

# 대규모 멀티모달 모델 기반 환경 인지 채널 예측

정지원, 문지훈, 심병호

서울대학교

{jwjeong, jhmoon}@islab.snu.ac.kr, bshim@snu.ac.kr

## Large Multimodal Model-based Environment-aware Channel Prediction

Jiwon Jung, Jihoon Moon, Byonghyo Shim

Seoul National Univ.

### 요약

대역폭과 커버리지의 균형을 달성하기 위해 6G 무선 통신은 upper mid-band (7-24 GHz)를 새로운 주파수 대역으로 활용할 예정이다. Upper mid-band 채널은 다수의 전파 경로로 구성되므로, 빔포밍 성능을 극대화하기 위해서는 기지국은 이동 단말의 다중경로 전파 환경을 추적해야 한다. 그러나 환경과 주변의 맥락 정보가 부족한 탓에 기존 RF 기반 채널 예측 기법은 불연속적인 채널 변화에 취약하다. 이에 채널을 연속적이고 정확하게 추적하기 위해 대규모 멀티모달 모델을 활용한 채널 예측 기법을 제안한다. 본 기법은 맥락 정보를 바탕으로 각 전파 경로를 추적한다. 이미지와 파일럿 측정 간의 공간 상관관계를 정량화함으로서 이동 단말과 신호 반사 지점의 움직임을 연속적으로 추적한다. 이때 대규모 멀티모달 모델의 자연어 처리 능력을 활용하여 세부적인 작업 지침을 내리고 네트워크 규격을 해석해서 채널 예측 정확도를 향상시킨다. 성능 평가 결과, 제안 기법은 기존 RF 기반 기법 대비 채널 예측 정확도에서 성능 향상 달성을 이뤄낸다.

### I. 서론

대용량 데이터 전송과 지연 민감성이 요구되는 다양한 애플리케이션을 지원하기 위해 주목을 받았던 millimeter-wave (mmWave) band 대역은 심각한 경로 손실로 인해 현실적인 상용화가 제한적이다. 이에 따라 새로운 주파수 대역대로 upper mid-band가 많은 관심을 받고 있다. Upper mid-band에서는 네트워크 용량과 커버리지를 향상시키기 위해 대규모 다중 입력 다중 출력 시스템 기반의 빔포밍 기법이 활용된다. 이를 위해 채널을 정확히 추정하고 예측하기 위한 다양한 기법들이 제안되어 왔다. 특히 5G NR에서는 기지국이 동기 신호를 전송하고 단말이 수신 전력이 최대가 되는 빔의 인덱스를 피드백하는 방식이 사용된다.[1] 또한 고주파 대역에서 빠르게 변화하는 채널을 추적하기 위해, 채널을 사전에 예측하는 기법들도 제안되었다.[2],[3] 하지만 기존 RF 기반 채널 예측 기법들은 채널의 급격한 변화를 예측하는 것에는 한계를 가진다. 이는 이동 단말과 산란체의 위치, 움직임과 같은 환경과 맥락 정보가 부재하기 때문에, 경로 차단, 사용자 이동 패턴 변화, 새로운 산란체의 등장 등으로 인해 발생하는 불연속적인 채널 변동을 효과적으로 예측하기 어렵다. 또한 upper mid-band에서는 다수의 전파 경로가 존재하므로, 기존 지배적인 단일 경로 추정에 초점을 맞추고 있는 채널 예측 기법들은 upper mid-band의 채널을 온전히 예측하기에는 제한적이다.

따라서 본 논문에서는 멀티모달 센싱 데이터와 대규모 멀티모달 모델을 활용하여 새로운 채널 예측 기법을 제안한다. 본 기법은 언어뿐 아니라 이미지, 비디오, 오디오 등 다양한 멀티모달 데이터를 활용해 정보를 효과적으로 획득할 수 있는 대규모 멀티모달 모델의 능력을 활용하여 비전 센싱과 RF 신호를 결합해 채널을 정확히 예측한다. 구체적으로 기지국은 과거 채널 정보로부터 채널 변화의 잠재적 패턴을 학습하며, 비전 센서를 활용해 물리적 환경을 직접 관측한다. 이를 통해 멀티모달 측정치를 분석함으로써 단말의 위치와 신호 반사 지점을 추적한다. 또한 비전 센싱 기반의

맥락 정보를 통해 전파 경로 변화에 필요한 단서 (단말과 산란체의 이동 방향)를 확보함으로써 다중경로 구조를 연속적으로 추적한다. 성능 평가 결과, 제안 기법은 기존 RF 기반 기법 대비 채널 예측 정확도에서 성능 향상 달성을 이뤄낸다.

