

6G 통신을 위한 멀티 모달 센싱을 활용한 빔포밍 기술 연구

변용석, 노태엽, 심병호
서울대학교

{ysbyun, tyroh}@islabs.snu.ac.kr, bshim@snu.ac.kr

Beamforming using Multi-modal Sensing for 6G Communications

Byun Yongsuk, Roh Taeyup, Shim Byonghyo
Seoul National University

요약

밀리미터 파 통신과 같은 핵심 기술을 사용하여 5G 및 6G 통신 시스템의 지속적으로 증가하는 데이터 속도 요구를 충족할 수 있다. 그러나 이러한 고주파 대역 통신은 심각한 경로 손실을 겪으며, 이는 다중 입력 다중 출력 시스템으로 실현된 빔포밍 기술을 사용하여 보상할 수 있다. 이러한 빔포밍 기술을 보다 효과적으로 수행하기 위해 다중 모달 센싱 기술을 이용한 빔 관리 기술을 제안한다. 컴퓨터 비전 기반 객체 검출기를 사용하여 다중 모드 감지 데이터에서 3D 위치 정보를 추출함으로써 기지국은 모바일의 위치를 향한 빔을 생성할 수 있다. 이는 일부 피드백 작업을 대체하여 빔 방향을 식별할 수 있기 때문에 제안기법은 데이터 속도 및 지연 시간 측면에서 성능 향상을 달성한다.

I. 서론

최근 밀리미터 파(millimeter wave, 이하 mmWave), 테라헤르츠(terahertz, 이하 THz) 대역 통신이 6G에서 데이터 집약적이고 지연에 민감한 애플리케이션을 지원하는 핵심 기술로 부상했다 [1]. 이 고주파 대역에서 잘 알려진 문제는 신호 전파, 반사, 확산 산란 및 대기 흡수로 인한 심각한 신호 경로 손실이다. 이러한 문제를 해결하기 위해 높은 방향성 빔을 생성하는 다중 입력 다중 출력 (Multiple-Input Multiple-Output, 이하 MIMO) 빔포밍 기술이 도입되었다. 빔포밍 이득을 최대한 활용하기 위해, 빔들은 신호 전파 경로에 적절하게 정렬되어야 하고, 따라서 기지국(base station, 이하 BS)에서 각도 형태의 채널 정보를 정확하게 획득하는 것이 빔포밍 기술의 성공에 중요하다. 채널 정보를 얻기 위해 5G NR에서 코드북 기반 빔 관리가 사용된다. 5G NR 빔 관리는 간단하고 구현이 용이한 반면, 기지국(base station, 이하 BS)과 모바일 간의 2 단계 핸드셰이크 메커니즘으로 인해 지연 및 자원 활용 효율 측면에서 오버헤드가 상당히 크다 [2].

이러한 채널정보를 획득하고 빔포밍을 수행하기 위해서 본 논문은 다중 모달 (multi-modal) 센싱 기술을 활용한 빔 관리 기술 (beam management using multi-modal sensing, 이하 BMMM)을 제안한다. 특히 다중 감지 모달리티(LiDAR 및 RGB 카메라)의 강점을 통합하고 활용하는 것은 큰 잠재력을 가지고 있으며, 다중 모달 감지 정보를 효과적으로 빔관리에 접목시킴으로써 실제로 빔포밍 이득 및 파일럿 신호 오버헤드 측면에서 상당한 이득을 달성함을 보인다. 구체적으로 본 논문에서 제안하는 BMMM은 컴퓨터 비전(computer vision, 이하 CV) 기반 객체 검출기를 사용하여 RGB 이미지에서 모바일 기기를 식별하고 BS에서 각도 방향을 추출한다. 그런 다음 우리는 라이다를 RGB 이미지에 투영하여 모바일의 거리를 추출한다. 각도 방향과 거리를 식별함으로써 모바일의

위치를 획득하고 모바일의 위치로 향하는 빔을 생성할 수 있다.

II. 제안 기법

제안된 BMMM의 주요 목표는 다중모달 센싱 기법을 사용하여 지향성 빔을 생성하는 것이다. RGB 이미지와 라이다 포인트 클라우드를 통합한 다음 CV 기반 객체 검출기를 활용하여 무선 객체(예: 휴대폰, 노트북, 태블릿)를 식별하고 위치를 추정할 수 있다. 제안 기법은 세 가지 주요 운영 단계로 구성된다: 1) 전경 추출, 2) CV-기반 모바일 기기 위치 추정, 3) 포지션 기반 빔 형성.

