

특집논문 (Special Paper)

방송공학회논문지 제26권 제6호, 2021년 11월 (JBE Vol.26, No.6, November 2021)

<https://doi.org/10.5909/JBE.2021.26.6.758>

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

다중 모델을 이용한 완전연결 신경망 기반 화면내 예측

문기화^{a)}, 박도현^{a)}, 김민재^{a)}, 권형진^{b)}, 김재곤^{a)†}

Intra Prediction Using Multiple Models Based on Fully Connected Neural Network

Gihwa Moon^{a)}, Dohyeon Park^{a)}, Minjae Kim^{a)}, Hyoungjin Kwon^{b)}, and Jae-Gon Kim^{a)†}

요 약

최근 딥러닝 기술을 비디오 부호화에 적용하는 다양한 연구가 진행되고 있다. 본 논문은 차세대 비디오 코덱인 VVC(Versatile Video Coding)에 채택된 신경망 기반의 기술인 MIP(Matrix-based Intra Prediction)를 확장한 완전연결계층(Fully Connected Layer) 기반의 다중 모델을 이용하는 화면내 예측 부호화 기법을 제시한다. 또한 다중 화면내 예측 모델을 위한 효율적인 학습기법을 제안한다. HEVC(High Efficiency Video Coding)에서의 성능검증을 위해 VVC의 MIP와 제안하는 완전연결계층 기반 다중 화면내 예측 모델을 HEVC의 참조 소프트웨어인 HM16.19에 추가적인 화면내 예측모드로 구현하였다. 실험결과 제안하는 방법이 HM16.19와 VVC MIP 대비 각각 0.47%과 0.19% BD-rate 성능향상이 있음을 확인하였다.

Abstract

Recently, various research on the application of deep learning to video encoding for enhancing coding efficiency are being actively studied. This paper proposes a deep learning based intra prediction which uses multiple models by extending Matrix-based Intra Prediction(MIP) that is a neural network-based technology adopted in VVC. It also presents an efficient learning method for the multi-model intra prediction. To evaluate the performance of the proposed method, we integrated the VVC MIP and the proposed fully connected layer based multi-model intra prediction into HEVC reference software, HM16.19 as an additional intra prediction mode. As a result of the experiments, the proposed method can obtain bit-saving coding gain up to 0.47% and 0.19% BD-rate, respectively, compared to HM16.19 and VVC MIP.

Keyword : Fully connected layer, HEVC, VVC, MIP, Multi-model network, Intra prediction

a) 한국항공대학교 항공전자정보공학부(Korea Aerospace University, School of Electronics and Information Engineering)

b) 한국전자통신연구원(Electronics and Telecommunications Research Institute)

† Corresponding Author : 김재곤(Jae-Gon Kim)

E-mail: jgkim@kau.ac.kr

Tel: +82-2-300-0414

ORCID: <https://orcid.org/0000-0003-3686-4786>

※ 이 논문의 연구 결과 중 일부는 한국방송 미디어공학회 “2021년 하계학술대회”에서 발표한 바 있음.

※ 본 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. 2017-0-00072, 초실감 테라미디어를 위한 AV부호화 및 LF미디어 원천기술 개발).

· Manuscript received September 17, 2021; Revised November 5, 2021; Accepted November 5, 2021.

Copyright © 2021 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.”

1. 서론

HEVC(High Efficiency Video Coding) 이후 UHD TV 및 360도 비디오 등 다양한 고화질 비디오 서비스의 확산으로 보다 효율적인 압축을 위한 새로운 비디오 부호화 표준의 필요성이 대두되었다^[1]. 이에 ITU-T VCEG(Video Coding Experts Group)와 ISO/IEC MPEG(Moving Picture Experts Group)의 공동 작업반인 JVET(Joint Video Experts Team)은 2020년 7월 차세대 비디오 표준인 VVC(Versatile Video Coding)/H.266 표준화를 완료하였다^[2]. VVC는 부호화 성능 향상을 위해 MIP(Matrix-based Intra Prediction) 및 LFNST(Low Frequency Non-Separable Transform)와 같은 학습된 신경망 모델을 이용한 부호화 기술들을 채택하였다^[3]. 또한 다양한 인공신경망 구조 및 학습기법이 연구되고 영상처리 분야에서도 뛰어난 성능을 보이면서 비디오 부호화에도 인공신경망을 접목시키는 연구의 필요성이 대두되었다. 이에 JVET은 VVC를 확장할 수 있는 신경망 기반의 비디오 부호화 기술의 잠재성을 확인하기 위해 NNVC AhG(Ad-hoc Group for Neural Network-based Video Coding)을 구성하고 관련 기술을 탐색하고 있다^[4]. 현재 NNVC AhG에서 논의되고 있는 딥러닝 기반의 부호화 기술들은 기존의 특정 압축 기술을 대체하거나 기존의 부호화 모드에 새로운 모드를 추가하는 방식으로 연구되고 있으며 인-루프(In-Loop) 필터링, RPR(Reference Picture Resampling), 그리고 화면내 예측 등의 세부 분야에서 주로 연구되고 있다^[5].