### II. 본론

#### 가. 채널 모델

본 논문에서는 다음과 같은 다중경로 채널 모델을 사용한다.

$$h_k = \sum_{l=1}^L \beta_{k,l} e^{j\alpha_{k,l}} a(\phi_{k,l}, \theta_{k,l}) \quad (1)$$

여기서  $L$ 은 경로의 개수,  $\beta_{k,l}$ 은  $l$ 번 째 경로의 경로 이득,  $\alpha_{k,l}$ 은 경로 지연, 반사, 도플러 시프트 등에 의해 유도되는 위상 변화이다. 또한  $a(\phi_{k,l}, \theta_{k,l})$ 은 방위각과 고각 방향의 송신 출발각에 따른 균일 평균 배열 응답 벡터를 나타낸다.

#### 나. 대규모 멀티모달 모델 기반 환경 인지 채널 예측

본 논문에서는 대규모 멀티모달 모델을 기반한 채널 예측 기법을 제안한다. 본 기법의 주요 목표는 이미지와 RF 측정값으로부터 채널 정보를 추출하여 미래의 채널을 예측하는 것이다. 이를 위해 본 기법에서는 대규모 멀티모달 모델을 기반으로 경로 추적 구조를 설계하며, 이동 단말 추적 네트워크와 반사 지점 추적 네트워크의 두 서브 네트워크를 통해 신호 전파 경로를 추적한다. 구체적으로 이동 단말 네트워크에서는 대규모 멀티모달 모델이 이미지에서 단말 위치를 검출한 뒤, 과거 위치 정보들을 기반으로 미래의 단말 위치를 예측한다. 다음으로 반사 지점 추적 네트워크에서는

산란체에 대한 regression을 수행해 신호 반사 지점을 식별한다. 이를 통해 대규모 멀티모달 모델은 미래의 전파 환경을 재구성한다.

### 이동 단말 추적 네트워크

이동 단말 추적 네트워크에서는 과거 단말 위치  $\{p_{mob}^{(t-t')}\}_{t'=1}^{T-1}$  및 이미지를 이용해 단말의 미래 위치를 예측한다. 구체적으로 먼저 비전 센싱 이미지에서 객체 탐지를 진행해 현재의 단말 위치  $p_{mob}^{(t)}$ 를 얻어낸다. 이후 과거 단말 위치를 활용해 다음의 조건부 확률을 학습함으로써 단말 위치  $p_{mob}^{(t+\Delta t)}$ 를 예측한다.

$$f(p_{mob}^{(t+\Delta t)} | p_{mob}^{(t-t')} \}_{t'=0}^{T-1}) \quad (2)$$

### 반사 지점 추적 네트워크

다음으로 대규모 멀티모달 모델은 반사 지점 집합을 예측한다. 대규모 멀티모달 모델은 비전 센싱 이미지와 RF 기반 채널 파라미터 추정 행렬을 사용해 시각적으로 주변 산란체 집합  $S_{sca}^{(t)}$  및 반사 지점  $P_{ref}^{(t)}$ 을 식별한다. 구체적으로 cross-attention 메커니즘을 이용해 이미지 픽셀과 RF 추정치 간의 공간적 연관성을 평가하고, 산란체가 포함된 영역에 높은 점수를 부여하여 산란체 객체를 선택한다. 이후 각 산란체에 대해 RF로 추정된 경로 각도에 대응하는 이미지 픽셀과 깊이값으로 반사 지점의 초기 추정치를 얻고, 후보 영역에서 표면 법선 정보를 이용해 반사 지점을 정밀 보정한다. 마지막으로 산란체 집합과 과거 반사 지점 위치를 바탕으로 다음의 조건부 확률을 학습함으로써 미래의 반사 지점 집합  $P_{ref}^{(t+\Delta t)}$ 을 예측한다.

$$f(P_{ref}^{(t+\Delta t)} | S_{sca}^{(t+\Delta t)}, \{P_{ref}^{(t-t')}\}_{t'=0}^{T-1}, p_{mob}^{(t+\Delta t)}) \quad (3)$$

## 4. 실험

### 시뮬레이션 환경 및 설정

본 실험에서는  $8 \times 8$  안테나로 구성된 균일 평면 배열을 갖는 기지국이 단일 안테나를 가지는 5명의 사용자를 서비스한다고 가정한다. 3D 환경에 위치한 기지국은 RGB-D 이미지를 연속적으로 획득하며, 채널 파라미터는 레이 트레이싱 기법을 통해 산출한다. 과거 측정값의 개수는 10, 측정 값 간의 간격은 10ms 예측 시점은  $\Delta t = 5$ 로 설정한다. 대규모 멀티모달 모델로는 LLaVA-NeXT-Interleave 7B를 사용한다.

