특집논문 (Special Paper)
방송공학회논문지 제25권 제3호, 2020년 5월 (JBE Vol. 25, No. 3, May 2020)
https://doi.org/10.5909/JBE.2020.25.3.313
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

VVC에서 360 비디오를 위한 랩-어라운드 움직임 벡터 예측 방법

이 민 훈[a)], 이 종 석[b)], 박 준 택[b)], 임 웅[c)], 방 건[c)], 심 동 규[b)], 오 승 준[a)‡]

Wrap-around Motion Vector Prediction for 360 Video Streams in Versatile Video Coding

Minhun Lee[a)], Jongseok Lee[b)], Juntaek Park[b)], Woong Lim[c)], Gun Bang[c)], Dong Gyu Sim[b)], and Seoung-Jun Oh[a)‡]

요 약

본 논문에서는 360 비디오의 특성을 이용하여 픽쳐 경계에서 코딩 효율을 증가시키는 움직임 벡터 예측 방법을 제안한다. VVC의 화면간 예측에서 움직임 벡터 후보군을 구성할 때 주변 블록의 위치가 픽쳐의 경계 바깥이면 후보군 구성 과정에서 제외되어 픽쳐 경계에서 코딩 효율이 감소하게 된다. 이를 해결하기 위해 360 비디오의 부/복호화를 위한 투영 방법의 특성을 이용하여 이미 복호화된 정보로부터 픽쳐의 경계에서 추가로 후보군을 구성하는 방법을 제안한다. 제안하는 방법의 성능 평가를 위하여 JVET-360 CTC의 임의 접근 환경에서 VTM6.0 및 360Lib9.1와 비교한다. 그 결과로써 제안하는 방법은 추가 계산 복잡도 없이 휘도 성분에서 평균 0.02%, 색차 성분에서 각각 평균 0.05%, 0.06%의 비트율 감소를 보이고, 픽쳐의 경계에서는 비트율이 휘도 성분에서 평균 0.29%, 색차 성분에서 각각 평균 0.45%, 0.43% 감소하였다. 그리고 DSCQS (Double Stimulus Continuous Quality-Scale) 방법을 통해 제안하는 방법의 주관적 화질 평가를 수행하여 MOS (Mean Opinion Score)값을 얻는다. MOS값은 평균 0.03 향상되었고, MOS값과 비트량을 이용하여 BD-MOS를 구한다. 그 결과, 제안하는 방법의 성능이 최대 8.78%, 평균 5.18% 향상하였다.

Abstract

In this paper, we propose a motion vector prediction method that increases the coding efficiency at the boundary of an image by utilizing the 360 video characteristic. In the current VVC method, the location of a neighbor block is excluded from the candidate list for inter prediction in case that it is out of boundary. This can reduce coding efficiency as well as subject quality. To solve this problem, we construct new candidates adding the location of the neighbor block at the boundary of the picture from already decoded information based on the projection method for 360 video coding. To evaluate the performance of the proposed method, we compare with VTM6.0 and 360Lib9.1 under Random Access condition of JVET-360 CTC. As a result, the coding performance shows a BD-rate reduction of 0.02% on average in luma component and 0.05%, 0.06% on average in chroma components respectively, without additional computational complexity. The coding performance at the picture boundary shows a BD-rate reduction of 0.29% on average in luma component and 0.45%, 0.43% on average in chroma components, respectively. Furthermore, we perform subjective quality test with the DSCQS method and obtain MOS values. The MOS value is improved by 0.03 value, and we calculate BD-MOS using MOS value and bit-rate. As a result, the proposed method improved performance by up to 8.78% and 5.18% on average.

Keyword : VVC, 360 Video, Wrap-around, Inter prediction, Motion vector candidate list

Copyright Ⓒ 2020 Korean Institute of Broadcast and Media Engineers. All rights reserved.
“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (http://creativecommons.org/licenses/by-nc-nd/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

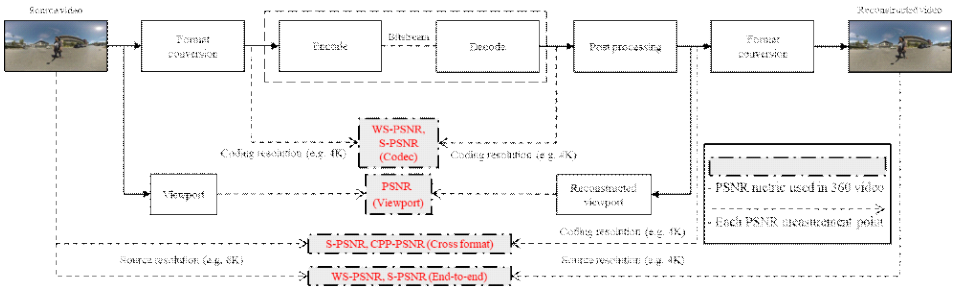

그림 1. 360 비디오의 부호화와 복호화를 위한 전체 흐름도 및 객관적 성능 평가를 위한 평가방법에 대한 개념도
Fig. 1. The overall flow chart for encode and decode and conceptual diagram for objective performance evaluation of 360 video

Ⅰ. 서 론

최근 들어 게임, 영화, 의료 서비스, 소셜 미디어 등 여러 분야에서 가상현실과 증강현실 응용 서비스가 대중화됨에 따라 360 비디오 기술에 대한 관심이 증가하고 있다[1]. 이에 따라 ITU-T VCEG (International Telecommunication Union Telecommunication Standardization Sector Video Coding Experts Group)과 ISO/IEC MPEG (International Organization for Standardization/International Electrotechnical Commission Moving Picture Experts Group)으로 구성된 JVET (Joint Video Experts Team)에서는 360 비디오와 HDR (High Dynamic Range) 비디오 등을 고려한 VVC (Versatile Video Coding)[2] 표준화 활동을 수행하고 있으며 2020년에 표준화가 완료될 예정이다. VVC 표준 기술의 압축 효율은 HEVC (High Efficiency Video Coding) 대비 약 2배 정도로 예상된다[3]. VVC 표준화를 위한 공통 실험을 위하여 참조 소프트웨어인 VTM (VVC Test Model)[4]을 구현하면서 360 비디오의 부/복호화를 위해 3차원에서 2차원 (3D-to-2D)으로 투영하는 (Projection) 다양한 방법들과 360 비디오의 객관적인 화질 평가를 위한 각 투영 방법들을 고려한 WS-PSNR[5], S-PSNR[6], CPP-PSNR[7] 등이 포함된 소프트웨어인 360Lib[8]도 함께 개발하고 있다. JVET에서 정한 360 비디오를 부/복호화하기 위한 과정은 그림 1과 같다.

