

# 채널 적응형 디지털 시맨틱 통신을 위한 공동 소스-채널 부호화

박주혁, 진요셉  
포항공과대학교

{joohyuk.park, yoseb.jeon}@postech.ac.kr

## Joint Source-Channel Coding for Channel-Adaptive Digital Semantic Communications

Joohyuk Park, Yo-Seb Jeon

Pohang Univ. of Science and Technology (POSTECH)

### 요약

본 논문은 다양한 채널 환경에 강인한 디지털 시맨틱 통신을 위한 공동 소스-채널 부호화(JSCC)를 제안한다. 기존의 시맨틱 통신 시스템에서는 디지털 변복조 방식의 비선형성, 다양한 채널 환경 및 변조 차수로 인해 종단간 기계학습을 JSCC에 적용하는 것이 어려웠다. 이를 해결하기 위해서 우리는 디지털 복조 후 출력의 불확실성을 개선할 수 있는 새로운 복조 방법을 개발한다. 나아가 인코더 출력과 디코더 입력 사이의 확률적 관계를 등가 채널 모형으로 모델링하여 JSCC의 종단간 기계학습이 가능하게 만든다. 추론 과정에서는 학습 과정에서 설정한 매개변수를 이용하여 과제 수행 성능을 높이고 통신 지연을 줄인다. 이미지 분류 과제에서 제안하는 기법이 기존의 JSCC 기법 대비 높은 성능을 도출함을 입증한다.

### I. 서론

시맨틱 통신(Semantic Communications)은 메시지의 의미를 송수신하는 과정으로 대용량 데이터 처리, 빠른 과제 수행을 목표로 한다. 이때 송신기의 목표는 수신기에서 목표 과제를 높은 성능으로 수행할 수 있도록 메시지의 의미를 전송하는 것이다. 현재까지 의미론적 정보(semantic information) 교환을 목적으로 딥러닝(Deep learning) 기반의 공동 소스-채널 부호화(JSCC)를 이용하는 기법이 활발히 연구되었다. 이중 대부분의 연구는 특정 통신 환경에서 JSCC 인코더 및 디코더를 종단간 학습(end-to-end learning)했다. 이는 학습 환경과 다른 신호 대 잡음비(SNR), 디지털 변조 방식을 이용하는 추론 환경에서 시스템의 성능을 보장할 수 없다는 단점이 있다. 이를 해결하기 위해 채널 적응형 JSCC 접근법도 제안되었으나, 이러한 연구는 디지털 변조 방식을 고려하지 않은 아날로그 통신 시스템을 기반으로 하기에 현대의 디지털 통신 시스템과 호환성이 떨어진다.

본 논문에서는 디지털 변조 차수(modulation order)를 유연하게 적용할 수 있고 여러 SNR에 강인하게 동작하는 디지털 시맨틱 통신을 위한 공동-소스 채널 부호화를 제안한다. 이를 위해 (i) 강인한 디지털 복조 방법, (ii) 강인한 학습 전략, (iii) 채널 적응형 변조 기법을 제시한다. (i)에서는 기존의 이진 출력(0,1) 대신 삼진 출력(0,0.5,1)을 이용하여 채널과 잡음에 의해 복조 출력의 불확실성이 발생할 때 0.5를 출력한다. (ii)에서는 인코더의 출력과 디코더의 입력 사이의 관계를 이진 대칭 이레이저 채널(BSEC)로 모델링하고 해당 모델에서 이용되는 매개변수들을 확률 분포에서 추출하여 이용한다. 이를 통해 다양한 채널 환경에 강인하고 변조 차수에 유연한 모델이 학습된다. (iii)에서는 채널 환경에 따라서 잠재 변수(latent variables)의 변조 차수를 결정함으로써 과제 성능을 유지함과 동시에 통신 지연을 낮추었다.

