

학습 가능한 잔여 벡터 양자화와 비균등 변조 인코딩을 이용한 디지털 시맨틱 통신

김원중, 이재인, 신원재*, 이정우
서울대학교, *고려대학교

dnjswnd116@snu.ac.kr, jilee@cml.snu.ac.kr, *wjshin@korea.ac.kr, junglee@snu.ac.kr

Learnable Residual Vector Quantization with Unequal Modulation Encoding for Digital Semantic Communications

Wonjung Kim, Jaein Lee, Wonjae Shin*, Jungwoo Lee
Seoul National Univ., *Korea Univ.

요약

본 논문에서는 정보의 중요도를 반영하여 비트율을 효과적으로 제어할 수 있는 디지털 시맨틱 통신 프레임워크인 ResUME 을 제안한다. 이 프레임워크는 잔여 벡터 양자화(RVQ)를 기반으로 계층적 구조를 형성하고, 각 단계의 정보 중요도에 따라 비균등 변조 방식을 적용함으로써 통신 자원의 효율적 할당을 가능하게 한다. 아울러, 양자화 오차와 채널 잡음을 동시에 고려하는 학습 기법인 CA-NSVQ 를 도입하여 다양한 채널 환경에서도 우수한 일반화 성능을 확보하였다. 제안된 프레임워크는 AWGN 채널 환경에서의 실험을 통해, 전 구간 SNR 조건에서 기존 방식보다 뛰어난 이미지 복원 성능을 나타냄으로써 디지털 통신 시스템에의 실용 가능성을 입증하였다.

I. 서론

최근 5G/6G 통신과 IoT 환경의 확산에 따라, 통신 시스템은 빠른 전송 속도와 높은 효율성을 동시에 요구받고 있다. 기존의 디지털 통신 방식은 Shannon 이론에 기반한 소스 및 채널 코딩을 통해 데이터의 정확한 전달을 추구하지만, 데이터의 의미를 고려하지 않기 때문에 비효율적인 경우가 많다. 이를 해결하기 위한 대안으로 시맨틱 통신이 주목받고 있으며, 특히 딥러닝 기반 Joint Source and Channel Coding(JSCC) 방식은 소스-채널 통합 최적화를 가능하게 한다 [1].

그러나 기존의 많은 연구들은 실제 디지털 통신 시스템과의 호환성이 부족한 구조를 제안해왔으며, 채널 잡음에 강건한 학습 전략이 미흡했다. 본 논문에서는 RVQ 기반의 계층적 양자화 구조와 단계별 비대칭 변조 방식을 결합한 ResUME 프레임워크를 제안하고, 추가적으로 양자화 및 채널 잡음을 통합적으로 고려한 학습 기법인 CA-NSVQ 를 통해 일반화 성능을 개선하였다.

II. 본론

ResUME 은 JSCC 구조와 채널 사이에 디지털화하는 과정을 RVQ 기반의 양자화 및 비대칭 변조 전략으로 통합한 프레임워크이다. 이는 그림 1 에 도시된 바와 같이, 계층적으로 중요한 정보부터 잔여 정보를 분리하여 양자화하며, 각 단계별 정보 중요도에 따라 적절한 변조

차수를 적용함으로써 전송 효율성과 복원 성능을 동시에 확보한다. 시스템은 SISO 기반 AWGN 채널 모델을 가정하고 수신단에서 완전한 CSI 정보를 활용한다. 그리고 JSCC 내 인코더는 단단계 컨볼루션 및 잔여 블록으로 구성되며 디코더는 인코더와 동일한 구조의 역과정을 포함하고, 잡음 혼합 인코더 및 디코더 구조를 통해 노이즈에 강인한 성능을 확보한다.

RVQ 단계별 정보 중요도는 데이터 처리 부등식을 기반으로 상호정보량의 점진적 감소로 수식화된다.

$$I(X; Q_1) \geq I(X; Q_2) \geq \dots \geq I(X; Q_L). \quad (1)$$

이 때, X 는 원본 데이터를, ℓ 은 단계를 나타내고 Q_ℓ 은 ℓ 번째 단계에 적용되는 양자화된 벡터를 의미한다. 이를 통해 제한된 총 전송 자원 내에서 단계별 최적 변조 차수를 선택하는 문제를 라그랑지안 최적화를 활용한다.