예를 들어, 먼저 8K ERP (Equi-Rectangular Projection) 영상이 초기 입력으로 사용되고 다운 샘플링 (Down-sampling) 과정과 함께 사용하고자 하는 투영 방법으로 변환한다. 이렇게 변환된 특정 투영의 4K 2D 영상이 부호화기의 입력 값이 된다. 그리고 부/복호화 과정 후에 나온 특정 투영의 4K 2D 영상은 바로 뷰포트 렌더러 (Viewport renderer)를 통해 3D 영상으로 변환되거나 혹은 원본과 같은 8K ERP 영상으로 업 샘플링 (Up-sampling)되어 출력되기도 한다. 360 비디오 영상은 전방위에서 동시에 영상을 취득하여 생성한 하나의 영상이고 본질적으로는 영상의 경계가 존재하지 않는다. 영상의 부/복호화를 위해 3차원에서 2차원으로 변환시키는 투영 과정에 의해 영상의 경계가 생기게 되고 이는 서로 연결되어 있다. 또한, 이 과정에서 화면 내 비활성 영역과 불연속성이 존재할 수 있기 때문에 코딩 효율이 감소하고 열화 현상이 발생할 수 있다. 이러한 문제를 보완하고자 투영 방법의 특성을 이용하는 많은 연구가 진행되고 있다[9],[10]. 현재 JVET에서는 투영 방법들로 ERP, PERP (Padded ERP), EAP (Equal-Area Projection), CMP (CubeMap Projection) 등 다양한 방법들이 채택되어 360Lib에 반영되어 있다. 여러 투영 방법 중 ERP는 가장

a) 광운대학교 전자공학과(Dept. of Electronic Engineering, Kwangwoon University)
b) 광운대학교 컴퓨터공학과(Dept. of Computer Engineering, Kwangwoon University)
c) 한국전자통신연구원(Electronics and Telecommunications Research Institute)
‡ Corresponding Author : 오승준 (Seoung-Jun Oh)
E-mail: sjoh@kw.ac.kr
Tel: +82-2-940-5102
ORCID: https://orcid.org/0000-0002-5036-376
※ 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 정보통신·방송 연구개발 사업의 일환 [출연처과제번호 2018-0-00207, 이머시브 미디어 전문연구실 (전문연구실)] 및 2019년도 광운대학교 교내 학술연구비 지원에 의해 연구되었음.
※ This work was supported by the ICT R&D program of MSIT/IITP [2018-0-00207, Immersive Media Research Laboratory] and the Research Grant of Kwangwoon University in 2019.
· Manuscript received March 23, 2020; Revised April 28, 2020; Accepted May 7, 2020.

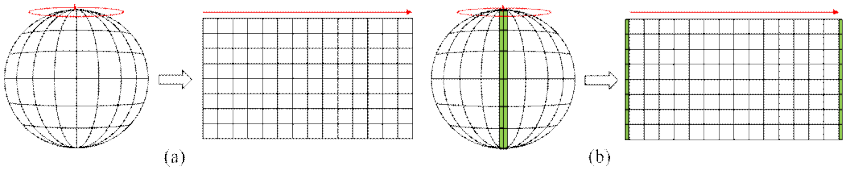

그림 2. 360 비디오의 투영 방법 (a) ERP; (b) PERP
Fig. 2. Projection method of 360 video (a) ERP; (b) PERP

많이 쓰이는 방법이다. PERP는 ERP로 투영된 2D 영상의 왼쪽과 오른쪽 경계에 패딩 (Padding) 과정을 거쳐 ERP를 적용한 2D 영상보다 픽쳐 (Picture)의 폭이 더 넓은 특성을 가지고 있다. ERP와 PERP를 이용하여 360 비디오를 2D영상으로 변환하는 과정은 그림 2와 같다. 여러 투영 방법 중 상기한 두 방법을 이용하여 360 비디오를 투영시키면 생성된 2D 영상에서 픽쳐의 왼쪽과 오른쪽 경계가 연결되는 특성을 가지게 되는 것을 그림 2에서 확인할 수 있다. 따라서 다른 투영 방법을 이용하여 얻은 2D 영상이나 일반 2D 영상보다 픽쳐 경계의 픽셀 (Pixel)간 유사도가 높다.

JVET에서는 상기한 특성을 반영한 기술로 랩-어라운드 (Wrap-around) 움직임 보상 기술이 채택되어 VTM6.0에 반영되어 있다[11]. 기존의 움직임 보상 과정에서 움직임 벡터가 가리키는 참조 픽쳐의 위치가 왼쪽 또는 오른쪽 경계 너머인 경우 움직임 벡터를 픽쳐의 경계로 클리핑 (Clipping)하여 해당 위치의 블록을 사용했지만, 이때 클리핑이 아닌 경계 너머의 부분은 수평 방향으로 미리 정의된 옵셋 (Offset)을 가감한 위치에서 가져와 사용하는 기술이다. 그리고 픽쳐의 왼쪽과 오른쪽이 연결되어 있으면 픽쳐의 경계에 픽셀의 유사하다는 특성뿐만 아니라 그림 3과 같이 움직임이 유사하다는 특성도 가지고 있다.

그림 3. ERP 또는 PERP로 변환한 360 비디오의 움직임의 유사도
Fig. 3. The similarity of 360 video's motion which is converted by ERP or PERP method

그리고 VVC의 화면간 예측은 현재 부/복호화하려는 픽쳐 내의 코딩 블록에 대한 예측값을 이전 시간에 이미 복호화된 픽쳐로부터 예측 블록을 생성하여 이를 통해 예측을 수행하는 방법이다. 이때 이미 복호화된 픽쳐와 현재 블록 간의 위치 관계를 표현한 것이 움직임 벡터이고 VVC에서는 이러한 움직임 벡터를 보다 효율적으로 코딩하기 위하여 현재 블록의 움직임 벡터를 이미 사용된 주변 블록들의 움직임 벡터 정보들로 후보군을 구성하고 최적의 움직임 벡터의 정보를 복호화기로 전송하고 있다. 이러한 과정에서 일반적인 2D 영상의 경우는 픽쳐의 경계에서 픽쳐 경계를 벗어나는 주변 위치 블록의 움직임 벡터를 사용하지 못한다. 따라서 픽쳐 경계에서 코딩 효율이 감소하는 문제가 발생한다.

이러한 문제점을 해결하기 위하여 본 논문에서는 360 비디오를 부/복호화할 때 사용되는 투영 방법 중 ERP와 PERP를 이용하여 변환한 360 비디오의 픽쳐의 왼쪽과 오른쪽 움직임이 유사한 특성을 이용한 랩-어라운드 움직임 벡터 예측 방법을 제안한다. 제안하는 방법은 기존 방법으로는 움직임 벡터 후보군 구성 과정에서 사용할 수 없던 주변 블록의 위치에 대해 반대편 위치의 블록을 이용하여 유의미한 후보군 탐색 과정을 수행할 수 있게 하여 픽쳐의 경계에서 예측 성능을 높이고자 하는 기술이다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 화면간 예측 기술과 제안하는 랩-어라운드 움직임 벡터 예측 기술에 관해 설명한다. 그리고 3장에서는 제안하는 방법의 성능을 객관적, 주관적 측면에서 평가하고 4장에서 결론을 맺는다.

