

딥 러닝 기반 이미지 압축 기술 동향

이순빈, 정종범, 류은석
성균관대학교 컴퓨터교육과
{soonbinlee, uof4949, esryu}@skku.edu

An Overview of Learned Image Compression

Soonbin Lee, Jong-Beom Jeong, Eun-Seok Ryu
Department of Computer Education, Sungkyunkwan University

요 약

딥 러닝 기반 학습 방법이 발전을 거듭함에 따라, 이미지 압축 기술에도 적용하려는 시도가 늘어나고 있다. 그러나 이러한 딥러닝 기반 이미지 압축 방법은 기존의 압축 기술과는 다른 양상을 나타내며, 압축 기술들이 적용되기 위해서는 여러 고려사항이 필요하다. 본 논문에서는 딥 러닝 기반 이미지 압축 기술에 대한 3가지 방향을 살펴보고, 딥 러닝 기반 이미지 압축 기술에서의 발전 방향과 고찰해볼 수 있는 요소들을 논의한다.

1. 서 론

멀티미디어 압축 기술은 그 중요성에 따라 상당수의 연구가 현재까지도 진행되고 있다. Moving picture experts group (MPEG)에서는 이미지와 영상 압축에 관련하여 현재까지도 지속적인 표준화 기술을 제정하고 진행중에 있으며 360 영상, 라이트 필드 등의 차세대 몰입형 미디어에 대한 수요에 따라 병렬 부복호화 등 다양한 여러 기술의 도입과 성능 탐색들이 이루어지고 있다.

그와 동시에 최근 딥 러닝 기술의 비약적인 발전으로, 미디어 압축에 대한 여러 시도가 이루어지고 있다. 딥 러닝 기반의 이미지 압축은 주어진 이미지 데이터셋에 대하여, 엔트로피 양을 최소화 하면서 복원된 이미지의 손실을 최소화할 수 있는 잠재 벡터(latent vector)를 출력할 수 있도록 하는 파라미터를 학습하는 방식으로 이루어지게 된다. 출력된 잠재 벡터는 엔트로피 코딩을 거쳐 적은 사이즈의 전송 가능한 형태가 되어 기존 코덱의 비트스트림(bitstream) 같은 기능을 하게 된다.

딥 러닝 기반의 이미지 압축의 방향은 (1) 학습하는 이미지 데이터 셋의 가능도(likelihood)를 명시적으로 계산하는 자기회귀(Autoregressive) 기반 방법, (2) generative adversarial network (GAN)을 이용하여 데이터 변환 분포를 직접적으로 배우는 방법, (3) 압축의 전 과정을 하나의 딥 러닝 모델로 대체하고자 하는 종단간 학습(End-to-end Learning)으로 구분할 수 있으며, 본 논문에서는 이러한 딥러닝 기반 학습 이미지 압축 모델의 동향에 대해 알아보고자 한다.

2. 관련 연구

2.1 미디어 부호화 표준화

HEVC의 최종 표준화 규격이 완성됨에 따라, HEVC 대비 2배 이상의 압축 효율 개선을 목적으로 하는 Versatile Video Coding (VVC) 표준화에 대한 작업이 진행중에 있다. 많은 기관들의 기고서 제안과 채택 논의를 통해, 현재의 비디오 코덱 기술은 높은 수준의

최적화를 달성하고 있으며 VVC 표준화와 함께 360 영상, 라이트 필드, 6DoF 몰입형 미디어의 특성에 맞는 표준 규격과 특화 기술들을 정의하기 위한 MPEG-I (Immersive) 표준화 또한 진행중에 있다 [1]. 현재의 딥러닝 기반 이미지 압축 기술은 압축 성능 탐색에 머물러 있는 단계이며, 다양한 종류의 미디어에 효율적으로 적용되기 위해서는 좀 더 많은 연구가 필요하다.

