

이미지 전송을 위한 시맨틱 및 비-시맨틱 하이브리드 통신

남혜린, 김성륜
연세대학교

hlnam@ramo.yonsei.ac.kr, slkim@yonsei.ac.kr

Hybrid Semantic and Non-semantic communication for Image transmission

Hyelin Nam, Seong-Lyun Kim

요 약

본 논문은 기존 이미지 복원을 위한 시맨틱 통신에서 발생하는 한계점을 완화하기 위하여 비-시맨틱 통신과 결합한 통신 시스템을 제안한다. 시맨틱 통신에서는 인공지능 (Artificial Intelligence, AI) 기반의 송신자와 수신자가 이미지 전송 시 시맨틱 신호로 인코딩, 전송 후 원본 이미지로 재생성하는 구조를 가진다. 이에 사용되는 AI 모델은 여러 개의 이미지로 시맨틱 신호 변환, 복원을 학습하기 때문에 특정 이미지에 대해 완벽한 복원을 하지 못하는 근본적인 한계점이 있으며, 이를 추가적인 비-시맨틱 정보를 전송함으로써 복원하는 알고리즘을 제안한다.

I. 서 론

시맨틱 통신은 전송 데이터가 가진 의미 혹은 대표할 수 있는 정보를 전송함으로써 의도된 task 를 잘 수행하는 것을 목표로 한다.[1] 이는 AI 의 발전과 함께 송, 수신자를 AI 로 구현함으로써 데이터로부터 의미 추출, 채널 신호로 변환까지의 일련의 과정을 수행할 수 있다. 그러나, 다량의 이미지로 이를 학습하기 때문에, 송, 수신자는 이미지들의 분포 특성을 반영하여, 개별 이미지에 대한 특징들을 구현하는데에 한계점이 존재한다.[2] 본 논문에서는 시맨틱 신호로 추출되지 않은 정보를 구성하여 추가 전송하며, 수직적 공간에 투영하는 수학적 기법을 통해 원본 데이터로 복원할 수 있는 하이브리드 시맨틱 통신 방법 (HSC, Hybrid Semantic Communication) 을 제안한다.

II. 본론

가. 시스템 구조

시맨틱 통신을 구성하는 송신자와 수신자는 AI 함수로 정의된다. 송신자는 인코더 모델을 가지고 있다. X 는 $L \times L$ 크기의 이미지이며, 인코딩 함수 f 를 통과하여 시맨틱 정보를 담고 있는 k 개 채널 심볼로 구성되어 있는 시맨틱 신호로 변환된다. 이후 이는 정량화 과정을 통해 전송 파워를 만족하는 신호 z 로 변환된다.

수신자는 디코더 모델을 사용하여 수신한 신호를 이미지로 복원한다. 이때 채널 노이즈에 영향을 받은 왜곡 신호는 z' 으로 표현한다. 수신자가 수행하는 디코딩 함수는 f^{-1} 이며, 결과적으로 $L \times L$ 크기의 이미지가 복원된다. 송, 수신자가 수행하는 인코딩, 디코딩 과정은 다음과 같다.

$$z = f(X), X' = f^{-1}(z') \quad (1)$$

나. 추가 통신을 통한 이미지 복원

수신자는 복원한 이미지 X' 를 가지며, 송신자는 원본 이미지 X 에 대한 비시맨틱 정보를 추가로 전송한다. 이에 수신자는 두 데이터를 종합하여 이미지 X^* 를 구성한다. 이는 복원 이미지와 원본 이미지 크기를 동일하게 유지하며, 복원 이미지보다 더 작은 픽셀 단위 차이를 발생시킨다. 이는 MSE 함수를 활용하여 아래 수식으로 측정되고, 본 논문에서 MSE 를 최소화하는 것을 목표로 한다.

$$MSE(X, X^*) = 1/L^2 \|X - X^*\|^2 \quad (2)$$

수신단에서 구성하는 이미지는 생성 이미지와 추가 수신한 데이터를 결합하기 위하여 수직 공간에 각 데이터를 투영하는 방식을 사용한다. 이에 A 라는 2 차원 행렬을 가정하면 A 의 열들의 선형 결합으로 구성된 벡터 공간이 구성된다. 이 공간은 수직, 보완적인 두 공간으로 나뉘며, 각각 열공간 (range space) 와 영공간 (null space) 로 나뉜다. 따라서, 송신자에서 다음과 같이 원본 이미지를 분해할 수 있다. [3]

$$X = A^+AX + (I - A^+A)X \quad (3)$$

원본 이미지가 가지는 비시맨틱 정보는 A 행렬을 이용한 선형변환으로 만들어진다고 정의하며, $y = AX$ 으로 나타낸다. 이에 식 (3) 와 같은 열, 영공간 분리 공식을 (Range-null decomposition) 활용하여 수신자가 가지는 생성 이미지는 열공간에 투영하고, 추가로 수신한 비시맨틱 정보는 영공간에 투영하여 결합한다. 구성된 최종 이미지는 다음과 같이 나타난다.

