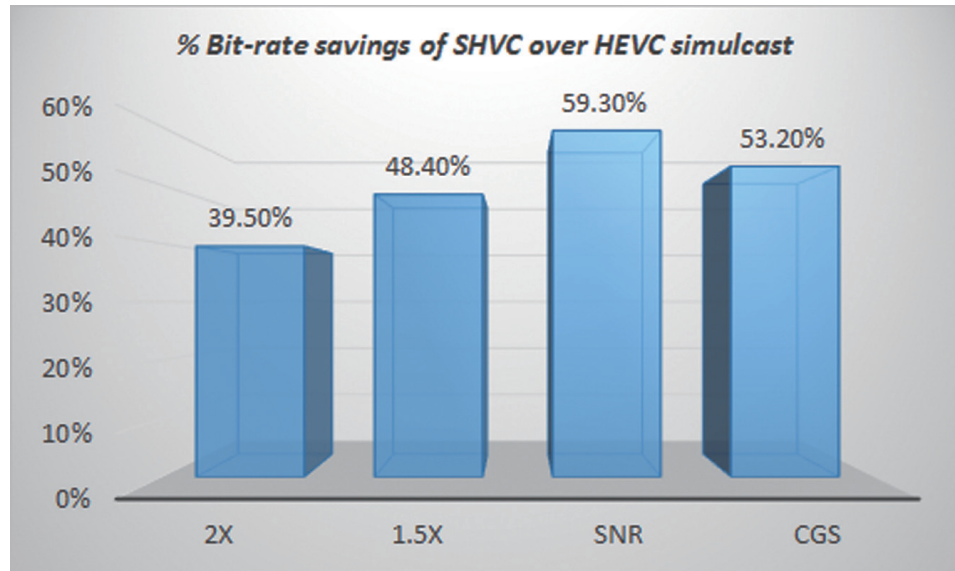


그림 6. HEVC simulcast 부호화 대비 SHVC의 평균 비트 감소율(MOS 기반)

맺음말

본 고에서는 HEVC 기반의 스케일러블 영상 압축 표준 기술인 SHVC를 소개하고 부호화 기술을 간략히 살펴보았다. SHVC는 참조 인덱스 접근 방식을 채택함에 따라 HEVC 코덱 설계의 많은 부분을 재사용하여 구현의 복잡도가 낮고, 스케일러빌리티 제공을 위한 최적화된 상위구조를 가진다. 최근 들어 멀티미디어 콘텐츠의 소비 환경이 이종망 및 다양한 단말을 통해 이루어지고, UHD, High Dynamic Range 및 Wide Color Gamut과 같은 새로운 비디오 포맷이 등장함에 따라, 스케일러빌리티 기능이 필요한 다양한 응용 분야에서 SHVC 표준 기술이 활용될 것으로 기대된다. 📺

[참고문헌]

- [1] H. Schwarz, D. Marpe, and T. Wiegand "Overview of the scalable video coding extension of the H.264/AVC standard," IEEE Trans. on Circuits and Systems for Video Technology 17(9), pp.1103-1120, July 2007.
- [2] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Trans. on Circuits and Systems for Video Technology 22(12), pp.1649-1668, December 2012.
- [3] J. Boyce, Y. Ye, J. Chen, and A-K. Ramasubramanian, "Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding Standard", IEEE Trans. on Circuits and Systems for Video Technology 26, pp.20-34, January 2016.
- [4] V. Seregin and Y. He "Common SHM test conditions and software reference configurations", JCT-VC document JCTVC-P1009, January 2014.
- [5] J. Chen, J. Boyce, Y. Ye, and M. M. Hannuksela "SHVC Test Model 7 (SHM 7) Introduction and Encoder Description," JCT-VC document JCTVC-R1007, July 2014.
- [6] Y. Ye, V. Baroncini, and Y-K. Wang, "SHVC verification test report", JCT-VC document JCTVC-W1004, February 2016