2.2 이미지 부호화 모델

이미지 부호화 모델은 심볼(symbol)을 표현하는 데에 필요한 비트 수인 $R(\text{Rate})$ 와, 원본과 복원된 이미지 간의 차이인 $D(\text{Distortion})$ 으로 표현에 필요한 비트 수가 적으면서도 복원된 이미지의 차이가 크지 않은 압축 모델이 보다 성능이 좋은 모델이라고 정의할 수 있으며, 다음 수식과 같이 나타낼 수 있다.

$$\min J = R + \lambda \cdot D$$

따라서 압축 네트워크의 손실 함수 또한 다음과 같은 수식을 최적화하는 것으로 정의될 수 있다.

$$J = -\mathbb{E}[\log_2 z] + \lambda \mathbb{E}[d(\mathbf{x}, \hat{\mathbf{x}})]$$

그러나 딥러닝 기반 모델에서는 이산적 심볼로 표현하여 엔트로피 코딩을 진행하고자 할 때, 양자화 과정에 있어서 역전파에 대한 미분값을 구하는 것에 따른 문제가 발생하게 된다. 따라서 벡터 양자화(Vector Quantization) 기법을 활용하거나[2], 균등 분포의 노이즈(Noise)를 더해주어[3] 네트워크가 잘 학습할 수 있도록 한다.

2.3 자기회귀(Autoregressive) 기반 이미지 압축 모델

자기회귀(Autoregressive) 기반 모델은 학습하고자 하는 데이터의 가능도(likelihood)를 명시적으로 계산해내는 생성 모델이다. 각 픽셀들의 분포를 하나씩 예측하기보다, 다음에 올 주변 픽셀간의 상관성을 함께 조건부 확률로 판단하여 이미지를 생성하는 기법으로[4], 픽셀 값과 같은 저차원적 정보도 보다 정확히 추론할 수 있다는 장점이 있다[5]. 낮은 해상도의 표현에서부터 계층적으로 복원 이미지

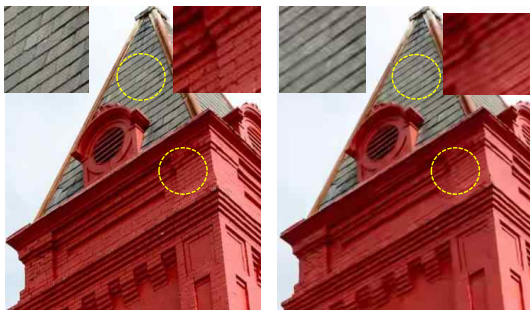
를 생성해내는 무손실(loseless) 압축 모델이 시도되었으며, GPU을 통한 병렬 계산 시 실시간 수준의 빠른 복원 속도를 보여준다[6].

2.4 GAN 기반 이미지 압축 모델

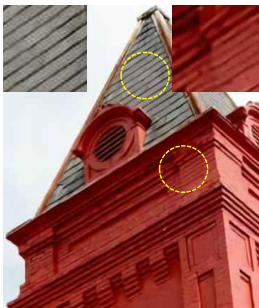
GAN 기반 이미지 압축 모델은 데이터의 가능도나 분포를 명시적으로 계산하지 않고, 임의의 랜덤 노이즈로부터 이미지로의 변환 자체를 학습하는 모델이다. GAN 기반 이미지 생성 모델은 다른 모델들보다 비교적 선명한 이미지를 생성해내는 장점이 있음이 알려져 있지만, 주요 이미지의 정보를 그대로 보존하지 못하는 치명적인 단점을 가지고 있다. [7]에서는 이미지 손실 평가 지표로 픽셀 오차 (Peak Signal-to-noise ratio, PSNR)뿐만이 아닌 인지적 정보 (perceptual loss)를 손실 함수에 포함하는 것을 제안했으며, [8]에서는 기존의 이미지 손실 평가 지표에는 불리하지만, 압축 성능과 사용자 평가 지표에서는 매우 좋은 결과를 얻었다는 것을 주장하였다.

