

# 고차 변조가 적용된 디지털 시맨틱 통신을 위한 파라미터 샘플링 기반 훈련 기법

김찬형, 이건희, 채민주, 임승찬<sup>\*</sup>

한경국립대학교, <sup>\*</sup> 홍익대학교

cksguddl841@hknu.ac.kr, gunny1226@hknu.ac.kr, cmj3254@hknu.ac.kr, sclim@hongik.ac.kr<sup>\*</sup>

## Parameter-Sampling-Based Training for Digital Semantic Communications under High-Order Modulation

Chan Hyung Kim, Geonhui Lee, Minju Chae, Seung-Chan Lim<sup>\*</sup>

Hankyong National University, <sup>\*</sup> Hongik University

### 요약

본 논문에서는 고차 변조 환경에서 디지털 시맨틱 통신 성능을 향상하기 위한 파라미터 샘플링 기반 훈련 기법을 제안한다. 주어진 변조 차수에 대한 실제적인 전송 특성을 반영하여 디지털 통신 과정을 이진 대칭 소실 채널로 등가 모델링하고 이를 활용해 종단 간 학습을 수행함으로써 강건한 시맨틱 통신 송수신기를 설계하였다.

### I. 서론

시맨틱 통신은 맥락에 맞는 의미를 전달하여 수신자의 과업 수행을 보장하는 패러다임이다 [1], [2]. 최근에는 디지털 파이프라인 상에서 의미를 전송하기 위한 디지털 시맨틱 통신을 설계하는 접근이 주목 받고 있다 [3], [4]. 기존 연구 [3]에서는 삼진 복조(ternary demodulation)를 통해 불확실한 잡재 변수를 소실(erasure)로 처리하고, 등가 이진 대칭 소실 채널(binary symmetric erasure channel, BSEC) 모델링을 활용한 종단 간 학습 방법을 제안하였다. 하지만, 기존 기법은 BSEC 확률에 4-QAM에 대한 근사식을 반영함에 따라 고차 변조 적용 시 성능 열화가 발생할 수 있다.

본 논문에서는 고차 변조가 적용된 디지털 시맨틱 통신의 성능 향상을 위해 두 가지 파라미터 샘플링 기반 훈련 기법을 제안한다: 1)  $M$ -QAM 확률 근사 기반 훈련 기법, 2) 최상위 비트(most significant bit, MSB) 지향 확률 모델링 기반 훈련 기법. 제안한 훈련 과정에서는 잡재 변수의 비트 반전 확률(bit-flip probability), 비트 정판정 확률(bit-correct probability), 비트 소실 확률(bit-erasure probability)을 근사적으로 분석하고, 이를 활용하여 디지털 통신 과정을 등가 BSEC 모델링함으로써 종단 간 학습을 수행한다. 시뮬레이션 결과, 256-QAM 전송 환경에서 제안 기법이 기존 기법 대비 개선된 이미지 전송 성능을 달성함을 검증하였다.

### II. 디지털 시맨틱 통신 시스템 모델

본 논문에서는 이미지 전송을 위한 디지털 시맨틱 통신 시스템을 고려한다. 송신단에서 원본 이미지  $\mathbf{u}$ 는 신경망 가중치  $\theta$ 로 매개 변수화된 인코더  $f_\theta(\cdot)$ 를 통해 길이  $N$ 의 잡재 벡터  $\mathbf{v} = f_\theta(\mathbf{u})$ 로 변환된다. 잡재 벡터는  $B$ -비트 균등 양자화기를 통해 길이가  $NB$ 인 이진 잡재 벡터  $\mathbf{b}$ 로 변환된다 [4]. 변조 차수  $M$ 에 대하여,  $\mathbf{b}$ 를  $\log_2 M$  비트씩 블록화하여  $T = NB/\log_2 M$  개의 부벡터  $\mathbf{b}_t = [b_{t,1}, \dots, b_{t,\log_2 M}]$ ,  $t \in \{1, \dots, T\}$ 로 나눈다. 각  $\mathbf{b}_t$ 는  $M$ -QAM 변조를 통해 심볼  $x_t$ 로 변환되며, 순차적으로 수신단으로 전송된다. 시간 슬롯  $t$ 에 대하여, 수신 신호는  $y_t = hx_t + n_t$ 로 표현된다. 여기서,  $h$ 는 블록 페이딩 채널 계수이며,  $n_t$ 는 평균이 0이고 분산이  $\sigma^2$ 인 가산 백색 가우시안 잡음이다. 송신 심볼의 평균 에너지가 1로 정규화되었을 때, 순시 신호 대 잡음비(signal-to-noise ratio, SNR)는  $\rho = |h|^2/\sigma^2$ 로 정의된다.

