

특집논문 (Special Paper)

방송공학회논문지 제25권 제3호, 2020년 5월 (JBE Vol. 25, No. 3, May 2020)

<https://doi.org/10.5909/JBE.2020.25.3.325>

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

스크린 콘텐츠를 위한 VVC 화면내 삼각형 분할 예측 방법

최재륜^{a)}, 권대혁^{a)}, 한희지^{a)}, 이하현^{b)}, 강정원^{b)}, 최해철^{a)†}

VVC Intra Triangular Partitioning Prediction for Screen Contents

Jaeryun Choe^{a)}, Daehyeok Gwon^{a)}, Heeji Han^{a)}, Hahyun Lee^{b)}, Jungwon Kang^{b)} and Haechul Choi^{a)†}

요약

VVC(Versatile Video Coding)는 ISO/IEC/ITU-T의 JVET(Joint Video Experts Team)에서 표준화 중인 새로운 비디오 부호화 표준으로 스크린 콘텐츠 부호화 툴을 포함한 다양한 기술을 채택하고 있다. 스크린 콘텐츠는 문자 영역과 같이 사선 방향 에지가 자주 발생하는 특징을 가지며, 이런 특징을 갖는 영상에 삼각형 형태의 분할 부호화를 적용하면 압축 효율이 증가할 수 있다. 본 논문에서는 스크린 콘텐츠를 위한 VVC 기반 화면내 삼각형 분할 예측 방법을 제안한다. 기존 VVC의 화면내 예측 부호화에서 삼각형 분할 예측을 지원하는 Triangular Prediction Mode 방법과 유사하게, 제안 방법은 화면내 예측 부호화에서 수직과 수평 방향 예측 모드와 주변 복원 참조 라인을 이용하여 두 개의 사각형 예측 블록을 생성하고 삼각형 모양의 마스크로 두 예측 블록을 가중합하여 최종 예측 신호를 만든다. 제안 방법의 실험 결과는 All Intra 스크린 콘텐츠 영상 실험에서 YUV 각각 평균 1.86%, 1.49%, 1.55% 부호화 성능 향상을 보이고, 자연 영상 실험 조건에서는 부호화 효율에 미미한 손실을 보였다. 결론적으로, 화면내 예측 부호화 모드에 제안 방법을 적용하여 압축 성능을 향상할 수 있었다.

Abstract

Versatile Video Coding (VVC) is a new video coding standard that is being developed by the Joint Video Experts Team of ISO/IEC/ITU-T and it has adopted various technologies including screen content coding tools. Screen contents have a feature that blocks are likely to have diagonal edges like character regions. If triangular partitioning coding is allowed for screen contents having such the feature, coding efficiency would increase. This paper proposes an intra prediction method using triangular partitioning prediction for screen content coding. Similar to the Triangular Prediction Mode of VVC that supports the triangular partitioning prediction, the proposed method derives two prediction blocks using Horizontal and Vertical modes and then it blends the predicted blocks applying masks with triangle shape to generate a final prediction block. The experimental results of the proposed method showed an average of 1.86%, 1.49%, and 1.55% coding efficiency in YUV, respectively, for VVC screen content test sequences.

Keyword : Video Coding, VVC, Intra Prediction, Triangular Partitioning, Screen Content Coding

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.”

I. 서론

비디오 콘텐츠의 다양화와 초고속 인터넷 보급으로 실시간 UHD/8K, 360도 비디오 및 HDR(High-Dynamic-range) 비디오 등의 초실감 미디어에 대한 산업계의 요구와 관심이 높아지고 있다. 이에 따라 비디오 및 오디오 등 멀티미디어 표준을 개발하는 ISO/IEC JTC1 SC29 WG 11 MPEG (Moving Picture Experts Group)와 ITU-T SG16 Q.1 VCEG (Video Coding Experts Group)은 공동으로 JVET(Joint Video Experts Team)을 구성하고 차세대 비디오 압축 표준인 VVC(Versatile Video Coding) 표준화를 진행하고 있다^{[1][3]}. VVC 표준화는 객관적 화질 기준으로 HD 영상과 UHD 영상에서 High Efficiency Video Coding(HEVC)/H.265^{[4][5]} 대비 38% BD-rate 이상의 부호화 효율을 달성하였고, 인코딩 시간과 디코딩 시간이 HEVC 대비 각각 8.9배 및 1.6배 증가하였다^[6].

최근 스트리밍 게임 서비스, 화면 공유 및 원격 데스크톱 응용 프로그램 등 스크린 콘텐츠에 대한 수요가 급증하고 있다. HEVC/H.265 표준화에서는 스크린 콘텐츠 부호화 기술을 확장(Amendment) 표준에서 별도 취급하였으나, VVC 표준화에서는 스크린 콘텐츠의 활성화를 반영하여 표준화 초기 단계에서부터 스크린 콘텐츠 압축 기술을 표준으로 채택하고 있다. 스크린 콘텐츠는 자연 영상과는 다른 신호 특성을 포함하며, 그중 한 가지 특성은 강인한 에지가 산재하다는 것이다. 이 중에는 수평과 수직 방향뿐만 아니라 사선 방향으로 나뉘는 두 영역에서 서로 상이한 텍스처를 갖는 블록도 많이 발생한다. 그러나 VVC를 포함한 기존 비디오 압축 표준의 화면내 예측에서는 사각형 분

할만을 허용하기 때문에 스크린 콘텐츠의 사선 방향으로 나뉘는 두 영역에서 서로 상이한 텍스처를 갖는 블록을 특별히 고려하지 않았다. 스크린 콘텐츠의 부호화 효율을 향상시키기 위하여, 본 논문에서는 사선 방향으로 나뉘는 두 영역에서 서로 상이한 텍스처를 갖는 블록이 있을 때, 이를 효과적으로 부호화할 수 있는 화면내 예측 방법을 제안한다. 제안 방법은 서로 다른 두 개의 화면내 예측 모드를 이용하여 두 개의 예측 블록을 생성하고 이를 서로 반대되는 삼각형 모양의 마스크를 이용해 혼합함으로써 최종 예측 블록을 생성한다. 이 최종 예측 블록은 기존 하나의 방향성 혹은 비방향성 화면내 예측 모드만을 이용해 생성한 예측 블록보다 사선 방향으로 나뉘는 두 영역에서 서로 상이한 텍스처를 갖는 블록을 더 정확히 예측할 수 있으며, 실험을 통해 제안 방법의 부호화 성능 개선 효과를 확인한다.

본 논문의 구성은 다음과 같다. 2장에서는 VVC 표준 개요 및 제안 방법과 관련한 VVC 예측 기술들을 소개하고, 3장에서는 ITP(Intra Triangular Partitioning) 방법을 제안하며, 4장에서는 제안 방법의 성능 실험을 보이고 이를 분석한다. 마지막으로 5장에서는 본 논문에 대해 결론짓는다.

II. VVC 표준화

1. VVC 표준화 개요

미디어 콘텐츠의 다양화, 실시간 스트리밍 서비스 확대, 8K 비디오 보급, 게임을 다운로드 받지 않는 스트리밍 게임 서비스 출현 등의 요인으로 인해 시장에서는 비디오 압축

a) 한밭대학교 정보통신전문대학원 멀티미디어공학과(Multimedia Engineering, Hanbat National University)

b) 한국전자통신연구원(ETRI)

‡ Corresponding Author : 최해철(Haechul Choi)

E-mail: choihc@hanbat.ac.kr

Tel: +82-42-821-1149

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

※ This work was supported by Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT)(No. 2016-0-00572, Development and Standardization of 5th Generation Video/Audio Coding Technology for Ultra High Quality Media Service).