기존 화면내 예측 기술은 예측하고자 하는 블록 주변의 복호화가 완료된 참조샘플들을 특정 각도의 방향성으로 패딩(padding)하는 방향성 모드, 참조샘플의 변화량을 반영하여 예측하는 Planar 모드 및 참조샘플의 평균으로 예측블록을 생성하는 DC 모드와 같이 미리 정의된 단순한 패턴을 이용하여 예측블록을 생성한다. 따라서, 기존 화면내 예측 모드는 부호화 블록 내부에 복잡한 패턴이 존재하거나 참조샘플과 예측샘플의 거리가 먼 경우에는 예측성능이 떨어지는 한계를 갖는다. 이러한 한계를 극복하고 화면내 예측 부호화 성능을 향상시키기 위하여 다중의 참조샘플 라인을 입력하는 학습된 인공신경망 기반의 화면내 예측 모드에 대한 연구가 활발하게 수행되고 있다^{[6][7]}.

본 논문에서 제안하는 화면내 예측 모델은 VVC의 MIP보다 깊은 완전연결계층(FC Layer: Fully Connected Layer)을 갖는 신경망 기반의 화면내 예측 모델을 사용하며 VVC MIP와 유사하게 같은 블록 크기에 대해 복수개의 화면내 예측 모델로 구성된다. 또한, 학습과정에서 각 모델들의 성능비교를 통해 최대의 성능을 갖는 학습 데이터와 모델을 분류하고 학습하여 각각의 모델들이 서로 다른 예측블록을 생성하도록 한다. 제안기법의 성능향상을 검증하기 위하여 VVC의 MIP와 제안하는 화면내 예측 모델을 HEVC의 참조 소프트웨어인 HM(HEVC reference software Model)에 화면내 예측의 추가적인 모드로 구현한 후 부호화 성능을 비교한다.

본 논문의 구성은 다음과 같다. 제2장에서는 VVC에서 채택된 MIP 기술과 완전연결계층 기반의 화면내 예측 모델을 소개하고, 제3장에서는 본 논문에서 제시하는 다중모델 학습기법을 제시한다. 제4장에서는 상세 실험내용과 실험 결과를 분석하며, 마지막으로 제5장에서 결론을 맺는다.

II. VVC MIP 및 완전연결계층 기반 다중 화면내 예측 모델

1. VVC MIP

MIP는 VVC에 새롭게 채택된 신경망 기반의 화면내 예측 기술이다. $W \times H$ 의 직사각형 블록샘플을 예측하기 위하여 블록 좌측의 H 개의 복원된 참조샘플 한 라인과 블록 상단의 W 개의 복원된 참조샘플 한 라인을 입력으로 사용한다. 예측신호를 생성하는 방법은 그림 1과 같이 평균화, 행렬 곱 그리고 선형보간 단계로 이루어진다. 참조샘플의 감소를 위한 평균화와 선형보간은 MIP의 복잡도 감소를 위한 것이며 감소된 참조샘플은 사전에 학습결과로 정의된 행렬과의 곱셈 곱으로 예측샘플을 생성한다. 입력되는 참조샘플의 크기는 예측하려는 블록의 크기에 따라 다르게 정의 되어 있는데 4×4 코딩 블록에 대해서는 30 개, 4×4 보다 크며 8×8 이하의 코딩블록에 대해서는 16 개, 8×8 보다 큰 블록에 대해서는 6 개의 MIP 예측 모드가 정의되어 있다.

