

시맨틱 이미지 전송에서 의미 보존과 전송 최적화를 위한 최신 접근법: 세분화와 디퓨전 기법을 중심으로

홍성훈, 원동욱, 이동현, 최성진, 이승찬, 조성래
중앙대학교 컴퓨터공학과

{shhong, dwwon, dhlee, sjchoi, sclee}@uclab.re.kr, srcho@cau.ac.kr

A Brief Survey on Semantic-Preserving Optimization in Semantic Image Transmission: From Segmentation-Based Allocation to Diffusion-Guided Reconstruction

Seonghun Hong, Dongwook Won, Donghyun Lee, Seongjin Choi, Seungchan Lee, and
Sungrae Cho

Department of Computer Science and Engineering, Chung-Ang University

요 약

시맨틱 이미지 전송은 데이터 압축 및 복원 기술을 넘어, 전송되는 정보의 ‘의미’를 중심으로 통신 효율을 극대화하고자 하는 차세대 통신 패러다임이다. 최근에는 전송 과정에서 의미적으로 중요한 정보를 선택적으로 전달하거나, 수신단에서 의미 구조에 기반한 고품질 복원을 수행하는 다양한 접근이 제안되고 있다. 본 논문에서는 이러한 시맨틱 이미지 전송 기술의 최신 연구에서 semantic segmentation 을 이용해 Region of Interest (ROI)을 사전에 추출하고 해당 영역에 전송 자원을 집중하는 방식과, DeepJSCC 의 복원 결과에 diffusion 모델을 후처리로 적용하여 시각적 품질을 보완하는 방식을 중심으로 조사한다. 이러한 접근들은 서로 다른 위치에서 의미 보존을 시도하며, 시맨틱 통신의 전송 효율성과 복원 신뢰성을 동시에 향상시키는 데 중요한 시사점을 제공한다.

I. 서 론

기존의 이미지 전송 시스템은 주로 원본 이미지의 화질을 가능한 손상 없이 전달하는 데 초점을 맞추어 왔다. 그러나 이러한 비트 기반 전송 모델은 채널 환경의 변화나 전송 자원 제약이 심한 상황에서 의미 있는 전송 품질을 보장하지 못한다 [1]. 반면, 시맨틱 통신은 이러한 한계를 극복하고자, 정보의 “의미”를 중심으로 하는 새로운 전송 방식으로 주목받고 있다. 특히 이미지 전송의 경우, 모든 픽셀 정보가 동일한 중요도를 가지지 않으며, 사용자가 궁극적으로 이해해야 할 정보는 일부 시맨틱 구조에 집중되어 있다는 점에서 의미 기반 전송의 필요성이 크다.

최근 연구들은 전송 이전에 의미 정보를 미리 추출하거나, 전송 이후 의미적 복원 능력을 강화하는 전략을 통해 시맨틱 이미지 전송의 품질을 향상시키고 있다 [2]. 이 가운데 본 논문에서는 세 가지 접근을 중심으로 시맨틱 전송 기술의 현재를 조사한다. 첫 번째는 semantic segmentation 을 통해 Region of Interest(ROI)와 Region of Non-Interest(RONI)를 구분하고, 의미적으로 중요한 영역에만 고대역폭 전송 자원을 집중시키는 구조이다. 두 번째는 전송된 이미지의 복원 결과를 diffusion 모델로 후처리하여 자연스럽게 고품질의 이미지를 생성하는 방식이며, 마지막은 diffusion 복원 과정 자체에 의미적 가이드

정보를 결합함으로써 의미 일치성이 높은 복원을 유도하는 접근이다. 이 세 가지 방법은 각각 전송의 앞단과 뒷단에서 의미 보존을 수행하며, 시맨틱 통신의 전체 파이프라인에 걸쳐 상호 보완적인 관점을 제공한다.

II. 본론

시맨틱 이미지 전송에서 가장 중요한 과제는 전송 자원 제약이나 채널 환경의 열악함 속에서도 의미의 본질을 어떻게 보존할 것인가이다. 본 논문에서는 공통적으로 다음의 세 가지 전략을 공유한다. 첫째, 전송 자원을 효율적으로 활용하기 위해 의미 중심의 정보 선별 또는 분리 처리 구조를 채택한다. 둘째, 전송 이후의 복원 단계에서 단순한 화질 회복을 넘어 사람이 이해할 수 있는 의미 구조를 복원하는 방향으로 발전한다. 셋째, 의미 정보를 복원하거나 생성하는 과정에 기계 학습 기반의 조건 또는 prior 를 적극적으로 활용함으로써, 불완전한 전송 결과로부터 의미를 재구성하려는 시도를 보여준다.

첫 번째 연구는 segmentation 기반 시맨틱 통신 모델이다 [3]. 이 연구는 전송 전에 이미지에서 의미적으로 중요한 영역인 ROI 를 semantic segmentation(U-Net 기반)을 통해 추출하고, 이를

고해상도 semantic encoder(SC2)를 통해 전송하며, RONI 는 저해상도 encoder(SC1)를 통해 처리함으로써 자원 효율성과 의미 보존의 균형을 달성하고자 하였다. 특히 θ -PSNR 이라는 새로운 평가 지표를 통해 ROI 의 재구성 정확도를 중점적으로 반영하였으며, 전송 자원을 ROI 에 집중 배분함으로써 압축률을 향상시키면서도 의미적 손실을 최소화하였다. 또한 Sparse Matrix 기반 압축 방식(COO)을 도입하여 네트워크 입력 시 발생하는 padding redundancy 를 줄였다는 점도 실용적 기여로 평가된다. 이 구조는 시맨틱 통신의 전처리 단계에서 의미 기반 전송 자원 배분의 필요성을 구체적으로 구현한 대표 사례로 볼 수 있다.