· Manuscript received March 25, 2020; Revised April 29, 2020; Accepted May 4, 2020.

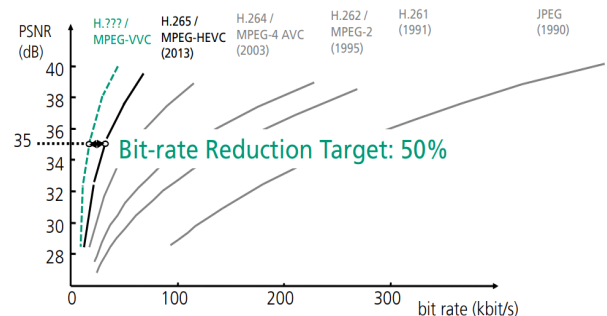


그림 1. 비디오 부호화 표준의 역사 및 성능^[7]

Fig. 1. History and performance of video coding standards^[7]

성능의 향상을 요구하고 있다^[1]. 그림 1^[7]은 역대 비디오 압축 표준의 압축률을 나타내며 새로운 표준이 제정될 때 그 이전 표준보다 약 50%의 비트 레이트(bit-rate) 감소를 보인다. ISO/IEC MPEG과 ITU-T VCEG의 공동팀인 JVET에서 개발하고 있는 차세대 비디오 부호화 표준 VVC도 근래 표준인 HEVC/H.265 표준 대비 50% 이상의 부호화 효율 향상을 목표로 한다. 이 목표 달성을 위해서 VVC는 다양한 압축 기술들을 채택하고 있다.

VVC의 화면내 예측은 HEVC의 화면내 예측 기술의 부호화 효율을 개선하거나 새로운 기술을 추가하였다^{[8][9]}. 이전 비디오 압축 표준인 HEVC/H.265는 DC, Planar, 35개 방향성 모드를 포함하여 37개 화면내 예측 모드를 지원하는 것에 반해, VVC는 두 배에 이르는 67개의 화면내 예측 모드로 확장하였으며 20개의 광각(Wide) 모드를 추가 허용함으로써 최대 87개의 화면내 예측 모드를 지원한다. VVC의 화면내 예측 모드를 나타내는 그림 2와 같이, VVC는 HEVC보다 더 다양한 방향성으로 정확한 예측을 수행할 수 있으며, 직사각형 형태의 화면내 예측 블록도 허용함에 따라 광각 방향성 모드를 사용하여 부호화 성능을 개선시켰다. 늘어난 개수의 화면내 예측 모드를 효과적으로 부호화하기 위해 MPM(Most Probable Mode) 개수도 6개로 증가시켰다. 또한 MIP(Matrix-based Intra Prediction), MRL

(Multiple Reference Lines)^[10], ISP(Intra Sub-Partitioning), PDPC(Position Dependent Intra Prediction Combination), CCLM(Cross-Component Linear Model) 모드 등의 새로운 화면내 예측 기술도 채택하였다^[9].

화면간 예측 방법과 관련하여, SKIP 및 MERGE 모드 측면에서는 HEVC와 유사한 MERGE 모드인 일반 MERGE(Regular MERGE) 모드에 추가로 MMVD(Merge with Motion Vector Difference) 모드, CIIP(Combined Inter and Intra Prediction) 모드, TRIANGULAR 모드^{[11][13]}, SbTMVP(Subblock-based Temporal Motion Vector Prediction)와 AFFINE MERGE 모드를 포함하는 서브 블록 기반 MERGE 모드가 채택되었다. AMVP(Advanced MVP) 모드 측면에서는 HEVC와 유사한 일반 AMVP(Regular AMVP) 모드에 SMVD(Symmetric MVD) 모드와 AFFINE AMVP 모드가 채택되었다. 또한, 이와 같은 MERGE 및 AMVP 모드 이외에 MERGE 혹은 AMVP에 추가로 적용하여 부호화 효율을 개선하는 디코더 기반 움직임 벡터 보정 기술인 DMVR(Decoder-side MV Refinement), BDOF(Bi-Directional Optical Flow), PROF(Prediction Refinement with Optical Flow)들이 추가되었다^[14].

다음 절에서 VVC의 채택 기술 중 본 논문에서 제안하는 방법과 관련 있는 TPM과 MRL에 대해 자세히 살펴본다.

2. TPM (Triangular Partitioning Mode)

TPM^{[11][13]}은 화면간 예측 모드의 하나로 서로 다른 움직임 정보로부터 획득한 두 개의 삼각형 모양을 가중합하여 하나의 예측 블록을 생성한다. 구체적으로 서로 다른 두 개의 참조 블록을 삼각형 모양의 마스크를 이용하여 블렌딩(Blending)함으로써 마치 삼각형으로 분할된 예측 신호를 생성한 효과를 얻을 수 있다. 블렌딩 연산은 두 개의 참조 블록을 하나의 예측 블록으로 만드는 과정으로 참조 블록을 그림 3과 같이 45도 또는 135도 사선 방향 중 한 가지 방향으로 나누어 가중치 연산을 통해 최종 예측 신호를 생성한다. 블렌딩 연산의 가중치는 탭(Tap) 수에 의해 정의된다. 예를 들어, 3-tap을 사용할 때 블렌딩 연산의 가중치는 6:2, 4:4, 2:6이 적용될 수 있고, 7-tap을 사용할 때는 7:1, 6:2, 5:3, 4:4, 3:5, 2:6, 1:7의 가중치가 적용될 수 있다. 휘도

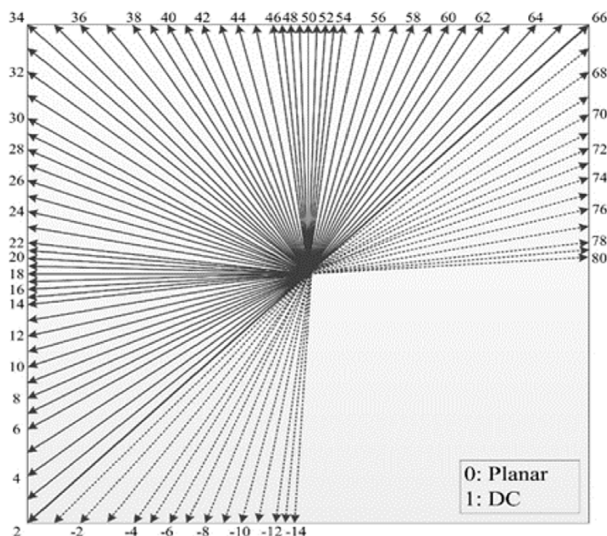


그림 2. VVC 화면내 예측 모드^[10]
 Fig. 2. Intra prediction mode of VVC^[10]

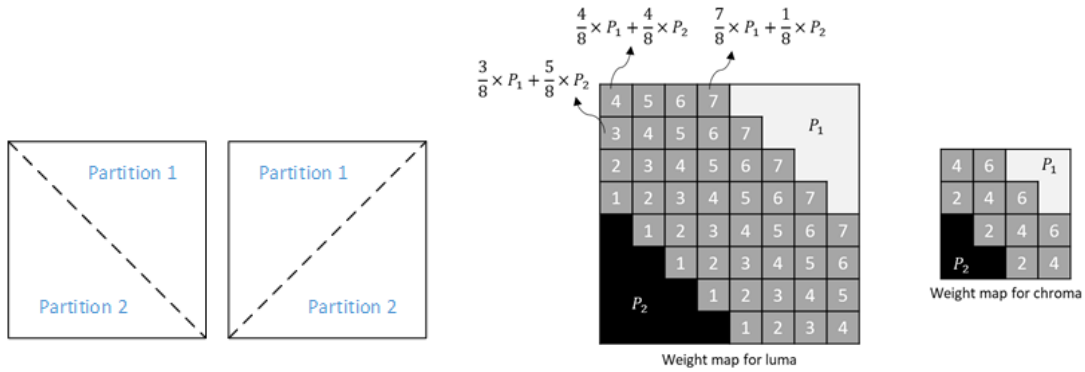


그림 3. 좌) TPM의 삼각형 분할 구조, 우) TPM의 가중치 값^[15]
 Fig. 3. Left) Triangular partitioning structure of TPM, Right) Weight value of the TPM^[15]

성분 블록의 경우 7-tap, 색차 성분 블록의 경우 3-tap을 사용하여 블렌딩 연산을 수행한다.

