

망 사용량 감축을 위한 딥러닝-기반 영상 융합 기술 및 활용 방안

주호성, 양현중*
포항공과대학교

zxcqa123@postech.ac.kr, *hyunyang@postech.ac.kr

A Study on the deep learning-based image fusion for reducing the overhead of communication systems

Hosung Joo, Hyun Jong Yang*
Pohang University of Science and Technology (POSTECH)

요 약

본 논문은 통신망에 가장 많은 영상정보에 대한 압축 및 복원 기술을 다룬다. 기존 영상 압축 코덱은 색상 정보를 압축하지 않았으나, 영상처리기술이 개선됨에 따라 색상 정보를 압축하더라도 복원이 가능하다. 본 논문에서는 하나의 고해상도 영상을 남겨 영상 융합 기술을 적용함으로써 딥러닝을 기반으로 영상의 해상도를 높이는 색상별 영상 복원 기술을 제안하고, 복원 기술에 대한 성능평가를 제시한다.

I. 서 론

Ericsson Mobility Report 에 따르면, 전체 이동통신 데이터 트래픽 중 69%는 영상정보에 해당한다. 이동통신 평균 통신량은 작년 월간 스마트폰 개당 10GB 수준에서 현재 월간 15GB 수준으로 연간 50% 증가하여 증가하는 추세가 매우 가파르며, 이와 같은 대역폭 요구의 증가는 AR(Augmented Reality; 증강현실), MR(Mixed Reality; 혼합현실) 및 VR(Virtual Reality; 가상현실)과 같은 XR 기술이 미래에 도입되어 영상정보가 증가하며 더 가속될 전망이다. [1]

최근 통신량의 급격한 증가에 따라 통신 망 사업자 및 ISP(Internet Service Provider)의 망 유지에 대한 부담이 증가하였고, ISP 를 위해 콘텐츠 사업자에게 망 사용료를 부과하는 일명 망 사용료 법안도 한국에서 발의되었다. 콘텐츠 사업자가 망 사용료를 추가 부담하면, 최종적으로 콘텐츠 소비자도 함께 사용료를 부담하게 될 가능성이 높다. 이러한 망 사용료는 망에 가해지는 부하에 비례할 것이므로, 증가 추세 역시 매우 가파를 것으로 예측된다. 그러나 지속적인 기술 발전을 통해 망 부담을 줄이면, 콘텐츠 소비자에게 가해질 망 사용료에 대한 부담도 함께 줄어들 것이다.

망 부담을 완화하기 위해선 특히 영상정보를 압축하는 기술이 추가적으로 개선되어야 할 필요가 있다. 현재 네트워크망을 통해 전송되는 영상정보 압축 방식 중 가장 많이 사용되는 기술을 살펴보면 크게 H.264(AVC) 및 H.265(HEVC) 코덱과 MPEG 코덱 등이 있는데, 이들 코덱은 색상을 나타내는 (Y, Cr, Cb)의 24bit 시공간계열 디지털 데이터를 각각 따로 압축하여 통신에 활용하는 기술이다. [2,3]

본 논문에서 활용하고자 하는 영상 융합 기술은 크게 멀티모달(multimodal) 영상 융합과 다중초점(multifocus) 영상 융합으로 나눌 수 있는데, 특히 멀티모달 영상 융합에서는 해상도가 낮은 영상을 해상도가 높은 영상에 정합하여 해상도를 향상시키는 SR(Super Resolution) 기술을 사용하고 있다. 만약 기존 영상정보 스트리밍 시스템에서 색상에 따라 다른 영상에 SR 과 같은 기술을 적용한다면 송신해야 할 정보를 최대 3 배까지 줄일 수 있다. 하지만, 아직 이와 같이 3 색의 색상 정보 중 하나의 정보만 남겨두고 나머지 2 색의 영상을 크게 압축한 뒤 차후 복원하는 방식의 코덱을 제시한 사례는 없었으며, 이와 같은 접근법이 영상의 가로세로 길이를 절반으로 압축하여 복구하는 방식을 활용해 최신의 저명 학술지에 소개되었다. [4]

본 논문에서는 압축된 영상을 송출하고 영상 융합 기술을 활용하여 영상을 복원하여, 콘텐츠 소비자가 망 자원의 절약에 동참하는 방법으로, 딥러닝 기반의 영상 복구 코덱을 제안하고 성능을 평가하고자 한다.