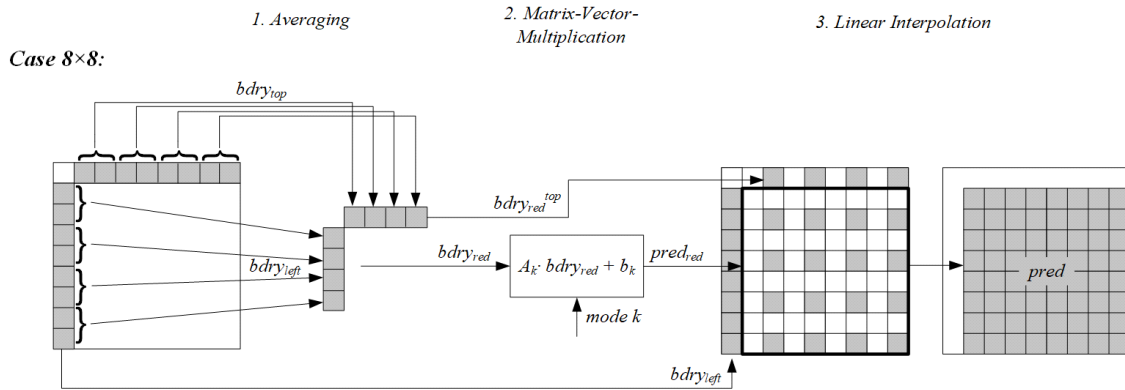


그림 1. 8 × 8 블록을 위한 VVC MIP
 Fig. 1. Illustration of VVC MIP for 8 × 8

2. 완전연결계층 기반 다중 모델

비디오의 공간적 중복성을 이용하여 복원된 주변 샘플을 통해 블록을 예측하는 화면내 예측은 시간적 중복성을 이용하는 화면간 예측과 비교하여 상대적으로 예측성능이 높지 않다. IPFCN(Intra prediction using fully connected network)은 기존 화면내 예측의 단점을 극복하기 위하여 인공신경망을 적용하였고, 화면내 예측의 부호화 성능을 높이기 위하여 완전연결계층으로 네트워크를 구성하고 다수의 참조샘플을 사용한다^[8]. 그림 2는 IPFCN의 네트워크 구조와 입출력 형태이다. L 은 참조샘플의 라인수를 의미하며, $N \times N$ 크기의 블록에 대해 총 $4NL + L^2$ 픽셀을 입력으로 사용한다. IPFCN의 출력은 N^2 크기의 벡터이며 네트워크

는 모두 d 개의 K_i 차원의 완전연결계층을 이용한다. 해당 논문에서는 네트워크 깊이에 대한 실험을 통해 계층이 깊어질수록 성능향상이 있음을 확인하였고, 각 층의 파라미터 개수가 증가할수록 예측성능은 높아지나 이에 따른 개선 폭은 크지 않음을 확인하였다. 최종적으로 8×8 크기의 블록에 대해서 128 개의 파라미터를 가지는 세 개의 완전연결계층을 이용하여 학습하였고, 해당 단일 모델을 화면내 예측의 추가적인 모드로 구현하여 부호화 성능을 확인하였다.

각 블록 크기에 대해 단일 모델을 사용하는 경우 코덱에 구현해야 할 모델의 수가 크게 감소하기 때문에 비디오 부호화의 복잡도 측면에서 높은 효율을 얻을 수 있다. 그러나 단일 모델을 적용할 때, 해상도나 블록 샘플의 특성에 따른

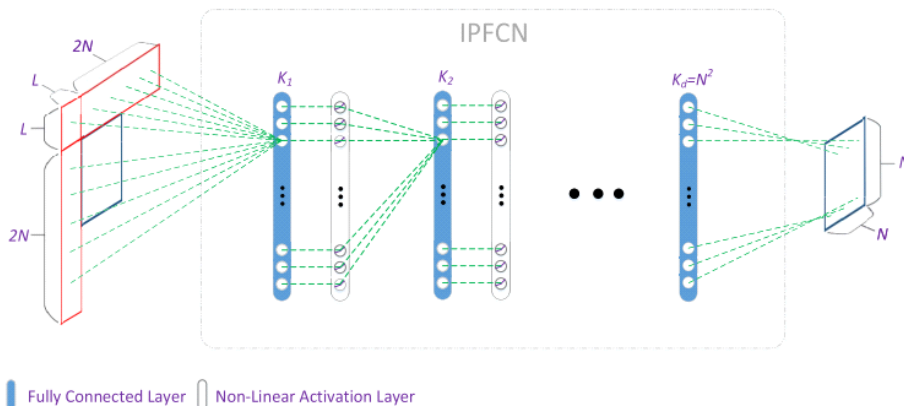


그림 2. 완전연결계층 기반 다중 모델 네트워크 (IPFCN)
 Fig. 2. Network architecture of multi-model based on fully connected layers

여러 특징들을 모두 반영하기에는 한계가 있다. 영상의 다양한 패턴을 모두 반영하여 학습하다 보면 특정 패턴을 위한 모델보다는 평균값이 적용된 예측영상을 생성하게 되며 이는 공개된 영상을 유도한다. 이러한 한계를 극복하고자 해당 본 논문에서는 서로 다른 특징을 가진 여러 개의 모델을 학습하여 화면내 예측의 추가적인 모드로 구현하였다.

