

특집논문 (Special Paper)

방송공학회논문지 제20권 제4호, 2015년 7월 (JBE Vol. 20, No. 4, July 2015)

<http://dx.doi.org/10.5909/JBE.2015.20.4.497>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

HEVC 기반 스케일러블 비디오 부호화 툴의 성능 분석

김용태^{a)}, 최진혁^{a)}, 최해철^{a)†}

Performance Analysis of Scalable HEVC Coding Tools

Yongtae Kim^{a)}, Jinhyuk Choi^{a)}, and Haechul Choi^{a)†}

요 약

최근 HD(High Definition)화질 및 UHD(Ultra High Definition)화질과 같은 고품질 방송 서비스가 등장하고, 무선 네트워크 기술의 발달로 스마트폰, 태블릿 PC 등과 같은 휴대용 멀티미디어 기기들이 보편화됨에 따라, 비디오 콘텐츠의 전달 및 소비 환경이 다양화되고 있다. 따라서 스케일러빌리티(Scalability)의 현실적 필요성이 점점 절실해졌으며, 시간, 공간, 화질 등의 확장성을 제공하는 Scalable Video Coding(SVC)이 등장하였다. 최근 ISO/IEC의 MPEG(Moving Picture Experts Group)와 ITU-T의 VCEG(Video Coding Experts Group)이 공동으로 결성한 Joint Collaborative Team on Video Coding(JCT-VC)에 의해 Scalable High Efficiency Video Coding(SHVC) 표준이 제정되었다. 본 논문은 공간적, 시간적, 화질적 스케일러빌리티를 제공하기 위하여 SHVC의 표준에 포함된 툴 뿐만 아니라 SHVC 표준화 과정에서 논의되었던 다양한 스케일러블 부호화 툴들을 소개하고, 기존 단일 계층 부호화 방식으로 부호화한 결과와 비교하여 SHVC의 부호화 효율에 대한 성능을 분석하였다. 이러한 성능 분석은 향후 부호화 성능 향상을 위한 알고리즘 개발 및 고속 부호화기 개발을 위한 부호화 툴의 선별 및 조정에 유용할 것으로 판단된다.

Abstract

Current communication networks consist of channels with various throughputs, protocols, and packet loss rates. Moreover, there are also diverse user multimedia consumption devices having different capabilities and screen sizes. Thus, a practical necessity of scalability on video coding have been gradually increasing. Recently, The Scalable High Efficiency Video Coding(SHVC) standard is developed by Joint Collaborative Team on Video Coding(JCT-VC) organized in cooperation with MPEG of ISO/IEC and VCEG of ITU-T. This paper introduces coding tools of SHVC including adopted and unadopted tools discussed in the process of the SHVC standardization. Furthermore, the individual tool and combined tool set are evaluated in terms of coding efficiency relative to a single layer coding structure. This analysis would be useful for developing a fast SHVC encoder as well as researching on a new scalable coding tool.

Keyword : scalable video, SHVC, video coding, performance analysis, HEVC

a) 한밭대학교 정보통신전문대학원(Hanbat National University)

† Corresponding Author : 최해철(Haechul Choi)

E-mail: choihc@hanbat.ac.kr

Tel: +82-42-821-1149

ORCID: <http://orcid.org/0000-0002-7594-0828>

* 이 논문은 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구개발사업[B0126-15-1013, 퍼즐형 Ultra-wide viewing 공간 미디어 생성 및 소비 기술 개발]과 2013년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(NRF-2013R1A1A1010344)의 일환으로 수행하였음.

· Manuscript received May 18, 2015; revised June 12, 2015; accepted June 12, 2015.

1. 서론

최근 다양한 멀티미디어 디바이스의 보급과 각종 멀티미디어 기술의 급속한 발전에 따라, 동영상 정보를 포함한 다양한 멀티미디어 서비스에 대한 소비자의 요구가 커지고 있으며, HD (High Definition) 화질 및 UHD (Ultra High Definition) 화질과 같은 고품질 방송 서비스가 등장하면서, 정보의 질적 향상에 더 많은 관심을 가지게 되었다.

멀티미디어 비디오 서비스에서 널리 쓰이고 있는 비디오 부호화 기술은 고정된 영상 포맷(format)과 비트율(bit-rate)로 압축 및 전송하는 단일 계층 압축 부호화(single layer video coding) 방식이다. 하지만 위와 같은 단일 계층 방식은 근래의 무선 네트워크 기술의 발달로 다양한 환경에서 시변의 전송조건, 망의 특성, 사용자 선호도, 그리고 단말의 다양한 성능 등의 조건들에 모두 적합하도록 미리 부호화하기 어려우므로 끊임 없는 품질의 멀티미디어에 대한 사용자의 욕구를 충족시킬 수 없다. 이러한 환경에서 실시간 비디오 전송 및 이종망(Heterogeneous) 서비스에 적합한 시간, 공간, 화질 등의 다양한 스케일러빌리티(scalability)를 지원하는 Scalable Video Coding(SVC) 기술은 멀티미디어의 전송 및 소비 환경 적응을 위한 매력적인 해결책 중의 하나이다^[1]. 이에 따라 ISO/IEC의 MPEG (Moving Picture Experts Group)와 ITU-T의 VCEG (Video Coding Experts Group)에 의해 결성된 Joint Collaborative Team on Video Coding (JCT-VC)은 Scalable High Efficiency Video Coding (SHVC) 표준화를 개발하였다^[2].

SHVC에서 요구하는 스케일러빌리티의 종류에는 공간적 스케일러빌리티(Spatial Scalability), 시간적 스케일러빌리티(Temporal Scalability) 그리고 화질적 스케일러빌리티(SNR Scalability)가 있다. 먼저 공간적 스케일러빌리티는 다양한 해상도의 영상을 계층으로 나누어 부호화한다. 하위 계층(Base Layer)은 기존 HEVC 기술만을 사용하여 부호화하고, 상위 계층(Enhancement Layer)은 움직임 정보, 잔차 신호 등 하위 계층과의 상관관계를 이용하여 부호화하며, SHVC에서는 1.5배와 2배 등 다양한 공간적 스케일러빌리티를 지원할 수 있다. 다음으로 시간적 스케일러빌리티는 같은 공간 해상도에서 프레임 주파수가 다른 것들

을 부호화 하는 방법이다. 이미 부호화한 시간적 하위 계층의 영상을 참조 영상으로 사용함으로써 부호화 효율을 증가시킬 수 있다. 마지막으로 화질적 스케일러빌리티의 경우 상위 계층은 하위 계층 보다 작은 양자화 계수 값으로 보정하여 화질을 계층별로 부호화하는 방법이다. 이로 인해 점차적으로 복원 영상의 화질이 향상되도록 비트스트림을 구성할 수 있다. 기본적으로 스케일러블 비디오 부호화는 각 계층별로 화질 또는 공간 해상도만 다를 뿐 동일한 영상이므로 계층 간의 상관성이 매우 높다. 그렇기 때문에 기본계층의 텍스처, 움직임 등의 부호화 정보들을 향상계층의 예측에 사용함으로써 부호화효율 향상을 기대할 수 있다.