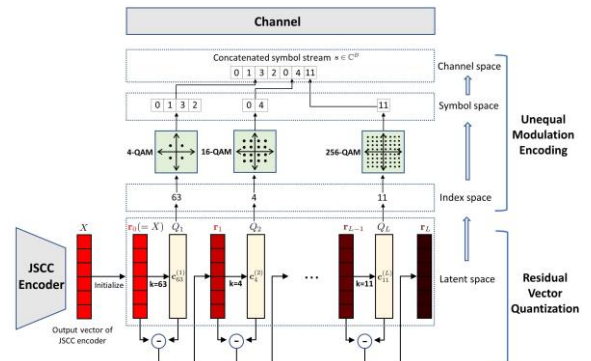


그림 1 ResUME의 동작 과정

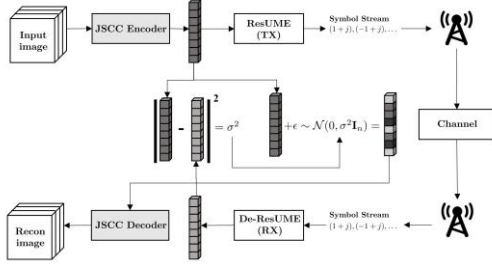


그림 2 CA-NSVQ의 적용 과정

$$\max_{M_1, \dots, M_L} \sum_{\ell=1}^L I(X; Q_\ell) (1 - p_e(M_\ell))^{\frac{b}{\log M_\ell}}. \quad (2a)$$

$$\text{subject to } \sum_{\ell=1}^L \frac{b}{\log M_\ell} \leq S, \quad M_\ell \in \mathbb{N}_{\geq 2}. \quad (2b)$$

이 때, M_ℓ 은 ℓ 단계에 적용되는 심볼 변조를 나타내고, $p_e(M_\ell)$ 는 변조 별 심볼 오류율을, b 는 보내는 비트 수를 나타내며 S 는 최대 허용 심볼 수를 나타낸다. 이 때 최적화를 위한 라그랑지안 승수 기법 및 KKT 조건을 통해 다음 결론을 이끌 수 있다.

$$M_1 \leq M_2 \leq \dots \leq M_L. \quad (3)$$

이를 통해 오름차순의 심볼 변조가 최적을 달성할 수 있음을 확인할 수 있다. 또한, 그림 2와 같이 제안하는 CA-NSVQ는 벡터 양자화 기반 시멘틱 통신을 위한 학습 기술로, 기존 NSVQ[2] 방식에 더해 양자화 잡음과 채널 잡음을 가우시안 잡음으로 근사하여 학습 시 AWGN 환경을 모사하고 일반화 성능을 향상시킨다.

손실 함수는 정보 병목 이론 및 변분 추론 기반으로 도출되며, VQ-VAE[3] 기반의 재구성 손실 \mathcal{L}_{fid} , 분산 손실 \mathcal{L}_{var} , Kullback-Leibler 발산 손실 \mathcal{L}_{KL} 의 세 가지 항을 포함한다. 각 손실 항은 정보 보존, 코드북 압축, 잠재 공간 정규화의 목적을 수행하며, 중단 간 학습 구조에서 복원 성능과 다양성의 균형을 이룬다.

$$\text{Loss} = \mathbb{E}_{\theta, \phi} \left[\underbrace{\frac{\mathcal{L}_{\text{rec}}}{\|\hat{\mathbf{x}} - \mathbf{x}\|_2^2} + \frac{\mathcal{L}_{\text{emb}}}{\lambda_1 \|\text{sg}(\mathbf{z}) - \hat{\mathbf{z}}\|_2^2} + \frac{\mathcal{L}_{\text{cmt}}}{\lambda_2 \|\mathbf{z} - \text{sg}(\hat{\mathbf{z}})\|_2^2}}_{\mathcal{L}_{\text{fid}}} + \underbrace{\lambda_3 \mathbb{E}_{\theta} [\|\mathbf{z}\|_2^2]}_{\mathcal{L}_{\text{var}}} + \underbrace{\lambda_4 D_{\text{KL}}(q_{\theta}(\hat{\mathbf{z}}) \|\mathcal{N}(0, \mathbf{I}_n))}_{\mathcal{L}_{\text{KL}}} \right] \quad (4)$$

앞서 제안된 프레임워크 및 학습 방식을 적용하여 RVQ 기반의 균등 변조와의 비교 실험을 진행하였다. 실험을 위해 학습용 데이터로 ImageNet 데이터셋을 128×128 크기로 크롭하여 활용하였고, 수치적 성능 평가는 PSNR[4] 및 MS-SSIM[5]으로 진행하였다. 모델의 압축률의 경우 채널-대역폭 비율 (CBR)[6]을 활용하여 CBR=0.0156의 값을 기준으로 성능비교를 진행하였고 그 결과는 그림 3과 같다.

