

Proceeding

ISSN 2287-4348

2026

한국스마트미디어학회 & 한국디지털산업학회

춘계학술대회

AX기반 스마트미디어
디지털 기술 공진화 패러다임

일시 | 2026년 4월 24일(금) ~ 25일(토)

장소 | 가천대학교 비전타워

2026

한국스마트미디어학회 & 한국디지털산업학회

춘계학술대회

AX기반 스마트미디어
디지털 기술 공진화 패러다임

일시 | 2026년 4월 24일(금) ~ 25일(토)

장소 | 가천대학교 비전타워



포스터 발표 P-6(학부생 논문)

○ 시간	04월 25일 (토) 15:10~16:40
○ 발표장	가천대학교 비전타워 B103 복도 B (유리 부착)
○ 좌장	서채연 교수(홍익대학교)

		Page
P6-01	제목: 소형 객체 탐지를 위한 밀도 기반 기법과 차영상 기반 기법 비교 저자: 박선준, 이상웅(가천대학교)	298
P6-02	제목: 치매 분류를 위한 PET 조영제별 3차원 의료 영상 분석 모델 성능 비교 실험 저자: 이정우, 안성윤, 이상웅(가천대학교)	300
P6-03	제목: 소프트 보팅 앙상블 기법을 활용한 악성 URL 이진 분류 시스템 설계 저자: 채지윤, 정세훈, 심춘보(순천대학교)	302
P6-04	제목: CV 및 센서 데이터 기반 토경 재배 AI 식물 관리 시스템 설계 저자: 정도영, 오현준, 김준형, 김장환, 서채연, 김영철(홍익대학교)	304
P6-05	제목: 기상 및 공간 데이터 이용한 초국소 열섬 지도 기반 이용자 맞춤형 안전 경로 탐색 서비스 저자: 조연후, 송강민, 이지선, 이신영, 서채연, 김영철(홍익대학교)	306
P6-06	제목: 비전 AI 및 RAG 이용한 복약 안전 관리 기반 메카니즘 저자: 강민준, 류강희, 윤성준, 서채연, 김영철(홍익대학교)	308
P6-07	제목: 가상환경내에서 제스처 및 의도 인식 기법 저자: 권성민, 서채연, 김영철(홍익대학교)	310
P6-08	제목: GeoAI 및 XAI 기반 도심 쓰레기 핫스팟 예측과 동적 라우팅 시스템 아키텍처 설계 저자: 김지환, 박태은, 이상현, 이재륜, 최영현, 최광훈(전남대학교)	312
P6-09	제목: 동적 이진 변환 시 코드 생성 최적화에 따른 에뮬레이션 성능 분석 저자: 이정윤, 김민재, 박신혁, 장하림, 민홍(가천대학교)	314
P6-10	제목: 건물 내 이동 경로 안내 프로그램 저자: 정준호, 신민섭, 김한일(제주대학교)	316

가상환경내에서 제스처 및 의도 인식 기법

권성민¹, 서채연², 김영철^{3*}

홍익대학교 세종캠퍼스 소프트웨어융합학과

e-mail : roda0817@gmail.com¹, {chaeyun², bob^{3*}}@hongik.ac.kr

Human Gesture and Intent Recognition Method with LLM Learning

Sungmin Kwon¹, Chaeyun Seo², R. Young Chul Kim^{3*}

¹Undergraduate Dept. of Software and Communications Engineering, Hongik University