SHVC 표준화 회의에서 논의 되었던 대표적인 부호화 구조는 시간적으로 동일한 위치에 있는 하위 계층의 대응영상을 상위 계층의 참조 리스트에 추가하여 상위 계층 부호화에서 화면간 예측과 같이 사용할 수 있도록 지원하는 Reference Index 기반 프레임워크와 기본 계층의 대응블록을 이용하여 상위 계층의 블록을 예측하는 IntraBL 기반 프레임워크이다. 각각의 프레임워크 모두 하위 계층의 정보를 상위 계층 부호화시에 참조하는 계층 간 예측(inter-layer prediction) 기술을 사용하고 있으며, 계층 간 예측 기술에는 상위 계층 부호화 시 하위 계층의 텍스처 정보를 참조하는 계층 간 텍스처 예측과 움직임 정보를 참조하는 계층 간 움직임 예측이 포함되어 있다. Reference Index 및 IntraBL 기반 프레임워크들은 각각 영상 및 블록 레벨에서의 틀이라는 차이점이 있다. 초기 SHVC 표준화 과정에서는 두 프레임워크를 함께 고려했으나, IntraBL 기반 프레임워크는 표준에서 결국 제외되었고 Reference Index 기반 프레임워크만이 표준으로 채택하였다. 이러한 이유 중 하나는 Reference Index 기반 프레임워크는 HEVC 표준의 high level syntax, 즉 영상 단위 정보 변경만으로 스케일러빌리티를 지원할 수 있는 반면, IntraBL 기반 방법은 블록 레벨에서 구문요소의 수정이 불가피했기 때문이다.

본 논문에서는 최근 완료된 SHVC 표준에 채택된 부호화 틀과 표준화에서 고려되었던 다양한 부호화 틀에 대한 성능 분석을 제시한다. 성능 분석은 기존의 HEVC로 하위 계층과 상위 계층을 각각 단일 계층(Single Layer) 부호화

한 성능과 다 계층 부호화인 SHVC로 부호화한 성능을 프레임워크별로 나누어 비교하였고, Reference Index 기반 프레임워크와 IntraBL 기반 프레임워크의 부호화 성능 비교뿐만 아니라 각각의 프레임워크에서 계층 간 예측 기술들의 성능을 비교하였다. 이러한 성능 분석은 SHVC가 Reference Index 기반 프레임워크로 표준화를 채택한 근거를 유추할 수 있으며, SHVC 고속 부호화기 개발뿐만 아니라 향후 새로운 스케일러블 비디오 부호화 툴의 개발에 유용할 것으로 판단된다.

본 논문의 2장에서는 SHVC에 대한 설명과 함께 표준 기술 및 표준화에서 고려되었던 기술에 대하여 소개하고, 3장에서는 2장에서 기술한 부호화 기술에 대한 성능을 평가를, 마지막으로 4장에서 결론을 맺으며 마무리한다.

II. SHVC 비디오 부호화 개요

1. 부호화 방식

공간적 스케일러빌리티를 위한 SHVC는 그림 1과 같이 낮은 공간 해상도의 비디오와 높은 해상도의 비디오를 각 공간적 계층에 입력하여 부호화하는 다 계층 방식에 기반한다. 각 공간적 계층들은 서로 독립적으로 부호화할 수도 있지만, 하위 계층에서 부호화한 정보를 이용하여 상위 계

층의 부호화 효율을 향상시킬 수 있다. 시간적 혹은 화질적 스케일러빌리티를 위한 구조에서도 동일한 방법을 적용할 수 있다. 다 계층 기반의 스케일러블 비디오 부호화 방식은 복호기의 구조에 따라 단일 루프 복호화 방식 또는 다중 루프 예측 복호화 방식으로 설계할 수 있다.

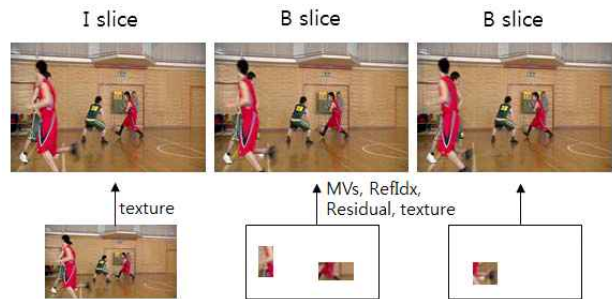


그림 2. SVC의 단일 루프 예측 방식

Fig. 2. single-loop decoding

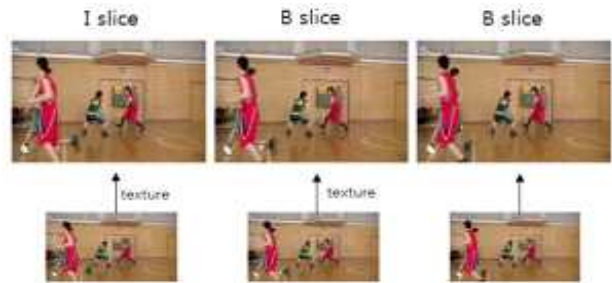


그림 3. SVC의 다중 루프 예측 방식

Fig. 3. multi-loop decoding

기존 H.264/AVC 기반의 SVC는 단일 루프 복호화 방식을 채택하였다. 단일 루프 예측 방식은 그림 2와 같이 복호화 대상이 되는 타겟 계층의 영상만을 완전히 복호화한다. 하위 계층 영상들은 전체 영상을 복호화하지 않고 계층간 Intra_BL로 참조된 블록, 부호화 정보 및 잔차(residual) 정보만을 선별적으로 복호화한다. 이와 달리 SHVC는 다중 루프 예측 방식에 기반한다. 다중 루프 예측 방식은 그림 3과 같이 복호화 대상이 되는 타겟 계층의 영상뿐만 아니라 참조하는 모든 하위 계층 영상을 복호화한다. 이러한 다중 루프 예측 방식은 단일 루프 예측 방식에 비해 압축의 성능은 높지만, 복잡도가 높다는 단점도 있다^[4]. SVC 표준화에

그림 1. SHVC 인코더의 high-level 블록 다이어그램^[3]
 Fig. 1. High-level block diagram of a SHVC encoder^[3]

서는 계산 복잡도의 과도한 증가를 감소시키기 위해서 단일 루프 복호화 방식을 채택하였으나, SHVC 표준화에서는 하드웨어 설계 측면에서 부/복호화기의 구조를 간소화시키고 블록 레벨에서의 HEVC 표준을 유지하기 위해 다중 루프 예측 복호화를 채택하였다.