1) 전경 추출

배경 포인트들 중에서 라이다의 레이저가 전송된 동일한 각도 방향의 과거 지점 중 라이다 센서와의 거리가 가장 긴 지점이 배경 지점에 해당한다. 그러면 백그라운드 포인트 클라우드 \mathbf{P}_{bg} 는 다음과 같이 표시된다.

$$\mathbf{P}_{bg} = \{\mathbf{p}_1^{(\tau_1)}, \dots, \mathbf{p}_{M_L}^{(\tau_{M_L})}\} \quad (1)$$

여기서 $\tau_i = \underset{\tau \in \{t-W, \dots, t-1\}}{\operatorname{argmax}} \|\mathbf{p}_i^{(\tau)}\|$ 는 길이 W 의 연속적인 라이다 시퀀스에서 최대 거리가 발생하는 시간 인덱스이다. 배경 포인트 클라우드 \mathbf{P}_{bg} 추출 후에 우리는 실시간 포인트 클라우드 $\mathbf{P}^{(t)}$ 에서 배경 포인트를 제거하여 전경 포인트 클라우드 \mathbf{P}_{fg} 를 추출한다. 구체적으로, 미리 계산된 배경점에서 유클리드 거리가 임계값 ϵ_{th} 보다 작은 점을 제거하여 배경 정보를 제거한다. 라이다의 측정에서 발생하는 노이즈로 인해 시간 단위가 다른 동일한 각도 방향에서 획득한 포인트는 약간 다를 수 있다. 따라서 서로 가까운 점을 동일한 점으로 간주하기 위해 임계값 ϵ_{th} 을 라이다의 거리 측정 오차로 설정한다 (우리의 경우 임계값 ϵ_{th} 를 0.001로 설정한다). 그러면 추출한 전경 포인트 클라우드는 다음과 같이 나타낼 수 있다.

$$\mathbf{P}_{fg} = \{\mathbf{p}_i^{(t)} \mid \|\mathbf{p}_i^{(t)} - \mathbf{p}_i^{(\tau_i)}\| > \epsilon_{th}, i = 1, \dots, M_L\} \quad (2)$$

전경 포인트 클라우드 \mathbf{P}_{fg} 를 추출한 이후에, 이미지 평면에서 전경 이미지 \mathbf{I}_{fg} 를 추출하기 위해 이미지 평면 \mathbf{I} 에 전경 포인트 클라우드 \mathbf{P}_{fg} 를 투영하고 투영된 포인트들이 존재하는 영역을 전경 이미지 \mathbf{I}_{fg} 라 한다.

2) CV-기반 모바일 기기 위치 추정

전경 추출 단계 이후, BS는 딥러닝(deep learning, 이하 DL) 기반의 객체 검출을 이용하여 소형 무선 객체의 객체 검출 및 3D 측위를 수행한다. DL 기반 객체 검출기의 입력으로 전경 이미지 \mathbf{I}_{fg} 를 사용하여 작은 크기의 무선 객체의 정확한 위치를 얻을 수 있다. DL 기반 객체 검출기는 백본(backbone), 넥(neck), 헤드(head)의 세 가지 주요 구성 요소로 구성된다. 백본은 입력 이미지에서 특징(예: 색상, 모양 및 얼굴)을 추출하는 역할을 하는 반면, 넥은 이러한 특징을 융합하고 공간 해상도를 줄이는 데 사용된다. 마지막으로 헤드는 이미지의 객체에 대한 경계 상자와 신뢰 점수를 예측하여 출력을 생성한다. 결론적으로, DL 기반 객체 탐지기는 각 \mathbf{I}_{fg} 내에서 N 개의 모바일 장치를 식별하고 위치를 추정한다. 구체적으로 DL 기반 객체 검출기는 경계 상자의 중심점으로 표현되는 객체의 클래스와 2D 위치 (u_n, v_n) 를 추정한다. 이때, 수평 각도 ϕ_n 및 수직 각도 θ_n 는 다음과 같이 나타낼 수 있다.

$$(\phi_n, \theta_n) = \left(\frac{u_n - c_x}{N_u} \phi_f, \frac{v_n - c_y}{N_v} \theta_n \right) \quad (3)$$

여기서 ϕ_f 와 θ_n 는 각각 수평 및 수직 화각(field of view)을 나타내며, (c_x, c_y) 는 이미지의 중앙 픽셀 좌표를 나타내고, N_u 와 N_v 는 각각 이미지의 총 수평 및 수직 픽셀 수를 나타낸다. 이후, 모바일의 3D 위치를 얻기 위해 (u_n, v_n) 에 해당하는 라이다 포인트에서 거리 정보를 추출한다.

3) 포지션 기반 빔 형성

안테나의 수가 많을 때 모바일의 채널 벡터는 서로 접근적으로 직교한다. 이것은 빔이 다른 모바일의 위치로 향함으로써 발생하는 간섭이 무시할 수 있음을 의미한다. 이 경우, 최적의 빔포밍 행렬 \mathbf{F}^* 는 다음과 같이 나타낼 수 있다.

$$\mathbf{F}^* = \mathbf{F}_{RF}^* \mathbf{F}_{BB}^* = [\mathbf{a}(\phi_1, \theta_1), \dots, \mathbf{a}(\phi_N, \theta_N)] \text{diag}(\sqrt{p_1}, \dots, \sqrt{p_N}) \quad (4)$$

여기서 $\mathbf{a}(\phi_n, \theta_n)$ 은 배열 응답 벡터(array response vector)를 나타내며, p_n 은 n 번째 빔의 송신 전력을 나타낸다.