예측 블록을 생성할 때 사용되는 두 개의 참조 블록은 메모리 접근 측면에서 복잡도를 줄이기 위해 단방향 예측(Uni-directional Prediction)만 허용된다. 그림 4와 같이 각 참조 블록의 단방향 움직임 정보는 현재 블록의 주변 위치 블록에서 유도한 움직임 정보 후보 리스트에서 선택한다^[13]. TPM과 관련된 구문요소(Syntax Element)는 45도 또는 135도 중 어느 방향으로 분할하는지를 알려주는 지시자와 예측 블록을 생성하기 위한 두 개의 움직임 정보를 지시하는 인덱스 2개를 포함한다.

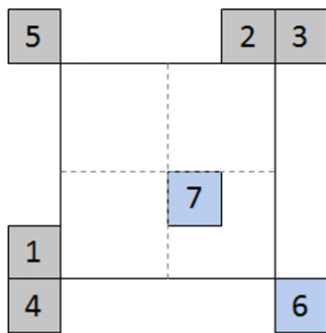


그림 4. TPM에서 움직임 벡터를 유도하기 위한 주변 블록 (1~7)^[13]
 Fig. 4. Neighboring blocks to derive motion vector for TPM^[13]

3. MRL (Multiple Reference Lines)

기존 비디오 표준에서 화면내 예측을 수행할 때 1 개의

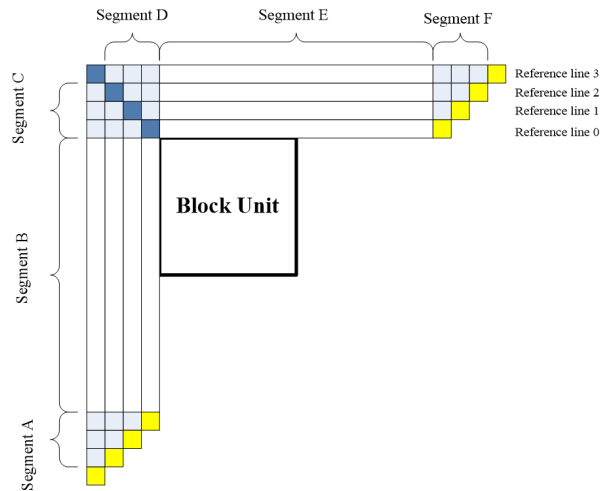


그림 5. MRL에서 사용하는 참조 라인^[9]
 Fig. 5. Reference lines of MRL^[9]

주변 참조 라인만을 사용한 것에 반해, MRL^[16]은 복수의 참조 라인 중 하나를 선택하여 예측하는 기술이다. 그림 5에서, 기존의 화면내 예측은 가장 가까운 0번 참조 화소 라인만 참조 가능했으며, MRL을 사용하면 0번 참조 화소 라인, 1번 참조 화소 라인, 혹은 3번 참조 화소 라인 중 한 가지를 선택적으로 사용할 수 있다. 1번 참조 화소 라인 또는 3번 참조 화소 라인이 선택되는 경우, 즉 0보다 큰 참조 라인 인덱스인 경우, 항상 MPM 중 하나의 화면내 예측 모드만 사용해야 한다. 따라서 MPM 모드가 아닌 잔여(Remaining) 모드는 사용할 수 없고, PDPC도 적용하지 않는다. 현재 CTU(Coding Tree Unit) 경계를 벗어날 때는 메

모리 감소를 위해 MRL을 적용할 수 없으며 이에 따라 CTU 경계에 인접하는 CU(Coding Unit)는 MRL을 사용할 수 없다.

III. Intra Triangular Partitioning

1. ITP (Intra Triangular Partitioning)

기존 화면내 부호화 방법은 사각형 블록 기반으로 부호화 효율 측면에서 최적의 화면내 예측 모드 하나를 선택하여 하나의 예측 블록을 생성한다. 블록의 화소가 대체로 유사하여 평활한 영상이거나 특정 방향성을 기준으로 유사한 화소를 가지는 경우 기존 화면내 예측 방법은 높은 예측 정확도를 가질 수 있다. 하지만 블록내 45° 혹은 135°의 사선 방향으로 두 개의 구분되는 영역이 존재하고 두 영역간 화소 유사성이 낮은 블록의 경우에는 부호화 효율이 낮아진다. 본 논문에서는 사각형 블록을 사선 방향으로 분할하여 각각의 삼각형 모양에 예측 신호를 별도 생성하여 부호화 효율을 높일 수 있는 화면내 예측 방법을 제안한다.

구체적으로, 제안하는 Intra Triangular Partitioning(ITP) 방법은 그림 6과 같이 이미 복원된 주변 상단 참조 라인과 좌측 참조 라인으로부터 서로 다른 두 개의 화면내 예측 모드를 이용하여 두 개의 예측 블록을 생성한다. 한 예측 블록에 사선 방향으로 점진적으로 낮은 값을 갖는 가중치 마스크를 적용하여 마치 삼각형 형태에만 값을 갖는 예측

블록(제1 예측 블록)을 생성한다. 다른 예측 블록에는 반대 방향으로 점진적으로 낮은 값을 갖는 가중치 마스크를 적용하여 예측 블록(제2 예측 블록)을 생성한다. 이 두 예측 블록의 합으로 최종 예측 블록을 획득한다.