그림 3과 같이, ResUME 모델이 균등 변조를 활용하는 RVQ+16QAM 모델보다 모든 SNR에 대해서 PSNR 및 MS-SSIM 성능 모두 앞섬을 확인할 수 있다. 그리고 동일 ResUME 모델 내에서 동일 CBR을 만족하는 다른 변조 ($M=[1,2,4,4]$ 및 $M=[0]$)에 대해서도 비교를 진행하였다. 이 때 M 의 각 인덱스는 변조 차수를 의미하며, 0부터 4는 각각 BPSK, QPSK, 16-QAM, 64-QAM, 256-QAM을 나타낸다. 이 경우, SNR이 낮은 경우에는 $M=[0]$ 가 높은 성능을 나타내지만 SNR이 높은 경우에는 $M=[1,2,4,4]$ 가 더 좋은 성능을 보임을 확인할 수 있다. 이를 통해 SNR 변화에 따라 비균등 변조 방식 수정을 통해 유동적으로 최적의 성능을 달성할 수 있음을 시사한다.

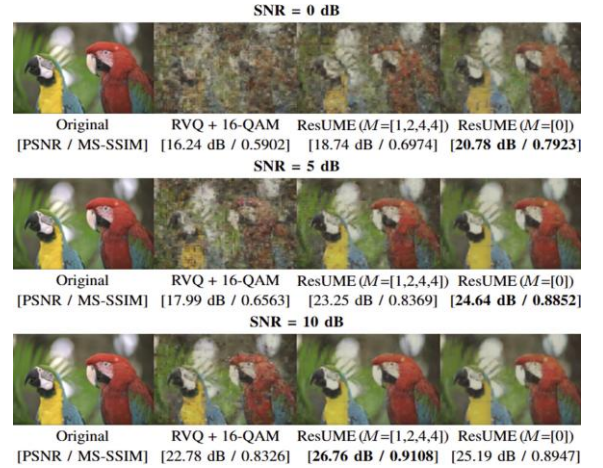


그림 3 다양한 SNR 조건에서 동일 압축률로 전송된 두 종류의 ResUME과 RVQ+16QAM 간의 복원 성능 비교.

III. 결론

본 연구는 비트율 제어 가능한 디지털화 프레임워크인 ResUME을 제안하고 일반화 성능 향상을 위한 CA-NSVQ 학습 기술을 제안함으로써, 정보 중요도 기반의 전송 전략과 잡음 강건한 학습 구조를 실현하였다. 또한 정보 병목 이론 및 변분 추론 기반의 손실 함수를 적용해 다양한 SNR에서 높은 복원 성능을 달성하였다. 향후에는 실제 무선 채널과의 적응성을 고려한 확장 연구가 필요할 것이다.

ACKNOWLEDGEMENT

본 연구는 정보통신기획평가원(RS-2024-00436887, RS-2021-0-00467)의 지원을 받아 수행된 연구임.

참고 문헌

- [1] E. Bourtsoulatz, D. B. 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, 2019.
- [2] M. H. Vali and T. Backstrom, "Nsvq: Noise substitution in vector quantization for machine learning," IEEE Access, vol. 10, pp. 13598-13610, 2022.
- [3] A. Van Den Oord, O. Vinyals et al., "Neural discrete representation learning," Advances in neural information processing systems, vol. 30, 2017.
- [4] A. Hore and D. Ziou, "Image quality metrics: Psnr vs. ssim," in 2010 20th international conference on pattern recognition. IEEE, 2010, pp. 2366-2369.
- [5] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003, vol. 2. Ieee, 2003, pp. 1398-1402.
- [6] J. Dai, S. Wang, K. Yang, K. Tan, X. Qin, Z. Si, K. Niu, and P. Zhang, "Toward adaptive semantic communications: Efficient data transmission via online learned nonlinear transform source-channel coding," IEEE Journal on Selected Areas in Communications, vol. 41, no. 8, pp. 2609-2627, 2023.