두 번째로 다룰 논문 diffusion 기반 시맨틱 전송 보완 기법이다[4]. 이 연구는 DeepJSCC 를 통해 전송된 저품질 이미지 복원 결과에 diffusion model 을 후처리로 적용함으로써, perceptual quality 를 비약적으로 향상시키는 방법을 제안한다. 기존 DeepJSCC 는 복원 이미지의 PSNR 이나 구조적 품질을 보장할 수 있으나, 사람 눈에 보기에는 어색한 결과가 나타날 수 있었다. diffusion 은 생성 기반의 모델로, 이미지의 전반적인 질감, 구조, 색감 등을 자연스럽게 복원할 수 있다는 장점이 있으며, 이러한 후처리 기술을 시맨틱 통신 시스템과 접목하여 수신 단 복원 품질을 획기적으로 끌어올린다. 이 방법은 별도의 의미 가이드 없이도 전송 손실이 있는 이미지를 자연스럽게 보완할 수 있다는 점에서 실용성과 구현 용이성이 높은 방식이다.

마지막 연구는 가장 진보된 의미 복원 전략을 포함하고 있는 SGD-JSCC(Semantics-Guided Diffusion for Deep Joint Source-Channel Coding)이다 [5]. 이 연구는 diffusion 기반 복원 모델에 텍스트 설명이나 엷지 맵과 같은 의미적 조건(guidance)을 주입하여, 복원된 이미지가 원본 의미와 구조를 더욱 정확히 반영하도록 설계하였다. 예를 들어 “a bus on the road”라는 텍스트 설명이 입력되면, diffusion 과정에서 해당 의미에 부합하는 이미지가 생성되며, 이는 단순한 픽셀 기반 복원을 넘어 의미 기반 생성이라는 차원을 제공한다. 실험에서는 guidance 없이 단독 diffusion 을 사용한 경우보다 PSNR, LPIPS 등 모든 지표에서 우수한 성능을 보였으며, 특히 SNR 이 낮아 전송 품질이 떨어지는 상황에서도 의미 정합성을 유지하는 강인한 복원 성능을 입증하였다. 이 방식은 시맨틱 통신의 수신단에서 의미 추론 능력을 적극적으로 활용한 가장 진보된 복원 구조로 평가된다.

III. 결론

본 논문에서는 시맨틱 이미지 전송에서 의미 보존과 전송 효율을 동시에 향상시키기 위해 최근 제안된 세 가지 대표적인 접근을 비교 분석하였다. segmentation 기반 접근은 전송 전 의미 정보를 선별하여 자원을 효율적으로 사용하는 구조이며, diffusion 기반 복원은 수신 측에서 이미지 품질을 시각적으로 개선하거나, 의미적 조건을 통해 복원 정확도를 향상시키는 기술이다. 세 논문은 각각 시맨틱 통신의 전송 단계, 복원 단계, 의미 생성 단계에 해당하는 기술이며, 상호 보완적으로 통합될 경우 강력한 의미 기반 이미지 전송 시스템을 구성할 수 있다. 향후 연구에서는 이러한 기술들의 결합뿐만 아니라, 의미 정보의 정량적 정의, RoI 선택 기준의 학습 기반 확장, 실시간 전송 환경 [6][7]에의 적용 가능성 등이 주요 과제로 남아 있을 것이다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-대학 ICT 연구센터(ITRC)의 지원(IITP-2025-RS-2022-00156353, 50%)과 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(No. RS-2024-00453301)을 받아 수행된 연구임

참 고 문 헌

- [1] G. Shi, Y. Xiao, Y. Li and X. Xie, "From Semantic Communication to Semantic-Aware Networking: Model, Architecture, and Open Problems," in *IEEE Communications Magazine*, vol. 59, no. 8, pp. 44-50, August 2021, doi: 10.1109/MCOM.001.2001239.
- [2] H. Xie, Z. Qin, G. Y. Li and B. -H. Juang, "Deep Learning Enabled Semantic Communication Systems," in *IEEE Transactions on Signal Processing*, vol. 69, pp. 2663-2675, 2021, doi: 10.1109/TSP.2021.3071210.
- [3] Jiale Wu, Celimuge Wu, Yangfei Lin, Tsutomu Yoshinaga, Lei Zhong, Xianfu Chen, Yusheng Ji, Semantic segmentation-based semantic communication system for image transmission, *Digital Communications and Networks*, Volume 10, Issue 3, 2024, Pages 519-527, ISSN 2352-8648, <https://doi.org/10.1016/j.dcan.2023.02.006>.
- [4] S. F. Yilmaz, X. Niu, B. Bai, W. Han, L. Deng and D. Gündüz, "High Perceptual Quality Wireless Image Delivery with Denoising Diffusion Models," *IEEE INFOCOM 2024 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Vancouver, BC, Canada, 2024, pp. 1-5, doi: 10.1109/INFOCOMWKSHPS61880.2024.10620904.
- [5] Maojun Zhang and Haotian Wu and Guangxu Zhu and Richeng Jin and Xiaoming Chen and Deniz Gündüz, "Semantics-guided diffusion for deep joint source-channel coding in wireless image transmission," 2025, *arXiv:2501.01138*.
- [6] M. Choi, J. Kim and J. Moon, "Wireless Video Caching and Dynamic Streaming Under Differentiated Quality Requirements," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 6, pp. 1245-1257, June 2018.
- [7] J. Kim, G. Caire and A. F. Molisch, "Quality-Aware Streaming and Scheduling for Device-to-Device Video Delivery," *IEEE/ACM Transactions on Networking*, vol. 24, no. 4, pp. 2319-2331, Aug. 2016.