^{2,3*}Dept. of Software and Communications Engineering, Hongik University

요 약

현재 메타버스 환경은 제스처 인식을 위해 HMD나 컨트롤러 같은 장비가 필수적이며, 기기를 착용하지 않으면 동작 인식이 불가능하다. 이를 보완하기 위해 WiFi 채널 상태 정보(CSI) 기반의 기술이 비접촉 센싱 방법으로 주목받고 있으나, 특정 환경에 학습되어 공간 배치가 변하거나 새로운 장소에서는 의도 추론 성능이 저하된다. 우리는 이러한 환경 의존성을 극복하고 다양한 공간 조건에서도 일반화된 의도 이해를 가능하게 하기 위해, 공간 문맥 정보를 활용한 새로운 의도 추론 방법을 제안한다. 제안하는 접근 방식은 제스처 인식 결과를 공간적 상황과 결합하여 해석함으로써, 단순한 패턴 분류를 넘어 문맥 기반의 의미 도출을 수행하도록 설계한다. 이를 통해 학습되지 않은 환경에서도 사용자의 행동 의도를 보다 유연하게 해석할 수 있을 것으로 기대한다.

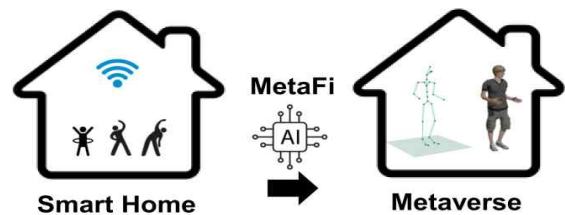
1. 서 론

최근 WiFi 채널 상태 정보(CSI)를 활용한 비접촉식 인간 행동 인식 기술은 프라이버시 보호와 낮은 설치 비용 덕분에 스마트 공간 및 메타버스 등 다양한 분야에서 주목받고 있다 [1, 2]. 그러나 기존 시스템은 특정 환경에서 수집된 신호 패턴에 의존하여 학습되는 경향이 있어, 공간 배치나 환경 조건이 변화하면 인식 성능이 급격히 저하되는 일반화의 한계를 지닌다. 또한, 사용자의 행동을 단순한 클래스 범주로 분류하는 데 그쳐 복잡한 상황에 적응하거나 의미적인 의도 수준까지 해석하기에는 어려움이 존재한다. 이러한 환경 의존성과 의미 해석의 한계는 실제 서비스 확장을 위해 반드시 해결해야 할 과제이며, 이에 따라 다양한 환경에서도 사용자의 의도를 안정적이고 유연하게 파악할 수 있는 일반화된 추론 방식이 요구되고 있다. 본 논문의 구성은 다음과 같다. 제2장에서는 관련연구인 MetaFi에 대해 언급하고, 제3장에서는 공간 문맥 정보를 활용한 사용자 의도 추론 시스템의 개념과 전체 시스템 구조를 제시한다. 제4장에서는 제안 방법의 구현 과정과 추론 절차를 설명하고, 다양한 환경 조건에서 수행한 실험 결과를 통해 제안 방법의 성능을 평가하고 기존 방법과 비교 분석한다. 마지막으로 제5장에서는 연구의 결론과 향후 연구에 대해 언급한다. 본 연구는 홍익대학교 소프트웨어융합학과 학부생들의 종합설계 프로젝트 결과물이며, 홍익대학교 메타버스 융합SW 아카데미 6기 교육생들의 프로젝트 결과물이다.

2. 관련 연구

WiFi 채널 상태 정보(CSI) 기반 행동 인식은 비가시선 환경

의 실용성을 입증한 WiSee와 WiGest를 거쳐, 최근 Widar 등 딥러닝을 활용한 정밀한 위치 독립적 모델로 발전해 왔다 [1-3]. 대표적 기술인 MetaFi는 무선 신호 분석과 교차모달 감독(Cross-modal Supervision)을 통해 카메라 방식의 사생활 침해 및 조명 문제를 해결하며 높은 자세 추정 정확도를 달성했다 [1]. 그림 1은 WiFi를 활용한 스마트 홈과 메타버스를 연결하는 MetaFi이다.