Ⅱ. 랩-어라운드 움직임 벡터 예측 방법

VVC의 화면간 예측은 크게 주변 블록과 현재 블록 간의 움직임 정보의 상관도를 이용하는 MERGE/SKIP 모드와 AMVP 모드가 있다. 두 모드에는 전송되는 비트량을 최소

화하면서 예측의 정확도를 높이기 위해 다양한 기술들이 채택되어 있고 현재 VTM6.0에 반영되어 있다. VVC에서는 크게 각각 일반 MERGE/SKIP 모드 (Regular MERGE/SKIP mode, 이하 R-MERGE/SKIP)와 일반 AMVP 모드 (Regular AMVP mode, 이하 R-AMVP), 서브블록 기반 MERGE/SKIP 모드 (Sub-block based MERGE/SKIP mode, 이하 Sb-MERGE/SKIP)와 서브블록 기반 AMVP 모드 (Sub-block based AMVP mode, 이하 Sb-AMVP)로 나뉜다. 이번 장에서는 기존 화면간 예측 기술들에 더하여, 본 논문에서 제안하는 방법에 관해 서술한다.

1. R-MERGE/SKIP과 R-AMVP

R-MERGE/SKIP과 R-AMVP에서 현재 CU (Coding Unit)의 움직임 벡터를 예측할 때 그림 4와 같이 현재 CU의 공간적, 시간적 주변 블록들을 이용하여 움직임 벡터 후보군을 구성하고 그중 최적의 움직임 벡터를 찾는 과정을 통해 예측을 수행한다.

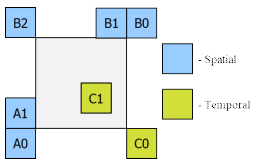

그림 4. 현재 CU의 공간적, 시간적 주변 블록의 위치
Fig. 4. The location of spatial and temporal neighbor blocks of current CU

현재 VVC에는 여기에 추가로 HMVP (History-based Motion Vector Prediction)와 PMVP (Pairwise Motion Vector Prediction)이 채택되어 VTM6.0에 포함되어 있다[11]. HMVP는 R-MERGE/SKIP과 R-AMVP에 모두 포함되어 있으며 부호화 순서상 현재 CU보다 이전 블록들을 예측한 움직임 벡터의 정보들로 테이블을 구성하여 후보군을 구성하는 방법이고, PMVP는 R-MERGE/SKIP에만 적용되는 기술로 이미 채워진 움직임 벡터 후보군 중 첫 번째와 두 번째 후보를 이용하여 후보군을 구성하는 방법이다. R-MERGE/SKIP과 R-AMVP에서는 이 과정을 기반으로 현재 CU의 움직임 벡터를 예측한다.

이러한 움직임 벡터 후보군 구성 과정에서 현재 CU가 픽쳐의 경계에 위치하는 경우 픽쳐의 왼쪽 경계에서 B2, A1, A0위치와 픽쳐의 오른쪽 경계에서 B0, C0위치는 후보군 구성 과정에서 사용할 수 없다. 그렇게 되면 최적의 움직임 벡터를 결정할 수 있는 후보 수가 다른 CU 대비 적기 때문에 예측 효율이 감소하는 문제가 발생한다. 따라서 본 논문에서는 이러한 점을 보완하기 위해 픽쳐의 경계에서 더욱 충분한 후보군을 구성하여 코딩 효율을 높이는 방법을 제안한다. 제안하는 방법의 R-MERGE/SKIP 후보군 구성 과정과 R-AMVP 후보군 구성 과정의 순서도는 그림 5와 6과 같다. ERP와 PERP를 이용하여 변환한 360 비디오는 픽쳐의 연결성 때문에 CU의 위치가 픽쳐의 왼쪽 경계일 때 B2위치에 수평 방향으로 적합한 옵셋 값을 더한 위치의 블록은 B2와 동일 위치의 블록이고 움직임 정보 또한 동일하다. 제안하는 방법은 해당 블록을 움직임 벡터 후보군 구성 과정에 사용하는 방법이다. 해당 블록은 픽쳐 내부에 존재하고 이미 복호화되어 있다면 해당 블록의 움직임 벡터

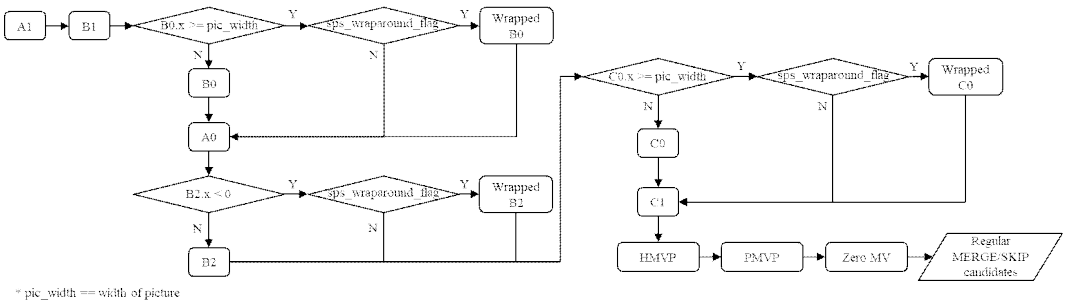

그림 5. 제안하는 방법의 R-MERGE/SKIP 후보군 구성 과정의 순서도
Fig. 5. Flow chart of R-MERGE/SKIP candidate list consisting process for proposed method

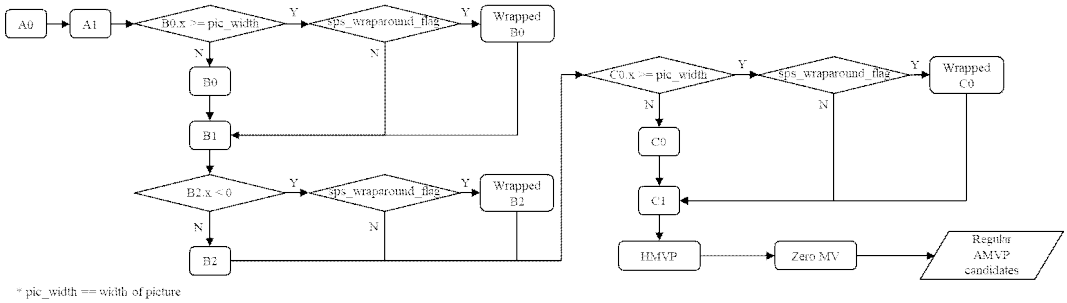

그림 6. 제안하는 방법의 R-AMVP 후보군 구성 과정의 순서도
Fig. 6. Flow chart of R-AMVP candidate list consisting process for proposed method

를 후보군 구성 과정에 사용하여 현재 CU가 픽처의 경계에 있을 때도 기존 방법보다 충분한 수의 후보들을 이용하여 예측의 효율을 높일 수 있다. 이때 A0, A1위치에서 옵셋을 더한 위치에 있는 블록은 어떠한 옵셋 값에 대해서도 아직 복호화되지 않았기 때문에 사용할 수 없다. 같은 방식으로 CU의 위치가 픽처의 오른쪽 경계일 때 B0, C0위치에 수평 방향으로 적합한 옵셋을 빼준 위치의 블록을 사용하여 각각 후보군 구성 과정에 사용할 수 있고 이에 따라 예측의 효율을 높일 수 있다.

