

## HEVC의 CNN 기반 색차신호 화면내 예측 기법

여충광, 문현철, 윤용욱, 김재곤

한국항공대학교

{cndrhkd17, hcmoon, yuyoon}@kau.kr, jgkim@kau.ac.kr

### A CNN-Based Method for Chroma Intra Prediction in HEVC

ChungKhang Yeo, HyeonCheol Moon, Yong-Uk Yoon, and Jae-Gon Kim  
Korean Aerospace University

#### 요 약

HEVC(High Efficiency Video Coding)의 색차신호 화면내 예측(Intra Prediction)은 복호화된 주변 화소로부터의 예측과 동일한 위치의 휘도신호의 예측 모드를 이용한 예측을 수행한다. 본 논문에서는 색차신호 화면내 예측의 성능 향상을 위하여 합성곱 신경망(CNN: Convolutional Neural Network) 기반의 색차신호 예측 기법을 제안한다. 제안하는 기법은 복원된 주변 블록의 휘도 및 색차신호로부터 CNN 을 이용하여 현재블록의 색차신호를 예측한다. 실험결과 제안한 CNN 기반의 색차신호 예측 기법이 HEVC 의 색차신호 화면내 예측보다 향상된 성능을 보임을 확인하였다.

#### 1. 서론

HEVC 는 복호화된 주변 참조샘플을 이용하여 다양한 방향성 모드로 화면내 예측을 수행한다. 휘도신호는 DC/Planar 및 33 가지의 방향성 모드를 포함한 35 모드에 대해 화면내 예측을 수행하며, 색차신호 화면내 예측은 4 가지 예측 모드와 참조하는 휘도신호의 모드를 이용한다. 다양한 모드의 화면내 예측은 예측의 정확도를 높여주지만, 여전히 참조샘플과 거리가 먼 화소의 예측은 정확도가 떨어진다. 특히, 색차신호는 제한된 예측 모드와 휘도성분의 예측 모드에 의존하기 때문에 예측 성능이 제한된다. 이에 따라 다양한 색차신호의 화면내 예측 기법이 제시되고 있으며, VVC(Versatile Video Coding)에는 CCLM(Cross Component Linear Model) 모드가 색차신호를 예측하는 모드로 채택되었다. CCLM 은 복호화된 휘도성분과 색차성분의 참조샘플간의 관계로부터 선형모델을 도출하여 현재블록에 대응되는 휘도신호로부터 색차신호를 예측한다. CCLM 은 높은 부호화 성능을 보여주며, 휘도신호와 색차신호와의 관계를 선형모델로 모델링할 수 있음을 보여준다.

본 논문에서는 색차신호의 화면내 예측 성능향상을 위하여 합성곱 신경망(CNN: Convolutional Neural Network)을 이용한 예측 기법을 제안한다. VVC 의 CCLM 과 유사하게 휘도신호와 색차신호의 주변 참조샘플들간의 관계를 CNN 을 이용하여 형성하고, 복호화된 주변화소들을 CNN 의 입력으로 하여 현재블록의 색차신호로 예측한다.

#### 2. 제안 기법

선형모델을 이용하여 휘도신호로부터 색차신호를 예측하는 VVC 의 CCLM 과 유사하게 휘도신호와 색차신호 간의 관계를 활용하는 기존의 CNN 기반의 색차신호 예측 기법[1]은 그림

1 과 같이 휘도성분과 색차성분을 CNN 의 입력으로 사용한다.

본 논문에서는 기존의 방법과 달리 그림 2 와 같이 주변 색차샘플만을 입력으로 하는 CNN 기반의 색차신호 화면내 예측기법을 제안한다. 제안한 CNN 구조는 16x16 블록 크기의 색차신호를 가정한다. 입력되는 Cb 와 Cr 의 주변블록으로 구성된 3 개 채널 데이터에 대해서 Cb 와 Cr 의 특성 차이를 유지하기 위하여 그림 2 와 같이 병렬적으로 각 3 개의 컨볼루션 층을 쌓는다. Cr, Cr 각각의 입력이 컨볼루션 층을 통과한 결과 특징 맵들을 채널별로 연결한다. 즉, 완전연결층(Fully Connected layer: FC layer)을 사용하기 위해 1 차원으로 플래트닝(flattening) 한다. 플래트닝 후 3 개의 FC 층을 사용하여 512 개 데이터를 출력하고 현재블록의 형태에 맞도록 16x16 2 채널로 재배열한다.

제안하는 CNN 을 학습시키기 위해 JVET CTC(Common Test Condition) [2]의 테스트 시퀀스 중 Class A, B, C 의 12 개 시퀀스, 테스트 데이터는 학습 데이터와 중복되지 않는 시퀀스 3 개를 사용하였다. 또한, 학습을 위한 복호화 된 주변샘플은 HM 16.0 을 사용하여 AI(All Intra) 모드로 부/복호화된 데이터를 사용하였다. 학습 데이터는 QP22, 테스트 데이터는 QP 22, 27, 32, 37 을 사용하여 실험을 진행하였다. 또한, 학습을 위한 손실 함수는 MAE(Mean Absolute Error)를 사용하였다.

