

무선 이미지 전송을 위한 딥러닝 기반의 효율적인 결합 소스-채널 부호화

임수빈[°], 김정현[°], 송홍엽^{*}순천향대학교[°], 연세대학교^{*}qlstn9150@gmail.com, kimjh@sch.ac.kr, hysong@yonsei.ac.kr

An Efficient Joint Source-Channel Coding Based on Deep Learning for Wireless Image Transmission

Subeen Leem[°], Junghyun Kim[°], Hong-Yeop Song^{*}Soonchunhyang Univ.[°], Yonsei Univ.^{*}

요 약

본 논문은 무선 통신 시스템에서 이미지 전송을 위한 결합 소스-채널 부호화 기법에 대해 다룬다. 기존의 딥러닝 기반 결합 소스-채널 부호화 기법인 DJSCC의 신경망 구조를 개선하여 향상된 성능을 갖는 새로운 딥러닝 기반의 결합 소스-채널 부호화 기법을 제안한다. 실험을 통해 제안한 기법이 기존의 DJSCC 기법보다 뛰어난 성능을 보임을 확인하였다.

I. 서 론

결합 소스-채널 부호화 기법은 소스 압축률과 채널 부호율을 가변적으로 적용하여 서비스 품질을 만족시키면서 대역폭 사용 효율을 높이는 부호화 기법이다. 전통적인 소스 부호화는 JPEG나 JPEG2000과 같은 압축 알고리즘을 통해 이미지 픽셀들 간의 상관관계를 줄이고 전송되는 정보량을 줄인다. 또한 채널 부호화는 LDPC, Turbo, Polar 부호와 같은 오류정정부호를 사용하여 무선 채널 환경에서 발생하는 오류를 극복하기 위한 탄력적인 전송을 가능하게 한다.

최근 인공 신경망을 이용하여 대부분의 채널 환경에서 기존의 소스-채널 부호화의 성능을 능가하는 딥러닝 기반의 결합 소스-채널 부호화 기법인 DJSCC(Deep Joint Source-Channel Coding)[1]이 제안되었다. 해당 기법은 채널의 변화에도 급격한 성능 저하를 겪지 않는 장점이 있다. 그러나 일부 채널 환경에서는 여전히 기존의 소스-채널 부호화의 성능보다 낮은 성능을 보이므로 여전히 개선의 여지가 남아있다.

본 논문에서는 DJSCC의 구조적 한계를 극복하여 더 나은 성능을 갖는 결합 소스-채널 부호화 기법을 제안한다. 특히 DJSCC에서 사용한 단순한 구조의 CNN에 층 간 연결을 추가한 DenseNet[2] 구조를 송수신기에 활용함으로써 이미지 특징 추출 기능을 향상시켰다.

II. 본론

DJSCC에서 사용한 신경망 구조는 송신기와 수신기 모두 기본 컨볼루션 층이 5개 쌓여있는 CNN 구조로 되어있다. DJSCC의 성능을 개선하기 위해 기존 송신기 구조에서 두 번째 층의 출력을 세 번째와 네 번째 층의 출력에 연결하고, 세 번째 층의 출력을 네 번째 층의 출력에 연결하였다. 또한 기존 수신기 구조에서 첫 번째 층의 출력을 두 번째와 세 번째 층의 출력에 연결하고, 두 번째 층의 출력을 세 번째 층의 출력에 연결하였다. 이러한 밀집한 신경망 구조를 통해 파라미터 수를 크게 증가시키지 않고도 이미지 특징 추출 기능을 향상시킬 수 있다.

제안하는 DJSCC-DN 기법의 구조가 그림 1에 표현되어 있다. 송신기에서는 $0 \sim 255$ 의 픽셀 값을 갖는 입력 이미지를 $0 \sim 1$ 로 정규화 한 후, 다섯 개의 컨볼루션 층으로 전달한다. 각 컨볼루션 층의 활성화함수로 비선형 활성화 함수인 PReLU[3]를 사용한다. 그림 1의 각 블록 위에 쓰여진 표기는 (필터 가로 $\times$ 필터 세로 $\times$ 필터 수 / 스트라이드 간격)을 의미하는데, 마지막 컨볼루션 층의 필터 수 c 는 학습 시에 정의하는 대역폭 압축률에 따라 조절된다. 이어서 평균 전력 제약 조건을 만족하게 하는 정규화[4]를 진행한 후, 잡음 채널을 통해 전송된다. 수신기는 다섯 개의 전치 컨볼루션 층과 역정규화 층으로 구성되어 송신기와 대칭되는 구조로 설계하였다. 마

