

합성곱 신경망 기반의 블로킹 노이즈 감소 기법

정지연, 김양우, 이영렬
세종대학교

jjjung@sju.ac.kr, ywkim@sju.ac.kr, yllee@sejong.ac.kr

Blocking noise reduction method based on Convolutional Neural Network

JiYeon Jung, Yang-Woo Kim, Yung-Lyul Lee
Sejong University

요 약

본 논문에서는 최신 비디오 압축 표준인 HEVC 를 기반으로, 영상의 부호화 과정에서 블록 사이에 형성되는 잡음을 감소시키기 위하여 합성곱 신경망을 학습시켜 영상의 품질과 예측 성능을 향상시키는 기법을 제안한다. 실험 결과 본 제안 기법은 기존의 HEVC 대비 평균 0.06dB 의 PSNR 개선을 보였고, 원본 값과 예측 값의 오차는 0.5%만큼 감소하였다.

1. 서론

통신망과 다양한 디바이스의 발전으로 4K 이상의 고해상도 영상들에 대한 비디오 압축 기술의 필요성이 대두되었고, 이에 따라 ITU-T VCEG(Video Coding Expert Group)와 ISO/IEC MPEG(Moving Picture Expert Group)에서 JCT-VC(Joint Collaborative Team on Video Coding) 그룹이 설립되었다. JCT-VC 그룹은 2010 년 1 월에 결성되었으며 2010 년 4 월 독일 드레스덴의 1 차 회의를 시작으로 최신 표준화 비디오 코덱인 H.265/HEVC(High Efficiency Video Coding)^[1]을 개발하였다.

HEVC 표준 기술에서는 CTU(Coding Tree Unit)로부터 최소 8x8 부터 최대 64x64 크기로 다양하게 분할될 수 있는 단위인 CU(Coding Unit)를 적용하였고, 이는 다시 예측 단위인 PU(Prediction Unit) 또는 변환 단위인 TU(Transform Unit)로 분할되어 화면 내/화면 간 예측 및 변환/양자화 과정을 수행한다. 이와 같은 과정에서 원본 영상과 예측 영상 간의 오차가 발생하게 되고 그로 인해 블로킹 현상(Blocking Artifacts) 및 링잉 현상(Ringing Artifacts)과 같은 영상 내 잡음 현상이 발생한다. 이러한 현상을 효과적으로 제거하기 위하여 HEVC 표준에서는 디블로킹 필터(De-blocking Filter)와 SAO(Sample Adaptive Offset)를 포함한 인루프 필터링 기법이 적용된다.

본 논문에서는 최근 활발히 연구되고 있는 딥러닝을 이용하여 영상 내 노이즈를 감소시키는 필터링 기법을 제안한다. 특히, 인접한 두 블록의 경계에 발생하는 블로킹 현상을 제거하기 위해 사용되는 기법인 디블로킹 필터를 대체하기 위한 합성곱 신경망 학습 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2 절에서는 제안하는 기법에 대해 살펴본 후, 3 절에서는 실험 결과 및 분석, 마지막으로 4 절에서는 본 논문에 대한 결론을 맺는다.

2. 제안하는 방법

기존의 HEVC 에서의 디블로킹 필터는 블록의 크기, 픽셀 값 및 QP 값 등에 따라 각각 다른 필터링 방식을 적용한다. QP 값이 비교적 클수록 강한 필터링, QP 값이 비교적 작을수록 약한 필터링이 적용되며, 블록의 크기와 픽셀 값의 특성에 따라 강한 필터링 혹은 약한 필터링을 적용시킨다.

제안한 방법은 이러한 개념을 토대로 하여 입력 값으로 들어온 블록의 크기에 따라 각각 다른 네트워크로 학습시킨다.

2.1 네트워크 구조

입력 영상은 기존의 HEVC 에서 디블로킹 필터링 과정을 거치지 않은 영상이 사용된다. 해당 영상에서 모든 TU 블록을 추출해내고, 각 블록의 크기에 따라 분류하여 입력 값으로 사용한다. 또한, 각 TU 블록에 대해서 현재 블록과 현재 블록에 인접한 좌측 및 상단의 픽셀들을 사용한다. 즉, 현재 블록과 현재 블록에 인접한 총 4 개 열, 4 개 행의 픽셀들을 입력 값으로 사용한다. QP 값은 큰 QP 값으로 고정하여, 현재 블록의 크기와 픽셀 값의 특성 및 주변 블록과의 상관관계 등을 고려한 네트워크를 사용하게 된다.

해당 입력 값을 통해 총 2 개의 네트워크를 학습시킨다. 첫 번째 네트워크는 작은 블록일 경우에 사용되는 네트워크이고, 두 번째 네트워크는 큰 블록일 경우에 사용되는 네트워크이다. 이때, 작은 블록이란 블록의 너비 및 높이가 8 이하일 경우의 블록을 의미하고, 큰 블록은 블록의 너비 및 높이가 16 이상일 경우의 블록을 의미한다.