III. 실험 결과

MMBM의 주요 작동은 감지 정보와 CV 기반 객체 감지를 사용하여 모바일의 위치를 식별하는 것이기 때문에 제안된 접근 방식의 효과를 평가하기 위한 데이터 세트가 필요하다. 이를 위해 Velodyne VLP-16 LiDAR와 Intel RealSense L515 RGB-D 카메라를 사용하여 데이터셋을 생성한다. 제안된 기법의 성능을 평가하기 위해서 컴퓨터 비전 작업의 효율성과 검출 성능이 우수한 트랜스포머 기반 모델의 변형인 Swin Transformer를 객체 검출기로 사용한다 [3]. 특히, 본 논문에서는 80개의 객체 클래스와 118,000개의 훈련 이미지로 구성된 MS-COCO 2017 데이터 세트에서 사전 학습된 모델을 사용한다 [4]. 통신 성능을 평가하기 위해서 저희의 시뮬레이션에서는 모바일 장치는 $10 \times 10 \text{ m}^2$ 의 정사각형 영역 내에 무작위로 분포한다. 신호 전파는 고주파 시스템에서 LoS 구성 요소에 의해 주로 결정되므로 28 GHz의 반송파 주파수를 가진 LoS 채널 모델을 사용한다.

우리는 위치 오차 측면에서 제안 기법의 성능을 평가하였으며, 제안 기법이 각도와 거리 정보 획득 모두에서 기존 방식을 능가하는 것을 보인다. 예를 들어, 5G-BM의

방향 오류는 7.9° 이지만 제안 기법의 경우 0.2° 로 97% 성능이 향상됨을 확인하였다. 또한 5G-BM의 경우 거리 오류가 128.5 cm인 반면 제안 기법의 경우 8.8 cm의 오차를 가진다. 이는 5G-BM에서 코드북 양자화 오차가 발생하는 반면에 제안 기법은 DL 기반 객체 검출기를 사용하여 연속 빔 방향을 직접 획득하여 불일치를 제거하기 때문이다.

그림 1에서는 안테나 개수에 대한 주파수 효율을 나타내며, 우리는 제안된 기술이 벤치마크 체계를 능가함을 보여준다. 특히, 안테나 수가 증가하면 성능 차이가 증가하는 것을 확인하였다. 예를 들어, 안테나 수가 64개에서 196개로 증가하면 제안 기법의 주파수 효율 이득은 7.2%에서 11.4%로 증가하며, 이는 좁은 빔이 위치 오차에 더 민감하기 때문이다.

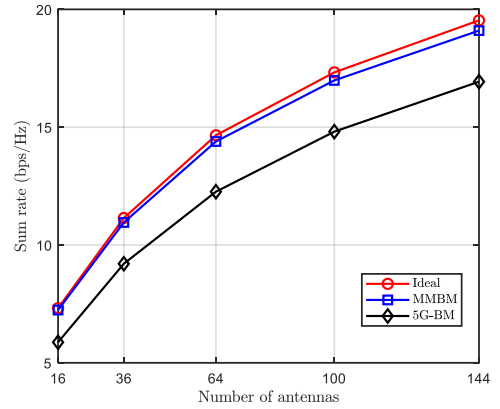


그림 1. 안테나 수에 따른 주파수 효율

IV. 결론

본 논문에서는 멀티 모달 센서(라이다와 RGB 카메라)를 이용한 새로운 빔 관리 기법을 제안하였다. 제안된 빔 관리 기법의 핵심 아이디어는 빔포밍 벡터 생성에서 다중 모달 센싱 정보로부터 획득된 위치 정보를 활용하는 것이다. 구체적으로 CV 기반 객체 검출기를 사용하여 RGB 이미지에서 모바일 기기를 식별하고 BS에서 각도 방향을 추출한다. 이후, 우리는 라이다를 RGB 이미지에 투영하여 모바일의 거리를 추출한다. 각도 방향과 거리를 식별함으로써 모바일의 위치를 획득하고 모바일의 위치로 향하는 빔을 생성한다. 그렇게 함으로써, 빔 포밍의 품질이 상당히 향상될 수 있고, 빔 트레이닝 오버헤드(자원 오버헤드 및 지연 시간)가 실질적으로 감소된다.

참고 문헌

- [1] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," Proc. IEEE, vol. 102, no. 3, pp. 366-385, 2014.
- [2] S. Kim, J. Park, J. Moon, and B. Shim, "Fast and accurate terahertz beam management via frequency-dependent beamforming," IEEE Trans. Wireless Commun., pp. 1-1, 2023.
- [3] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), 2021, pp. 9992-10002.
- [4] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in Proc. European Conf. Comput. Vis. (ECCV), 2014, pp. 740-755.