예측 방식 측면에서, SHVC 표준화에서는 다 계층 기반의 부호화에서 부호화 효율을 향상시키기 위한 다양한 계층 간 예측 방식이 제안되었다. 서론에서 기술한바와 같이 SHVC에서는 Reference Index 기반과 IntraBL 기반의 프레임워크가 고려되었다. 각 프레임워크에서 계층 간 부호화를 위한 틀들을 2-2, 2-3장에서 소개한다.

2. Reference Index 기반 SHVC

Reference Index 기반 SHVC는 하위 계층 영상을 다중 계층의 해상도에 맞춰 업샘플링하여 만든 계층 간 참조 영상(Inter-layer Reference Picture, ILP)을 참조 영상 리스트에 추가하는 방법으로 계층 간 예측을 지원한다.

2.1 계층 간 텍스처 예측(Inter-layer texture prediction)

Reference Index 기반 SHVC에서 계층 간 텍스처 예측 구조는 HEVC의 구조에서 block level syntax와 복호화 과정은 유지하고, 오직 high level syntax만 변경하였다. 즉, 현재 영상과 동일한 POC 값을 갖는 하위 계층의 영상을 현재 영상의 부호화를 위한 참조 영상으로 사용할 수 있도록 구조화 하였다. 그렇기 때문에 상위 계층은 동일 계층의 시간적 참조영상 뿐만 아니라 하위 계층의 계층 간 참조영상을 모두 참조 영상 리스트에 포함한다. SHVC는 참조 영상 리스트를 그림 4와 같이 구성한다. 참조 영상 리스트 L0는 시간적 참조 영상인 과거 참조 영상과 미래 참조 영상 사이에 계층 간 참조 영상을 추가하여 구성한다. 참조 영상 리스트 L1은 HEVC의 참조 영상 리스트 구성과 같은 방법으로 시간적 참조영상인 미래 참조 영상, 과거 참조 영상 순서로 구성되며, 그 뒤에 long term 참조 영상과 계층 간 참조 영상으로 구성한다. 만약 항상 계층 영상을 P-Slice로 부호화한다면, 계층 간 참조 영상은 참조 영상 리스트 L0에만 추가되고, 항상 계층 영상을 B-Slice로 부호화한다면, 참

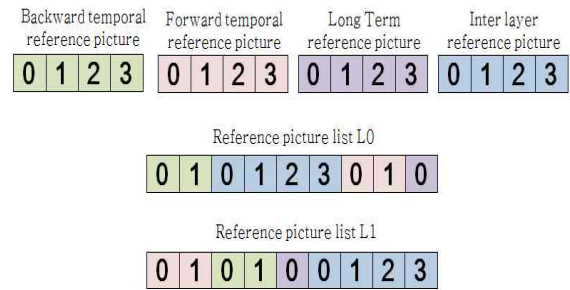


그림 4. 참조 영상의 종류 및 구성
Fig. 4. type and construction of the reference picture

조 영상 리스트 L0와 L1에 모두 계층 간 참조 영상을 추가한다. 이와 같이, SHVC는 하위 계층의 복원된 영상을 참조 리스트에 삽입함으로써 화면간 예측과 동일한 방식으로 계층 간 참조를 지원한다. 이러한 구조를 Reference Index 기반 프레임워크에서의 계층 간 텍스처 예측이라고 지칭한다. 계층 간 텍스처 예측과 관련된 부호화 틀은 다음과 같다.

- REF_IDX_ME_ZEROMV^[5]

계층 간 참조 영상에서 유사한 텍스처 블록을 찾기 위한 PU의 움직임 벡터(Motion Vector, MV) 탐색 과정은 기본적으로 HEVC와 동일하다. AMVP(Advanced Motion Vector Prediction) 과정에 의해 얻어지는 MVP(Motion Vector Predictor)를 움직임 탐색 시에 중심점(혹은 시작점)으로 사용한다. 하지만 주변 블록으로부터 획득한 MVP는 현재 영상과 시간적으로 차이를 갖는 동일 계층의 복원 영상을 참조했을 수 있기 때문에 시간적으로 동일한 시점의 계층 간 참조 영상에 이를 직접 중심점으로 적용하는 것은 계층 간 예측 효율을 떨어뜨릴 수 있다. 따라서 계층 간 참조 영상이 항상 계층의 참조 영상으로 선택 되었을 경우에는 대응되는 PU의 움직임 벡터를 0(ZeroMV)로 고정한다.

- ENCODER_FAST_MODE^[6]

SHVC AI(All Intra mode)의 경우 simulcast보다 두 배의 인코딩 시간을 요구한다. 그렇기 때문에 하위 계층이 I_Slice이고 상위 계층이 P_Slice로 부호화할 때, 상위 계층에는 오직 한 개의 참조 영상이 있으므로 움직임벡터는

ZeroMV로 고정한다. 이 경우, 서로 다른 화면간(inter) 예측 모드에 대한 비용 계산에서 많은 중복성이 있기 때문에 2Nx2N 이외의 화면간 예측 모드에 대한 비용 계산을 시도할 필요가 없다. 중복 계산을 피하기 위해서 모든 화면간 예측 후보 탐색 및 모드 결정을 비활성화하고, ZeroMV인 2Nx2N 계층 간 참조 영상 단방향 예측(uni-prediction) 모드를 추가한다.

• MAX_TID_REF_PRESENT_FLAG^[7]

하위 계층 영상이 random access picture(RAP)인 경우에 기본 계층을 제외한 모든 향상계층을 sub-layer라 하며, 계층간 예측에 사용하지 않게 선언된 sub-layer의 영상은 간단하게 입력 비트스트림에서 제거하거나 디코딩 과정에서 생략 가능하다. 2개 이상의 서로 다른 계층에서 계층 간 예측의 허용을 제어하는 것은 추가적인 복잡도를 줄일 수 있다.

• MAX_TID_FOR_REF_LAYERS^[8]

2개 이상의 상위계층이 존재할 경우, 각 계층 별로 참조하는 계층을 설정할 수 있다. 이 때, 참조하는 계층을 직접 참조 계층이라 하며, MAX_TID_FOR_REF_LAYERS는 직접 참조 계층의 temporal sub-layer(Tid) 영상이 상위 계층의 계층 간 예측에 사용되었는지 여부를 판단하는데 사용되는 flag이다. 이 정보는 계층 간 예측에 이용되지 않는 sub-layer에 속한 영상을 제거하는 것에 유용하다. 예를 들어 그림 5와 같이 직접 참조 계층의 Tid=2는 계층 간 예측으로 Layer ID=1은 사용할 수 있지만, Layer ID=2는 사용할 수 없게 설정할 수 있다. 이는 HEVC의 다 계층 확장에서 계산 복잡도, 부호화 효율, 메모리 대역폭 등에 trade-off를 제공한다. 특히 실시간 응용 분야에서 매우 유용할 수 있다.