네트워크 구조는 그림 1 과 같다. 블록의 크기에 따라 각각 다른 구조를 사용하게 되는데, 블록의 크기가 작을 경우에는 원본 영상과의 오차 값이 커지는 경향이 있기 때문에 깊은

그림 1. 제안하는 네트워크 구조

네트워크 구조를 사용하여 학습시키고, 오차 값이 비교적 적은 큰 블록에 대해서는 얇은 네트워크 구조를 사용하여 학습시킨다. 그림 1 은 작은 블록일 경우에 사용되는 깊은 네트워크 구조를 나타낸 것이며, 얇은 네트워크 구조는 그림 1 에서 첫 번째 dense block layer 와 3 번째 convolution layer (빨간 점선으로 표시된 부분)를 제외한 네트워크 구조이다.

2.2 학습 방법

학습에 사용된 시퀀스는 총 6 개로, 총 데이터의 개수는 300,000 개이며, 테스트에 사용된 시퀀스는 총 3 개이고, 총 데이터의 개수는 30,000 개이다. 실험에 사용된 각각의 시퀀스에 따라 사용된 데이터 개수는 표 1 과 같다.

네트워크는 원본 영상과 네트워크를 통해 학습된 복원 영상에 대한 MAE(Mean Absolute Error)를 이용하여 Adam Optimizer 로 학습하였다. 학습 시 배치 사이즈는 6 으로 고정하여 학습하였고, 총 10 번의 epoch 을 통해 학습이 진행되었다. 200,000 번 째 학습 시에는 learning rate 를 감소시켜, 학습률을 향상시켰다. 네트워크 학습을 위해 Tensorflow Python API 를 사용하였다.

	Sequence	resolution	The number of Data
Training set	Kimono	1920x1080	50,000
	Cactus	1920x1080	50,000
	BasketballDrive	1920x1080	50,000
	BQTerrace	1920x1080	50,000
	BQMall	832x480	50,000
	PartyScene	832x480	50,000
Test set	ParkScene	1920x1080	10,000
	BasketballDrill	832x480	10,000
	RaceHorses	832x480	10,000

표 1. 실험에 사용된 데이터 개수

3. 실험 결과 및 분석

본 논문의 실험에 사용된 참조 소프트웨어는 HM-16.20 버전이며, JCT-VC CfP(Call for Proposal) 단계에서 결정된 공통 실험 조건^[2]에 맞추어 진행하였다. 실험은 All intra(AI) configurations 에서 진행하였고, B, C 클래스들을 사용하여 총

9 개의 시퀀스들을 사용하였다. 각 시퀀스마다 37 의 QP 로 설정하여 실험 결과 및 분석을 진행하였다. 실험 결과는 기존의 HEVC 에서의 더블록킹 필터링 과정을 거친 영상과의 비교를 통해 이루어졌고, 객관적 화질을 나타내는 PSNR (Peak Signal-to-Noise Ratio) 및 기존 영상과의 MSE(Mean Squared Error) 로 표현하였다. 실험 결과 작은 블록에 대한 네트워크는 기존의 HEVC 영상의 PSNR 이 31.31dB 이고, 학습된 네트워크 영상의 PSNR 은 31.39dB 로, 총 0.08dB 의 PSNR 개선을 보였고, MSE 는 0.7%만큼 감소하였다. 큰 블록에 대한 네트워크는 기존의 HEVC 영상의 PSNR 이 32.59dB 이고, 학습된 네트워크 영상의 PSNR 은 32.62dB 로, 총 0.03dB 의 PSNR 개선을 보였고, MSE 는 0.3%만큼 감소하였다.

4. 결론

본 논문에서는 최신 비디오 압축 표준인 HEVC 를 기반으로, 영상의 부호화 과정에서 블록 사이에 형성되는 잡음을 제거하기 위하여 합성곱 신경망을 학습시켜 영상의 품질과 예측 성능을 향상시키는 기법을 제안한다. 제안한 기법을 HM-16.20 버전에 구현하여 성능을 확인하고 분석하였고, 실험 결과는 xx 을 통해 확인할 수 있다. 실험 결과 본 제안 기법은 기존의 HM-16.20 대비 평균 0.06dB 의 PSNR 개선을 보였고, 원본 값과 예측 값의 오차는 0.5%만큼 감소하였다. 본 논문의 제안 방법을 토대로 블록의 특성 및 블록 간의 상관관계를 고려하여 화질 개선 및 부호화 향상 등의 추가적인 방법에 대한 연구가 필요하다고 보여진다.

5. 감사의 글

이 논문의 일부는 2019 년도 정부(과기정통부)의 재원으로 한국여성과학기술인지원센터의 지원을 받아 수행된 기초연구사업임 (No. 20190623)

6. 참고문헌

- [1] <https://hevc.hhi.fraunhofer.de/trac/hevc/browser/tags/HM-16.20>
- [2] http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=7281