III. 제안 다중 예측 모델의 학습기법

본 장에서는 제2장에서 기술한 완전연결계층 기반의 다중 화면내 예측 모델을 위한 학습기법을 제안한다. 학습에서 사용된 화면내 예측 모델의 구조 및 학습기법에 대해 상세히 기술하고, 학습된 다중 화면내 예측 모델의 예측영상을 통해 각 모델이 다른 특성을 갖도록 학습되는 것을 보인다.

1. 네트워크 구조

VVC에 채택된 MIP에서 사용된 벡터 값은 한 개의 완전연결계층으로 학습된 행렬이다. 제안기법은 네 개의 완전연결계층을 이용해 보다 깊은 네트워크를 구성하였다. 또한 입력은 예측하려는 블록에서 가장 가까운 위치의 샘플들로 참조샘플을 구성하는 기존 MIP 방식과 다르게 주변 참조샘플의 문맥적 특성을 반영함으로써 방향성 모드로 예측하기 어려운 부호화 블록에 대해서도 예측성능을 높이기 위하여 좌측과 상단의 네 개 라인을 참조샘플로 사용하였다. 따라서, 8×8 크기의 블록에 대해 총 144 개의 참조샘플을 입력으로 마지막 계층을 제외하고 모두 1024의 깊이를 가진 완전연결계층을 사용한다. 마지막 계층은 8×8 영상이 될 수 있도록 총 64 개의 예측샘플을 출력하도록 한다. 제안하는 네트워크 구조는 총 8 개의 완전연결계층 기반의 화면내 예측 모델로 구성되어 있으며, 학습과정에서 각 모델이 서로 다른 특징의 영상에 대해 학습될 수 있도록 도움을 주기위해 최적의 모드를 판별해주는 모드예측 모델을 사용한다. 모드예측 모델의 입력은 화면내 예측 모델의 입력과 같으며 정의된 모드의 수에 맞는 8 개의 확률을 출력한다. 학습과정에서 모드예측 모델은 학습데이터에 대해

손실이 최소인 화면내 예측 모델의 인덱스에 높은 확률을 출력하도록 학습된다.

2. 학습기법

손실함수는 높은 예측성능을 유지함과 동시에 비디오 부호화의 효율을 향상시키는 방향으로 식 1과 같이 정의한다 [9]. \check{k} 는 8 개의 모드 중 최적의 비용을 갖는 모드를 의미하며 각 \check{k} 에 대한 손실인 $loss_{pred}(\check{k})$, 모드예측 모델 결과에 대한 크로스 엔트로피(cross entropy) 손실인 $loss_{mode}(\check{k})$, 그리고 실제 모드를 전송할 때 발생하는 비트수인 $loss_{bit}(\check{k})$ 를 더해서 최종 손실함수로 사용한다. 최적의 모드(\check{k})는 $loss_{pred}(k)$ 의 값이 최소가 되는 모드로 선택된다.

$$L(\check{k}) = loss_{pred}(\check{k}) + loss_{mode}(\check{k}) + loss_{bit}(\check{k}) \quad (1)$$

$$\check{k} = \underset{k}{\operatorname{argmin}} loss_{pred}(k) \quad (2)$$

$loss_{pred}(k)$ 는 비디오 부호화의 변환 및 양자화를 반영하는 손실함수이며 식 3과 같이 정의한다. 식 4에서 T 는 2D DCT-2 변환함수, O 와 P 각각 원본화소와 예측화소를 의미한다. 즉, C 는 원본블록과 예측블록의 잔차블록에 대해 2D DCT 변환을 수행하여 얻은 변환계수를 의미하며, 각 변환계수에 대한 절대값의 합인 SATD(Sum of Absolute Transformed Difference)를 손실함수로 사용한다. SATD와 함께 변환계수가 0에 빠르게 수렴할 수 있도록 로지스틱(logistic) 함수(g)를 이용하여 각 모드의 예측에 대한 손실함수를 정의하여 학습을 진행한다. $loss_{bit}(k)$ 는 비디오 부호화 과정에서 전송해야 하는 비트의 수가 줄어들어 압축에서 이득을 얻는 것을 반영한다. 모드예측 모델로 얻은 모드와 최적의 모델의 모드 값의 차이에 대한 비트 값을 손실함수로 정의한다.

