

UAV 통신을 위한 멀티모달 센싱 기반 빔 예측

여예린, 김정현

세종대학교

yealin0817@gmail.com, j.kim@sejong.ac.kr

Multi-Modal Sensing-assisted Beam Prediction for UAV Communications

Yerin Yeo, Junghyun Kim

Sejong Univ.

요약

본 논문에서는 동적인 변화가 많은 무선 통신 시스템에서 효율적인 빔포밍이 가능하도록 카메라 이미지 데이터와 GPS 데이터를 모두 사용해 최적의 빔을 예측하는 모델을 제안한다. 기존 연구에서는 카메라 이미지 데이터와 GPS 데이터 각각을 사용하는 싱글모달 모델을 설계하여 성능의 한계를 보였다. 본 논문에서는 이러한 한계를 극복하기 위해 두 가지 데이터를 융합해 데이터 간의 정보를 고려할 수 있는 새로운 멀티모달 빔 예측 모델을 제안한다.

I. 서론

밀리미터파(mmWave) 통신 시스템은 대형 안테나 어레이와 함께 빔포밍 기술을 활용하여 탁월한 데이터 전송 속도를 달성한다. 빔포밍에 좁은 빔을 사용하면 노드 간 간섭이 최소화되고 대상 수신 전력이 최대화된다. 그러나 좁은 빔을 최적으로 관리하려면 특히 고속 이동 환경에서 상당한 오버헤드가 필요하므로 빔 관리에 상당한 어려움이 따른다 [1]. 이 문제를 극복하기 위한 효과적인 접근 방식 중 하나는 대역 내 무선 신호 대신 멀티모달 센싱 데이터를 사용하여 최적의 빔을 예측하는 것이다.

본 논문에서는 동적인 변화가 많은 무선 통신 시스템에서 효율적인 빔포밍이 가능하도록 이미지 데이터와 GPS 데이터를 이용해 최적의 빔을 예측하는 모델을 제안한다. 기존 논문 [2]에서는 각각의 데이터만 사용하여 성능을 비교했다. 하지만 GPS 데이터만 사용한다면 정확한 위치가 아닐 때 발생하는 오류에 치명적일 수 있고, 이미지 데이터만 사용한다면 조명과 날씨 등 여러 조건에 민감할 수 있다. 이를 극복하기 위해 [3]에서는 4가지 센서에서 수집한 데이터를 트랜스포머 기반의 블록 [4]으로 융합해 빔을 예측하는 모델을 제안하였다. 본 논문에서는 각 데이터의 특징뿐만 아니라 [3]에서 제안된 모델을 활용해 2가지 센서 데이터 간의 정보를 고려하는 새로운 빔 예측 모델을 제안한다.

II. 본론

본 논문에서 사용한 데이터 셋은 DeepSense 6G [5]에서 제공하는 시나리오 23으로, 이 데이터 셋은 실제 무선 환경에서 드론과 고정 기지국을 이용해 측정하였다. 기지국에서 측정된 전력 값 중 가장 큰 값을 갖는 빔의 인덱스 값과 카메라 이미지 데이터, 드론의 실시간 위치를 측정한 GPS 데이터 등으로 구성된다. ResNet50를 적용하여 이미지 데이터의 특징을 추출하였고, GPS 데이터는 min-max 정규화를 적용한 뒤, 트랜스포머 블록의 임베딩 크기와 맞추기 위해 Fully Connected (FC) Layer를 적용하였다.

이미지 데이터와 GPS 데이터는 각각 ResNet 블록과 FC Layer를 통과해 Embedding Layer를 거쳐 크기가 변환된 후 합쳐진다. 이후, 시퀀스 정보를 고려하기 위한 Positional Embedding 벡터와 더해져 8개의 Layer로 구성된 트랜스포머 블록으로 전달되고, 이를 통해 데이터 간의 관계를 학습하게 된다. 합쳐진 데이터를 다시 나누고, 크기를 변환하여 다음 ResNet 블록 및 FC Layer로 전달하는데, 이때 각 데이터의 출력 값과 이전 입력 값을 더해져 전달한다. 이 과정은 트랜스포머 블록의 임베딩 크기를 각각 64, 128, 256, 512로, Interp. Layer에서 사용하는 Scale Factor를 각각 8, 4, 2, 1로 변경하여 진행한다. 마지막으로, 이미지 데이터는