• FAST_INTRA_SHVC^[9]

향상 계층의 CU와 기본 계층의 대응 CU, PU가 모두 화면내(intra) 예측인 경우, 같은 화면내 예측 방향성을 가질 가능성이 높다. 따라서 향상 계층의 화면내 예측 모드에서 angular 모드의 후보수를 기존 33 가지에서 2 또는 4 가지의 경우로 제한하여 부호화기의 복잡도를 줄인다.

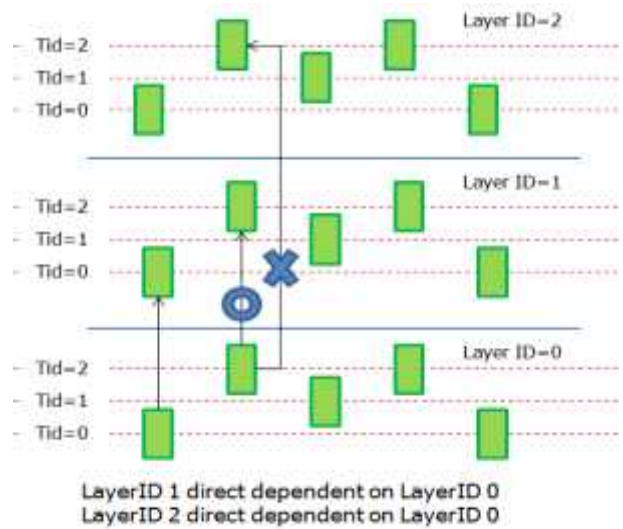


그림 5. 직접 참조 계층을 이용한 계층 간 예측
 Fig. 5. Inter layer prediction using direct reference layer

2.2 계층 간 움직임 예측(Inter-layer motion prediction)

Reference Index 기반의 SHVC에서 계층 간 움직임 예측 구조는 참조 영상으로 계층 간 참조 영상을 사용하여 Merge/AMVP의 TMVP(Temporal Motion Vector Predictor)를 유도함으로써 움직임을 예측하는 구조이다. 표 1과 같이 TMVP 후보 결정시에 시간적 참조 영상을 대신하여 ILP를 사용한다.

표 1. SHVC의 TMVP후보
 Table 1. TMVP candidates of SHVC

Merge candidate list	AMVP candidate list
Left	Left
Above	Above
Above right	TMVP(ILP)
Left bottom	Zero candidates
Above left	
TMVP(ILP)	
Combined candidates	
Zero candidates	

공간적인 스케일러빌리티의 경우, 향상 계층과 참조 계층 사이에 TMVP를 유도하기 위해 추가적인 block level 복호화 과정의 수정 없이 Motion Field Mapping(MFM) 과

정이 수행된다.

- REF_IDX_MFM^[10]

ILP의 움직임 정보는 기본 계층 영상의 움직임 정보를 기반으로 얻을 수 있다. ILP의 16x16 블록에 움직임 벡터 및 참조 영상 인덱스와 같은 움직임 정보와 예측모드는 기본 계층 영상의 대응하는 블록에 움직임 정보와 예측모드로부터 유도한다.

- POSITION_ROUNDING_OFFSET^[11]

그림 6과 같이 ILP의 움직임 벡터를 유도하는 과정에서 ILP의 블록과 대응하는 기본 계층의 블록 위치를 (xRL, yRL)이라 한다. (xRL, yRL)은 16x16크기의 1번 블록 내부에 위치하고 있다. 하지만 (xRL, yRL)은 4번 블록과 더 가깝기 때문에 4번 블록의 움직임과 더 높은 상관성을 가질 확률이 높다. 따라서 식 (1)과 같이 반올림 연산을 통해 기본 계층 블록에서 더 가까이 있는 움직임벡터를 사용하도록 유도하여 움직임 벡터의 정확성을 높인다. 여기서 R=4, S=4로 정의한다.

$$\begin{aligned} xRL &= (xRL + R \gg S) \ll S \\ yRL &= (yRL + R \gg S) \ll S \end{aligned} \quad (1)$$

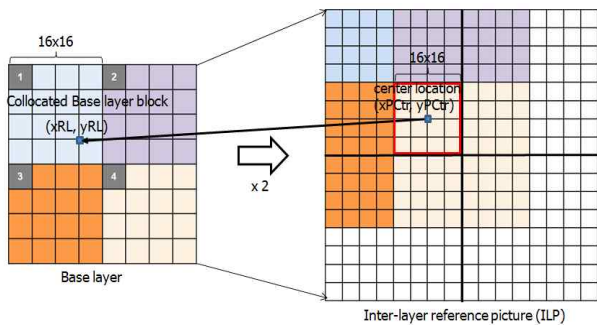


그림 6. 기본 계층을 이용한 향상 계층의 움직임 유도
Fig. 6. motion derivation using collocated base layer

3. IntraBL 기반 SHVC

IntraBL 기반 SHVC는 기본 계층 영상을 업샘플링하여

만든 복원 영상의 블록을 이용하여 현재 향상 계층 영상의 블록을 예측하는 방법이다.

3.1 계층 간 텍스처 예측(Inter-layer texture prediction)

IntraBL 기반 SHVC에서, 계층 간 텍스처 예측은 부호화 기본단위인 CU 레벨에서 수행하며, 기존 HEVC의 예측모드와 더불어 IntraBL 모드라는 새로운 예측모드를 추가하였다. 예측모드로 IntraBL 모드를 선택 할 경우, 업샘플링한 기본 계층의 블록은 상위 계층의 블록을 부호화하기 위한 예측으로 사용한다.

- INTRA_BL_DST4x4^{[12][13]}

IntraBL 모드로 예측한 4x4 luma TU(Transform Unit)에 Discrete Sine Transform을 적용하고, 8x8 이상의 luma TU에는 Discrete Cosine Transform을 적용한다.

- NO_RESIDUAL_FLAG_FOR_BLPRED^[14]

HEVC의 TU블록에 변환과 양자화가 수행되면, 많은 경우에 대해 계수가 0으로 나타나게 된다. 이러한 점을 토대로 분할된 TU에 대해 양자화 이후에 하나 이상의 변환 계수가 존재하는지의 여부를 알려주는 cbf(coded block flag) 플래그를 전송한다. SHVC에서도 계층 간 텍스처 예측으로 부호화된 TU에 대해 HEVC와 같은 방법을 적용한다.