수신단에서는 등화를 통해  $\tilde{y}_t = x_t + h^* n_t/|h|^2$ 를 얻을 수 있다. 각 등화 심볼  $\tilde{y}_t$ 는 순차적으로 삼진 복조기를 거쳐 길이  $\log_2 M$ 의 삼진 부벡터  $\hat{\mathbf{b}}_t = [\hat{b}_{t,1}, \dots, \hat{b}_{t,\log_2 M}]$ 로 복원된다. QAM 신호 성좌의 인접한 두 점 간 최소 거리를  $d_{\min}$ 이라고 하고, 삼진 복조를 위한 소실 마진 비율을 0.5로 설정한다. 삼진 복조에서는 인접한 두 점의 중점으로부터 등화 심볼  $\tilde{y}_t$ 까지의 거리가  $d_{\min}/4$ 보다 크면 대응되는 비트 0 또는 1로 경판정하여 삼진 변수  $\hat{b}_{t,k}$ 를 결정하고, 작다면 불확실로 간주하여 소실을 의미하는 0.5의 값으로 삼진 변수  $\hat{b}_{t,k}$ 를 결정한다 [3]. 모든 심볼에 대해 복조된 부벡터들을 결합하여 얻은 삼진 벡터를  $\hat{\mathbf{b}}$ 라고 할 때, 역양자화기를 통해 복원된 잡재 변수들을 원소로 갖는 벡터  $\hat{\mathbf{q}}$ 를 얻은 뒤 [4], 신경망 가중치  $\phi$ 로 매개 변수화된 디코더  $f_\phi(\cdot)$ 를 통해 이미지  $\hat{\mathbf{u}} = f_\phi(\hat{\mathbf{q}})$ 를 복원한다.

디지털 통신 과정의 미분 불가능 연산이 포함되어 종단 간 송수신기 공동 학습이 어렵기 때문에, 변조-무선 채널 전송-등화-삼진 복조의 과정을 BSEC로 등가 모델링하여 인코더와 디코더 신경망을 단일 목적함수로 종단 간 학습한다 [3]. 입력 이진 잡재 변수  $b_{t,k}$ 에 대한 출력 삼진 변수  $\hat{b}_{t,k}$ 를 갖는 BSEC는 비트 반전 확률  $\alpha_{t,k}$ , 비트 정판정 확률  $\beta_{t,k}$ , 비트 소실 확률  $\gamma_{t,k}$ 로 정의되며,  $\alpha_{t,k} + \beta_{t,k} + \gamma_{t,k} = 1$ 을 만족한다. 종단 간 훈련의 순전과 수행 시에는 원본 이미지  $\mathbf{u}$ 를 인코딩 및 양자화하여 얻은 이진 잡재 벡터  $\mathbf{b}$ 를 병렬 BSEC 채널을 통해 전송함을 가정하여 확률적 샘플링으로 삼진 벡터  $\hat{\mathbf{b}}$ 를 생성한다. 이후  $\hat{\mathbf{b}}$ 를 역양자화 및 디코딩함으로써 이미지  $\hat{\mathbf{u}}$ 를 복원한다. 손실 함수는 원본 이미지  $\mathbf{u}$ 와 복원 이미지  $\hat{\mathbf{u}}$  간의 평균 제곱 오차로 두고 역전과를 수행한다.

### III. 제안 훈련 기법

다양한 변조 차수 및 채널 조건에 일반화된 단일 모델을 설계하기 위해서는, 시간 슬롯  $t \in \{1, \dots, T\}$ 와 비트 인덱스  $k \in \{1, \dots, \log_2 M\}$ 에 대해 BSEC 확률  $\{\alpha_{t,k}, \beta_{t,k}, \gamma_{t,k}\}$ 을 샘플링하여 훈련하는 파라미터 샘플링 전략을 고려한다 [3]. 모든  $t$ 와  $k$ 에 대해  $\alpha_{t,k}$ 를 샘플링한 뒤  $\beta_{t,k}$ 와  $\gamma_{t,k}$ 를 계산하고, 이렇게 얻은 확률값으로 총  $NB$  개의 병렬 BSEC를 구성하여 종단 간 훈련을 수행한다. 기존 방법 [3]에서는 4-QAM을 전제로

BSEC를 모델링하여 훈련을 진행하기 때문에, 고차 변조 적용 시 성능 열화가 발생될 수 있다. 본 연구에서는 주어진 변조 차수  $M$ 에 대해 고신뢰 디지털 시맨틱 통신 송수신기 설계를 위해, 다음의 두 가지 파라미터 샘플링 기반 훈련 기법을 제안한다.