블록을 사선으로 나누는 방법은 좌상측에서 우하측으로 혹은 우상측에서 좌하측으로 나누는 두 가지 방향이 가능하다. 본 논문에서는 다음 두 가지 이유로 좌상측에서 우하측으로 나누는 경우만을 고려한다. 첫째, 두 가지 방향을 모두 고려하면 사선 방향을 지시하기 위한 지시자를 시그널링(Signaling)해야 한다. 이는 화면내 예측을 사용하는 블록인 경우 항상 전송해야 하므로 비트 오버헤드로 인해 부호화 효율이 감소할 수 있다. 둘째, 우상측에서 좌하측으로 나누는 경우 우하 삼각형 영역은 좌측 혹은 상측 참조 라인과 거리가 멀어진다. 화면간 예측의 경우 참조 영상으로부터 현재 블록 위치의 우측과 하측 영역도 참조할 수 있지만, 화면내 예측의 경우 현재 블록의 좌측과 상측의 영역만 참조할 수 있으므로 먼 공간적 거리를 갖게 된다. 이에 따라 화소 유사성이 떨어지며 예측 정확도도 낮아진다. 공간적 거리가 멀더라도 평활한 영상과 같이 예측이 쉬운 영상의 경우에는 정확한 예측이 될 수도 있다. 그러나 평활한 영상의 경우에는 기존의 화면내 예측 기술들도도 충분하며 허용하는 경우의 수가 늘면 더 정확한 예측으로 압축 왜곡(Distortion)이 낮아질 수 있는 것과 함께, 늘어난 경우의 수 중 하나를 알리기 위한 지시자의 비트율(Rate)도 증가하기에 Rate-Distortion Optimization(RDO) 측면에서 경우의 수를 늘리는 것은 부호화 효율의 감소로 이어질 수 있다. 따라

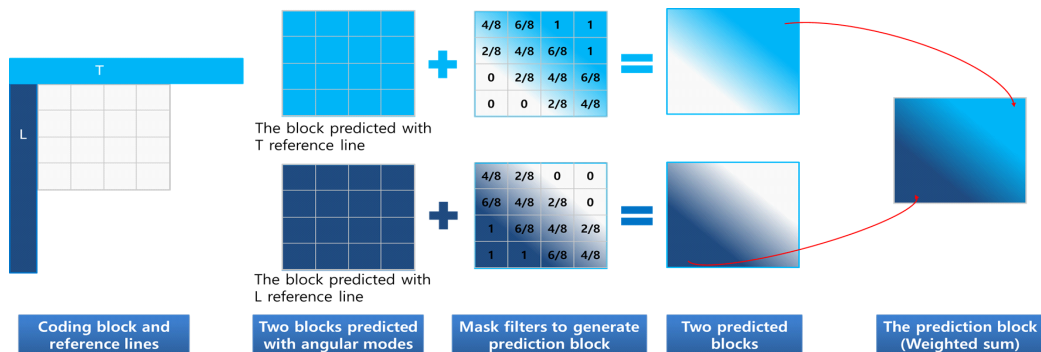


그림 6. ITP의 예측 신호 생성 방법
 Fig. 6. Prediction signal generation method of ITP

서 본 논문에서는 실험을 통해 좌상측에서 우하측으로 나누는 한 가지 경우만 허용하기로 설계하였다.

2. ITP 예측 신호 생성 방법

블록에서 사선으로 나뉘지는 두 영역에 대해 각각 어떤 화면내 예측 모드를 적용할 것인지 고찰이 필요하다. 이를 위해 임의의 두 화면내 예측 모드를 사용할 수 있도록 허용하거나 미리 정의된 화면내 예측 모드를 사용하는 방법을 생각해 볼 수 있다. 만약 임의의 두 화면내 예측 모드를 사용하면, 한 CU에 대해 두 개의 화면내 예측 모드 지시자를 전송해야 한다. 이에 따라 MPM 모드와 연동한 화면내 예측 모드 지시자의 부호화 방법을 설계해야 하며 이는 VVC 구문요소가 크게 수정되어 다른 부호화 기술에 큰 영향을 미칠 수 있으며 추가 구문요소의 증가로 인해 부호화 효율이 크게 감소할 수 있다. 따라서 본 논문에서는 화면내 예측 모드에서 많이 선택되는 수평과 수직 방향 예측 모드로 고정하여 두 영역의 예측 신호를 생성하도록 설계하였다. 일반적으로 화면내 예측 모드에서는 Planar와 DC 모드가 가장 많이 선택되지만, 이 두 모드는 블록 내의 화소들이 모두 유사한 경우에 적절하며 상측과 좌측의 참조 라인을 함께 이용하므로 블록을 사선으로 나누어 별도로 예측 신호를 생성하는데 적절하지 않다. 결론적으로, 제안하는 ITP는 블록을 좌상측에서 우하측 방향의 사선으로 나누고 좌하 삼

각형 영역은 가까운 좌측 참조 라인으로부터 수평 방향으로 예측 신호를 생성하고 우상 삼각형 영역은 가까운 상측 참조 라인으로부터 수직 방향으로 예측 신호를 생성하였다.

그림 7과 같이, 수평과 수직 방향 예측 모드를 이용한 ITP 예측 신호 생성 방법은 삼각형으로 분할된 제1 예측 블록을 좌측 참조 라인(L)으로부터 수평 방향 모드로 생성하고, 제2 예측 블록을 상측 참조 라인(T)으로부터 수직 방향 모드로 생성한다. 이후 마스크를 이용한 블렌딩 과정을 통해 삼각형 모양으로 결합된 최종 예측 신호를 생성한다.

3. ITP와 MRL의 연동

수평 방향과 수직 방향으로 예측 신호를 생성할 때, MRL 기술을 연동하여 예측 신호의 정확성을 높일 수 있다. 그림 8과 같이 MRL은 기존 화면내 예측 모드에서 사용하는 0번 참조 화소 라인 뿐만 아니라 1번 참조 화소 라인과 3번 참조 화소 라인을 지원한다. MRL의 사용 여부에 따라 1 개의 참조 라인 또는 3 개의 참조 라인 중 하나를 선택할 수 있다. RDO 측면에서 MRL의 사용이 결정되면 3 개의 참조 라인 중 하나를 선택하여 예측 정확도를 높인다. 본 논문에서 제안하는 ITP도 제1 및 제2 예측 블록 생성을 위해 수평과 수직 방향 예측 시 MRL을 허용한다. 즉 현재 CU를 ITP만 이용하여 부호화할 수 있으며, ITP와 MRL을 함께 연동

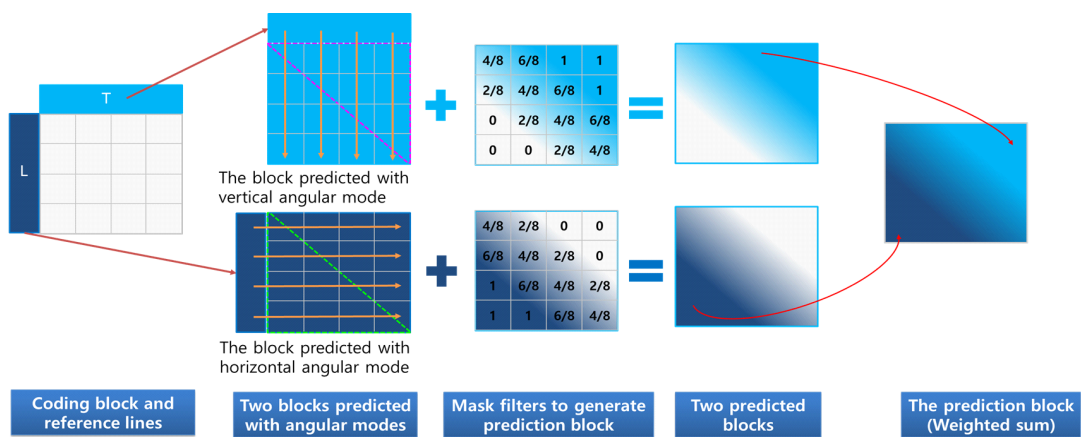


그림 7. 수평과 수직 방향 예측 모드를 이용한 ITP의 예측 신호 생성 방법
Fig. 7. Prediction signal generation method of ITP using Horizontal and Vertical prediction modes

