

CNN 기반 HEVC 루프 필터의 성능 비교

이소윤, 홍진형, 오병태
한국항공대학교

olnayoung@kau.kr, hong_1993@kau.kr, byungoh@kau.ac.kr

Performance evaluation of CNN-based in-loop filter for HEVC

So Yoon Lee, Jin Hyung Hong, Byung Tae Oh
Korea Aerospace University

요 약

본 논문에서는 HEVC in-loop filter 부분에 CNN 기반 기술을 사용한 알고리즘들을 소개하고, 이들 기술들의 성능을 비교실험을 통해 분석한다. 현재 HEVC 의 in-loop filter 에는 잡음을 제거하는 deblocking filter 와 오프셋을 보상하는 sample adaptive offset 필터가 적용되어 있는데, 두 알고리즘이 수행하고 있는 역할을 CNN 기반 filter 로 대체한 기술들을 선택적으로 선별하여 비교실험을 진행하고, 그 결과를 비교 및 분석한다.

In this paper, we introduce the CNN-based in-loop technology for HEVC, and analyze the performance of these algorithms through comparative experiments. The current in-loop filters in HEVC are composed of a deblocking filter that removes noise and a sample adaptive offset filter that compensates for signal offsets. A couple of CNN-based filters replacing the roles of these two algorithms are selected and compared.

1. 서론

동영상이 일상 생활에 많이 쓰임에 따라 동영상 압축 기법의 중요도가 높아지고 있다. 최근 개발된 동영상 압축 표준인 High efficiency video coding(HEVC)는 이전 표준 코덱인 H.264/AVC와 비교해서 약 50%의 비트율 감축을 달성하는 데에 성공했다.

HEVC 내부는 여러 기능을 하는 블록들로 구성되어 있다. 이중 in-loop filter는 압축 과정 중에 발생하는 여러 잡음을 없애는 역할을 하고 있다. In-loop filter는 Deblocking Filter (DF), Sample Adaptive Offset (SAO)의 두 가지 기술로 구성된다. DF는 Coding Unit (CU)의 경계에서 생기는 blocking artifact를 완화시키는 데에 사용된다. SAO는 영역 내 화소 간의 edge 특성을 고려하여 복원 영상의 왜곡을 감소시킨다.

최근 영상처리 분야에서 Convolutional Neural Networks

(CNN) 기반의 방식이 큰 관심을 받고 있다. CNN은 Super-resolution (SR)이나 classification과 같은 문제를 해결하는 데에 훌륭한 성능을 보여주고 있다 [1-3]. 초창기에 CNN 기반으로 SR 문제를 푼 SRCNN 기술이 소개되었는데, [1] SRCNN이 좋은 결과를 가져오자 여러 다른 CNN들도 개발되기 시작했다 [2]. 그러다 SR이나 classification외에도 JPEG 및 HEVC의 압축 효율 향상을 위해 CNN이 사용되기 시작되어 ARCNN 등이 개발되기 시작했다. [3-5].

제안 논문에서는 압축 성능을 향상시키기 위해 HEVC의 in-loop filter 부분을 CNN 네트워크로 교체한 여러 알고리즘들을 소개하고 압축성능을 비교 및 분석해 보고자 한다.

2. 알고리즘

A. CNN-based in-loop filtering (IFCNN)

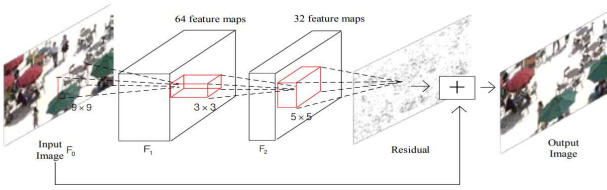


그림 1. IFCNN의 네트워크 구조.

IFCNN의 기존 SRCNN의 구조를 기반으로 residual learning 을 도입한 구조이다. 단순히 3개의 convolution layer 로 구성되어 있고, 각 층 사이에 ReLU 층을 삽입했다.

$$F_i = \max(0, W_i * F_{i-1} + B_i), i = 1, 2 \quad (1)$$

$$Y = W_3 * F_2 + B_3 \quad (2)$$

수식 1에서 W_i 와 B_i 는 각각 $(i - 1)$ 번째 층과 i 번째 층 사이의 convolution filters와 biases들이다. 수식 2에서 마지막 층의 출력은 Y라는 것을 알 수 있다. 여기서 Y는 IFCNN의 최종 출력을 뜻한다. 각각 W_1 은 $9 \times 9 \times 64$, W_2 는 $64 \times 3 \times 3 \times 32$, W_3 은 $32 \times 5 \times 5 \times 1$ 의 크기를 가진다.

IFCNN은 SAO만 CNN으로 대체 하였지만, open loop로 실행되는 All-intra mode의 경우에는 DF와 SAO부분을 모두 CNN으로 대체했다.

B. Variable-filter-size Residue-learning CNN (VRCNN)

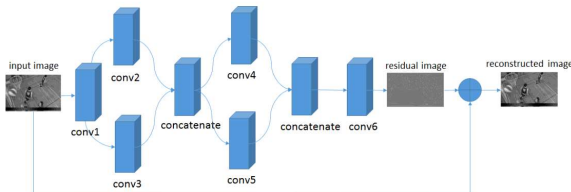


그림 2. VRCNN의 네트워크 구조.