II. 본론

영상을 압축하는 방법에는 크게 엔트로피 기반 압축과 다운샘플링(down-sampling)이 있다. 이론적으로 무손실 압축인 엔트로피 기반 압축에 비해 다운샘플링은 손실이 발생하지만 영상 압축 수준을 쉽게 제어할 수 있다. 더하여, SR 등의 복구 기술을 활용하면 다운샘플링에 의한 손실을 효과적으로 줄이고 영상을 크게 압축하여 통신망에 도움을 준다.

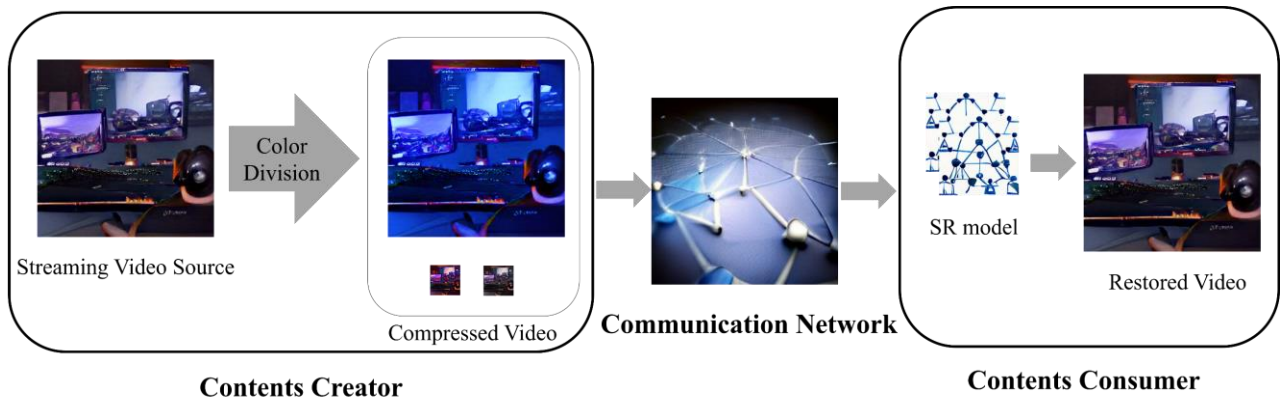


그림 1. 망 사용량 감축을 위한 딥러닝-기반 융합 영상 기술 중 SR 모델의 활용 방안

SR을 구현하는 방법엔 크게 업샘플링(up-sampling)된 도메인에 생성 모델(generative model)을 활용하는 방법과 매 프레임마다 많은 데이터를 모은 후 통계적인 추정을 하는 방법이 있다. 그러나 통신 부하가 심한 상황에서는 하나의 프레임에 대한 많은 데이터를 모을 기회가 거의 없으므로, 통계적 추정은 적용하기 어렵다.

그러므로 생성 모델을 활용하는 방법을 고려하면, 요즘 가장 자주 활용되는 고성능 생성 모델의 종류로 VAE (Variational AutoEncoder), Diffusion model, 그리고 GAN (Generative Adversarial Network)이 있으며, 영상 분야에서 최근에 제안되어 가장 성능이 좋을 것으로 기대되는 대용량 및 고비용의 생성 모델로서 ViT(Visual Transformer)가 있다. 그러나 망 사용자에게 너무 많은 연산을 필요로 하는 생성 모델을 활용하도록 하기엔 아직 무리가 있으므로, 본 논문에서는 VAE를 응용하여 딥러닝 기반의 저비용 영상 융합 기술을 제안한다. (그림 1 참조) 알맞은 모델을 구현하는 방법에는 여러 가지의 방법이 있겠지만, fully-connected 인공신경망을 통해 모델을 구현하고 성능을 평가하면 기준선으로 생각할 수 있다.