1)  **$M$ -QAM 확률 근사 기반 훈련 기법:** 변조 차수  $M$ 에 대해, 높은 SNR 영역에서 확률의 근사식을 다음과 같이 분석할 수 있다 [3]:

$$\alpha_{t,k}^{(M)} \approx \frac{4}{\log_2 M} \left(1 - \frac{1}{\sqrt{M}}\right) \Phi \left( \frac{3}{2} \sqrt{\frac{3\rho_{t,k}^{(M)}}{M-1}} \right), \quad (1)$$

$$\beta_{t,k}^{(M)} \approx 1 - \frac{4}{\log_2 M} \left(1 - \frac{1}{\sqrt{M}}\right) \Phi \left( \frac{1}{2} \sqrt{\frac{3\rho_{t,k}^{(M)}}{M-1}} \right), \quad (2)$$

$$\gamma_{t,k}^{(M)} \approx 1 - \alpha_{t,k}^{(M)} - \beta_{t,k}^{(M)}. \quad (3)$$

여기서  $\Phi(\cdot)$ 는 가우시안 Q-함수를 나타낸다. 제안 기법에서 비트 반전 확률의 상한은 동작 SNR의 최소값  $\rho$ 을 식 (1)에 대입하여  $\bar{\alpha}_{t,k}^{(M)}(\rho)$ 로 설정한다. 이어, 샘플링된  $\alpha_{t,k}^{(M)} \sim \text{Unif}(0, \bar{\alpha}_{t,k}^{(M)})$  이용하여 순시 SNR  $\rho_{t,k}^{(M)}$ 을 다음과 같이 계산한다:

$$\rho_{t,k}^{(M)} \approx \frac{M-1}{3} \left( \frac{2}{3} \Phi^{-1} \left( \frac{\alpha_{t,k}^{(M)} \sqrt{M \log_2 M}}{4(\sqrt{M}-1)} \right) \right)^2. \quad (4)$$

얻어진  $\alpha_{t,k}^{(M)}$ 와  $\rho_{t,k}^{(M)}$ 을 식 (2)와 (3)에 대입하여 각각  $\beta_{t,k}^{(M)}$ 와  $\gamma_{t,k}^{(M)}$ 를 구하고, 모든  $t$ 와  $k$ 에 대해  $\{\alpha_{t,k}^{(M)}, \beta_{t,k}^{(M)}, \gamma_{t,k}^{(M)}\}$ 을 바탕으로 병렬 BSEC를 모델링한 뒤 종단 간 훈련을 수행한다.

2) **MSB 지향 확률 모델링 기반 훈련 기법:** 송신 변조 심볼을 구성하는  $\log_2 M$  개의 비트 중에서 MSB에 발생하는 오류는 복원 오차에 큰 영향을 미칠 수 있기 때문에 신중한 보호가 필요하다. 변조 차수  $M$ 에 대해 MSB가 전송되는 등가 BSEC 확률의 근사식을 다음과 같이 분석한다:

$$\alpha_{t,k}^{\text{MSB}} \approx \frac{2}{\sqrt{M}} \Phi \left( \frac{3}{2} \sqrt{\frac{3\rho_{t,k}^{\text{MSB}}}{M-1}} \right), \quad (5)$$

$$\beta_{t,k}^{\text{MSB}} \approx 1 - \frac{2}{\sqrt{M}} \Phi \left( \frac{1}{2} \sqrt{\frac{3\rho_{t,k}^{\text{MSB}}}{M-1}} \right), \quad (6)$$

$$\gamma_{t,k}^{\text{MSB}} \approx 1 - \alpha_{t,k}^{\text{MSB}} - \beta_{t,k}^{\text{MSB}}. \quad (7)$$

본 제안 기법에서는  $\alpha_{t,k}^{\text{MSB}}$ 의 상한을 설정하기 위해 동작 SNR의 최소값  $\rho$ 을 식 (5)에 대입하여  $\bar{\alpha}_{t,k}^{\text{MSB}}(\rho)$ 를 구한다. 이를 이용하여  $\alpha_{t,k}^{\text{MSB}} \sim \text{Unif}(0, \bar{\alpha}_{t,k}^{\text{MSB}})$ 의 샘플링을 통해  $\alpha_{t,k}^{\text{MSB}}$ 를 얻으면, 역함수의 관계를 통해 순시 SNR  $\rho_{t,k}^{\text{MSB}}$ 을 다음과 같이 계산할 수 있다:

$$\rho_{t,k}^{\text{MSB}} \approx \frac{M-1}{3} \left( \frac{2}{3} \Phi^{-1} \left( \frac{\alpha_{t,k}^{\text{MSB}} \sqrt{M}}{2} \right) \right)^2. \quad (8)$$

