

GAN 을 이용한 압축된 동영상 품질 향상

김용성, 이유진, 김범윤, 전병우
 성균관대학교 전자전기공학부
 {asdk4253, yjlee0816, kbumyoon, bjeon}@skku.edu

GAN-based Quality Enhancement of Compressed Video

Yongseong Kim*, Yujin Lee, Bumyoon Kim, and Byeungwoo Jeon
 School of Electronic and Electrical Engineering*,
 Sungkyunkwan University, Korea

요 약

본 논문에서는 딥러닝의 주요 기법 중 하나인 GAN 을 활용하여 압축된 영상의 품질을 개선하는 방법을 제안한다. 제안하는 GAN 의 생성자는 U-Net 과 ResNet 을 기반으로 구성되었으며, 판별자는 합성곱층과 전연결층으로 구성하였다. 네트워크의 학습은 HEVC (High Efficiency Video Coding)의 테스트 모델인 HM16.25 를 사용하여 RA (Random Access) 구성하에 양자화 계수 37 로 압축된 영상을 입력으로 하여 수행되었다. 제안하는 네트워크의 성능 확인을 위해 학습 시와 동일한 조건으로 압축된 다른 영상을 입력으로 하여 실험하였다. 실험 결과 영상의 평균 PSNR 은 34.20dB 에서 34.24dB 로 0.04dB 의 품질 향상이 이루어진 것을 확인할 수 있었다.

1. 서론

동영상 압축 기술은 동영상 데이터의 통신 트래픽 증가 및 저장 공간 부족의 문제를 해결하기 위한 핵심 기술 중 하나이다 [1]. 그러나 널리 사용하는 손실 압축방식의 경우, 압축 과정 중 양자화 과정에서 데이터의 손실이 발생하기 때문에 압축 이후 복원된 영상은 원본 영상 대비 품질이 저하된다는 문제점이 존재한다. 따라서 이러한 문제를 해결하기 위해 부호화 및 복호화 된 동영상에 대하여, 동영상 데이터의 크기를 증가시키지 않으면서 영상의 품질 향상을 위한 방법으로 후처리를 통한 방법이 주로 연구되고 있으며 최근에는 딥러닝을 이용한 후처리 방법이 많이 연구되고 있다. 본 논문에서는 U-Net [2]과 ResNet [3]에 기반한 생성자 네트워크 모델과 합성곱층과 전연결층 기반의 판별자 모델로 이루어진 GAN (Generative Adversarial Network)을 사용하여 영상의 품질을 향상시키는 방법을 제안한다.

2. 제안 방법

GAN 은 게임이론에 기초하여 생성자와 판별자를 동시에 학습하는데 [4], 생성자는 판별자가 진위 여부를 판단하기 어려운 가짜 영상을 생성하도록 학습하며, 판별자는 생성자가 만들어낸 영상을 원본 영상과 구별할 수 있도록 학습한다. 이 과정이 반복되면서 생성자는 원본 영상과 가까운 복원영상을 만들 수 있기 때문에 열화 된 영상을 효율적으로 원본과 최대한 비슷하도록 보정할 수 있다. GAN 에서 생성자와 판별자는 다양한 종류의 네트워크를 사용할 수 있지만 제안 방법에서는 <그림 1 (a)>와 같이 U-Net, ResNet 을 기반으로 생성자를 설계하고, <그림 1 (b)>와 같이 일반적인 합성곱층과 전연결층을 기반으로 판별자를 설계하였다. <그림 1>에서 Conv k, c 는 $k \times k$ 크기의 합성곱 커널을 사용하고, 출력 채널이 c 인 합성곱층을 의미한다.

U-Net 은 수축 경로와 확장 경로로 구성되어 수축 경로의 출력이 확장 경로의 입력으로 들어갈 수 있는 구조로 설계된 인공 신경망이다 [2]. 수축 과정을 통해 영상을 세분화할 수 있으며 이를 통해 지역적 특징을 잘 추출할 수 있다.

본 연구는 과학기술정보통신부 재원의 한국연구재단의 사업(NRF-2020R1A2C2007673)의 연구결과로 수행되었음.