2. Sb-MERGE/SKIP과 Sb-AMVP

VVC의 Sb-MERGE/SKIP과 Sb-AMVP의 기술로는 현재 어파인 (Affine) 움직임 예측 방법과 SbTMVP (Sub-block Temporal Motion Vector Prediction) 모드가 채택되어 VTM6.0에 포함되어 있다[11]. 어파인 움직임 예측 방법은 현재 CU의 좌측 상단, 좌측 하단, 우측 상단, 우측 하단의 CPMV (Control Point Motion Vector)를 기반으로 어파인 움직임 모델을 표현한 것으로, 두 개의 CPMV를 사용하는 4-파라미터 어파인 움직임 모델 (4-parameter affine motion model)과 세 개의 CPMV를 사용하는 6-파라미터 어파인 움직임 모델을 사용하여 현재 CU를 서브블록 단위로 예측을 수행하는 방법이다.

이때 CPMV를 구하는 과정에서 그림 7과 같이 현재 CU의 공간적, 시간적 주변 블록을 이용한다. 어파인 MERGE/SKIP 모드에서는 상속 (Inherited) 후보군 구성과 조합 (Constructed) 후보군 구성을 통해 어파인 MERGE/SKIP 후보군을 구성하고, 어파인 AMVP 모드에서는 상속 후보

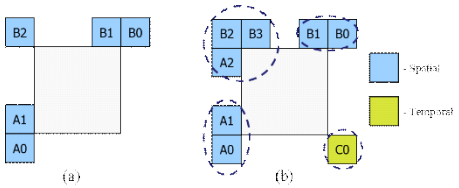

그림 7. (a) 상속 어파인 움직임 예측자의 위치 (b) 조합 어파인 모드를 위한 각 CPMV 후보의 위치
Fig. 7. (a) Locations of inherited affine motion predictors (b) Locations of each CPMV candidate for constructed affine mode

군 구성, 조합 후보군 구성, 평행이동 (Translational) 후보군 구성을 통해 어파인 AMVP 후보군을 구성한다. 상속 후보군 구성 과정에서 사용되는 주변 블록은 그림 7 (a)와 같고 주변 블록이 어파인 모드로 예측되었다면 해당 CU의 CPMV로부터 현재 CU의 CPMV를 추론 (Extrapolated)하여 후보군에 추가한다. 조합 후보군 구성 과정에서 사용되는 주변 블록은 그림 7 (b)와 같고 현재 CU의 주변 블록의 움직임 벡터 정보를 이용하여 현재 CU의 각 CPMV를 결정한 후 표 1과 같이 미리 정의된 2개 또는 3개의 CPMV의 조합들을 통하여 순서대로 후보군에 추가하는 방법이다.

표 1. 조합 어파인 모드에 사용되는 CPMV들의 조합
Table 1. Combinations of CPMVs are used to constructed affine mode

Combinations of CPMVs are used to constructed affine mode
1. {CPMV0, CPMV1, CPMV2}
2. {CPMV0, CPMV1, CPMV3}
3. {CPMV0, CPMV2, CPMV3}
4. {CPMV1, CPMV2, CPMV3}
5. {CPMV0, CPMV1}
6. {CPMV0, CPMV2}

평행이동 후보군 구성 방법은 조합 후보군 구성 과정에서 구한 CPMV들을 각각 현재 CU의 모든 CPMV로 정하여 후보군에 추가하는 방법이다. 어파인 움직임 예측 방법도 현재 CU의 주변 블록의 정보를 가지고 예측을 수행하는 기술이기 때문에 마찬가지로 현재 CU가 픽쳐의 경계에 위치한 경우 기존 방법으로는 사용할 수 없던 픽쳐 경계 바깥의 주변 블록들에 대하여 제안하는 방법을 적용할 수 있다. 제안하는 방법을 적용하면 현재 CU의 위치가 픽쳐의 왼쪽 경계일 때 좌측 상단, 현재 CU의 위치가 픽쳐의 오른쪽 경계일 때 우측 상단과 우측 하단에 위치하는 주변 블록에 대해 적합한 옵셋 값을 가감한 위치에 있는 블록을 상속 후보군 구성, 조합 후보군 구성, 평행이동 후보군 구성 과정에 적용할 수 있고 이에 따라 예측의 효율을 높일 수 있다. 제안하는 방법의 어파인 모드 후보군 구성 과정의 순서도는 그림 8과 같다.

SbTMVP 모드는 Sb-MERGE/SKIP에만 적용되는 기술로써, 현재 CU의 대응블록을 찾는 과정과 대응블록의 움직임 벡터로부터 현재 CU의 움직임 벡터를 서브블록 단위로 예측을 수행하는 과정을 통하여 예측을 수행한다.

현재 CU의 좌측 (그림 4의 A1위치)의 참조픽쳐가 대응참조픽쳐와 동일한 경우 좌측 블록의 움직임 벡터를 이용하고 그렇지 않은 경우 제로 움직임 벡터를 이용하여 찾은 대응참조픽쳐 내의 대응블록으로부터 8x8의 서브블록 단위로 현재 CU의 움직임 벡터 예측을 수행한다. 이때 사용할 대응블록의 위치가 픽쳐의 경계를 넘어가면 픽쳐의 경계로 클리핑하여 해당 위치의 블록을 대응블록으로 사용하여 예측을 수행하게 된다. 그렇게 되면 픽쳐의 경계에서 벗어나는 정도와 상관없이 모두 클리핑하여 사용하기 때문에