### 성능 평가 결과

그림 1은 수신 signal to noise ratio (SNR)의 함수로서 예측 채널의 정규화 제곱평균오차를 나타낸다. 제안한 대규모 멀티모달 모델 기반 채널 예측 기법은 기존 기법들에 비해 채널 예측 정확도를 큰 폭으로 향상시키는 것을 확인할 수 있다. 예를 들어 SNR = 20dB에서 본 기법은 Kalman filter 기반 채널 예측[4], prony-based angular-delay domain(PAD) 기반 채널 예측[5], Transformer 기반 채널 예측[6] 대비 정규화 제곱평균오차를 각각 65.4%, 65.7%, 51.7% 이상 감소시킨다.

## III. 결론

본 논문에서는 upper mid-band 대역 시스템을 위한 대규모 멀티모달 모델 기반 환경 인지 채널 예측 기법을 제안하였다. RF 신호 기반 채널 추정에만 의존하는 기존 접근과 달리, 본 기법은 센싱 정보와 과거 채널 정보

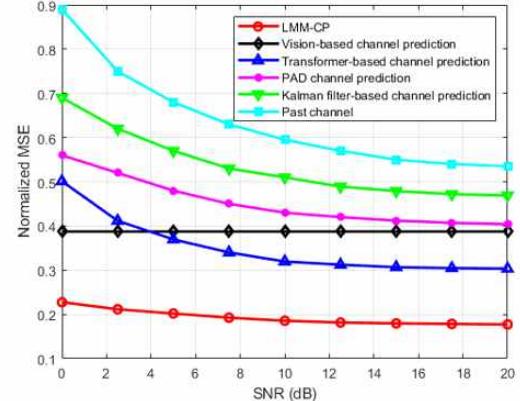


그림 1. 수신 SNR에 따른 채널 정규화 제곱평균오차

를 결합해 환경을 정밀하게 인지한다. 특히 대규모 멀티모달 모델은 단말 및 산란체의 이동 방향을 획득함으로써 전파 경로 환경의 급격한 변화를 보다 정확하게 예측할 수 있다. 성능 평가를 통해 본 기법이 채널 예측 성능 향상에 매우 효과적임을 확인하였다.

## ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 지원으로 한국연구재단의 지원을 받아 수행된 연구임 (RS-2022-NR070834)

## 참 고 문 헌

- [1] R. M. Dreifuerst and R. W. Heath, "Massive MIMO in 5G: How beamforming, codebooks, and feedback enable larger arrays," *IEEE Commun. Mag.*, vol. 61, no. 12, pp. 18 - 23, 2023.
- [2] Q. Xue, C. Ji, S. Ma, J. Guo, Y. Xu, Q. Chen, and W. Zhang, "A survey of beam management for mmWave and THz communications towards 6G," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 3, pp. 1520 - 1559, 2024.
- [3] J. Moon, K. A. Ngo, B. Lee, and B. Shim, "Multi-modal sensing-aided channel prediction for 6G mmWave massive antenna systems," *IEEE Trans. Cogn. Commun. Netw.*, 2025.
- [4] H. Kim, S. Kim, H. Lee, C. Jang, Y. Choi, and J. Choi, "Massive MIMO channel prediction: Kalman filtering vs. machine learning," *IEEE Trans Commun.*, vol. 69, no. 1, pp. 518 - 528, 2020.
- [5] H. Yin, H. Wang, Y. Liu, and D. Gesbert, "Addressing the curse of mobility in massive MIMO with prony-based angular-delay domain channel predictions," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 12, pp. 2903 - 2917, 2020.
- [6] H. Jiang, M. Cui, D. W. K. Ng, and L. Dai, "Accurate channel prediction based on transformer: Making mobility negligible," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2717 - 2732, 2022.