$$X^* = A^+y + (I - A^+A)X \quad (4)$$

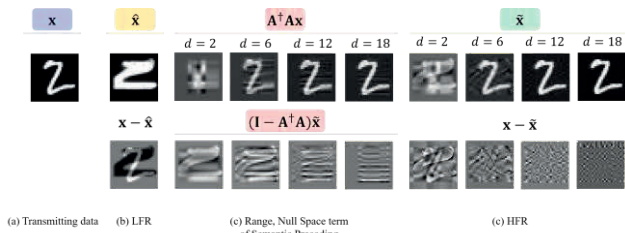


그림 1. d 에 따른 열, 영공간 투영 이미지 및 복원 이미지

식 (3)와 (4)을 참고하여 식 (2)의 픽셀 단위 차이점을 유도하면 $\text{Tr}((I - A^+A)(X - X')^T(X - X'))$ 와 같다. 이를 최소화하는 투영 함수 및 비시맨틱 추출함수인 A는 원본 이미지와 복원 이미지의 차이로 만들어진 행렬, $(X - X')^T(X - X')$ 의 고유 벡터 (eigen vector)로 구성하였을 때와 같다. 즉, A가 dxL 크기일 때, d개의 고유 벡터로 구성되며, 다음과 같이 구성하였을 때와 같다. d는 추가전송하는 비시맨틱 정보 y의 크기를 결정하며, d가 커짐에 따라 통신량이 증가한다. 또한, 원본 이미지 특정 고유 특징을 반영하여 MSE는 줄어든다. d=L일 경우 MSE=0으로 동일한 이미지로 복원이 가능하게 된다.

다. 실험 결과

2차원 이미지 데이터인 MNIST를 사용하였고, 각 이미지는 $L \times L = 28 \times 28$ 크기이다. 인코더와 디코더는 각각 4개의 FC (fully-connected) 신경망 레이어로 구성되어 있으며, 인코더에서 출력하는 신호는 $k=128$ 개의 채널 심볼로 이루어져 있다.

그림 1은 d를 증가시킴에 따른 복원 이미지의 결과를 보여준다. 추가 전송을 하지 않았을 시 시맨틱 신호를 이용해 복원한 이미지는 원본 이미지와 의미적으로 유사하나, 디테일 및 고유 특징은 복원하지 못한다. 그러나, d에 비례한 y 추가통신 후 식 (4)를 통한 복원 이미지는 원본 이미지의 특징을 반영함을 보인다.

그림 2는 채널 환경에 따른 최종 복원 이미지와 원본 이미지의 차이인 MSE를 나타낸다. 추가 전송을 하지 않고 기존 시맨틱 통신에 의한 복원 이미지와의 차이점은 'SC'로 나타낸다. 본 채널환경에서 y는 JPEG 압축 및 LDPC 코드, 16QAM 기법을 사용하여 채널 심볼로 변환하여 전송한다. 시맨틱 인코더와 디코더가 채널 노이즈가 있는 환경에서 학습된 경우는 'noise robust'라고 일컫는다. 이를 신호 대 잡음비 (SNR)에 대해 실험하였을 때, 제안하는 HSC 기법은 SC에 비해 MSE가 작음을 확인할 수 있다. SC는 noise robust의 경우, 인코더와 디코더의 이미지 복원 성능이 모든 채널 환경에서 최적화되도록 학습되기 때문에, SNR이 높은 채널 환경에서 성능이 떨어짐을 알 수 있다. 반면, HSC는 채널 환경이 우수할수록 비시맨틱 정보의 추가 통신을 통해 MSE가 개선된다.

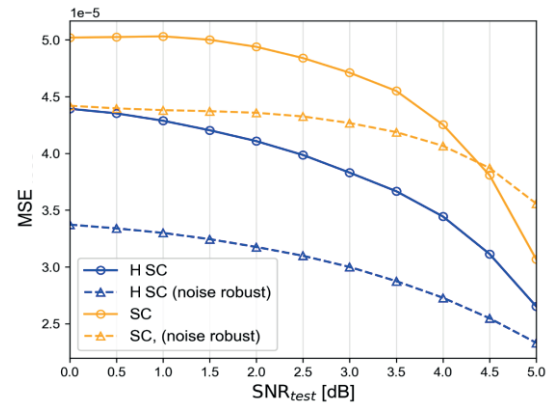


그림 2. 채널 환경에 따른 MSE 차이점

III. 결론

본 논문에서는 시맨틱 통신의 한계점을 극복하고, 통신량에 대해 원본 이미지와 복원 이미지의 차이를 감소시키기 위하여 비시맨틱 정보를 추가 전송하는 하이브리드 시맨틱 통신 방식, HSC를 제안한다. 이는 d로 통신량을 조절할 수 있으며, 이에 대해 복원 성능이 높아짐을 보였다. 또한, 다양한 채널 환경에서 실험 시, 복원 이미지와 원본 이미지의 차이가 작은 경향을 보인다.

ACKNOWLEDGMENT

이 논문은 2025년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2021-0-00347, 6G 통신을 위한 Post MAC, No.2022-0-00420, 6G

중단간 초정밀 네트워킹을 위한 핵심기술 개발)

참고 문헌

- [1] Choi, J., Nam, H., Park, J., Ko, S.-W., Choi, J., Bennis, M. and Kim, S.-L. (2025). Interoperable Semantic Communication. In Foundations of Semantic Communication Networks (eds W. Saad, C. Chaccour, C. Kurisummootil Thomas and M. Debbah). <https://doi.org/10.1002/9781394247912.ch10>
- [2] G. Cicchetti, E. Grassucci, J. Park, J. Choi, S. Barbarossa and D. Comminiello, "Language-Oriented Semantic Latent Representation for Image Transmission," 2024 IEEE 34th International Workshop on Machine Learning for Signal Processing (MLSP), London, United Kingdom, 2024, pp. 1-6, doi: 10.1109/MLSP58920.2024.10734812.
- [3] Grassucci, Eleonora, et al. "Rethinking multi-user semantic communications with deep generative models." arXiv preprint arXiv:2405.09866 (2024).