### II. 본론

본 논문에서는 특정 기계학습 과제를 수행하는 JSCC 인코더와 디코더로 구성된 디지털 시맨틱 통신 시스템을 고려한다. 송신기는 입력 데이터를 이진 잠재 변수로 변환한다. 이후 디지털 변조를 적용하면  $2^M$ -QAM 심볼이 얻어지고 무선 감쇠 채널을 거쳐 수신기에 도착한다. 수신기에서는 채널 이퀄라이제이션을 거쳐 다음의 신호를 얻는다.

$$\tilde{y}[n] = hx[n] + \tilde{v}[n] \quad (1)$$

이때  $\tilde{v}[n] \sim CN(0, 1/\text{SNR})$ ,  $\text{SNR} = |h|^2/\sigma^2$ 이다. 위 신호가 디지털 복조를 거쳐 JSCC 디코더의 입력을 만든다. 디코더의 출력은 기계학습 과제를 수행하기 위한 신경망의 입력이 된다.

본 논문에서 고려하는 강인한 디지털 복조 방법의 경우, 각 잠재 변수에 대한 결정의 신뢰도 수준을 로그 공산비(LLR)를 기반으로 측정하여 중간값 0.5를 할당하는 기준을 결정한다. 시간 슬롯  $n$ 에서 전송된 심볼  $x[n]$ 과 관련된  $m$ 번째 이진 잠재 변수를  $b_{nm}$ 이라 하면  $b_{nm}$ 의 LLR을 다음과 같이 계산할 수 있다.

$$\mathcal{L}_m(\tilde{y}[n]) = \ln \left( \frac{\mathbb{P}(b_{nm} = 0 | \tilde{y}[n])}{\mathbb{P}(b_{nm} = 1 | \tilde{y}[n])} \right) \quad (1)$$

만약  $\mathbb{P}(b_{nm} = 0 | \tilde{y}[n]) \approx \mathbb{P}(b_{nm} = 1 | \tilde{y}[n])$  라면 복조 과정에서  $b_{nm}$ 의 불확실성이 증가하고 이에 대응되는 LLR이 0에 근접한다. 따라서 수신 신호  $\tilde{y}[n]$ 을 기반으로  $b_{nm}$ 을 결정할 때 LLR의 크기를 신뢰 수준의 척도로 이용할 수 있다.  $\tau_{nm} > 0$ 을 중간값 0.5를 적용하기 위한 문턱값(threshold)이라고 하면 제안하는 복조 방법  $\text{Demod}_M(\tilde{y}[n]) = [\hat{b}_{n_1}, \hat{b}_{n_2}, \dots, \hat{b}_{n_M}]$ 은 다음과 같이 유도된다.

$$\hat{b}_{nm} = \begin{cases} 0, & \text{if } \mathcal{L}_m(\tilde{y}[n]) > \tau_{nm} \\ 0.5, & \text{if } |\mathcal{L}_m(\tilde{y}[n])| \leq \tau_{nm} \\ 1, & \text{if } \mathcal{L}_m(\tilde{y}[n]) < -\tau_{nm} \end{cases} \quad (3)$$

본 논문에서는 예측 불가능한 채널 환경 및 변조 차수, 디지털 변복조의 비선형성을 반영하여 end-to-end

learning 을 수행하는 것의 어려움을 극복하고자 이러한 효과를 확률적으로 나타내는 BSEC model 을 도입했다. 해당 모델에서  $i$  번째 JSCC encoder output  $s_i$  와  $i$  번째 JSCC decoder input  $\hat{b}[i]$ 의 관계는 다음과 같다.

$$p_{\text{noisy}}(\hat{b}_i | s_i, \tilde{\mu}_i, \tilde{d}_i, \theta) = \begin{cases} \tilde{\mu}_i s_i + (1 - \tilde{d}_i - \tilde{\mu}_i)(1 - s_i), & \text{if } \hat{b}_i = 0 \\ \tilde{d}_i, & \text{if } \hat{b}_i = 0.5 \\ \tilde{\mu}_i(1 - s_i) + (1 - \tilde{d}_i - \tilde{\mu}_i)s_i, & \text{if } \hat{b}_i = 1 \end{cases} \quad (4)$$

이때  $1 + \lfloor (i-1)/M \rfloor = n$ 이 성립하며 위 식에서  $\tilde{\mu}_i$ 와  $\tilde{d}_i$ 는 각각  $i$  번째 BSEC model 의 bit-flip 확률, bit-erasure 확률을 나타내는 매개변수들이다. 서로 다른 채널 환경에 강인하게 동작하기 위해 학습과정에서  $\tilde{\mu}_i$ 를 균등 분포에서 추출한다 ( $\tilde{\mu}_i \sim \text{Uniform}[0, \alpha_i]$ ), 여기서  $\alpha_i$ 는 목표로 하는 강인함 수준(robustness level)이다. 잡재 변수마다 다른  $\alpha_i$ 를 설정하여 추론 과정에서 유연하게 변조 차수를 선택할 수 있게 한다.