- INTRA_BL_CTX_CHANGE^[15]

IntraBL 기반 프레임워크에서 계층 간 텍스처 예측 시에 IntraBL flag를 전송하여 계층 간 텍스처 예측을 알린다. 이 경우, HEVC에서 인접 영역의 참조 신택스 요소를 사용하는 CABAC의 확률정보는 메모리가 크게 증가한다는 단점이 있다. 게다가 인접 영역의 이용 가능성 검사를 위한 별도의 파싱 과정이 필요하다. 때문에 IntraBL 기반 프레임워크에서 계층 간 텍스처 예측 시에 CABAC 구문은 주변 지역 정보의 독립적으로 CU 크기에 의존적인 유도 방법을 사용한다. 더 정확하게 말하면, 현재 CU의 크기(ctDepth)는 문맥 인덱스(ctxIdxInc) 유도를 위해 사용된다. 표 2는 ctxIdxInc와 CtDepth사이의 관계를 나타낸다.

표 3. CU 크기에 의존적인 CABAC 문맥 초기화
 Table 3. CABAC context initialization in CU-size dependent

CtDepth	ctxIdxInc	Note (corresponding CU size)
0, 1(MaxCuDepth-1)	0	64x64/32x32
2(MaxCuDepth-1)	1	16x16
3(MaxCuDepth)	2	8x8

3.2 계층 간 움직임 예측(Inter-layer motion prediction)

HEVC는 움직임 벡터 부호화를 위해서 merge mode와 AMVP를 사용한다. Merge에서, PU의 움직임 벡터, 참조 영상 인덱스 등의 움직임 파라미터는 motion vector difference(mvd) 없이 주변 블록으로부터 상속한다. IntraBL 기반 프레임워크에서는 표 3과 같이 merge의 첫 번째 후보로 항상 계층 블록과 대응하는 기본 계층 블록의 움직임 벡터를 사용한다. 이는 주변 블록보다 하위 계층과의 움직임 유사성이 더 높기 때문이다.

표 4. IntraBL 기반 프레임워크에서 merge 후보 유도 방법
 Table 4. merge candidates derivation for IntraBL based framework

Merge candidate list
Base layer MV
Left
Above
Above right
Left bottom(if candNum<4)
Above left(if candNum<4)
TMVP
Combined candidates
Zero candidates

• IL_MRG_SIMPLIFIED_PRUNING^[16]

Merge 모드로 부호화시 각 후보들마다 처음의 후보와 중복되는지 여부를 확인한다. 하지만 이는 계산 복잡성이 늘어나는 단점이 있다. 이러한 복잡성을 줄이기 위해, 표 4와 같이 두 번째(A1), 세 번째(B1) 후보에 대해서는 첫 번째(colBase) 후보와 중복검사를 수행하고 하위 3개의 후보(B0, A0, B2)에 대해서는 중복검사를 수행하지 않는다.

표 5. merge모드 후보 결정시 간소화된 중복성 검사
 Table 5. simplified redundancy check for merge mode candidate

	SHVC Merge candidate construction
0	colBase
1	A1(Left) : if(A1! = colBase)
2	B1(Above) : if(B1!=A1)&&(B1!=colBase))
3	B0(Above right) : if(B0!=B1)
4	A0(Left bottom) : if(candNum<4) if(A0!=A1)
5	B2(Above left) : if(candNum<4) if(B2!=B1&&B2!=A1)

3.3 계층 간 잔차 예측(Inter-layer residual prediction)

GRP(Generalized Residual Prediction)^[17]는 기본 계층의 텍스처 정보를 사용하여 항상 계층의 잔차 신호를 예측하는 계층 간 예측 기술이다. 이전의 계층 간 예측은 predictor로 항상 계층의 블록에 대응하는 업샘플링한 기본 계층의 블록을 사용하지만, GRP를 이용하는 계층 간 예측은 항상 계층의 predictor로 잔차 신호를 추가함으로써 예측을 향상시킨다.

그림 7에서 현재 영상의 부호화하려는 블록(BEL)과 동일 계층에서 움직임 예측을 통해 찾은 예측 블록(PEL), 두 블록간의 움직임 벡터(MVEL)를 찾는다. 예측을 통해 찾은

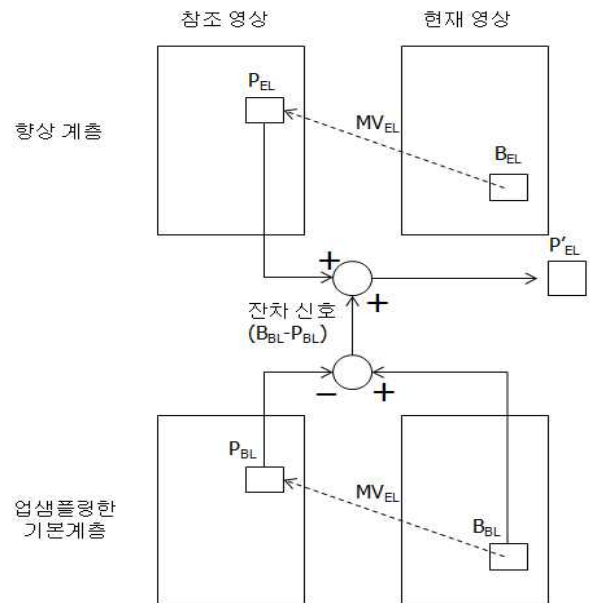


그림 7. Generalized Residual Prediction
 Fig. 7. Generalized Residual Prediction

MVEL을 업샘플링한 기본 계층에 적용하여, BEL과 대응하는 블록(BBL)으로부터 예측 블록(PBL)을 찾는다. BBL과 PBL을 이용하여 잔차 신호(BBL-PBL)를 생성하고, 이 잔차 신호에 가중치($\omega = 0, 0.5, 1$)를 항상 계층의 predictor에 추가하여 예측 신호(P'EL)를 생성한다.

계층 간 잔차 예측은 SHVC 표준에 포함되지 않았다. 예측의 효율보다는 복잡도 증가, 신택스 추가 등의 문제가 있었기 때문이다.

III. 실험 결과 및 분석

본 논문에서는 SHVC 성능 분석을 위해 SHVC 참조 소프트웨어인 SHM 6.0을 사용하였고, 표준에서 제외된 IntraBL 기반 프레임워크는 SHM 2.0을 사용하였다. simulcast의 성능은 HEVC Test Model (HM)14.1를 사용하여 기본 계층과 항상 계층을 각각 독립적으로 부호화하여 성능을 분석하였다. 성능 분석에 사용된 PC의 사양은 2.4 GHz zeon E5 2665, 128GB RAM이다. 테스트 영상으로는 SHVC Common Test Condition(CTC)^{[18][19]}에서 제공하는 클래스 A 중 traffic 영상과 클래스 B 영상 중 Basketball-Drive 영상을 사용하였다. 공간적 스케일러빌리티의 경우, 클래스 A의 실험에는 기본 계층에 1280x800 영상을, 항상 계층에 2560x1600 영상을 사용하여 실험하였고, 클래스 B의 실험에는 기본 계층에 960x540 영상을, 항상 계층에 1920x1080 영상을 사용하였다. QP(Quantization Parameter) 값은 기본 계층과 항상 계층 모두 CTC에서 제공하는 22, 26, 30, 34를 적용하였다. 화질적 스케일러빌리티의 경우, 클래스 A의 실험에는 기본 계층과 항상 계층에 2560x1600 영상을, 클래스 B의 실험에는 1920x1080 영상을 사용하였다. QP 값은 기본 계층은 26, 30, 34, 38을, 항상 계층은 기본 계층의 QP값보다 6 낮은 20, 24, 28, 32을 사용하였다. 부호화 구조는 Random-Access mode(RA), Low-Delay B mode(LD), All-Intra mode(AI)에 대해 실험하였다. 표 4는 설명한 실험 조건을 나타낸다.