최종적으로,  $\alpha_{t,k}^{\text{MSB}}$ 와  $\rho_{t,k}^{\text{MSB}}$ 을 식 (6)과 (7)에 대입하여 각각  $\beta_{t,k}^{\text{MSB}}$ 와  $\gamma_{t,k}^{\text{MSB}}$ 를 구하면, 모든  $t$ 와  $k$ 에 대해 도출된  $\{\alpha_{t,k}^{\text{MSB}}, \beta_{t,k}^{\text{MSB}}, \gamma_{t,k}^{\text{MSB}}\}$ 을 바탕으로 병렬 BSEC를 모델링한 뒤 종단 간 훈련을 수행한다.

#### IV. 시뮬레이션 결과

제안 기법의 성능 검증을 위해 8-비트 양자화 및 256-QAM이 적용된 디지털 시맨틱 통신 시스템을 고려하였다. 실험 데이터는 CIFAR-10를 사용하였으며, 모든 이미지를  $[0, 1]$  범위로 정규화하였다. 이미지 복원 성능

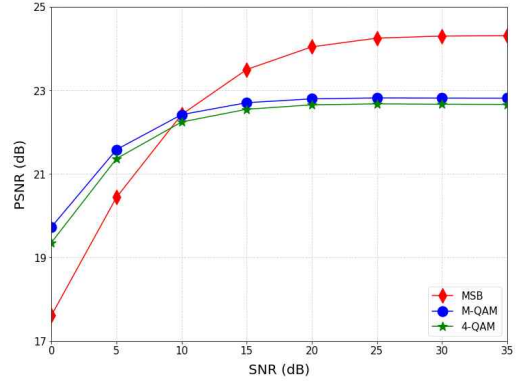


그림 1. 제안한 훈련 기법의 이미지 복원 성능.

평가는 최대 신호 대 잡음비(peak signal-to-noise ratio, PSNR)로 수행하였다. 신경망 인코더와 디코더는 [3]의 구조를 기반으로 설계하였고, 잠재 공간 차원은 528로 설정하였다. 종단 간 학습은 Adam 옵티마이저로 학습률 0.001, 배치 크기 256, 에포크 수 20으로 진행하였다.

그림 1은 제안한 훈련 기법과 기존 4-QAM 기반 훈련 기법으로 설계한 디지털 시맨틱 통신 시스템의 이미지 복원 성능을 비교한다. 제안한  $M$ -QAM 확률 근사 기반 훈련은 기존 대비 전 SNR 구간에서 일관된 PSNR 향상을 보였다. MSB 지향 확률 모델링 기반 훈련은 0-10 dB 구간에서는 열화되지만, 10-35 dB 환경에서는 1dB 이상의 PSNR 성능 이득을 달성함을 확인하였다. 이는 비트 반전 확률의 샘플링 상한값 설정에 따라 학습에서 직면하는 무선 통신 환경(유효 SNR 분포)가 달라질 뿐만 아니라, 병렬 BSEC 등가 모델을 정의하는 세 가지 확률의 설정에 따라 훈련되는 시맨틱 통신 송수신기의 이미지 복원 성능이 크게 달라짐을 시사한다.

#### V. 결론

본 논문에서는 고차 변조 환경에서 디지털 시맨틱 통신 성능 향상을 위해 파라미터 샘플링 기반 훈련 기법을 제안하였다. 제안 기법은 주어진 변조 차수에서 실제 무선 전송 특성을 반영한 확률 분석에 기반해 학습을 수행함으로써, 강력한 이미지 복원 성능을 달성함을 검증하였다.

#### ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-대학ICT연구센터(ITRC)의 지원(IITP-2025-RS-2023-00259991, 50%)과 2025년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초사업임(RS-2025-25432451, 50%).

#### 참고 문헌

- [1] D. Gunduz et al., "Beyond transmitting bits: Context, semantics, and task-oriented communications," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 5-41, Jan. 2023.
- [2] 김찬형, 박성찬, 임승찬, "시맨틱 통신 기술 개요 및 연구 동향," 2023년도 한국통신학회 추계종합학술발표회, pp. 332 - 333, 2023. 11.
- [3] J. Park, Y. Oh, S. Kim, and Y.-S. Jeon, "Joint source-channel coding for channel-adaptive digital semantic communications," *IEEE Trans. Cogn. Commun. Netw.*, vol. 11, no. 1, pp. 75 - 89, Feb. 2025.
- [4] Y. Oh, J. Park, J. Choi, J. Park and Y.-S. Jeon, "Blind training for channel-adaptive digital semantic communications," to appear in *IEEE Trans. Commun.*, 2025.