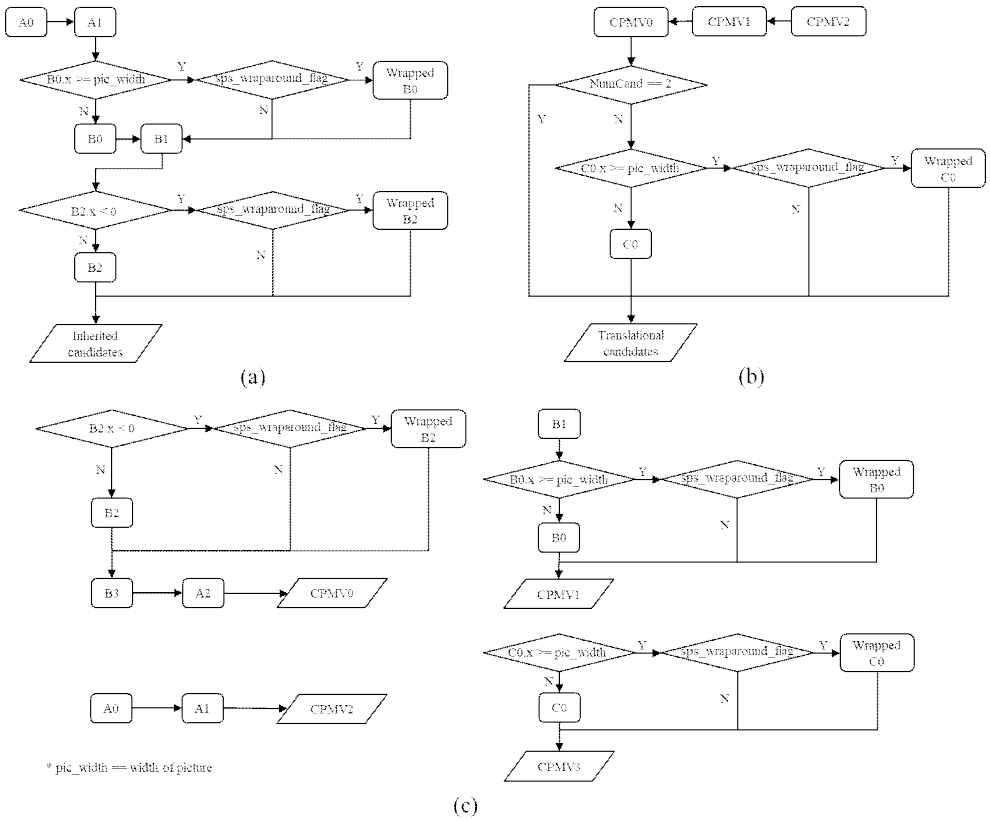

그림 8. 제안하는 방법의 어파인 모드 후보군 구성 과정의 순서도 (a) 상속 후보군; (b) 평행이동 후보군; (c) 조합 후보에 사용되는 각 CPMV
Fig. 8. Flow chart of affine mode candidate list consisting process for proposed method (a) Inherited candidate list; (b) Translational candidate list; (c) Each CPMV used in constructed candidate

예를 들어, 한 픽쳐 내의 서로 다른 두 CU를 각각 SbTMVP 모드로 예측을 수행할 때 픽쳐의 경계를 벗어난 두 개의 블록이 수평 또는 수직으로 벗어난 정도가 같은 경우 서로 다른 위치의 대응블록을 사용하려 했으나 클리핑 과정을 통해 서로 같은 대응블록을 사용하는 문제가 발생하게 된다. 제안하는 방법은 이 과정에서 클리핑이 아닌 대응블록의 위치가 픽쳐의 왼쪽 경계 바깥인 경우 옵셋을 더한 위치의 블록을 대응블록으로, 대응블록의 위치가 픽쳐 오른쪽 경계 바깥인 경우 옵셋을 빼준 위치의 블록을 대응블록으로 사용하여 원래 가져오려 했던 위치의 블록을 이용함으로써 예측의 효율을 높이는 방법이다. 제안하는 방법을 적용한 SbTMVP 모드 예측 과정의 순서도는 그림 9와 같다.

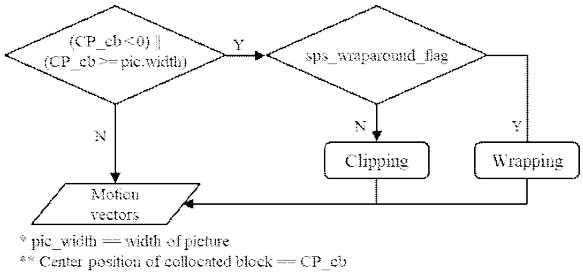

그림 9. 제안하는 방법의 SbTMVP 모드 예측 과정의 순서도
Fig. 9. Flow chart of SbTMVP mode prediction process for proposed method

제안하는 방법에서 사용하는 옵셋 값은 랩-어라운드 움직임 보상에서 사용되는 옵셋 값과 동일하고, 제안하는 방법의 사용 여부를 판별하는 플래그 (Flag)는 랩-어라운드 움직임 보상 기술의 사용 여부 플래그와 동일하게 사용한다. 사용되는 투영 방법이 ERP인 경우, 픽쳐의 폭을 옵셋 값으로 사용하고 PERP의 경우도 마찬가지이다. PERP는 ERP에서 픽쳐의 왼쪽과 오른쪽 경계부분에 기하학 패딩 (Geometry padding)을 사용하기 때문에 옵셋 값으로 픽쳐의 폭이 아닌 ERP를 사용하여 얻은 2D 영상의 픽쳐 폭을 사용한다. 또한, 제안하는 방법의 개념은 어떠한 투영 방법을 이용하더라도 영상의 경계가 연결되어 있다는 특성에 의해 모든 360 비디오 영상에 대하여 적용이 가능하고 서브픽쳐 (Subpicture), 슬라이스 (Slice) 등의 영상 분할 구조를 통해 픽쳐를 부분적으로 독립 부/복호화를 수행할 때도 적용할 수 있다.

Ⅲ. 실험 환경 및 결과

본 논문에서 제안하는 랩-어라운드 움직임 벡터 예측 방법의 객관적, 주관적 성능 평가를 위하여 VTM6.0 및 360Lib9.1에 구현하였고 JVET 360-CTC (Common Test

표 2. 제안하는 방법의 VTM6.0 및 360Lib9.1 대비 성능
Table 2. Performance of the proposed method over VTM6.0 and 360Lib9.1

Random Access PERP format		End-to-end WS-PSNR			EncT	DecT
Class	Sequence	Y	U	V		
S1	SkateboardInLot	-0.07%	-0.26%	-0.50%	100%	99%
	ChairliftRide	0.00%	-0.08%	-0.01%	101%	100%
	KiteFlite	-0.01%	0.00%	0.06%	100%	99%
	Harbor	0.00%	0.02%	-0.03%	100%	99%
	Trolley	-0.01%	0.04%	0.06%	99%	99%
	Gaslamp	-0.02%	-0.01%	0.00%	100%	98%
S2	Balboa	0.00%	-0.12%	-0.18%	100%	99%
	Broadway	-0.03%	0.01%	0.08%	100%	99%
	Landing2	-0.02%	-0.21%	-0.04%	101%	100%
	BranCastle2	0.00%	0.01%	0.03%	100%	99%
S1		-0.02%	-0.05%	-0.07%	100%	99%
S2		-0.01%	-0.08%	-0.03%	100%	99%
Average		-0.02%	-0.06%	-0.05%	100%	99%

표 3. 픽쳐 경계에서 제안하는 방법의 VTM6.0 및 360Lib9.1 대비 성능
Table 3. Performance of the proposed method of picture boundary over VTM6.0 and 360Lib9.1