학습 과정 중에 손실 값의 수렴을 가속화하기 위해 입력 데이터의 평균을 0 으로 근사하게 만든다[3]. 이를 위해 입력 되는 학습 데이터에 대한 평균  $\alpha$  를 구하여 식 (1)과 같이 입력 데이터 평균 값을 0 으로 근사화한 후 입력한다. 식 (2)와 같이 입력 전 계산한 평균을 전송하여 신경망 출력결과에 전송된 평균을 더한 후 화소 값 범위로 클리핑(clipping) 한다.

$$X_c = X - \alpha \quad (1)$$

$$Y = \min(\max(Y_c + \alpha), 0, 255) \quad (2)$$

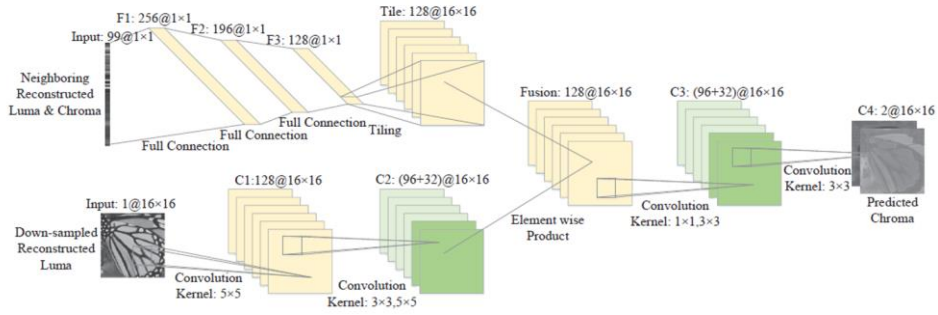


그림 1. 색차신호 예측 네트워크 구조[1]

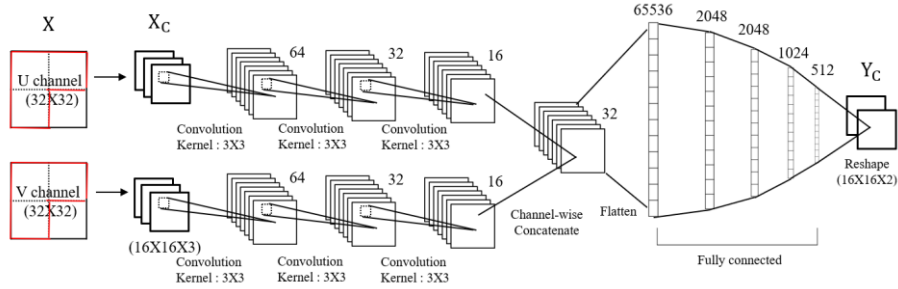


그림 2. 제안하는 CNN 구조(예측블록 크기: 16x16)

#### 4. 실험결과

제안한 CNN 기반의 색차신호 예측 기법의 성능을 그림 1의 기존의 CNN 기반 예측 기법[2]과 비교하였다. 즉, 두 CNN 기반의 예측 기법의 성능을 HEVC 색차신호 예측성능 기준으로 평균적인 PSNR 향상을 비교하였다. 표 2 실험결과의 Delta PSNR 은 HEVC 의 색차신호 예측보다 예측성능이 개선된 모든 블록들만의 PSNR 이득의 총합을 전체블록 수로 나눈 평균 PSNR 이득을 나타낸다. 실험결과 표 2 와 같이 제안 기법이 HEVC 화면내 색차신호 예측방법 보다 Cb, Cr 에 대해서 각각 0.51, 0.38dB 의 PSNR 이득을 보였으며 이는 기존의 CNN 기법 대비 약 0.22, 0.19dB 향상된 성능이다.

표 1. 학습 데이터 구성(16x16)

Class	HM 16.0 in AI	
	Training (QP 22)	Test (QP 22, 27, 32, 37)
Class A	Campfire, CatRobot, Tango FoodMarket, ParkRunning	DaylightRoad
Class B	BasketballDrive, Cactus MarketPlace, RitualDance	BQTerrace
Class C	BasketballDrill PartScene, eHorse	BQMall

표 2. 제안된 기법의 실험 결과

Methods	Delta PSNR (dB) over HM 16.0	
	Cb	Cr
Hybrid [1]	0.29	0.19
제안 기법	0.51	0.38

#### 5. 결론

본 논문에서는 색차신호의 화면내 예측성능 향상을 위한 CNN 기반의 색차신호 예측기법을 제안하였다. 제안한 CNN 구조는 현재블록 주변의 복호화된 색차신호 블록을 사용한다. 실험결과 제안기법이 HEVC 화면내 색차신호 예측 대비 Cb, Cr 각각 0.51, 0.38 dB PSNR 이득이 있음을 확인하였고, 이는 기존 CNN 기반의 기법보다 약 0.22, 0.19dB 개선된 것이다.

#### 감사의 글

본 논문은 2017 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. 2017-0-00072)

#### 참고 문헌

- [1] Y. Li, Z. Li, J. Yang, N. Xu, D. Liu, and Liu, "A Hybrid Neural Network for Chroma Intra Prediction," In Proc., ICIP 2018, Oct. 2018.
- [2] F. Bossen, J. Boyce, X. Li, V. Seregin, and K. Suhring, "JVET Common Test Conditions and Software Reference Configurations for SDR Video," JVET-K1010, Jul. 2018.
- [3] T. Dumas, A. Roumy, and C. Guillemot, "Context-Adaptive Neural Network-Based Prediction for Image Compression, IEEE Trans. Image Process., vol. 29, 2020.