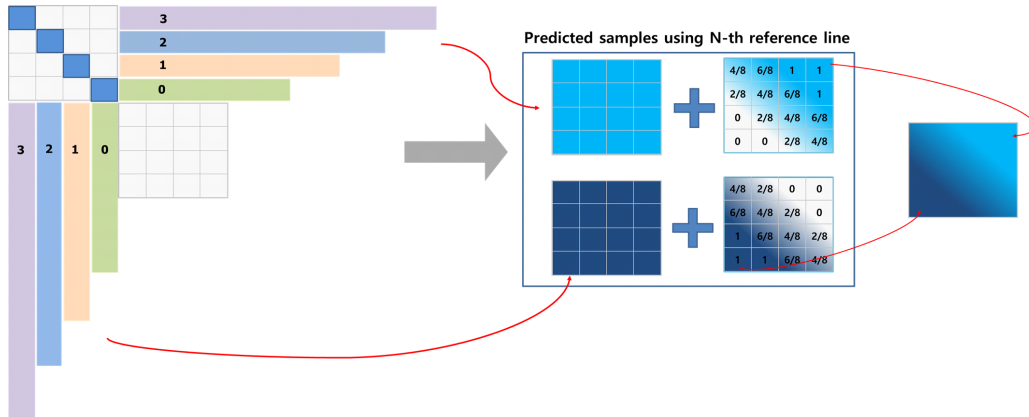


그림 8. ITP와 MRL의 연동방법
 Fig. 8. Cooperation with ITP and MRL

하여 부호화할 수도 있다. ITP와 MRL을 함께 연동하면, 3 개의 참조 라인 중 하나를 선택하고 그 선택된 참조 라인으로부터 제1 및 제2 예측 블록을 위한 수평과 수직 예측을 수행한다.

4. 블렌딩 방법

제안 방법은 블렌딩 마스크를 이용하여 그림 9와 같이 좌하측(LT) 영역과 우상측(AT) 영역의 예측 신호를 혼합하여 하나의 사각형 예측 신호를 생성한다. 마스크의 탭 수는 블록 크기 및 모양에 상관없이 3개로 간소화하였다. 스크린 콘텐츠는 에지가 선명한 특징이 있으므로 에지를 잘 보존하기 위해서, VVC의 화면간 예측 모드 부호화 기술 중 하나인 TPM 모드에서 사용하는 7-tap보다 작은 3-tap을 사용

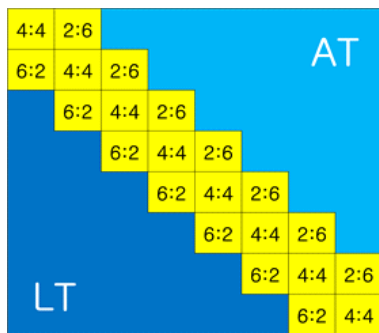


그림 9. 제안 방법에서 사용된 3-tap 블렌딩 마스크
 Fig. 9. 3-tap blending mask for the proposed method

하였다. 이 3-tap 마스크에서 블렌딩 연산의 가중치는 6:2, 4:4, 2:6를 적용한다. 예를 들어, 가중치가 2:6인 경우, LT 영역 즉 제1 예측 블록에는 2/8의 가중치, AT 영역 즉 제2 예측 블록에는 6/8의 가중치를 곱한 후 합함으로써 최종 예측 블록을 생성한다.

5. ITP 모드 시그널링 방법

제안 방법은 VVC 기반으로 현재 CU가 ITP 모드인지를 알리기 위한 구문요소인 `cu_itp_flag`를 정의한다. 구문요소 `cu_itp_flag`의 발생 확률에 기반하여 문맥 모델(Context model)을 설계하고 CABAC(Context Adaptive Binary Arithmetic Coding)을 이용하여 엔트로피 부호화한다.

구문요소 `cu_itp_flag`의 시그널링 방법은 그림 10과 같다. VVC는 `pred_mode_flag`가 화면내 예측 모드를 지시하는 경우 MIP 모드 여부를 알리기 위한 `intra_mip_flag`를 전송한다. MIP 모드인 경우, `cu_itp_flag`를 포함한 나머지 화면내 예측 모드의 구문요소를 시그널링하지 않고 MIP 모드로 부호화한다. MIP 모드가 아니며 MRL을 허용하는 경우, 참조 라인의 인덱스를 알리는 `intra_luma_ref_idx`를 시그널링한다. 제안 ITP 모드는 MRL과 연동할 수 있다. 따라서 ITP는 MRL의 Reference Line 0, Reference Line 1, Reference Line 3에 모두 적용될 수 있으므로 `intra_luma_ref_idx`이 무슨 값이든 `cu_itp_flag`를 시그널링한다. 이 `cu_itp_flag`의 값이 ITP 모드임을 알리면, 현재 블록은 제안 방법인

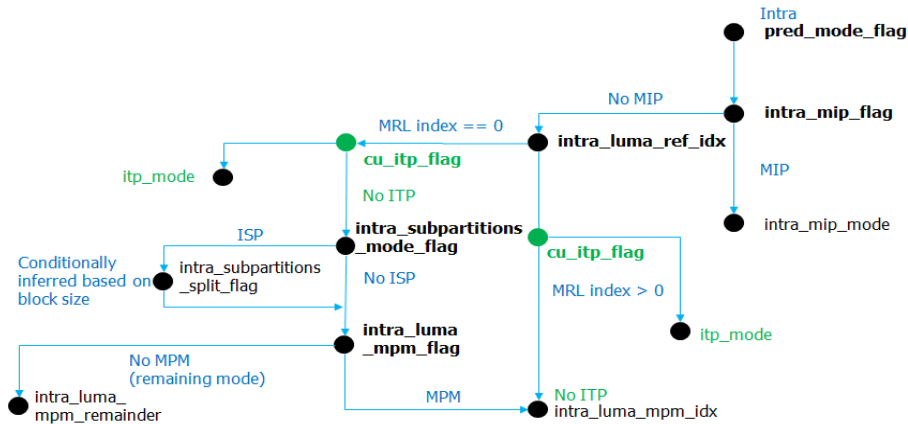


그림 10. VVC 기반 ITP 모드 시그널링 방법
Fig. 10. VVC based ITP mode signaling method

ITP 모드로 부호화된다. 이 `cu_itp_flag`의 값이 ITP 모드임가 아님을 알리면, 기존 VVC의 구문요소 시그널링 방법과 같이 현재 블록을 부호화한다. 즉 ISP 모드를 알리는 `intra_subpartitions_mode_flag`가 전송되며, 이 플래그의 값에 따라 `intra_luma_mpm_flag`가 전송될 수 있으며, 또한 `intra_luma_mpm_flag` 값에 따라 `intra_luma_mpm_remainder`가 추가 시그널링될 수 있다. 앞서 설명한 ITP 시그널링 방법은 VVC의 기존 시그널링 방법을 크게 변경하지 않고 ITP 모드를 허용할 수 있다.