$$loss_{pred}(k) = |C_k| + ag(\beta|C_k| - \gamma) \quad (3)$$

$$c_k = T(O - P_k) \quad (4)$$

$$g(x) = 1/(1 + e^{-x}) \tag{5}$$

$$loss_{mode}(q(r), k) = -\log\left(\frac{\exp(q(r)_k)}{\sum_i \exp(q(r)_i)}\right) \tag{6}$$

학습은 8 × 8 블록에 대해 8 개의 모델로 진행한다. COCO 데이터셋^[10]을 학습 데이터셋으로 사용하였고, 각 영상을 128 × 128 CTU 단위로 분리하고 배치 크기는 64로 설정한다. 본 논문에서 제안하는 화면내 예측 모델을 학습하기 위하여 Adam 최적화기를 이용하고 학습이 거듭될수록 파라미터들을 미세 조정해주기 위한 학습률 감소 기법을 적용하였다.

3. 학습 모델의 주관적 성능평가

그림 3과 그림 4는 제안방법으로 학습된 모델이 생성한 예측블록을 보인 것이다. 그림 3의 (a)는 BVI-DVC의 ‘CDowntownHongKongVideo’ 시퀀스의 일부 128 × 128 원본영상이다. 그림 3의 (b)는 (a)에 해당하는 8 × 8 블록에 대해 PSNR 값을 기준으로 최적으로 선택된 다중 화면내

예측 모델들에 대한 예측블록이다. 그림 4는 제안하는 다중 화면내 예측 모델에 대한 각 모델 별 예측블록을 보인다. 그림 4의 모델 3은 DC와 유사하게 방향성이 거의 없고, 모델 4는 수직 수평의 방향성을 보이며 모델 2는 대각선 방향의 패턴으로 예측블록을 생성하는 등 각 모델들이 서로 다른 패턴 및 특징을 갖으며 예측블록을 생성함을 확인할 수 있다.

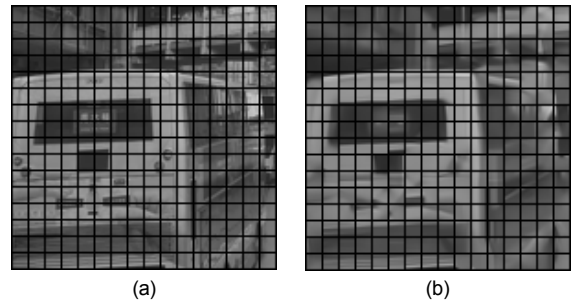


그림 3. 8 × 8 블록을 위한 제안기법 모델의 예측블록 결과 영상(휘도샘플)
 (a) 원본영상 (b) 블록 별 최적의 모델 예측영상
 Fig. 3. Prediction block image result of the proposed model for 8 × 8 blocks
 (a) Original image (b) Optimal model prediction image for each block