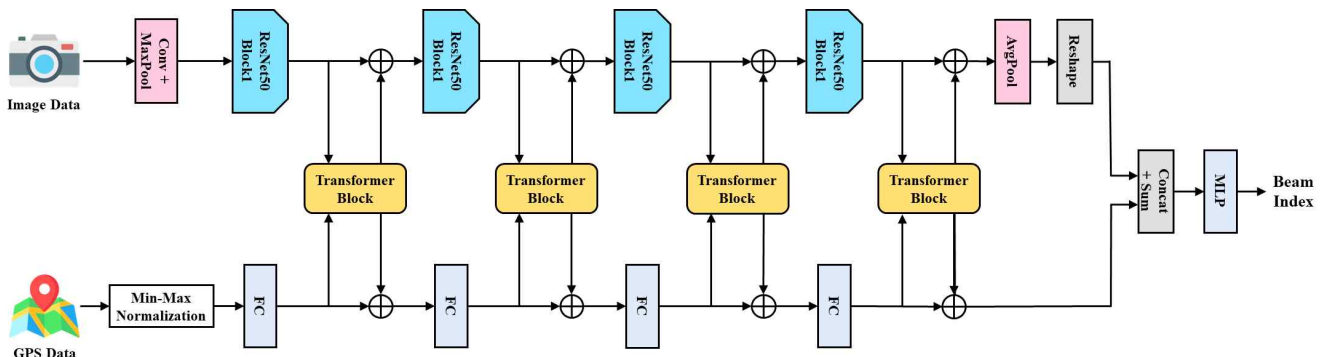


그림 1. 트랜스포머 기반의 멀티모달 빔 예측 모델.

Pooling Layer와 Reshape Layer를 거친 후, GPS 데이터와 합쳐지고 3개의 FC Layer로 구성된 Multi-layer Perceptron (MLP)을 통해 최적의 빔 인덱스를 예측하게 된다. 본 논문에서 제안하는 빔 예측 모델과 트랜스포머 블록은 각각 그림 1과 그림 2에 나타나 있다.

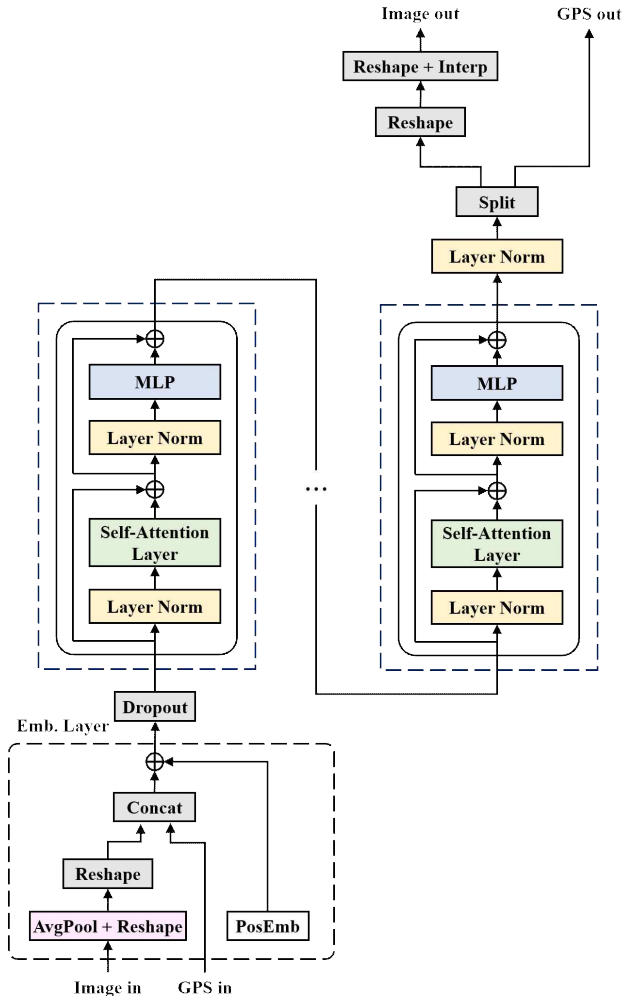


그림 2. 트랜스포머 블록 구조.