IV. 실험 및 성능 분석

본 논문에서는 제안 ITP 방법을 VVC 참조 소프트웨어인 VTM6.0^[17]를 기반으로 구현하였고, 성능 평가를 위해 VVC 표준화에서 사용 중인 공통 테스트 조건^[18]을 사용하였다. VVC의 공통 테스트 영상 중 스크린 콘텐츠 영상인 SCC 클래스, 자연 영상과 컴퓨터 그래픽스가 섞여 있는 F 클래스, 자연 영상인 A1, A2 클래스 영상에 대해 공통 테스트 조건에 명시된 양자화 계수 4개의 값에 대한 비트스트림을

표 1. 실험 영상
Table 1. Test sequences

Common Test Condition / All Intra Configuration					
Class	Resolution	Sequences	Bit depth	Frames	Frame Rate
SCC	1920x1080	ChineseEditing	8	600	60
		Console	8	600	60
		Desktop	8	600	60
		FlyingGraphics	8	300	60
F	832x480	BasketballDrillText	8	500	50
	1920x1080	ArenaOfValor	8	600	60
	1280x720	SlideEditing	8	300	30
		SlideShow	8	500	20
A1	3840x2160	Tango2	10	294	60
		FoodMarket4	10	300	60
		Campfire	10	300	30
A2	3840x2160	CatRobot	10	300	60
		DaylightRoad2	10	300	60
		ParkRunning3	10	300	50

생성하여 부호화 효율을 비교하였다. 실험 영상의 구체적인 정보는 표 1과 같다. VVC 표준화는 SCC 클래스와 일반 영상(A1, A2, F 클래스)에 대해 다른 공통 실험 조건을 적용하고 있다. Intra Block Copy(IBC)^[3] 등의 스크린 콘텐츠만을 위한 기술은 SCC 클래스의 실험에서는 허용하지만 일반 영상의 실험에서는 사용하지 않는다.

표 2는 공통 테스트 조건 중 하나인 All_intra Configuration에서 제안 방법의 성능을 나타낸다. 제안 방법은 SCC 클래스영상 ChoenseEditing, Console, Desktop, Flying-Graphics 대해서 Y BD-rate^[19]는 각각 -0.2%, 0.03%, 0.00%, 0.02%의 결과를 보였다. 평균 -0.04%의 다소 낮은 비트율 감소이다. 반면 F 클래스 BasketballDrillText, ArenaOfValor, SlideEditing, SlideShow에 대해 Y BD-rate는 각각 0.00%, 0.05%, -1.09%, -0.74%로 평균 -0.45%의 유의미한 비트율 감소를 보였다. 제안하는 ITP는 스크린 콘텐츠의 특징을 반영한 부호화 툴이다. 그럼에도 불구하고 제안 방법이 클래스 F에 대해서는 유의미한 부호화 효율을 보였지만 SCC 클래스에 대해서는 부호화 효율이 오히려

미비하다. 그 이유는 앞서 설명한 것과 같이 공통 실험 조건에서는 SCC 클래스에 대해서만 스크린 콘텐츠를 위한 부호화 툴들이 허용되기 때문이다. IBC와 같은 스크린 콘텐츠의 특성을 고려한 툴들이 압축률을 이미 상당히 높였기에 제안하는 ITP가 부호화 효율을 높일 수 있는 여유가 적다. F 클래스에 대해서는 VVC의 스크린 콘텐츠 부호화 툴들이 허용되지 않기 때문에 ITP가 비교적 높은 부호화 효율을 보일 수 있었다. 자연 영상인 A1, A2 클래스에 대해서는 평균 0.01%의 미비한 비트율 증가를 보였다. ITP의 허용으로 인해 구문요소를 추가적으로 시그널링하여 비트 오버헤드가 증가했음에도 불구하고 자연 영상에 대해서 부호화 손실을 거의 유발하지 않는다. 이는 ITP로 인한 비트 오버헤드와 ITP로 인한 예측 정확도 향상의 효과가 서로 비슷했기 때문으로 판단된다.

SCC 클래스 영상에서 ITP의 성능을 다시 검증하기 위해서, VVC의 대표적인 스크린 콘텐츠 부호화 툴인 IBC를 허용하지 않고 실험을 진행하였다. 표 3과 같이, IBC_off 조건에서 ChineseEditing, Console, Desktop, FlyingGraphics

표 2. VVC 공통 실험 조건에서의 ITP 모드 부호화 효율
 Table 2. ITP mode performcne under the VVC common test condition

Class	Sequence	Proposed method (BD-rate %)		
		Y	U	V
SCC	ChineseEditing	-0.20	-0.17	-0.17
	Console	0.03	-0.08	-0.03
	Desktop	0.00	-0.04	-0.01
	FlyingGraphics	0.02	0.02	0.03
	Average	-0.04	-0.06	-0.04
F	BasketballDrillText	0.00	0.09	0.08
	ArenaOfVlovor	0.05	0.05	0.04
	SlideEditing	-1.09	-0.87	-1.02
	SlideShow	-0.74	-0.70	-0.60
	Average	-0.45	-0.36	-0.37
A1	Tango2	0.04	-0.10	-0.01
	FoodMarket4	0.00	-0.01	0.08
	Campfire	-0.01	0.00	0.09
	Average	0.01	-0.03	0.05
A2	CatRobot	0.03	-0.06	-0.05
	DaylightRoad2	0.00	0.19	0.08
	ParkRunning3	0.00	0.02	0.00
	Average	0.01	0.03	0.11

표 3. IBC_off 조건에서의 ITP 모드 부호화 효율
Table 3. Coding efficiency of ITP mode under IBC_off condition

Class	Sequence	Proposed method (BD-rate %)		
		Y	U	V
SCC	ChineseEditing	-2.37	-1.97	-2.01
	Console	-2.08	-1.57	-1.62
	Destktop	-1.84	-1.37	-1.44
	FlyingGraphics	-1.15	-1.05	-1.14
Average		-1.86	-1.49	-1.55

영상에 대해 Y Bd-rate가 각각 -2.37%, -2.08%, -1.84%, -1.15%의 상당히 높은 비트율 감소를 보였다. IBC를 허용하지 않을 때 제안하는 ITP는 SCC 클래스에 대해 평균 1.86%의 높은 부호화 효율을 가짐을 알 수 있다. 즉, 본 논문에서 제안하는 ITP는 IBC와 유사한 효과를 가지는 부분이 존재하는 것으로 보인다. 다만, 복잡도 측면에서 IBC와 ITP는 상당한 차이가 있다. 디코더 측에서 ITP는 참조 영역을 위한 추가적인 메모리를 필요로 하지 않는 반면, IBC는 한 개의 CTU 영역만큼 추가적인 메모리를 필요로 한다. 또한 ITP는 기존 화면내 예측 모드의 연산을 재활용하여 예측 블록을 획득하지만, IBC는 기존 화면내 부호화에서 수행하지 않는 움직임 예측 및 보상을 위한 복잡한 연산을 추가적으로 수행해야만 한다. 특히 IBC는 VVC 참조 소프트웨어

의 인코더 복잡도를 62%나 증가 시킨다^[20]. 이처럼 ITP는 IBC보다 낮은 계산 복잡도로 상당한 부호화 효율을 제공할 수 있기 때문에, 저 복잡도를 요구하는 응용에서 ITP는 IBC를 대체할 수 있는 스크린 콘텐츠용 화면내 부호화 틀이 될 수 있다.

앞선 표 3의 실험에서 ITP가 선택되는 경향을 확인하기 위해서, ITP의 부호화 효율이 가장 높은 영상인 ChineseEditing과 가장 낮은 부호화 효율을 갖는 FlyingGraphics에 대해 ITP가 선택된 블록을 그림 11과 같이 가시화하였다. 그림 11에서 박스로 그려진 블록은 RDO 관점에서 최적 모드로 ITP가 선택된 CU이다. ITP의 부호화 효율이 높은 만큼, 그림 11에서도 다양한 크기와 많은 수의 CU들이 ITP 모드로 부호화되었음을 확인할 수 있다. 그림 11에서 보이는 바와 같이 ChineseEditing 영상은 문자가 많다. 이 문자들은 사선 에지를 많이 포함하고 있기 때문에 ITP가 많이 선택되었으며, FlyingGraphics 영상은 ChineseEditing 영상보다 적은 문자 영역을 가지고 있고 사선 에지의 분포가 적어 비교적 낮은 수의 CU들이 ITP로 부호화되었다. 그림 11에서 ITP로 부호화된 CU 중 사선 에지가 없는 아주 평활한(Homogeneous)한 블록도 찾을 수 있다. 현재 CU의 상측과 좌측의 참조 샘플이 모두 동일한 값을 갖는다면 ITP와 67개 화면내 예측 모드 모두 동일한 예측 신호를 생성한다.