VRCNN은 HEVC에서 CU의 크기가 다양하다는 점을 사용하기 위해, 다양한 filter size를 사용한 후 이를 다시 합치는 방식으로 네트워크를 구성하였다. VRCNN의 filter size는 5×5 , 3×3 , 1×1 인데, 이는 최근 SR에 적용되는 CNN은 작은 크기의 filter를 사용하고 층을 깊게 쌓는다는 점에서 착안했다.

실제 구현에서 VRCNN은 DF와 SAO 모두 CNN으로 대체하였으며, All-intra mode에서 Post-processing filter로서 VRCNN 네트워크를 적용하였다.

3. 실험 결과

두 방식의 공정한 실험 및 평가를 위해 HEVC의 DF와 SAO를 모두 CNN으로 대체하는 방식을 택하였다. 즉,

CNN 네트워크의 입력으로 DF에 입력되는 영상을 사용하였다. 학습 영상으로는 416×240 크기의 BasketballDrill, BlowingBubbles, BQMall, BQSquare, PeopleOnStreet, Traffic의 일부 영상을 사용하였으며, 테스트 영상은 BasketballDrive, BQTerrace, Cactus, Kimono, ParkScene의 101~200번째 frame을 사용하였다. 실험은 All-Intra 모드와 Low Delay P 모드 2가지 조건에서 실험을 진행하였다.

학습을 진행할 때 Quantization Parameter (QP)를 고정하여 학습을 진행해 보기도 하였으며, 여러 QP를 혼합하여 학습을 진행하는 방식도 진행하였다. QP를 고정하는 경우에는 QP 22, 27, 32, 37만을 이용하였으며, 여러 QP를 이용할 경우에는 QP 20~29, QP 30~39의 두 종류로 학습을 진행하였다.

표 1. ALL-INTRA 의 PSNR 결과 비교 (고정 QP 학습)

	IFCNN	VRCNN
QP 22	-0.43	-0.69
QP 27	+0.38	+0.41
QP 32	+0.51	+0.65
QP 37	+0.37	+0.50
AVERAGE	+0.19dB	+0.22dB

표 2. ALL-INTRA 의 PSNR 결과 비교 (혼합 QP 학습)

	IFCNN	VRCNN
QP 22	+0.33	+0.40
QP 27	+0.52	+0.61
QP 32	+0.45	+0.58
QP 37	+0.48	+0.61
AVERAGE	+0.45dB	+0.55dB

표 3. LOW DELAY P 의 BDBR 결과 비교 (혼합 QP 학습)

	IFCNN	VRCNN
BasketballDrive	-2.1	-4.0
BQTerrace	+14.4	+32.8
Cactus	+11.3	+13.2
Kimono	-0.6	+1.4
PS	+7.7	+11.8
AVERAGE	+5.22%	+5.18%

표 1, 2 의 All-intra mode의 결과는 5개의 테스트 영상을 QP별로 평균을 낸 것이다. 위 결과와 같이 All-intra의 경우, 학습 QP의 종류에 관계 없이 평균적으로 IFCNN과 VRCNN 모두 0.2 ~ 0.5dB 정도의 성능 향상을 보였다.

반면 표 3에서의 Low Delay P 모드에서는 성능이 저하되었다. 실제 post-processing filter로 설계된 필터인 VRCNN은 몇몇 시퀀스에서는 지나치게 좋지 않은 성능을 가짐을 확인할 수 있었다. 추후 학습조건 및 압축 환경 등 다양한 환경을 고려하여 보다 면밀한 실험이 필요할 것으로 보인다.

4. 결론

본 논문은 HEVC의 압축 효율 향상을 위해 CNN을 활용한 여러 알고리즘들을 비교실험 하였다. 실험 결과 All-intra 모드에서는 큰 성능 향상을 보였지만, Low Delay P 모드에서는 성능이 크게 저하되는 것을 확인할 수 있었다. 하지만 BasketballDrive에서는 $-2.1\sim-4.0\%$ 의 성능 향상을 보였지만 나머지 영상에서는 오히려 성능이 나빠지는 등, 영상 별 결과의 차이가 크기 때문에 학습 영상의 종류 및 환경을 보다 면밀히 분석한 후, 향후 보다 개선된 필터를 제안하고자 한다.

감사의 글

본 연구는 경기도의 경기도 지역협력 연구센터 사업(GRRC)의 일환으로 수행하였음 [2017-B02, 이동형 미디어 디바이스를 위한 영상처리 및 UI 플랫폼 기술연구].

참조문헌

- [1] Dong, C., Loy, C. C., He, K., and Tang, X. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 2016, 38.2: 295-307.
- [2] Kim, J., Kwon Lee, J., and Mu Lee, K. Accurate image super-resolution using very deep convolutional networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016. p. 1646-1654.
- [3] Dong, C., Deng, Y., Change Loy, C., and Tang, X. Compression artifacts reduction by a deep convolutional network. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2015. p. 576-584.
- [4] Park, W. S., and Kim, M. CNN-based in-loop filtering for coding efficiency improvement. In: *Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, 2016 IEEE 12th. IEEE, 2016. p. 1-5.
- [5] Dai, Y., Liu, D., and Wu, F. A convolutional neural network approach for post-processing in hevc intra coding. In: *International Conference on Multimedia Modeling*. Springer, Cham, 2017. p. 28-39.