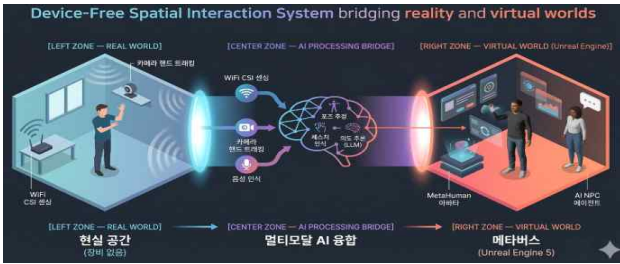


(그림 1) WiFi를 활용한 스마트 홈과 메타버스를 연결하는 MetaFi

그러나 이러한 방식은 여전히 주변 환경 변화에 민감하고 신호 간섭 시 개별 움직임 식별이 어렵다는 한계가 있다. 이를 극복하기 위해 최근 Few-shot 학습과 Chain-of-Thought(CoT) 기법을 갖춘 대규모 언어 모델(LLM)을 활용하여 추론 성능을 높이려는 시도가 이어지고 있다 [3, 4]. 하지만 WiFi CSI 인식 결과와 공간 맥락을 결합해 의도 해석까지 확장한 연구는 미비한 실정이다. 이에 본 연구는 LLM의 추론 능력을 결합하여, 미학습 환경에서도 사용자의 움직임을 정확히 파악하고 의도를 유연하게 해석하는 통합적 접근 방안을 제안한다.

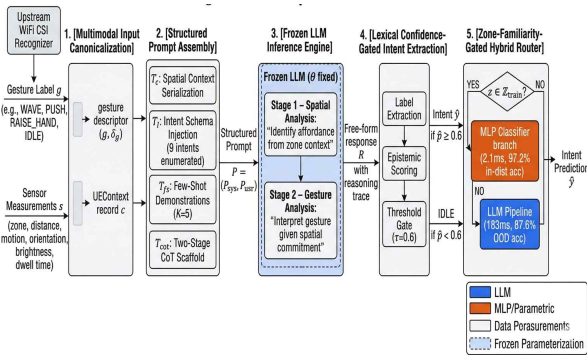
3. WiFi CSI 기반 LLM 제스처 의도 추론 기법

제한한 WiFi CSI 기반 LLM 제스처 의도 추론 기법은 학습되지 않은 환경에서도 사용자의 움직임을 정확히 파악하고 새로운 의도를 유연하게 해석할 수 있는 방법이다. WiFi CSI 신호는 업스트림 인식기를 거쳐 네 가지 제스처 ({WAVE, PUSH, RAISE_HAND, IDLE})로 분류된다. 이후 ContextIntent는 인식된 제스처와 구역 정보를 결합하여 의도를 추론한다. 이때 학습된 구역은 MLP 분류기(제스처 4차원 및 구역 원-핫 6차원 입력, 의도 9차원 출력)를 통해 결과를 도출하며, 미학습 구역은 LLM을 활용하여 의도를 추론한다. 그림 2는 ContextIntent의 개념도이다.



(그림 2) ContextIntent 개념도

현실공간에서의 제스처가 장비없이 가상공간에서 인식되는 개념이다.



(그림 3) ContextIntent 전체 시스템 구조

그림 3은 ContextIntent 전체 시스템 구조이다. WiFi CSI 정보를 활용한 제스처 인식 기술은 멀티모달 AI 융합을 통해 메타버스 공간에서의 제스처 인식 및 상호작용 기술로 확장 가능할 것으로 기대한다.

공간 우선 2단계는 Two-Stage CoT로 추론한다. 추론 순서를 명시적으로 강제하는 것이 핵심이다.

- 1단계: 공간분석

제스처를 고려하기 전, 현재 위치에서 가능한 상호작용 목표를 먼저 파악한다.

- 2단계: 제스처 분석

공간 맥락 위에서 관찰된 제스처를 해석한다.