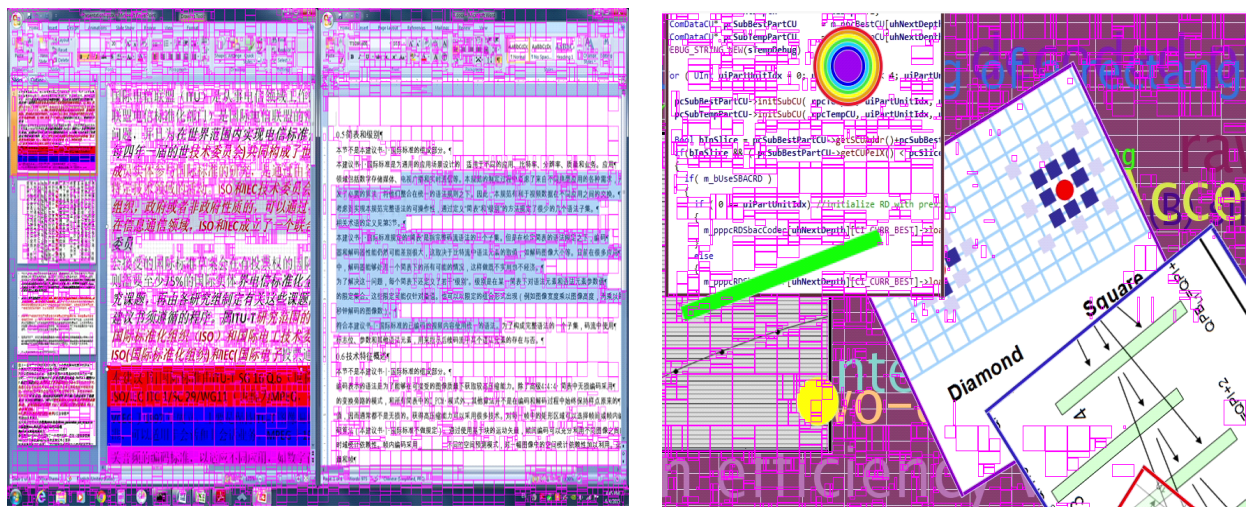


그림 11. ITP가 최적 모드로 선택된 CU, (좌) ChineseEditing 영상, (우) FlyingGraphics 영상
Fig. 11. CUs coded by ITP mode as the best mode, (left) ChineseEditing sequence, (right) FlyingGraphics sequence

표 4. 블렌딩 마스크 tap 수에 따른 ITP 부호화 효율
 Table 4. Coding efficiency of ITP according to the number of blending mask taps

Class	Sequence	3-tap (BD-rate, %)			7-tap (BD-rate, %)		
		Y	U	V	Y	U	V
SCC	ChineseEditing	-2.37	-1.97	-2.01	-1.48	-1.20	-1.25
	Console	-2.08	-1.57	-1.62	-1.19	-0.85	-0.86
	Desktop	-1.84	-1.37	-1.44	-1.47	-1.05	-1.09
	FlyingGraphics	-1.15	-1.05	-1.14	-0.31	-0.33	-0.41
Average		-1.86	-1.49	-1.55	-1.33	-0.85	-0.86

그런데 ITP는 cu_itp_flag만 1로 시그널링하고 더 이상의 구문요소를 시그널링하지 않지만, 화면내 예측 모드는 cu_itp_flag를 0으로 시그널링한 후 특정 화면내 예측 모드를 MPM 혹은 Remainer로 추가 시그널링해야 한다. 즉 구문요소 비트양에서 ITP가 유리하기 때문에 그림처럼 평활한 블록에서도 ITP가 선택되는 경우가 발생한다.

ITP는 블렌딩 과정에서 3-tap 마스크를 적용한다. 반면 ITP와 유사한 화면간 부호화 틀인 TPM은 최대 7-tap의 마스크를 사용한다. ITP의 3-tap 마스크의 적정성을 검증하기 위해 ITP 방법에 3-tap과 7-tap을 적용할 때의 부호화 효율을 비교해 보았다. 표 4는 그 실험 결과를 나타낸다. 블렌딩 마스크 3-tap을 적용한 경우, 실험 영상 ChineseEditing, Console, Desktop, FlyingGraphics에 대해서 Y 부호화 효율이 평균 1.86% 향상한 결과를 보이며, 7-tap을 적용한 경우에는 평균 1.33% 향상으로 3-tap 대비 낮은 결과를 보인다. 스크린 콘텐츠는 에지가 선명하고 색의 분포가 자연 영상에 비해 낮으므로 7-tap보다는 3-tap의 적은 수 화소에만 필터링을 가하는 것이 에지 보전 및 블러링을 막는데 효과적이라는 것을 확인할 수 있다.

V. 결론

AR/VR, 스트리밍 게임, 온라인 강의, 화상 회의 등의 활성화로 인해 컴퓨터에 의해 생성된 스크린 콘텐츠 압축이 많은 관심을 끌고 있다. 본 논문에서는 현재 표준화 중인 VVC 표준에 기반한 스크린 콘텐츠 부호화 틀을 제안했다. 기존 화면내 예측 부호화에서는 사각형 분할만을 지원하는 한계가 있으며, 문자와 그래픽 등으로 인해 사선 방향 에지

가 산재한 스크린 콘텐츠의 특성을 고려하지 않았다. 제안 방법은 화면내 부호화에서 두 개 삼각형 예측 신호를 생성함으로써 스크린 콘텐츠의 사선 방향 에지를 효과적으로 부호화한다. 이는 대표적 스크린 콘텐츠 부호화 기술인 IBC보다 낮은 복잡도로 평균 1.86%의 상당히 높은 부호화 효율을 갖기에 저복잡도를 요구하는 응용에서 IBC를 대체할 수 있는 부호화 기술이 될 수 있을 것으로 판단된다. 제안 방법은 고정된 두 개의 화면내 예측 모드를 이용하여 예측 신호를 생성한다. MPM과 주변 블록의 부호화 정보 등을 이용하여 적응적으로 화면내 예측 모드를 선택하는 연구를 추진한다면 스크린 콘텐츠 환경에서의 부호화 성능을 더 향상시킬 수 있을 것이라 기대한다.