그림 1에는 콘텐츠 생산자에게서 영상 원본을 얻고 색상별로 데이터를 분리한 뒤, 한 개의 색상을 선택하고 나머지는 압축해 통신망을 통해 전송하는 과정이 나타나 있다. 콘텐츠 소비자는 압축된 영상과 딥러닝 기반의 SR 모델을 통해 원본 영상에 가까운 영상을 얻어낸다.

SR 모델의 성능을 평가하는 방법으로 주로 PSNR (Peak Signal-to-Noise Ratio)를 확인한다. PSNR은 원본 영상과 복원 영상 사이의 평균 오차인 MSE (Mean of Squared Error)를 영상에 알맞게 dB 단위로 환산해주는 도구로, 높을수록 복원이 잘 된 것이며, 수식 (1)으로 표현할 수 있다. 여기서 Q 는 영상의 양자화 레벨을 뜻한다.

$$\text{PSNR} = 10 \log_{10} \frac{Q^2}{\text{MSE}} \quad (1)$$

모델을 사용해 복원한 데이터는 활용이 가능해야 한다. 따라서, 본 모델을 활용해 복원한 데이터가 실제로 다른 업무(task)에 얼마나 잘 작동되는지 확인할 필요가 있다. 업무에 대한 강건성을 보이기 위해 자주 적용하는 분류(classification) 문제에 적용하여 분류정확도 차이를 확인할 수 있다.

또한, 통신망을 통해 영상정보를 전송하므로 전송된 비트 에러에 대한 강건함도 확인하면 차후 코덱의 표준화에 도움이 될 것이다. MSE, PSNR을 압축율에 따라 보일 수 있으며, 마지막으로, 통신망의 SNR 환경에 따라 비트에러율(BER)을 가지고 전송된 영상 데이터에 대해 모델의 성능 변화를 확인할 수 있다.

III. 결론

본 논문에서는 통신망에 가장 많은 영상정보에 대한 압축 기술 및 복원 기술을 다루었다. 기존 영상 압축 코덱은 색상 정보를 압축하지 않았으나, 영상처리기술이 개선됨에 따라 색상에 따른 정보를 압축하더라도 복원이 가능하다. 본 논문에서는 하나의 고해상도 영상을 남겨 영상 융합 기술을 적용함으로써 복원 기술을 제안하였다. 영상의 해상도를 높이는 색상별 영상 복원 기술로서 딥러닝 기반의 SR(Super Resolution) 기술을 사용할 것을 제안하였고, 그 구현 방법으로서 VAE(Variational AutoEncoder)를 응용한 fully-connected 인공신경망 구조를 제안하였다. 마지막으로 복원 기술에 대한 성능평가 방법을 제시하였지만, 상세한 구현 및 성능평가는 지면이 부족하여 차후 시행한다.

ACKNOWLEDGMENT

This work was conducted as a part of the research projects of "Development of automatic screening and hybrid detection system for hazardous material detecting in port container" financially(1525012204) supported by the Ministry of Oceans and Fisheries, Korea.

참고 문헌

- [1] Ericsson Mobility Report, (2021 June), and (2022 June).
- [2] Barman, Nabajeet, et al. "Codec Compression Efficiency Evaluation of MPEG-5 part 2 (LCEVC) using Objective and Subjective Quality Assessment." arXiv preprint arXiv:2204.05580, (2022).
- [3] Barman, Nabajeet, and Maria G. Martini. "User Generated HDR Gaming Video Streaming: Dataset, Codec Comparison, and Challenges," IEEE Transactions on Circuits and Systems for Video Technology 32.3, pp. 1236-1249, (2021).
- [4] Yeh, Chia-Hung, et al. "Deep learning-based compressed image artifacts reduction based on multi-scale image fusion," Information Fusion 67, pp. 195-207, (2021).