Random Access PERP format		End-to-end WS-PSNR			Codec WS-PSNR			PSNR		
Class	Sequence	Y	U	V	Y	U	V	Y	U	V
S1	SkateboardInLot	-0.37%	-1.22%	-0.71%	-0.44%	-1.35%	-0.82%	-0.43%	-1.33%	-0.26%
	ChairliftRide	-0.16%	-0.05%	-0.25%	-0.25%	-0.14%	-0.28%	-0.23%	-0.13%	-0.13%
	KiteFlite	-0.09%	-0.01%	-0.05%	-0.09%	-0.05%	-0.02%	-0.10%	-0.03%	0.01%
	Harbor	-0.15%	-0.29%	-0.36%	-0.15%	-0.27%	-0.29%	-0.16%	-0.28%	-0.31%
	Trolley	-0.16%	-0.08%	-0.28%	-0.19%	-0.07%	-0.27%	-0.19%	-0.11%	-0.28%
	Gaslamp	-0.20%	-0.19%	-0.24%	-0.18%	-0.11%	-0.20%	-0.19%	-0.15%	-0.20%
S2	Balboa	-0.48%	-1.45%	-0.98%	-0.61%	-1.23%	-0.84%	-0.65%	-1.10%	-0.89%
	Broadway	-0.40%	0.60%	0.02%	-0.67%	0.00%	-0.21%	-0.70%	0.21%	-0.25%
	Landing2	-0.46%	-0.83%	-1.06%	-0.52%	-1.10%	-1.31%	-0.45%	-1.17%	-1.17%
	BranCastle2	-0.40%	-1.00%	-0.45%	-0.44%	-0.98%	-0.50%	-0.45%	-0.95%	-0.53%
S1		-0.19%	-0.31%	-0.31%	-0.22%	-0.33%	-0.31%	-0.22%	-0.34%	-0.20%
S2		-0.43%	-0.67%	-0.62%	-0.56%	-0.83%	-0.71%	-0.56%	-0.75%	-0.71%
Average		-0.29%	-0.45%	-0.43%	-0.36%	-0.53%	-0.47%	-0.35%	-0.50%	-0.40%

Condition)[12]의 임의 접근 (Random Access) 환경에서 실험을 진행하였다. 또한 360 비디오의 부/복호화를 위한 투영 방법은 PERP를 이용하였다. 이번 장에서는 제안하는 방법의 객관적인 평가와 주관적인 평가에 대해 다룬다.

360 비디오의 성능 평가를 위한 BD-rate[13]를 계산하기 위하여 사용되는 여러 PSNR의 측정 위치는 서로 다르고 각각의 측정 위치에 대한 개념도는 그림 1과 같다. 표 2와 표 3은 제안하는 방법에 대한 BD-rate 결과이다.

제안하는 방법은 픽쳐의 왼쪽과 오른쪽 경계에만 적용하는 기술임에도 불구하고 표 2를 보면 전체 픽쳐에 대하여 추가적인 계산 복잡도 없이 End-to-end WS-PSNR을 기준으로 휘도 성분에서 최대 0.07%, 평균 0.02%의 비트 감소율을 보였고, 색차 성분에서 각각 평균 0.06%, 0.05%의 비트 감소율을 보였다. 또한, 제안하는 방법은 상기 서술한 대로 픽쳐의 왼쪽과 오른쪽 경계에 위치한 각 CU에 대해 적용되는 방법이기 때문에 픽쳐의 왼쪽과 오른쪽 경계에서 추가로 성능 평가를 하였다. 여기서 픽쳐의 경계 범위는 픽쳐의 왼쪽과 오른쪽의 경계에서부터 1개의 CTU (Coding Tree Unit) 너비까지로 정하였다. CU의 최대 너비는 CTU의 너비이므로 픽쳐 경계의 범위를 위와 같이 정하였다. 표 3은 제안하는 방법의 성능 평가를 픽쳐의 왼쪽과 오른쪽 경계에서 진행한 결과이다. 표 3을 보면 제안하는 방법은 픽쳐의 왼쪽과 오른쪽 경계에 대하여 End-to-end WS-PSNR을 기준으로 휘도 성분에서 최대 0.48%, 평균 0.29%의 비트 감소율을 보였고 색차 성분에서 각각 평균 0.45%, 0.43%의 비트 감소율을 보여 제안하는 방법이 적용되는 픽쳐의 경계에서 성능이 전체 픽쳐에 대한 성능보다 10배 이상 좋음을 보였다. 또한 End-to-end PSNR뿐만 아니라 Codec WS-PSNR과 PSNR에 대하여도 휘도 성분과 색차 성분에서 각각 0.36%, 0.53%, 0.47%와 0.35%, 0.50%, 0.40%의 비트율 감소를 보였다.

또한, 제안하는 방법에 대한 주관적 화질 평가를 수행하였다. 주관적 화질 평가를 위하여 20명의 일반인으로 구성된 평가단을 구성하였고, 평가단은 실험의 신뢰성을 위하여 10명씩 A, B 두 조로 나뉘어 각각 수행하였다. 평가 영상으로는 JVET-360CTC에 정의된 방법에 따라서 각 영상의 동적 뷰포트 (Dynamic Viewport)를 추출하여 얻는 해상도의 영상을 사용하였다. 평가 방법으로는 ITU-R BT. 500-13 권고문서[14]의 화질 평가 방법 중 DSCQS 방법을 이용해 MOS값을 얻었다. 화질 평가 결과는 그림 10과 같이 막대그래프로 나타냈으며, 각각 평가된 MOS값의 평균값 또한 함께 선형 그래프로 나타냈다. MOS값은 전체적으로 기존 방법 대비 제안하는 방법에서 높게 측정되었으며, 평균적으로 0.03 높은 MOS 결과를 보였다. 또한, MOS값의 최대, 최소를 비교했을 때에도 몇몇 영상에서는 낮거나 같은 화질로 평가받았으나 전반적으로 높은 화질로 평가되

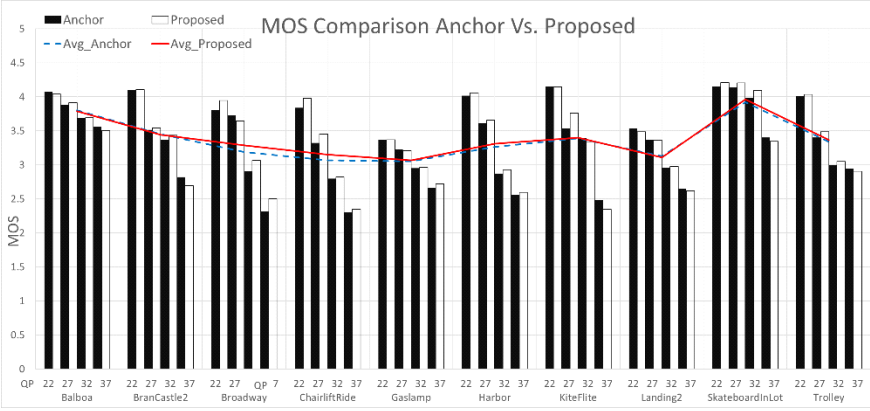

그림 10. 기존 방법과 제안하는 방법의 주관적 화질 비교 결과
Fig. 10. Experimental result of perceived visual quality for conventional and proposed method

었다. 그리고 각 영상의 MOS값과 비트량으로 BD-MOS[15]를 구하여 주관적 화질 성능을 비교하였다. 표 4는 각 영상에 따른 BD-MOS 성능이다. 표 4를 보면 각 영상에 대하여 제안하는 방법이 최소 0.90%, 최대 8.78%, 평균 5.18%의 성능 향상을 보였다.

