

AI 글라스: 소통공간 확장

AI Glass: Extending the Communication Space

송기봉 (K.B. Song, kbsong@etri.re.kr)	콘텐츠융합연구실 책임연구원
최영재 (Y.J. Choi, scholarchoi@etri.re.kr)	콘텐츠융합연구실 선임연구원
김지완 (J.W. Kim, jiwankim@etri.re.kr)	콘텐츠융합연구실 연구원
지형근 (H.-K. Jee, hkjee@etri.re.kr)	콘텐츠융합연구실 책임연구원/실장
정일권 (I.K. Jeong, jik@etri.re.kr)	콘텐츠연구본부 책임연구원/본부장
이정익 (J.I. Lee, jiklee@etri.re.kr)	초실감메타버스연구소 책임연구원/소장

ABSTRACT

Artificial intelligence (AI) Glass, in the form factor of traditional glasses, is expected to become a new user interface that surpasses the touch interface. In this paper, we introduce the technological trends of extended reality (XR) devices and analyze the reasons for the stagnation of the XR device market. Additionally, we examine key sub-technologies relevant to AI Glass development, including spatial computing, large language model-based generative AI, and AI agents. Finally, we briefly outline the technologies that require further development for AI glass in the future.

KEYWORDS AI 글라스, AI 에이전트, XR 디바이스, 거대언어모델, 공간컴퓨팅, 멀티모달 LLM

I. 서론

터치 인터페이스 기반의 스마트폰은 현재까지도 대다수의 사람이 이용하는 모바일 컴퓨팅의 의사소통 도구이다. 그러나, 많은 기술자는 포스트 스마트폰 시대의 주요한 디바이스는 XR 디바이스가 될 것이라 예측하고 있으며, 또한 글로벌 ICT 기업들은 관련 기술 및 시장을 독점적으로 장악하고 있는 실정이다.

그러나 이러한 XR 디바이스도 사용자 편의성과 터치 인터페이스의 대안기술 한계 돌파가 늦어짐으로 인하여 일반인이 대중적으로 사용하기에는 가격 대비 성능 및 기술의 낮은 만족도로 시장의 정체기가 계속되고 있는 실정이다.

최근 Apple VISION PRO에 의해 다시 한번 상극하게 된 공간컴퓨팅 기술과 OpenAI에 의해 촉발된 거대언어모델(LLM) 기반의 생성형 AI 기술은 'AI 글라스'라는, XR 디바이스의 시장 정체기를 탈피할

* DOI: <https://doi.org/10.22648/ETRI.2025.J.400206>

* 이 논문은 한국전자통신연구원 연구운영비지원사업(기본사업) 오감·감성체험 초실감 상호작용 기술연구개발과제(25ZC1200)의 논문입니다.



수 있는, 새로운 인터페이스 기술을 제시하고 있다.

공간컴퓨팅 기술은 공간에서 사용자의 위치와 움직임의 방향을 센서로 추적하고, 추적한 정보를 바탕으로 사용자가 바라보는 실제 세계에 생성한 콘텐츠를 증강하여 사용자와 디지털 정보 간의 상호작용을 보다 직관적이고 몰입감 있게 만드는 기술이다.

거대언어모델 기반의 생성형 AI 기술은 인코더와 디코더에서 제공되는 데이터인 텍스트에서 의미를 추출하고 텍스트 내의 단어와 단어 간의 관계, 구문 간의 관계를 이해하는 기술을 가지는, 방대한 양의 데이터로 사전 학습된 초대형 딥러닝 모델이다[1].

본고에서는 XR 디바이스 기술, 공간컴퓨팅 기술, 생성형 AI 기술 등에 대해 알아보고 공간컴퓨팅과 생성형 AI 기술이 AI 글라스의 발전에 미치는 기술 기능 및 구조에 대해 언급하고자 한다.

상세하게는 II장에서 XR 디바이스 소개, 발전 과정과 구성기술에 대해 언급하며, III장에서 AI 글라스의 태동 과정을 XR 디바이스가 시장 정체를 가지게 되는 원인과 공간컴퓨팅과 생성형 AI 태동기의 관점에서 분석해 본다. IV장에서는 AI 글라스 기업 동향 및 구성기술을 분석하며, V장에서는 AI 글라스 구성기술과 공간컴퓨팅과 생성형 AI 기술과의 상호기술적 관계성을 분석하며 결론을 맺는 순서로 소개한다.