실험은 Reference Index 기반과 IntraBL 기반 프레임워크로 구별하여 각 프레임워크의 성능과 각 프레임워크에 속하는 개별 톨의 성능을 평가한다. 추가로 simulcast와 프

레이믹워크들과의 성능을 비교함으로써 스케일러블 부호화 방법의 장점을 제시한다.

표 6. SHVC 실험 조건

Table 6. SHVC experimental conditions

scalability	테스트 영상		해상도		QP	
			기본 계층	항상 계층	기본 계층	항상 계층
2x spatial scalability	class A	Traffic	1280 x 800	2560 x 1600	22	22
	class B	Basketball Drive	960 x 540	1920 x 1080	26	26
SNR scalability	class A	Traffic	2560 x 1600		30	26
	class B	Basketball Drive	1920 x 1080		34	30
					38	34

1. Reference Index 기반 SHVC 성능 평가

가장 먼저 Reference Index 기반 프레임워크의 성능을 실험하였다. Reference Index 기반 프레임워크는 하위 계층의 복원 영상을 상위 계층의 참조 리스트에 삽입함으로써 IntraBL 기반 프레임워크의 방법에서 texture prediction처럼 하위 계층의 복원된 영상을 참조 할 수 있다. 이에 따라 Reference Index 기반 프레임워크에서의 성능 평가 실험은 하위 계층의 복원 영상 참조를 허용한 상태를 기준으로 나머지 개별 톨의 성능을 평가한다.

Reference Index 기반 프레임워크에 대한 실험이 표 6과 같다. 먼저, texture prediction과 관련 톨에 대한 성능을 실험하였다. Texture prediction의 관련 톨들을 모두 사용한 결과와 비교하여 AI 조건에서는 부호화 성능은 0.1% 향상되었고, 부호화 속도는 42.3%로 감소되었다. 이렇게 거의 동일한 부호화 성능으로 부호화 속도를 대폭 감소시킨 원인은 ENCODER_FAST_MODE의 영향이 때문이다. 그리고 RA 및 LD 조건에서도 REF_IDX_ZEROMV로 인해 부호화 성능 및 부호화 속도가 개선되었다. REF_IDX_ZEROMV은 기본 계층의 영상이 항상 계층의 참조영상으로 선택 하는 경우가 많이 발생하기 때문에 MVP의 시작점을 AMVP 과정으로 얻지 않고, 제로(Zero) 움직임 벡터로 고정시켜 부호화 속도의 향상이 이루어 졌음을 알 수 있다.

ENCODER_FAST_MODE는 AI의 경우 기본 계층이 I 슬라이스이고 향상 계층이 P 슬라이스로 구성되기 때문에 향상 계층 부호화 시에 모든 화면간 모드 결정을 비활성화한다. 그로 인해 화면간 모드 결정 과정에서 발생하는 많은 중복성이 제거됨으로써 부호화 속도 향상을 가져왔다.

다음으로 motion prediction에 대한 성능을 실험하였다. motion prediction은 MFMM을 사용하여 부호화하는 방식이므로 MFMM을 사용하는 것이 motion prediction을 사용하는 것과 같다. 화면내 부호화만을 수행하는 AI 부호화 구조를 제외한 모든 경우에서 부호화 성능이 향상되었다. 이는 상위 계층 블록의 움직임이 하위 계층의 대응 블록 움직임과 높은 상관성을 갖기 때문이다.

마지막으로 Reference Index 기반 프레임워크의 texture prediction 및 motion prediction 성능을 평가한다. 각 pre-

diction을 비롯한 prediction 관련 툴들을 사용하여 종합적으로 부호화 성능 증가를 초래한 것으로 보인다.

2. IntraBL 기반 SHVC 성능 평가

표준에서 제외된 IntraBL 기반 프레임워크의 실험을 위해 texture prediction과 motion prediction으로 나누어 실험하였다. 실험 결과는 표 7과 같다.

먼저, texture prediction과 관련 부호화 툴에 대한 실험을 하였다. 실험은 motion prediction과 관련 툴들을 사용하지 않고, texture prediction만을 사용한 채로 관련 툴을 모두 사용하지 않은 것을 비교 대상 기준으로 성능을 평가한다. INTRABL_DST_4X4의 성능은 4x4 luma TU에 대해서만 DST를 적용하기 때문에 x2 스케일러블리티의 AI에서 비

표 7. Reference Index 기반 프레임워크의 툴 성능 비교

Table 7. Performance Comparison for tool of Reference Index based Framework

Prediction		Sequence	2x			SNR	
			AI BDRate [%]	RA BDRate [%]	LD BDRate [%]	RA BDRate [%]	LD BDRate [%]
Texture Prediction	REF_IDX_ZEROMV + ENCODER_FAST_MODE	Traffic	-0.1	-0.3	0.0	-0.5	-0.1
		BasketballDrive	-0.1	-1.2	-0.4	-1.3	-0.4
		Enc. time	42.3	92	95.1	94.5	96.6
	REF_IDX_ZEROMV	Traffic	-0.5	-0.2	0.1	-0.3	0.0
		BasketballDrive	-0.3	-0.4	-0.2	-0.2	0.1
		Enc. time	90.1	90.4	92.9	93.3	95.3
	ENCODER_FAST_MODE	Traffic	0.7	0.3	0.1	-0.1	0.0
		BasketballDrive	1.6	-0.8	-0.2	-1.2	-0.4
		Enc. time	44.3	101	101.8	100.8	101.2
Motion Prediction	REF_IDX_MFM	Traffic	0.0	-0.3	-0.1	-1.4	-1.3
		BasketballDrive	0.0	-1.9	-0.6	-2.3	-1.2
		Enc. time	100.1	98.4	99	97.8	99.1
	POSITION_ROUNDING_OFFSET	Traffic	0.0	-0.9	-0.7	-1.4	-1.3
		BasketballDrive	0.0	-2.0	-0.7	-2.3	-1.2
		Enc. time	100	98.4	99.1	97.6	99
texture + motion prediction		Traffic	-0.1	-1.1	-0.7	-1.8	-1.4
		BasketballDrive	-0.1	-2.8	-0.9	-3.1	-1.4
		Enc. time	42.5	89.6	92.9	91.5	95.1

교적 높은 부호화 성능 향상이 있었다. NO_RESIDUAL_FLAG_FOR_BLPRED은 공간적 스케일러빌리티의 AI 구조에서 성능 향상이 있었으나, 화질적 스케일러빌리티의 경우 오히려 부호화 성능 감소가 있었다. 이는 cbf가 오히려 비트스트림 낭비가 됐다는 것을 의미한다. INTRABL_CTX_CHANGE의 성능은 class A에 속한 시퀀스들처럼 큰 영상에서는 부호화 성능이 증가하였지만, class B에 속한 비교적 작은 영상에서는 부호화 성능이 감소하였다. 이는 큰 영상에서 CU가 64x64 또는 32x32 크기의 경우가 많기 때문이다.