표 4. 각 시퀀스에 따른 BD-MOS 성능
Table 4. Performance of BD-MOS for each sequence

Class	Sequence	BD-MOS
S1	SkateboardInLot	-7.95%
	ChairliftRide	-8.78%
	KiteFlite	-6.12%
	Harbor	-6.70%
	Trolley	-5.22%
	Gaslamp	-0.90%
S2	Balboa	-2.19%
	Broadway	-6.72%
	Landing2	-1.78%
	BranCastle2	-5.44%
Average		-5.18%

그림 11은 주관적 화질이 개선된 예시이다. 효과적으로 확인하기 위하여 비교할 부분을 잘라서 확대하여 표시하였고 YUView[16]을 사용하여 해당 CTU의 분할구조를 시각화하여 함께 표시하였다. 노란색으로 표시된 블록이 화면간 예측을 사용한 CU, 파란색으로 표시된 블록이 화면내 예측을 사용한 CU이다.

그림 11의 (a) ~ (c)는 SkateboardInLot 시퀀스의 임의 접근 환경, QP = 27, 97번째 프레임에서 각각 원본과 기존

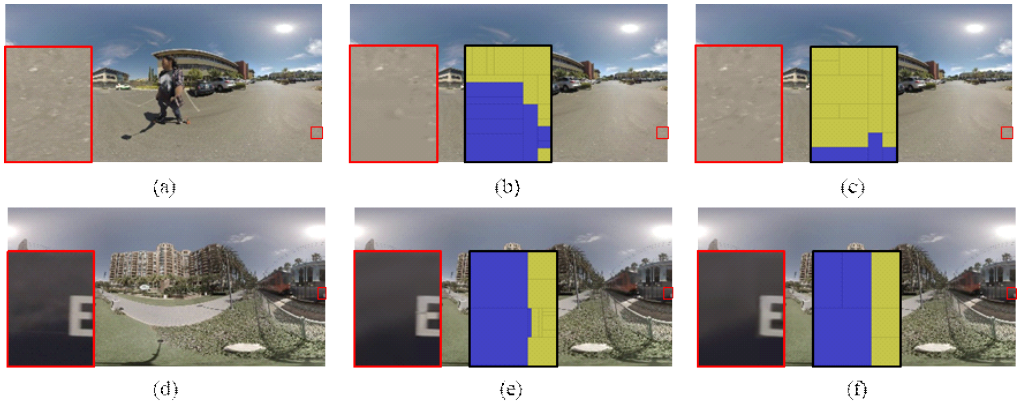

그림 11. 주관적 화질이 개선된 예시 - SkateboardInLot 시퀀스의 임의 접근 환경, QP=27, 97번째 프레임 (a) 원본; (b) 기존 방법; (c) 제안하는 방법; Trolley 시퀀스의 임의 접근 환경, QP=37, 200번째 프레임 (d) 원본; (e) 기존 방법; (f) 제안하는 방법
Fig. 11. Examples of improved subjective quality - SkateboardInLot with RA configuration, QP=27, 97th frame (a) Original; (b) Existing method; (c) Proposed method; Trolley with RA configuration, QP=37, 200th frame (d) Original; (e) Existing method; (f) Proposed method

표 5. 그림 11에서 비교한 CTU에 대한 PSNR과 전송되는 비트량 비교
Table 5. The comparison of PSNR and transmitted Bit-rate for CTU of figure 11

	SkateboardInLot (In figure 10)		Trolley (In figure 10)	
	Existing method	Proposed method	Existing method	Proposed method
PSNR (dB)	37.04	37.96	37.38	36.93
Bit-rate (bits)	2715	2584	214	191

방법과 제안하는 방법을 나타낸 것이다. 기존 방법에서 화면내 예측으로 선택되었던 많은 블록이 화면간 예측으로 선택되면서 예측 성능이 향상되어 바닥 쪽의 텍스쳐(Texture) 표현이 기존 방법 대비 더욱 잘 된 것을 확인할 수 있다. 이로 인하여 표 5에서와같이 전송되는 비트량도 기존 방법 대비 제안하는 방법에서 줄어든 것을 확인할 수 있다. 그림 11의 (d) ~ (f)는 Trolley 시퀀스의 임의 접근 환경, QP = 37, 200번째 프레임에서 각각 원본과 기존 방법과 제안하는 방법을 나타낸 것이다. 제안하는 방법에서 기존 방법 대비 CTU 내의 'E'의 표현이 더 매끄러움을 확인할 수 있다. 기존 방법에서는 해당 부분에서 여러 CU로 분할되어 예측되었고, 제안하는 방법에서는 하나의 CU로 예측됨으로써 예측 성능이 향상되어 효과를 얻음을 보였다. 마찬가지로 표 5를 보면 제안하는 방법에서 전송되는 비트량 감소 효과를 얻었다.

Ⅳ. 결 론

본 논문에서는 360 비디오에 대한 VVC의 화면간 예측의 성능 개선을 위하여 랩-어라운드 움직임 벡터 예측 방법을 제안하였다. 제안하는 방법은 360 비디오의 부/복호화를 위한 투영 방법 중 ERP와 PERP를 적용하여 변환한 360 비디오의 픽쳐의 왼쪽과 오른쪽의 움직임이 유사하다는 특성을 반영하여 화면간 예측 과정에서 움직임 벡터 후보군을 구성할 때 이를 이용하여 픽쳐의 경계 바깥 위치에서도 이미 복호화된 정보로부터 후보군을 구성하는 방법이다. 또한, 제안하는 방법의 개념은 모든 360 비디오 영상과 영상 분할 구조를 통해 독립 부/복호화를 수행하는 경우에도 적용할 수 있다. 제안하는 방법의 성능 평가를 위하여 JVET 360-CTC의 임의 접근 환경에서 VTM6.0 및 360Lib9.1에 PERP를 사용하여 구현하고 제안하는 방법과 기존 방법을 객관적, 주관적으로 비교하였다. 실험 결과로써 제안하는 방법은 휘도 성분에서 최대 0.07%, 평균 0.02%의 비트율 감소와 색차 성분에서는 각각 0.06%, 0.05%의 비트율 감소를 보였고, 픽쳐의 왼쪽과 오른쪽 경계에서의 부호화 효율은 휘도 성분에서 최대 0.48%, 평균 0.29%의 비트율 감소와 색차 성분에서 각각 0.45%, 0.43%의 비트율 감소를 보였다. 그리고 DSCQS 방법을 통해 제안하는 방법의 주관적 화질 평가를 수행하여 MOS값을 얻는다. MOS값은 평균 0.03 향상되었고, MOS값과 비트량을 이용하여 BD-MOS를 구하였다. 그 결과, 제안하는 방법의 성능이 최대 8.78%, 평균 5.18%의 향상을 보였다.