본 논문에서는 추론 과정에서 과제 성능을 최대화하기 위해 채널 환경과 학습 과정에서 설정한 강인함 수준에 따라 변조 차수를 변경하는 방법을 제시한다. 가장 직관적인 방법은 채널 환경에 상관없이 가장 낮은 변조 차수(4QAM)를 설정하는 것인데, 이는 SNR 이 충분히 높을 때 시멘틱 통신의 스펙트럼 효율성(Spectral efficiency)을 떨어뜨린다는 단점이 존재한다. 제안하는 기법은 JSCC 인코더와 디코더가 bit-flip 확률이  $\alpha_i$ 일 때까지는 성능을 보장할 수 있도록 학습했기 때문에 추론 과정에서도 이진 에러가 발생할 확률이 특정값 이하 ( $\tilde{\mu}_i \leq \beta_M \alpha_i$ )가 되도록 변조 차수를 결정한다. 이때  $\beta_M < 1$ 은 비트오류율(BER) 근사치의 영향을 보정하기 위한 조정 상수이다. 이러한 기준에 따라 변조 차수를 결정하기 위해서  $2^M$ -QAM 의 BER 을 근사하면 다음을 얻는다.

$$\tilde{\mu}_i = \frac{4}{M} \left(1 - \frac{1}{\sqrt{2^M}}\right) Q\left(\frac{(1 + \alpha_i) \sqrt{3\text{SNR}}}{\sqrt{2^M - 1}}\right) \quad (5)$$

위 식에서  $Q(\cdot)$ 은 Q 함수이고,  $\alpha_i \in [0, 1]$ 는 중간값 0.5 를 할당하기 위한 결정 경계(decision boundary)로 앞서 서술한 threshold 와 다음의 관계가 있다.

$$\tau_i = \frac{1}{\tilde{\sigma}^2} \left\{ \left( \frac{(1 + \alpha_i) d_M}{2} \right)^2 - \left( \frac{(1 - \alpha_i) d_M}{2} \right)^2 \right\} \quad (6)$$

$$= \frac{d_M^2 \alpha_i}{\tilde{\sigma}^2} = \frac{6\text{SNR}}{2^M - 1} \alpha_i$$

위 식에서  $d_M = \sqrt{6/(2^M - 1)}$  은 정규화된  $2^M$ -QAM 의 심볼 간 최소 거리이다. 학습 과정에서는  $\tau_i = \text{SNR}$ ,  $M = 2$ 로 설정하여 모든 잡재 변수에 대해  $\alpha_i = 0.5$ 를 이용했으며 추론 과정에서도 동일한 값을 가지게 설정했다. 예를 들어, (5)의 BER 근사를 기반으로 최적의 변조 차수를 4QAM, 16QAM, 64QAM 중에서 결정하면 다음과 같다.

$$M_i^* = \begin{cases} 2, & \text{if } \tau_{2,i} \leq \sqrt{\text{SNR}} \leq \tau_{4,i} \\ 4, & \text{if } \tau_{4,i} \leq \sqrt{\text{SNR}} \leq \tau_{6,i} \\ 6, & \text{if } \tau_{6,i} \leq \sqrt{\text{SNR}} \end{cases} \quad (7)$$

위 식에서  $\tau_{2,i} = \frac{1}{1 + \alpha_i} Q^{-1}(\beta_2 \alpha_i)$ ,  $\tau_{4,i} = \frac{\sqrt{5}}{1 + \alpha_i} Q^{-1}\left(\frac{4\beta_4 \alpha_i}{3}\right)$ ,  $\tau_{6,i} = \frac{\sqrt{21}}{1 + \alpha_i} Q^{-1}\left(\frac{12\beta_6 \alpha_i}{7}\right)$ 이고  $Q^{-1}(\cdot)$ 는 Q 함수의 역함수이다. (7)을 통해  $\alpha_i$ 가 작을수록  $M_i^*$ 도 작아진다는 것과 SNR 이 좋을수록 각 잡재 변수에서 더 높은 변조 차수를 선택함을 알 수 있다. 잡재 변수마다  $\alpha_i$ 를 다르게 설정했기 때문에 각기 다른 변조 차수를 이용하여 심볼로 변환된다. 이러한 특징은 넓은 SNR 영역에서 과제 성능은 유지하면서 평균적인 스펙트럼 효율성을 각 SNR 에서 극대화시킨다.