다음으로 motion prediction과 관련된 툴에 대한 성능을 실험하였다. 실험은 texture prediction의 부호화 툴을 모두 사용하지 않은 것을 비교 대상 기준으로 삼았다. motion prediction과 관련된 모든 툴들을 모두 사용한 경우는 전반

적으로 부호화 속도의 큰 차이 없이 부호화 성능이 감소하였다. 이는 IL_MRG_SIMPLIFIED_PRUNING의 merge모드 후보 결정시 첫 번째 후보와의 중복성 검사를 간소화한 것이 오히려 성능 저하를 초래한 것으로 보인다.

IntraBL 기반 프레임워크의 마지막 실험으로 texture prediction 및 motion prediction의 관련된 툴들을 사용하지 않은 것을 비교 대상으로 삼고, 모든 툴들을 사용한 것을 비교하였다. IntraBL 기반 프레임워크의 prediction 및 관련 툴들의 성능은 RA에서 평균 20%의 부호화 효율 증가를 보였다.

3. simulcast와 SHVC의 각 프레임워크 성능 평가

Simulcast와 각 프레임워크를 비교 실험하였다. 그림 8과

표 8. IntraBL 기반 프레임워크의 툴 성능 비교
Table 8. Performance Comparison for tool of IntraBL based Framework

Prediction		Sequence	2x			SNR	
			AI BDRate [%]	RA BDRate [%]	LD BDRate [%]	RA BDRate [%]	LD BDRate [%]
Texture Prediction	INTRABL_DST_4X4 + NO_RESIDUAL_FLAG_FOR_BLPRED + INTRABL_CTX_CHANGE	Traffic	-0.3	-0.1	-0.1	0.1	-0.3
		BasketballDrive	-0.5	0.0	-0.1	0.1	0.0
		Enc. time	101.6	99.8	99.6	98.6	99.5
	INTRABL_DST_4X4	Traffic	-0.2	-0.1	0.0	-0.3	-0.3
		BasketballDrive	-0.3	-0.1	-0.1	-0.2	-0.1
		Enc. time	100.1	100.0	100.0	99.4	99.5
	NO_RESIDUAL_FLAG_FOR_BLPRED	Traffic	-0.3	0.1	0.1	0.3	0.1
		BasketballDrive	-0.3	-0.1	-0.3	0.1	-0.1
		Enc. time	101.5	99.9	99.8	99.3	99.8
	INTRABL_CTX_CHANGE	Traffic	0.0	-0.1	0.0	0.0	-0.2
		BasketballDrive	0.1	0.2	0.2	0.1	0.1
		Enc. time	100.3	99.9	99.9	98	99.5
Motion Prediction	IL_MRG_SIMPLIFIED_PRUNING	Traffic	0.0	0.2	0.3	0.3	0.4
		BasketballDrive	0.0	0.2	0.1	0.2	0.2
		Enc. time	100.0	100.0	100.1	100.6	100.2
texture + motion prediction		Traffic	-33.8	-17.1	-0.5	-22.6	-8.2
		BasketballDrive	-20.3	-19.4	-13.8	-23.9	-17.3
		Enc. time	114.3	102	101	100.6	100.9

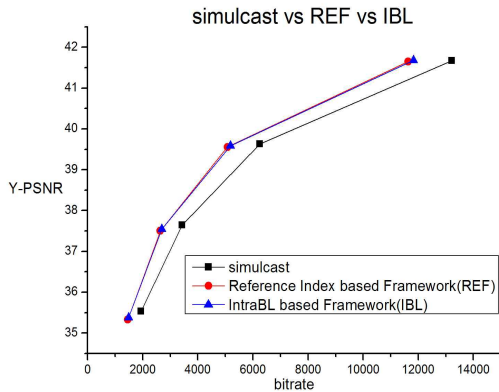


그림 8. simulcast와 SHVC 프레임워크 성능 비교 Traffic(class A)-x2-RA
 Fig. 8. Traffic(class A)-x2-RA, Performance comparison of simulcast and SHVC framework

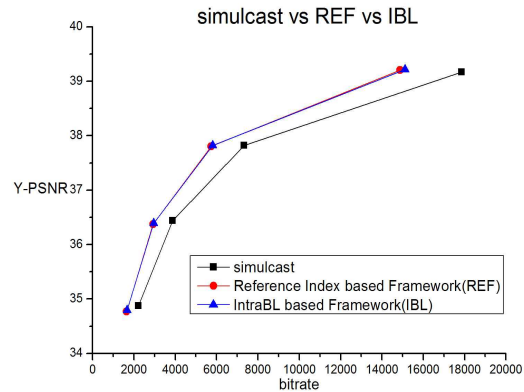


그림 9. simulcast와 SHVC 프레임워크 성능 비교 BasketballDrive(class B)-x2-RA
 Fig. 9. BasketballDrive(class B)-x2-RA, Performance comparison of simulcast and SHVC framework

그림 9에서와 같이 Reference Index 기반 프레임워크와 IntraBL 기반 프레임워크의 성능은 비슷하였고, SHVC의 성능은 simulcast보다 약 20% 정도 향상되었음을 알 수 있다. SHVC 표준화는 HEVC의 구조를 토대로 high level syntax만 변경하는 Reference Index 기반 프레임워크를 채택하여 HEVC 프로세스의 변화를 최소화하고, 핵심 구조를 재사용하여 계층 간 예측이 가능하게 구조를 설계하는 것을 목표로 하였다. 이는 3D-HEVC와의 호환도 쉽게 가능하다.

V. 결론

본 논문에서는 SHVC의 주요 기술들을 살펴보고 이에 따른 성능 평가를 제시하였다. 성능 평가에 대한 의미는 현재 SHVC에서 사용되는 각각의 기술들이 기존의 방법보다 어떠한 점에서 효율적인지 보이고자 하였다. 이러한 분석은 향후 부호화 성능 향상을 위한 알고리즘 개발뿐만 아니라 고속 부호화기 개발을 위한 알고리즘 개발에도 유용할 수 있을 것으로 판단된다. SHVC는 다양한 멀티미디어 소비자 기기와 이종 네트워크 등의 혼재된 환경에 적응적으로 대처할 수 있게 해주어 더 좋은 품질의 비디오 서비스를 제공할

수 있다. 따라서 스케일러블 비디오 부호화 방식은 비용과 효율의 생산성 측면에서 매우 매력적인 기술이 될 것이다.