참 고 문 헌 (References)

[1] C. Kim, "A Comparative Study for Virtual Reality 360° Contents Shooting Equipments Based on Real World", Journal of Broadcast Engineering, Vol.21, No. 5, pp.714-725, Sep. 2016.
[2] J. Lee, J. Park, H. Choi, J. Byeon, and D. Sim, "Overview of VVC", Broadcasting and Media Magazine, Vol.24, No.4, pp.10-25, Oct. 2019.
[3] F. Bossen, X. Li, and K. Suehring, JVET-P0003, "AHG report: Test model software development (AHG3)", Oct. 2019.
[4] VTM, https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM
[5] Y. Sun, A. Lu, and L. Yu, JVET-D0040, "AHG8: WS-PSNR for 360 video objective quality evaluation", Oct. 2016.
[6] M. Yu, H. Lakshman, and B. Girod, "A framework to evaluate omni-directional video coding schemes", in 2015 IEEE International Symposium on Mixed and Augmented Reality, pp. 31 – 36, Oct. 2015.
[7] V. Zakharchenko, K. P. Choi, E. Alshina, and J. H. Park, "Omnidirectional video quality metrics and evaluation process", in 2017 Data Compression Conference, pp. 472 – 472, Apr. 2017.
[8] 360Lib, https://jvet.hhi.fraunhofer.de/svn/svn_360Lib/
[9] S. Yoon and G. Park, "Improve Compression Efficiency of 360degree VR Video by Correcting Perspective in Cubemap Projection", Journal of Broadcast Engineering, Vol.22, No.1, pp.136-139, May. 2017.
[10] Y. Yoon, Y. Ahn, D. Sim, and J. Kim, "Efficient methods of inactive regions padding for segmented sphere projection (SSP) of 360 video", IEICE TRANSACTIONS on Information and Systems, Vol.101, No.11, pp.2836-2839, Nov. 2018.

[11] B. Bross, J. Chen, and S. Liu, JVET-O2001, “Versatile Video Coding (Draft 6)”, Jul. 2019.
[12] P. Hanhart, J. Boyce, K. Choi, and J.-L. Lin, JVET-L1012, “JVET common test conditions and evaluation procedures for 360° video”, Oct. 2018.
[13] G. Bjøntegaard, VCEG-M33, “Calculation of average PSNR differences between RD-curves”, Apr. 2014.
[14] ITU-R BT.500-13, “Methodology for the subjective assessment of the quality of television pictures”, Jan. 2012.
[15] P. Hanhart, M. Rerabek, F. De Simone, and T. Ebrahimi, “Subjective quality evaluation of the upcoming HEVC video compression standard”, Applications of digital image processing XXXV, Vol.8499, pp.1–13, Aug. 2012.
[16] YUView, https://https://github.com/IENT/YUView

저 자 소 개

이 민 훈

- 2019년 2월 : 광운대학교 수학과 학사
- 2019년 3월 ~ 현재 : 광운대학교 전자공학과 석사과정
- ORCID : https://orcid.org/0000-0001-8165-5380
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전

이 종 석

- 2016년 2월 : 광운대학교 전자공학과 학사
- 2018년 2월 : 광운대학교 전자공학과 석사
- 2018년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- 2020년 1월 ~ 현재 : 디지털인사이트 재직 중
- ORCID : https://orcid.org/0000-0001-8045-0244
- 주관심분야 : 영상압축, 컴퓨터비전, 고해상도 위성영상 처리

박 준 택

- 2017년 2월 : 광운대학교 컴퓨터공학과 학사
- 2019년 2월 : 광운대학교 컴퓨터공학과 석사
- 2019년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- ORCID : https://orcid.org/0000-0001-5647-7260
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전

임 웅

- 2008년 : 광운대학교 컴퓨터공학과 학사
- 2008년 3월 ~ 2010년 2월 : 광운대학교 컴퓨터공학과 석사
- 2010년 3월 ~ 2016년 : 광운대학교 컴퓨터공학과 박사
- 2012년 2월 ~ 2013년 : Simon Fraser Univ. 방문연구원
- 2016년 4월 ~ 현재 : 한국전자통신연구원 미디어부호화연구실 선임연구원
- ORCID : https://orcid.org/0000-0002-1772-0683
- 주관심분야 : 영상처리, 비디오 압축, 기계 학습, 컴퓨터 비전

저 자 소 개

방 건

- 2014년 2월 : 고려대학교 컴퓨터학과 박사 졸업
- 2008년 ~ 2013년 : MPEG FTV AhG 그룹 EE 코디네이터
- 2011년 9월 ~ 2012년 10월 : MIT RLE ATSP 그룹 방문 연구원
- 2002년 10월 ~ 2006년 2월 : ATSC T3/S2 ACAP 데이터방송 표준화 에디터
- 2000년 4월 ~ 현재 : 한국전자통신연구원 미디어부호화연구실 책임연구원
- ORCID : https://orcid.org/0000-0003-4355-599X
- 주관심분야 : 2D/3D 비디오 부호화, 영상처리, 컴퓨터비전, 인공지능

심 동 규

- 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 : 바로비젼 선임연구원
- 2002년 4월 ~ 2005년 2월 : University of Washington Senior research engineer
- 2005년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 교수
- ORCID : https://orcid.org/0000-0002-2794-9932
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전

오 승 준

- 1980년 2월 : 서울대학교 전자공학과 학사
- 1982년 2월 : 서울대학교 전자공학과 석사
- 1988년 5월 : 미국 Syracuse University 전기/컴퓨터공학과 박사
- 1982년 3월 ~ 1992년 8월 : 한국전자통신연구원 멀티미디어연구실 실장
- 1986년 7월 ~ 1986년 8월 : NSF Supercomputer Center 초청 학생연구원
- 1987년 5월 ~ 1988년 5월 : Northeast Parallel Architecture Center 학생연구원
- 1992년 3월 ~ 1992년 8월 : 충남대학교 컴퓨터공학부 겸임교수
- 1992년 9월 ~ 현재 : 광운대학교 전자공학과 교수
- 2002년 3월 ~ 2017년 12월 : SC29-Korea 전문위원회 대표위원
- ORCID : https://orcid.org/0000-0002-5036-3761
- 주관심분야 : 비디오데이터처리, 영상압축, 컴퓨터비전