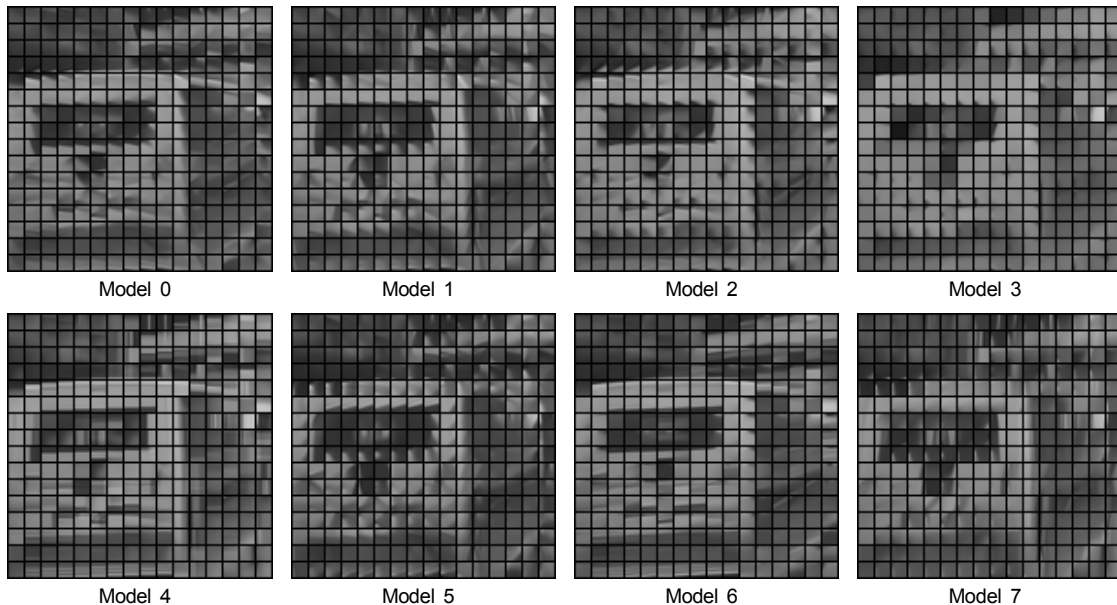


그림 4. 8 × 8 블록에 대한 다중 모델의 예측블록 결과 영상(모드 0~7)
 Fig. 4. Prediction block image result of the 8 multi-model (Mode 0 ~ 7) for 8 × 8 block

IV. 비디오 부호화 실험결과

본 논문의 실험은 VVC의 MIP와 제안하는 모델의 성능 비교를 위하여 각각 HEVC 참조 소프트웨어 HM16.19를 사용하였다. 최근 차세대 비디오 코덱으로 표준화가 완료된 VVC에는 MIP 이외의 비디오 부호화 성능향상을 위한 여러 가지 모드가 채택되어 있다. 인공신경망 기반으로 학습된 VVC의 화면내 예측기법인 MIP와 제안하는 학습기법을 반영한 화면내 예측 모델의 성능향상을 단순 비교하기 위하여 각 모드들을 HEVC 화면내 예측의 추가적인 예측 모드로 구현하였다.

MIP는 기존 VVC에 정의되어 있는 MIP 행렬 값들의 곱을 이용하여 예측하였고, 제안하는 기법은 8 개의 학습된 모델이 HEVC 내에서 직접 수행되도록 하였다. 두 실험 모두 JVET CTC(Common Test condition) 의 A, B, C, D 클래스 시퀀스를 사용하였고, AI(AI-intra) 부호화 모드에서

22, 27, 32, 37의 QP 로 동일하게 진행하였다^[11]. 또한 부호화 과정에서 수행되는 블록의 크기를 8 x 8로 제한하여 실험을 진행하였으며, 각 클래스의 결과는 각 클래스에 해당하는 모든 시퀀스들의 BD-rate 평균값이다.

표 1은 VVC의 MIP 기술을 HM16.19에 구현하였을 때 HM16.19 대비 부호화 성능이며 휘도성분(Y)에서 0.27%의 BD-rate 성능향상을 보였다. 표 2는 제안하는 완전연결계층 기반의 다중 화면내 예측 모델에 대한 HM16.19 대비 부호화 성능이며, Y에서 0.47% 성능향상을 보였다. 표 1과 표 2의 비교를 통해, VVC MIP 대비 제안하는 기법이 Y, Cb, Cr 성분에 대해 각각 0.19%, 0.08%, 0.02% 성능향상이 있음을 확인하였다.

그림 5는 VVC MIP 및 제안기법의 주관적 성능을 보여주는 영상으로 JVET CTC 중 D 클래스(416 x 240)의 'BasketballPass'의 한 프레임에 대한 HM16.19, VVC MIP를 구현한 HM16.19 및 제안기법의 부호화 결과 영상이다.