참고 문헌 (References)

- [1] Haechul Choi, Kyung Il Lee, Jung Won Kang, Seong-Jun Bae, and Jeong-Ju Yoo, "Overview and Performance Analysis of the Emerging Scalable Video Coding", JBE Vol.12, no. 6, November 2007.
- [2] J. Chen, J. Boyce, Y. Ye, M. M. Hannuksela, G. J. Sullivan and Y.-K. Wang, "High efficiency video coding (HEVC) scalable extension Draft 6", Q-1008_v5, 17th JCT-VC meeting, Valencia, Mar. 2014.
- [3] Jianle Chen, Jill Boyce, Yan Ye, Miska M. Hannuksela, "SHVC Test Model 2 (SHM 2)", Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, 13th Meeting, JCTVC-M1007, Incheon, KR, 18-26 Apr. 2013.
- [4] Hyomin Choi, Junghak Nam, Donggyu Sim, and Ivan V. Bajic, "Scalable Video Coding Based on High Efficiency Video Coding (HEVC)", Communications, Computers and Signal Processing (PacRim), 2011 IEEE Pacific Rim Conference on.
- [5] Xiaoyu Xiu, Yuwen He, Yong He, Yan Ye, "TE A2: Inter-layer reference picture placement", JCTVC-L0051, Geneva, CH, 14-23 Jan. 2013
- [6] Peng Yin, Taoran Lu and Tao Chen, "Non-TE2: Inter-layer reference picture placement", JCTVC-L0174, Geneva, CH, 14-23 Jan. 2013
- [7] Hahun Lee, Jung Won Kang, Jinho Lee, Jin Soo Choi, "MV-HEVC/SHVC HLS: Signalling for Inter-Layer prediction indication", JCTVC-N0120, Vienna, AT, 25 July - 2 Aug. 2013
- [8] Krishna Rapaka, Ye-kui Wang, Jianle Chen, Vadim Seregin, Hendry, Marta Karczewicz, "MV-HEVC/SHVC HLS: On inter-layer RPS deri-

- vation and sub-layer inter-layer dependency”, JCTVC-O0225, Geneva, CH 23 Oct. - 1 Nov. 2013
- [9] Edouard Francois, ShaSha Shi, Christophe Gisquet, Guillaume Laroche, Patrice Onno, “Non-SCE1: simplification of intra mode coding in SHVC”, JCTVC-M0115, Incheon, KR, 18-26 Apr. 2013
- [10] Jianle Chen, Vadim Seregin, Liwei Guo, Marta Karczewicz, “Non-TE5: on motion mapping in SHVC”, JCTVC-L0336, Geneva, CH, 14-23 Jan. 2013
- [11] Christophe Gisquet, Patrice Onno, Edouard Francois, Guillaume Laroche, “SCE2: Results on test 2.1”, JCTVC-N0139, Vienna, AT, 25 July - 2 Aug. 2013
- [12] Chuohao Yeo, Yih Han Tan, “On transform selection for IntraBL mode in SHVC”, JCTVC-L0067, Geneva, CH, 14-23 Jan. 2013
- [13] Vadim Seregin, Tzu-Der Chuang, Do-Kyoung Kwon, Fabrice Le Leannec, Patrice Onno, “AHG11: DST transform for Intra-BL mode”, JCTVC-L0204, Geneva, CH, 14-23, Jan. 2013
- [14] Tzu-Der Chuang, Yu-Wen Huang, Shawmin Lei, Chulkeun Kim, Joonyoung Park, Byeongmoon Jeon, “TE2-3.1.7: Inter-layer intra prediction signaling with residual skip flag”, JCTVC-L0069, Geneva, CH, pp. 14-23, Jan. 2013
- [15] Tomoyuki Yamamoto, “On TextureRL flag context”, JCTVC-M0075, Incheon, KR, 18-26 Apr. 2013
- [16] Tzu-Der Chuang, Yu-Wen Huang, Shawmin Lei, “SCE5.1.5: Inter-Layer motion information prediction”, JCTVC-M0124, Incheon, KR, 18-26 Apr. 2013
- [17] Jianle Chen, Krishna Rapaka, Xiang Li, Vadim Seregin, Liwei Guo, Marta Karczewicz, Geert Van der Auwera, Joel Sole, Xianglin Wang, Chengjie Tu, Ying Chen, “Description of scalable cideo coding technology proposal by Qualcomm”, JCTVC-K0035, Shanghai, CN, 10-19 Oct. 2012
- [18] Xiang Li, Jill Boyce, Patrice Onno, Yan Ye, “Common SHM test conditions and software reference configurations”, JCTVC-M1009, Incheon, KR, 18-26 Apr. 2013
- [19] Vadim, Yong He, “Common SHM test conditions and software reference configurations”, JCTVC-Q1009, Valencia, Es, 27 March - 4 April 2014

저 자 소 개

김 용 태



- 1998년 2월 : 경북대학교 전자공학과 학사
- 2000년 2월 : 경북대학교 전자공학과 (공학석사)
- 1999년 12월 ~ 2000년 6월 : 하이닉스 3G 비동기시스템팀 대리
- 2000년 8월 ~ 2010년 2월 : 한국전자통신연구원(ETRI) 네트워크연구소 선임연구원
- 2010년 4월 ~ 2012년 1월 : ㈜이프리시스템 친환경에너지연구소 연구소장
- 2012년 11월 ~ 현재 : 탑알앤디㈜ 대표이사
- ORCID: <http://orcid.org/0000-0001-8847-3061>
- 주관심분야 : 영상통신, 비디오 부호화, 시스템 설계

최 진 혁



- 2012년 2월 한밭대학교 멀티미디어공학과 (학사)
- 2013년 3월 ~ 2015년 2월 한밭대학교 멀티미디어공학과 (공학석사)
- ORCID: <http://orcid.org/0000-0002-0221-8092>
- 주관심분야 : 비디오 부호화, 영상처리, 패턴 인식

최 해 철



- 1997년 2월 : 경북대학교 전자공학과 학사
- 1999년 2월 : 한국과학기술원 전기및전자공학과 (공학석사)
- 2004년 8월 : 한국과학기술원 전기및전자공학과 (공학박사)
- 2004년 9월 ~ 2010년 2월 : 한국전자통신연구원(ETRI) 방송미디어연구부 선임연구원
- 2010년 3월 ~ 현재 : 한밭대학교 정보통신공학과 부교수
- ORCID: <http://orcid.org/0000-0002-7594-0828>
- 주관심분야 : 영상통신, 비디오 부호화, 패턴 인식