II. XR 디바이스 기술동향

증강현실(AR), 가상현실(VR), 혼합현실(MR) 기술을 포함하고 있는 XR 디바이스 기술인 확장현실(XR)은 OCULUS의 VR HMD, 구글의 구글 글라스, MS의 홀로렌즈, 매직리프, 쉘컴 XR 스냅드래곤 프로세스 등 하드웨어뿐만 아니라 소프트웨어, 서비스 등에 대한 XR 산업의 생태계 조성을 위해 대다수

표 1 XR 디바이스 구성요소

구성요소	기술정의	기술목표
디스플레이	XR 콘텐츠를 표시하는 기술	실물과 같은 XR 콘텐츠 표시
인식 및 트래킹	사용자의 데이터를 실시간 추적하는 기술	증강영상 움직임이 사용자의 움직임과 실시간 매칭
영상합성 렌더링	XR 콘텐츠를 고해상도, 고화질로 생성시켜 실물과 정합하는 HW/SW 기술	가상영상이 실물위치에 실시간으로 정합
인터랙션 및 인터페이스	XR 콘텐츠를 지각, 인지, 조작, 입력할 수 있도록 돕는 기술	쉽고 편하고 정확하게 인터랙션 기능 제어
통신네트워크	XR 콘텐츠의 실시간 대용량 데이터 통신기술	지연 없는 실시간 데이터 스트리밍의 실감효과 극대화

의 글로벌 ICT 기업들이 앞다투어 기술을 개발하고 있는 차세대 컴퓨팅 플랫폼의 일종이다.

2019년 이후 최근의 경우 MS의 Hololens 2, VUZIX의 BLADES 2, 쉘컴의 XR2는 5G 및 Wi-Fi 7 등의 통신소자 발전과 함께 대용량 XR 콘텐츠 서비스를 가능하게 하고 있다.

최근 공간컴퓨팅 제안의 Apple VISION PRO와 안경형 폼 팩터에 근접한 메타 오라이언(Orion)은 새로운 인터페이스의 가능성을 보여주는 계기가 되고 있다.

XR 디바이스는 표 1과 같이 디스플레이, 인식 및 트래킹, 영상합성 렌더링, 인터랙션 및 인터페이스, 통신네트워크로 구성되어 있다[7].

XR 디바이스의 구성요소에서 알 수 있듯이 XR 디바이스는 현실과 가상의 결합이 지연 없이 표현하기 위한 ICT 전체 기술을 포함하고 있으며, 현재 개발되고 있는 기술로는 VR, AR, MR을 포함하고 있는 디바이스, 글라스로는 AI 글라스를 포함하고 있다.

본고에서는 전통적인 안경형 폼 팩터의 모습만을 글라스로 정의하고 있으며, 디바이스는 ICT 기술을 포함하는 기능들을 포함하는 시스템으로 정의하고

있다. 참고로 AR 디바이스, AI 글라스를 구분하는 방법은 IV장에서 구체적으로 언급하였다.

앞서 보여준 바와 같이 XR 디바이스의 경우 기술 구성의 복잡성으로 인한 사용자 편의성 부족으로, 소비자가 만족할 만한 수준의 기술개발은 아직도 미진한 실정이며 시장의 정체를 돌파하는 새로운 인터페이스 제시에 대한 논의가 이루어지고 있다.

III. XR 디바이스 대중화 해결 사항

XR 디바이스의 시장 정체와 AI 글라스의 시장 확산은 공간컴퓨팅과 생성형 AI 활용의 차이에 기인한다.

XR 디바이스는 공간컴퓨팅을 활용해 3D 공간 맵핑과 AR/VR/MR 콘텐츠를 제공했으나, 실질적이고 일상적인 활용 사례가 부족하다. 즉, 콘텐츠의 제한성과 정밀한 공간 데이터를 필요로 하는 환경에서의 제약이 소비자에게 실질적인 가치를 제공하지 못하고 있다.

또한, 소비자가 사용하기에는 기술적인 성능에 비해 기기의 가격 만족도가 높지 않아 대중화를 통한 시장 확산이 이루어지지 않는 실정이다.