참고 문헌 (References)

- [1] A Segall, V Baroncini, J Boyce, J Chen, T Suzuki, "Joint Call for Proposals on Video Compression with Capability beyond HEVC," ITU-T/ISO/IEC JVET, JVET-H1002, Oct. 2017. <https://mpeg.chiariglione.org/standards/exploration/future-video-coding/n17195-joint-call-proposals-video-compression-capability>
- [2] B. Bross, J. Chen, S. Liu, and Y.-K. Wang, "Versatile Video Coding (Draft 8)," ISO/IEC/ITU-T JVET, JVET-Q2001, Jan. 2020. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9675
- [3] J. Chen, Y. Ye, and S. Kim, "Algorithm Description for Versatile Video Coding and Test Model 8 (VTM 8)," ISO/IEC/ITU-T JVET, JVET-Q2002, Jan. 2020. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9676
- [4] High Efficiency Video Coding (HEVC), Rec. ITU-T H.265 and ISO/IEC 23008-2, Jan. 2013. <https://www.itu.int/itu-t/recommendations/rec.aspx?rec=12905>
- [5] Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, and Thomas Wiegand, "Overview of the High Efficiency Video Coding (HEVC)

- Standard,” IEEE Trans. Circuits Syst. Video Technol., Vol. 22, No. 12, pp. 1649-1668, Dec. 2012. <https://doi.org/10.1109/TCSVT.2012.2221191>
- [6] F. Bossen, X. Li, and K. Sühring, “JVET AHG Report: Test Model Software Development (AHG3),” ISO/IEC/ITU-T JVET, JVET-Q0003, Jan. 2020. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9363
- [7] B. Bross, “Versatile Video Coding (VVC) on the Final Stretch,” ITU workshop on the future of media, Geneva, Switzerland, Oct. 2019. <https://www.itu.int/en/ITU-T/Workshops-and-Seminars/20191008/Pages/programme.aspx>
- [8] Jong-Seok Lee, Jun-Taek Park, Han-Sol Choe, Ju-Hyeong Byeon, and Dong-Gyu Sim, “Overview of VVC,” Broadcasting and Media Magazine, Vol. 24, No. 3, pp.10-25, Oct. 2019.
- [9] Heeji Han, Jaeryun Choe, Daehyeok Gwon, and Haechul Choi, “VVC Intra Prediction and Coding Tools,” Broadcasting and Media Magazine, Vol. 24, No. 3, pp.39-54, , Oct. 2019.
- [10] Liang Zhao, Xin Zhao, Shan Liu and Xiang Li, “CE3-related: Unification of angular intra prediction for square and non-square blocks,” ITU-T/ISO/IEC JVET, JVET-L0279, Oct. 2018. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4374
- [11] D. Park, Y. Yoon, J.-G. Kim, J. Lee, and J. Kang, “CE10-related: Simplification of Triangular Partitions,” ISO/IEC/ITU-T JVET, JVET-M0352, Jan. 2019. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=5159
- [12] F. Chen and Li Wang, “CE10-related: Simplified Triangle Prediction Unit Mode,” ISO/IEC/ITU-T JVET, JVET-M0328, Jan. 2019. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=5135
- [13] X. Wang, Y.-W. Chen, X. Xiu, and T.-C. Ma, “CE4-related: An Improved Method for Triangle Merge List Construction,” ISO/IEC/ITU-T JVET, JVET-N0340, March. 2019. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=6061
- [14] Jungwon Kang and Hahyun Lee, “VVC Inter Prediction Coding Technologies,” Broadcasting and Media Magazine, Vol. 24, No. 3, pp.55-70, , Oct. 2019.
- [15] Ru-Ling Liao and Chong Soon Lim, “CE10.3.1.b: Triangular prediction unit mode,” ITU-T/ISO/IEC JVET, JVET-L0124, Oct. 2018. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4205
- [16] B. Bross, P. Keydel, H. Schwarz, D. Marpe, T. Wiegand, L. Zhao, X. Zhao, X. Li, S. Liu, Y.-J. Chang, H.-Y. Jiang, P.-H. Lin, C.-C. Kuo, C.-C. Lin, and C.-L. Lin, “CE3: Multiple Reference Line Intra Prediction (Test 1.1.1, 1.1.2, 1.1.3 and 1.1.4),” ISO/IEC/ITU-T JVET, JVET-L0283, Oct. 2018. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4379
- [17] J. Chen, Y. Ye, and S. H. Kim “Algorithm Description for Versatile Video Coding and Test Model 6 (VTM 6),” ISO/IEC/ITU-T JVET, JVET-O2002, July 2019. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=7826
- [18] F. Bossen, J. Boyce, X. Li, V. Seregin, and K. Sharing, “JVET Common Test Conditions and Software Reference Configurations for SDR Video,” ISO/IEC/ITU-T JVET, JVET-N1010, Mar. 2019. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=6643
- [19] G. Bjontegaard, “Calculation of Average PSNR Differences Between RD Curves,” ITU-T SG16 Q.6 VCEG, VCEG-M33, Apr. 2001. http://wftp3.itu.int/av-arch/video-site/0104_Aus/VCEG-M33.doc
- [20] W.-J. Chien, V. Seregin, and M. Karczewicz, “CE8-related: CPR Mode Signaling and Interaction with Inter Coding Tools,” ISO/IEC/ITU-T JVET, JVET-M0483, Jan. 2019. http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=5293

저 자 소 개



최 재 루

- 2018년 : 한밭대학교 정보통신공학과 학사
- 2020년 : 한밭대학교 멀티미디어공학과 석사
- ORCID : <https://orcid.org/0000-0002-3250-4153>
- 주관심분야 : 비디오 부호화, 영상처리, 패턴인식, 기계학습

저 자 소 개

권 대 혁



- 2012년 : 한밭대학교 멀티미디어공학과 학사 졸업
- 2014년 : 한밭대학교 멀티미디어공학과 석사 졸업
- 2020년 : 한밭대학교 멀티미디어공학과 박사 졸업
- ORCID : <https://orcid.org/0000-0003-2771-1921>
- 주관심분야 : 비디오 부호화, 영상처리, 패턴인식

한 희 지



- 2018년 : 한밭대학교 정보통신공학과 학사
- 2020년 : 한밭대학교 정보통신전문대학원 멀티미디어공학과 석사
- 2020년 ~ 현재 : 한밭대학교 정보통신전문대학원 멀티미디어공학과 박사과정
- ORCID : <https://orcid.org/0000-0001-8187-2553>
- 주관심분야 : 비디오 부호화, 영상처리, 패턴인식, 기계학습, 딥러닝

이 하 현



- 2002년 : 한국항공대학교 항공전자공학과 학사
- 2005년 : 과학기술연합대학원대학교(UST) 이동통신 및 디지털방송공학전공 공학석사
- 2008년 ~ 현재 : 한국전자통신연구원(ETRI) 선임연구원
- ORCID : <https://orcid.org/0000-0001-7043-7564>
- 주관심분야 : 영상 압축 및 영상신호처리

강 정 원



- 1993년 : 한국항공대학교 항공전자공학과 학사
- 1995년 : 한국항공대학교 항공전자공학과 공학석사
- 2003년 : Georgia Institute of Tech. ECE 공학박사
- 2003년 ~ 현재 : 한국전자통신연구원 책임연구원 및 표준전문위원
- ORCID : <https://orcid.org/0000-0003-4003-4638>
- 주관심분야 : 영상 압축, 영상신호처리

최 해 철



- 1997년 : 경북대학교 전자공학과 학사
- 1999년 : 한국과학기술원 전기및전자공학과 석사
- 2004년 : 한국과학기술원 전기및전자공학과 박사
- 2004년 9월 ~ 2010년 2월 : 한국전자통신연구원(ETRI) 방송미디어연구부 선임연구원
- 2010년 3월 ~ 현재 : 한밭대학교 정보통신공학과 교수
- ORCID : <http://orcid.org/0000-0002-7594-0828>
- 주관심분야 : 영상처리, 비디오 부호화, 패턴인식, 기계학습, 딥러닝