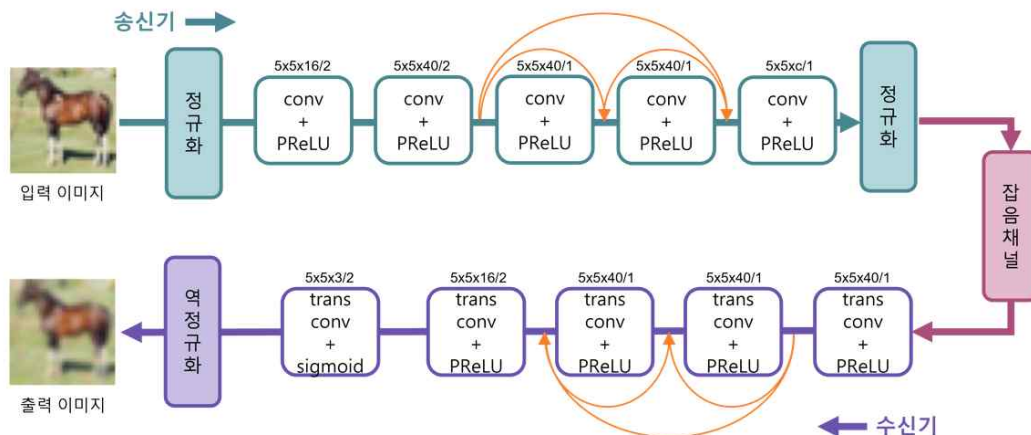


그림 1. DJSCC-DN의 구조

지막 역정규화 층은 전치 컨볼루션 층의 출력 값을 원본 이미지의 픽셀 값의 범위인 0~255로 맞춰주는 역할을 한다.

학습 및 테스트에 사용한 데이터셋은 32x32크기의 컬러 이미지들로 구성된 CIFAR-10이다. 이 중 50,000장의 이미지는 학습 데이터로, 10,000장의 이미지는 테스트 데이터로 분류하여 사용했다. 아래 표는 DJSCC-DN의 학습에 사용된 하이퍼 파라미터를 나타낸 것이다.

파라미터	설정값
손실 함수	MSE(Mean Squared Error)
배치 크기	100
최적화 방법	Adam, 학습률=0.001
반복 학습 횟수	20
학습 방법	SNR=0,10,20dB마다 학습

송신된 이미지와 수신된 이미지 간의 손실 정보를 평가하기 위해 최대 신호 대 잡음비(Peak Signal-to-Noise Ratio, PSNR)[5]를 성능 지표로 사용하였다. 이는 다음의 식 (1)과 같이 정의된다.

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

여기서 MAX는 이미지의 한 픽셀이 가질 수 있는 최댓값이고, MSE는 송신 이미지와 수신 이미지의 각 픽셀 오차를 제공하여 평균 낸 값이다. 즉, PSNR값이 클수록 손실이 적은 이미지라고 할 수 있다.

그림 2는 기존 DJSCC 기법과 제안하는 DJSCC-DN 기법의 대역폭 압축률에 따른 PSNR 성능을 비교한 것이다. 여기서 대역폭 압축률은 입력 이미지가 n차원 실수 벡터로 표현되고 채널 입력이 k차원 복소수 벡터로 표현될 때 이들 간의 비율인 k/n로 정의된다. 학습 및 테스트를 위한 채널 환경으로 신호 대 잡음비(Signal-to-Noise Ratio, SNR)를 0dB, 10dB, 20dB로 고려하였다. 또한 대역폭 압축률은 0.06, 0.26, 0.49인 경우를 고려하였다. 고려한 각 채널 환경에서 모든 대역폭 압축률에 대해 제안하는 DJSCC-DN 기법이 기존 DJSCC 기법 대비 같거나 더 나은 성능을 보이는 것을 확인할 수 있다.