본 논문에서는 제안하는 모델의 성능을 확인하기 위해 각각 32개, 64개 빔을 예측하는 2가지 실험을 진행하였고, 기존 연구 [2]에서 더 나은 성능을 갖는 이미지 데이터 기반 빔 예측 모델과 성능을 비교하였다. 이때, epoch는 40, batch size는 32, learning rate는 0.0001로 시작해 epoch 10, 20, 30에서 0.1씩 곱해지는 스케줄링을 적용하여 실험을 진행하였다. 또한, 성능 지표로는 Top-k Accuracy를 사용하였는데, 이는 예측된 상위 k개 빔 내에 있는 샘플 중, 실제 정답 값을 나타내는 빔에 해당하는 샘플의 비율을 의미한다. 즉, 본 논문에서 평가하는 Top-3 정확도는 상위 3개 예측 빔에 해당하는 샘플 중 정답 빔에 해당하는 샘플의 비율을 나타낸다. 먼저 32개의 빔으로 다운 샘플링 하여 최적의 빔을 예측한 실험 결과는 표 1에 나타나 있다. 실험 결과, 트랜스포머 블록을 적용한 제안 모델이 기존 모델 대비 Top-1, 2, 3 정확도 모두에서 향상된 성능을 보였다.

표 1. 32개 빔에 대한 예측 성능 비교.

	Top-1	Top-2	Top-3
기존 모델 [2]	0.8653	0.9710	0.9903
제안 모델	0.8841	0.9894	0.9964

다음으로, 64개 빔에 대해서 최적의 빔을 예측한 결과, 32개 빔과 마찬가지로 Top-1, 2, 3 정확도 모두에서 기존 모델 대비 제안 모델의 성능이 높은 것을 확인할 수 있었다. 또한, 이 경우는 32개 빔을 사용하는 경우보다 레이블의 수가 많아져 예측하기 까다로워지므로 모델 성능이 전체적으로 낮아진 것을 확인할 수 있다. 그럼에도 불구하고 제안 모델의 Top-3 성능은 약 99%의 높은 성능을 보였다. 64개 빔에 대해 예측한 실험 결과는 표 2에 나타나 있다.

표 2. 64개 빔에 대한 예측 성능 비교.

	Top-1	Top-2	Top-3
기존 모델 [2]	0.7647	0.9376	0.9850
제안 모델	0.7750	0.9438	0.9903

III. 결론

본 논문에서는 한 가지 센서 데이터만을 사용하는 기존의 딥러닝 기반 빔 예측 모델을 개선하기 위해, 8개의 Layer로 구성된 트랜스포머 블록을 통해 두 가지 데이터를 융합하여 각 데이터 간의 관계까지 고려하는 새로운 빔 예측 모델을 제안하였다. 32개 빔과 64개 빔을 예측하는 2가지 실험에서 제안 모델이 기존 모델 대비 Top-1, 2, 3 정확도 모두에서 높은 성능을 보이는 것을 확인하였다. 특히 제안 모델의 Top-3 정확도는 두 가지 실험 모두에서 거의 100%에 가까운 정확도를 보였다. 이러한 결과를 바탕으로 추후 LiDAR 및 Radar 같은 다양한 센서 데이터를 추가로 활용한다면 더욱 높은 정확도를 갖는 빔 예측 모델을 구축할 수 있고, 다양한 시나리오에서 수집된 실측 데이터를 활용하여 학습하면 실제 통신 환경에서 효과적으로 사용될 수 있을 것으로 기대된다.

참고 문헌

- [1] J. Kim and J. Kim, "Augmenting Beam Alignment for mmWave Communication Systems via Channel Attention," *Electronics*, vol. 12, no. 20, pp. 1-15, Oct. 2023.
- [2] G. Charan, A. Hredzak, C. Stoddard, B. Berrey, M. Seth, H. Nunez, and A. Alkhateeb, "Towards Real-World 6G Drone Communication: Position and Camera Aided Beam Prediction," in *GLOBECOM 2022- 2022 IEEE Global Communications Conference*. IEEE, 2022, pp. 2951 - 2956.
- [3] Y. Tian, Q. Zhao, Z. Kherroubi, F. Boukhalfa, K. Wu, and F. Bader, "Multimodal Transformers for Wireless Communications: A Case Study in Beam Prediction," *arXiv preprint arXiv:2309.11811*, 2023.
- [4] K. Chitta, A. Prakash, B. Jaeger, Z. Yu, K. Renz, and A. Geiger, "Transfuser: Imitation with Transformer-based Sensor Fusion for Autonomous Driving," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 11, pp. 12878 - 12895, Nov. 2023.
- [5] A. Alkhateeb, G. Charan, T. Osman, A. Hredzak, and N. Srinivas, "DeepSense 6G: Large-Scale Real-World Multi-Modal Sensing and Communication Datasets." 2022. [Online]. Available: <https://www.DeepSense6G.net>