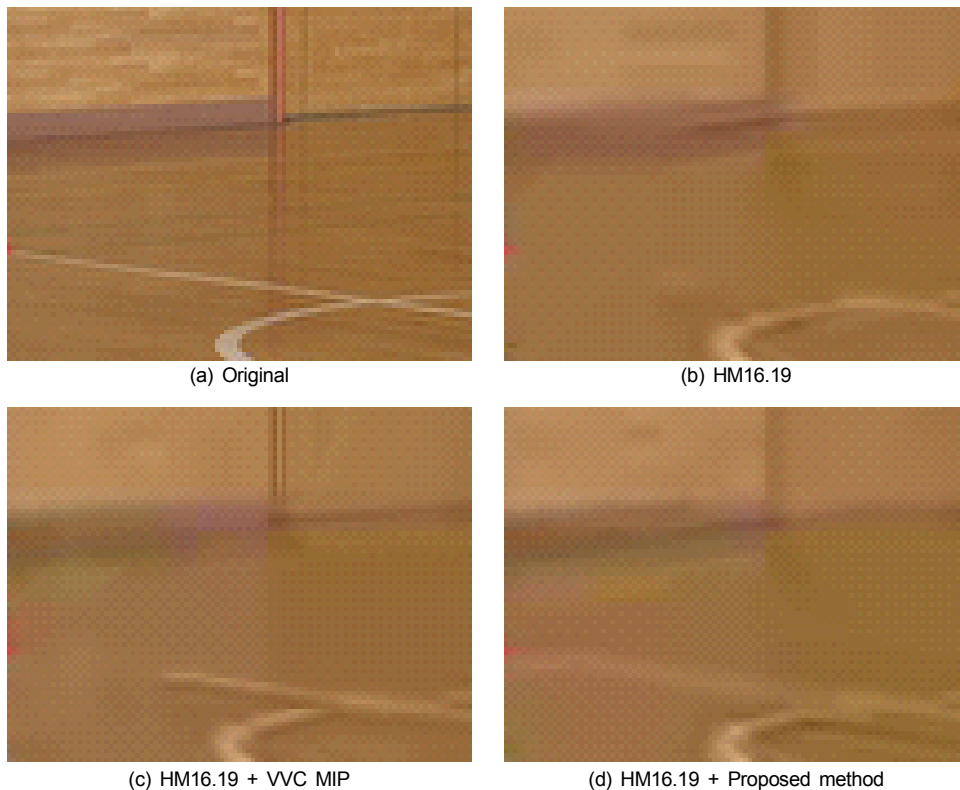


그림 5. VVC MIP 및 제안기법에 대한 주관적 화질 비교(BasketballPass, QP 37)

Fig. 5. Subjective quality comparison of the reconstructed images for BasketballPass with AI configuration

표 1. HM16.19 대비 VVC MIP 부호화 성능
Table 1. BD-rate coding performance of VVC MIP (Anchor: HM16.19)

	All Intra Main10		
	Over HM16.19		
	Y	U	V
Class A1	-0.38%	0.11%	0.06%
Class A2	-0.27%	0.06%	0.21%
Class B	-0.30%	0.12%	0.14%
Class C	-0.22%	0.30%	0.24%
Class D	-0.21%	0.02%	0.00%
Overall	-0.27%	0.12%	0.13%

표 2. HM16.19 대비 제안기법의 부호화 성능
Table 2. BD-rate coding performance of prediction models proposed method (Anchor: HM16.19)

	All Intra Main10		
	Over HM16.19		
	Y	U	V
Class A1	-1.01%	-0.18%	-0.09%
Class A2	-0.52%	0.11%	0.01%
Class B	-0.45%	0.02%	0.09%
Class C	-0.27%	0.23%	0.21%
Class D	-0.24%	0.02%	-0.03%
Overall	-0.47%	0.05%	0.11%

HEVC(b)와 VVC MIP(c) 대비 제안기법(d)의 결과 영상에서 농구장의 라인 부분이 원본과 유사하게 예측되는 등 유의미한 부분의 블록에 대한 예측을 잘 수행하는 것을 확인할 수 있었다. 따라서 제안하는 방법의 화면내 예측 모델이 HEVC 와 VVC MIP 대비 객관적 성능과 주관적 성능 측면에서 모두 개선이 있음을 확인하였다.

V. 결론

차세대 비디오 코덱으로 최근 표준화가 완료된 VVC에 처음으로 인공지능망 기술을 비디오 부호화 기술에 적용하는 기술인 MIP가 채택되었다. 본 논문에서는 부호화 성능향상을 위해서 MIP를 확장한 완전연계층 기반의 화면내 예측 모델과 그 학습기법을 제안하였다. 제안기법을 HEVC 화면내 예측의 추가적인 모드로 적용하여 실험한 결과 HM16.19 대비 0.47% BD-rate 비트 절감의 성능향상이 있음을 확인하였고, VVC MIP 대비 0.19% BD-rate 성

능향상이 있음을 확인하였다. 특히, 이는 기존 VVC MIP 기술에서 8 x 8 에 대해 총 16 개의 모드를 사용한 반면 제안하는 기법은 8 개의 모드만을 추가하여 구현하였기에 유의미한 성능결과로 볼 수 있다. 그러나, 본 논문에서는 8 x 8의 한정적인 블록 크기에 대해서만 모델을 학습하고 부호화 성능을 확인하였다. 추가적인 블록 크기에 대한 학습과 성능확인에 대한 연구가 필요하다.