본 연구에서는 제안된 JSCC 기법의 우수성을 입증하기 위해 MNIST 데이터셋을 이용하여 이미지 분류 과제를 수행하는 모의 실험을 수행했다. 성능 비교를 위해 학습 과정동안 이진 대칭 채널(BSC)의 bit-flip 확률이 모든 잡재 변수에 대해 동일한 NECST [1] 기법을 이용했다. 이는 채널 적응형 기법이 아님을 알 수 있다.

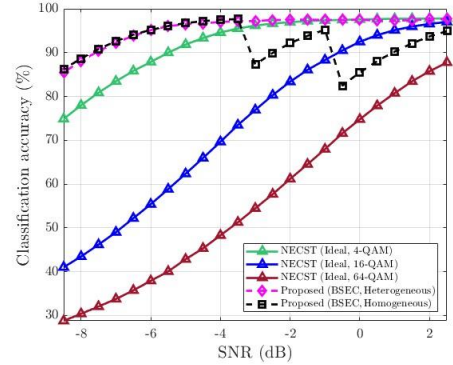


그림 1. 제안된 JSCC 기법과 NECST 기법의 정확도 비교

그림 1 에서 NECST 는 다양한 훈련 모델 중 성능이 잘 나온 점들만 나타낸다. 제안기법 중 heterogeneous 는  $\alpha_i$ 가  $i$ 에 따라 선형 증가하는 경우 ( $\alpha_1 = 0.29, \alpha_N = 0.45$ )이고 homogeneous 는  $\alpha_i = 0.4, \forall i$  인 경우이다. Heterogeneous 는 전체 SNR 범위에서 고정 변조 차수를 이용한 NECST 기법보다 항상 좋은 성능을 보임을 알 수 있다. Homogeneous 는 높은 변조 차수로 변할 때 성능 감소가 일어남을 알 수 있다. 이는 모든 잡재 변수들이 같은 변조 차수를 가지기 때문에 발생한 현상이다. 예를 들어  $-3\text{dB} \leq \text{SNR} \leq -1\text{dB}$  일 때 제안기법은 16QAM 을 사용하기 때문에 4QAM 을 이용한 NECST 기법 대비 성능 저하가 나타난다. 그럼에도 NECST 와 같은 변조 차수를 이용할 때 성능을 비교하면 더 좋은 결과를 낸다. 우리의 결과는 잡재 변수 간에 서로 다른  $\alpha_i$ 를 할당하는 것이 제안된 JSCC 접근 방식에 있어 변조 차수 선택의 유연성뿐만 아니라 과제 성능을 극대화하기 위해 필수적임을 명확히 보여준다.

### III. 결론

본 논문에서는 채널 적응형 디지털 시멘틱 통신을 가능하게 하는 새로운 JSCC 접근 방식을 제안했다. 이를 위해 우리는 먼저 이진 잡재 변수의 빈번한 bit-flip 을 방지하면서 복조 출력의 표현력을 향상시키는 강건한 복조 방법을 개발했다. 그 후에는 JSCC 인코더와 디코더의 end-to-end learning 을 용이하게 하는 동시에 다양한 채널 상태와 변조 차수에 대한 그들의 강인함과 유연성을 향상시키는 훈련 전략을 개발했다. 또한 통신 지연 시간을 줄이면서 과제 성능을 유지할 수 있는 채널 적응형 변조 차수 변경 기술을 고안했다. 모의 실험을 통해 제안된 접근 방식이 통신 지연 시간과 과제 성능 측면에서 기존의 JSCC 접근 방식을 능가함을 보여주었다.

### ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2022R1C1C1010074).

### 참고 문헌

- [1] K. Choi, K. Tatwawadi, A. Grover, T. Weissman, and S. Ermon, "Neural joint source-channel coding," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Long Beach, CA, USA, Jun. 2019, pp. 1182–1192.