학습 구역의 5가지 구역-제스처 조합에 대한 2단계 추론 예시를 Few-Shot으로 제공한다. 추론 순서의 강제는 공간 맥락을 고려하지 않고 제스처의 표면적 의미만으로 판단하는 오류를 효과적으로 방지한다. 실제 WiFi CSI 하드웨어를 사용하지 않으며, 제스처-구역 조합을 텍스트로 구성된 가상 시나리오로 의도 추론 단계만 독립적으로 검증한다. 가상 거실(벽난로, 책장, 문)의 9가지 조합을 in-distribution으로, 미학습 5개 구역(TV, 냉장고, 창문, 소파, 에어컨)의 15가지 조합을 OOD 테스트셋으로 구성한

다. 각 조건당 50회 시행하고(모델: qwen/qwen 3.5-9b), 엄격(Strict)·관대(Lenient) 두 가지 기준을 병행한다.

표 1. 방법별 정확도 비교, 각 50회 시행

방법	In-Dist	OOD (lenient)	95% CI (OODLenient)
MLP (학습 기반)	82.0%	0.0%	-
LLM Zero-shot (v1)	64.0%	26.0%	10-32%
LLM Few-shot K=3 (v1)	78%	20.0%	10-32%
Two-Stage CoT K=0 (v2)	66.0%	20.0%	10-32%
Two-Stage CoT K=5 (v2)	90.0%	60.0%	46-73%

각 방법에 실험을 수행한 결과(qwen/qwen3.5-9b 기준), Two-Stage CoT(K=5)는 In-Distribution 환경에서 90%, OOD환경에서 60%의 정확도를 기록하며 모든 비교 방법 중 가장 우수한 성능을 보인다. CoT K=0(20%)이 Zero-shot(26%)보다 낮은 결과는 2단계 구조가 예시 없이는 제약으로 작용함을 시사하며, K=5에서 성능이 급등(20%→60%)하는 것은 CoT 구조가 충분한 예시와 결합될 때 공간 어포던스 추론이 효과적으로 활성화됨을 보여준다. 구역 정보 기여도 분석 결과, 구역 정보를 제거할 경우 OOD정확도가 60%에서 8%로 크게 감소하여 구역 맥락 의도 해석의 핵심적인 신호임을 정량적으로 입증한다.

조건	OOD lenient
구역 정보 포함 (CoT K=5)	60.0%
구역 정보 제거 (제스처만)	8.0%

표 2. 구역 정보 유무에 따른 OOD lenient 정확도

5. 결 론

우리는 기존 WiFi CSI 기반 제스처 인식 시스템이 학습되지 않은 구역에서 의도 해석 성능이 급격히 저하되는 문제를 해결하기 위해, LLM의 상식 추론 능력을 결합한 ContextIntent 프레임워크와 공간 우선 2단계 추론(Two-Stage CoT) 기법을 제안한다. 제안된 방법은 단순히 제스처 패턴을 분류하는 방식에서 벗어나, 공간 맥락을 우선적으로 해석한 뒤 의도를 도출함으로써 환경 변화에 따른 일반화의 한계를 효과적으로 극복한다. 이는 공간 맥락 정보가 사용자의 제스처 의도를 정확히 해석하는 데 핵심적인 요소임을 입증한 결과이다. 향후 다양한 LLM 모델을 대상으로 일반화 성능을 추가 검증하는 한편, 실제 WiFi CSI 하드웨어 환경에서의 End-to-End 통합 실험을 통해 실생활 환경에 즉시 적용 가능한 지능형 인터랙션 기술로 발전시킬 계획이다.

참고 문헌

[1] H. Liu et al., "MetaFi: Device-Free Pose Estimation via Commodity WiFi for Metaverse Avatar Simulation," IEEE WF-IoT, 2022.
 [2] F. Yang et al., "DensePose From WiFi," arXiv:2301.00250, 2023.
 [3] T. Brown et al., "Language Models are Few-Shot Learners," NeurIPS, 2020.
 [4] J. Wei et al., "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," NeurIPS, 2022.