참고 문헌 (References)

- [1] High Efficiency Video Coding, Version 1, Rec. ITU-T H.265, ISO/IEC 23008-2, Jan. 2013.
- [2] Versatile Video Coding, ISO/IEC FDIS 23090-3, Jul. 2020.
- [3] J. Chen, Y. Ye, S. Kim, "Algorithm description for Versatile Video Coding and Test Model 13 (VTM 13)," Joint Video Experts Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29, JVET-V2002, Apr. 2021.
- [4] S. Liu, E. Alshina, J. Pfaff, M. Wien, P. Wu and Y. Ye, "JVET AHG report: Neural-network-based video coding," Joint Video Experts Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29, JVET-V0011, Apr. 2021.
- [5] Alshina, S. Lui, W. Chen, F. Galpin, Y. Li, Z. Ma, H. Wang, "EE1: Summary of Exploration Experiments on Neural Network-based Video Coding," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29, JVET-W0023, Jul. 2021.
- [6] "Use cases and requirements for Deep Neural Networks based Video Coding," ISO/IEC JTC 1/SC 29/WG 2, N22, Oct. 2020.
- [7] T. Dumas, A. Roumy and C. Guillemot, "Context Adaptive Neural Network Based Prediction for Image Compression," IEEE Trans. Image Proc., vol. 29, Aug. 2019.
- [8] J. Li, B. Li, J. Xu and R. Xiong, "Intra prediction using fully connected network for video coding," In Proc. IEEE International Conference on Image Processing (ICIP) 2017, IEEE, Sept. 2017.
- [9] P. Helle, J. Pfaff, M. Schäfer, R. Rischke, H. Schwarz, D. Marpe, and T. Wiegand, "Intra Picture Prediction for Video Coding with Neural Networks," In Proc. DCC 2019, IEEE, Mar. 2019.
- [10] T. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. Zitnick, and P. Dollár, "Microsoft COCO: Common Objects in Context," 2015, arXiv:1405.0312.
- [11] J. Boyce, K. Suehring, X. Li, and V. Seregin, "JVET common test conditions and software reference configurations," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-J1010, Apr. 2018.
- [12] M. Kim, G. Moon, D. Park, H. Kwon and J. Kim, "Intra Prediction Using Multiple Models Based on Fully Connected Layer," In Proc. KIBM E Annual Summer Conf. June. 2021.

저 자 소 개



문 기 화

- 2021년 2월 : 한국항공대학교 소프트웨어학과 학사
- 2021년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학부 석사과정
- ORCID : <https://orcid.org/0000-0001-6727-7790>
- 주관심분야 : 비디오 부호화, 영상처리, 딥러닝



박 도 현

- 2016년 2월 : 국립한밭대학교 멀티미디어공학 학사
- 2018년 2월 : 한국항공대학교 항공전자정보공학과 석사
- 2018년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 박사과정
- ORCID : <https://orcid.org/0000-0002-5873-0132>
- 주관심분야 : 비디오 부호화, 영상처리, 360 VR 비디오, 머신러닝



김 민 재

- 2016년 3월 : 한국항공대학교 항공전자공학부 입학
- 2022년 2월 : 한국항공대학교 항공전자공학부 졸업예정
- ORCID : <https://orcid.org/0000-0002-1635-0780>
- 주관심분야 : 전자공학, 멀티미디어



권 형 진

- 1997년 2월 : KAIST 전기 및 전자공학과 학사
- 2001년 2월 : KAIST 전기 및 전자공학과 석사
- 2001년 3월 ~ 현재 : 한국전자통신연구원 책임연구원
- ORCID : <https://orcid.org/0000-0002-2888-3154>
- 주관심분야 : 영상 신호처리, 영상 압축, 기계학습



김 재 곤

- 1990년 2월 : 경북대학교 전자공학과 학사
- 1992년 2월 : KAIST 전기 및 전자공학과 석사
- 2005년 2월 : KAIST 전기 및 전자공학과 박사
- 1992년 3월 ~ 2007년 2월 : 한국전자통신연구원(ETRI) 선임연구원/팀장
- 2001년 9월 ~ 2002년 7월 : Columbia University 연구원
- 2015년 12월 ~ 2016년 1월 : UC San Diego, Visiting Scholar
- 2007년 9월 ~ 현재 : 한국항공대학교 항공전자정보공학부 교수
- ORCID : <http://orcid.org/0000-0003-3686-4786>
- 주관심분야 : 비디오 부호화 표준, 비디오 신호처리, Immersive Video, Deep Learning