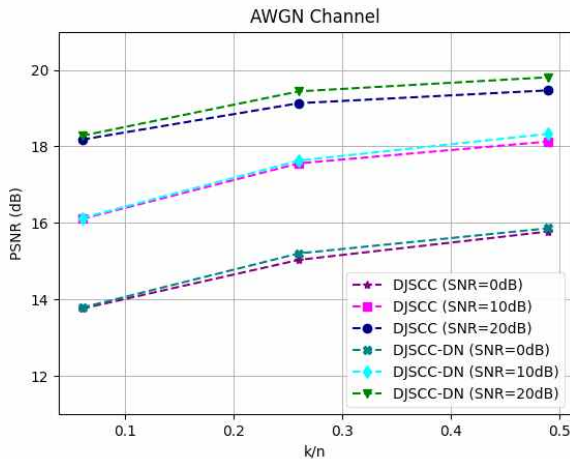


그림 2. 대역폭 압축률(k/n)에 따른 PSNR 성능 비교

그림 3은 대역폭 압축률을 0.49로 고정하고 다양한 테스트 채널 환경에서 기존 DJSCC 기법과 제안하는 DJSCC-DN 기법의 PSNR 성능을 비교한 것이다. 신호 대 잡음비가 0dB, 10dB, 20dB인 환경을 타겟으로 학습된 각 모델에 대해 타겟 채널 환경과 유사한 테스트 환경에서 제안하는 DJSCC-DN 기법이 기존 DJSCC 기법 대비 같거나 더 나은 성능을 보이는 것을 확인할 수 있다. 특히 채널 환경이 좋을수록 이 격차는 크게 증가하였다.

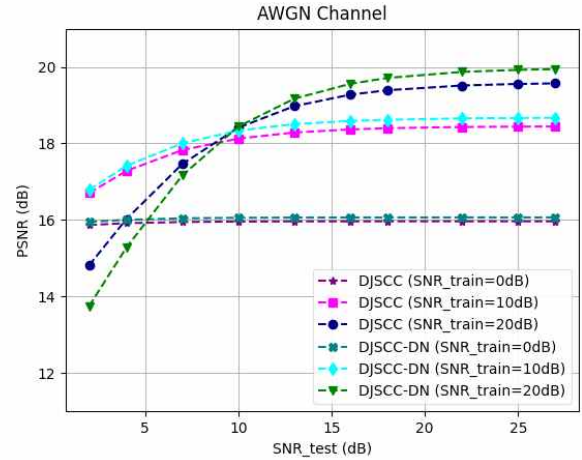


그림 3. 테스트 환경에서의 신호 대 잡음비에 따른 PSNR 성능 비교

III. 결론

본 논문에서는 무선 통신 시스템에서 이미지 전송을 위한 딥러닝 기반 결합 소스-채널 부호화 기법인 DJSCC의 구조를 개선하여 DJSCC-DN을 제안하였다. 다층으로 구성된 CNN 기반 송수신기 구조에 층 간 연결을 추가하여 특징 추출 기능을 향상시켰다. 향후 추가 성능 개선을 위한 연구 방법으로 새로운 신경망 구조 활용과 PSNR 뿐만 아니라 SSIM[5]을 성능 지표로 활용한 학습 방법과 다양한 채널 환경에서의 성능 비교 및 분석 등을 고려할 수 있다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.2020R1A2C201196911).

참 고 문 헌

- [1] E. Bourtsoulatz, D. Burth Kurka, and D. Gündüz, "Deep Joint Source-Channel Coding for Wireless Image Transmission," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 567-579, Sep. 2019.
- [2] G. Huang, Z. Liu, L. Van Der Maaten, and K. Weinberger, "Densely Connected Convolutional Networks," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017 pp. 2261-2269.
- [3] T. Jiang and J. Cheng, "Target Recognition Based on CNN with LeakyReLU and PReLU Activation Functions," in *Proc. International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)*, 2019, pp. 718-722.
- [4] I. Kozintsev and K. Ramchandran, "Robust Image Transmission over Energy-Constrained Time-Varying Channels using Multiresolution Joint Source-Channel Coding," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 1012-1026, Apr. 1998.
- [5] D. Ziou and A. Horé, "Image Quality Metrics: PSNR vs. SSIM," in *Proc. International Conference on Pattern Recognition (ICPR)*, Istanbul, Turkey, 2010 pp. 2366-2369.