2.5 종단간 학습 기반 이미지 압축 방법

종단간 학습 기반 이미지 압축 방법은 [3]에서 제안된 기법을 기반으로, HEVC의 인트라(Intra) 예측과 같은 역할을 하는 사전 정보를 활용하는 등 압축 성능 개선을 위한 후속 연구가 이루어졌다[9]. Super-Resolution 기법과 함께 학습하여 제안된 최근 결과는 VVC intra 압축 성능에 근접한 만큼 현재 가장 높은 압축 성능을 보이고 있는 방법이다[10]. [3]에서는 학습을 통해 데이터 분포를 가장 잘 보존할 수 있는 반응 필터(reponse filter)의 최적의 scale과 bias 파라미터 값을 찾는 generalized divisive normalization (GDN) 방법을 기반으로 하는 특징을 갖고 있다. 이미지 압축과 같은 low-level 비전의 경우 세부적인 이미지 정보를 보존하도록 풀링(pooling) 과정 없이 다운샘플링(downsampling)을 이용하는 등 객체 검출과 같은 컴퓨터 비전 task와는 다른 접근이 필요함을 시사한다.



Original HEVC
PSNR: 29.82



[3]
PSNR: 29.38

(그림 1) HEVC와 학습 모델 [3]을 통해 복원된 이미지의 예시
그림 1은 종단간 학습 기반 이미지 압축 방법에서 나타나는 Rate

조절에 따라 발생하는 품질 손실의 형태(artifact)를 나타낸다. 블록 기반으로 양자화가 진행되어 block artifact가 일어나는 기존의 코덱과는 달리 영상 전체에서 세부 정보가 생략되는 다른 양상으로 나타나게 되며, 여러 딥러닝 기반 압축 모델에 있어서 현재 기존의 손실 평가 지표인 PSNR과 주관적 화질 평가의 불일치에 대한 문제가 지적되고 있다[11]. 따라서 이러한 평가 지표에 대한 고찰과 딥러닝 기반 압축 모델에 대한 연구가 지속적으로 이루어져야 하며, 추후 동영상 뿐만 아니라 다시점 영상과 360 영상 등 차세대 몰입형 미디어에 특화된 코딩 툴들도 함께 연구되어야 할 필요가 있다.

ACKNOWLEDGMENT

본 연구는 한국전력공사의 2016년 선정 기초연구개발과제 연구비에 의해 지원되었음 (과제번호 : R17XA05-68). 이 논문은 또한 2018년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.2018-0-00765, 6DoF지원 초고화질 몰입형 비디오의 압축 및 전송 핵심 기술 개발)

참고문헌

- [1] ISO/IEC JTC1/SC29/WG11, "AHG on MPEG-I Visual", 127th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG/m49228, 2019.
- [2] Agustsson, Eirikur, et al. "Soft-to-hard vector quantization for end-to-end learning compressible representations." Advances in Neural Information Processing Systems. 2017.
- [3] Ballé, Johannes, Valero Laparra, and Eero P. Simoncelli. "End-to-end optimized image compression." arXiv preprint arXiv:1611.01704 (2016).
- [4] Van den Oord, Aaron, et al. "Conditional image generation with pixelcnn decoders." Advances in neural information processing systems. 2016.
- [5] Mentzer, Fabian, et al. "Conditional probability models for deep image compression." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
- [6] Mentzer, Fabian, et al. "Practical full resolution learned lossless image compression." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019.
- [7] Santurkar, Shibani, David Budden, and Nir Shavit. "Generative compression." 2018 Picture Coding Symposium (PCS). IEEE, 2018.
- [8] Agustsson, Eirikur, et al. "Generative adversarial networks for extreme learned image compression." Proceedings of the IEEE International Conference on Computer Vision. 2019.
- [9] Ballé, Johannes, et al. "Variational image compression with a scale hyperprior." arXiv preprint arXiv:1802.01436 (2018).
- [10] Lee, Jooyoung, Seunghyun Cho, and Munchurl Kim. "A Hybrid Architecture of Jointly Learning Image Compression and Quality Enhancement with Improved Entropy Minimization." arXiv preprint arXiv:1912.12817 (2019).
- [11] Blau, Yochai, and Tomer Michaeli. "The perception-distortion tradeoff." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.