특히, II장에서 언급한 XR 디바이스를 구성하고 있는 요소기술 중에서 마이크로 디스플레이 가격은 선뜻 소비자가 구입하여 사용하기에는 적정 수준의 가격이 아니기에 시장 정체의 주요한 요인 중의 하나로 판단된다. 따라서 기술의 수준을 만족하는 적정한 가격의 영상표현 디스플레이 소자는 필수적으로 해결해야 하는 문제가 될 것으로 생각된다.

또한, 오픈소스 기반 통상적인 딥러닝이나 머신러닝 기반의 AI 기술개발 및 특정 콘텐츠 저작의 도구에 함몰되어 있어 사용자 맞춤형 콘텐츠 생성이나 실시간 대화형 인터페이스 제공에서 한계를 드러내었다.

표 2 XR 디바이스와 AI 글라스의 시장 접근

비교요소	XR 디바이스	AI 글라스
공간 컴퓨팅	공간 데이터 활용에 제한적, 인프라 부족	공간 인식 및 실질적 정보 제공, 일상 통합
생성형 AI	활용 부족, 제한적 콘텐츠 제공	실시간 생성형 AI 통합, 맞춤형 콘텐츠 제공
소비자 효용성	특정 산업용에 초점	일상적인 문제 해결과 실질적 가치 제공
대중화 가능성	제한적, 고비용, 소수 활용	실용적, 직관적, 광범위한 수요 충족

반면에 AI 글라스는 공간컴퓨팅을 기반으로 실시간 정보 제공, AR 내비게이션, 개인화된 데이터 피드백 등 일상생활에 직접적인 가치를 제공한다. 예를 들어, 사용자가 보는 환경을 분석해 실시간 정보를 제공하거나 업무를 지원하는 방식으로 활용된다. 그리고 생성형 AI를 적극적으로 통합하여 실시간 콘텐츠 생성, 자연어 기반 대화형 인터페이스, 개인화된 학습 및 행동 추천 등을 제공함으로써 사용자 경험을 극대화하고 있다.

결론적으로 표 2와 같이 AI 글라스에서 공간컴퓨팅과 거대언어모델 기반의 생성형 AI를 실질적으로 결합하여 사용자에게 직접적이고 실용적인 가치를 제공하며 대중화에 성공하고 있는 반면, XR 디바이스는 주로 기술 중심적 접근에 머무르며 특정 분야에 제한된 사용 사례와 기술의 낮은 만족도로 대중화에 실패하고 있다[7].

따라서 현시점에서 공간컴퓨팅과 거대언어모델 기반의 생성형 AI를 실질적으로 결합하여 사용자에게 직접적이고 실용적인 가치를 제공하고 시장의 한계상황을 돌파할 것으로 기대되고 있는 AI 글라스 기술에 대해 상세하게 분석하고자 한다.

IV. AI 글라스 기업 동향 및 구성기술

앞서 언급한 바와 같이 최근 Apple VISION PRO의 공간컴퓨팅 기술과 레이벤 메타의 거대언어모델

표 3 AI 글라스 기술개발 현황

국가/기업/제품명	
주요 기능	비고
미국/메타/Orion	
<ul style="list-style-type: none"> • 눈, 손 추적, 제스처 EMG 감지밴드 • u_LED Waveguide • 생성형 AI/Meta AI-Llama3 	
미국/메타/Ray-Ban Meta	
<ul style="list-style-type: none"> • 카메라, 영상통화 • 오디오/라이브스트리밍 • 음성인식/음성명령/Meta AI-Llama3 	
미국/구글/(비공개)	
<ul style="list-style-type: none"> • AI 에이전트 프로젝트 '아스트라' • 안경형 폼 팩터(영상제공) 	
중국/Locktech/Locktech AI glasses	
<ul style="list-style-type: none"> • 13MP 카메라/2K 비디오 • 음성인식/음성명령/GPT-4o/Gemini 	
중국/-/XFlip glass	
<ul style="list-style-type: none"> • u_LED Waveguide(녹색) • OpenAI/AI Whisper 음성제어 	49g
중국/Zhuhai Mojie/MOJIE	
<ul style="list-style-type: none"> • CES 25/AR + AI 	49g
중국/슈퍼헝사/지에후안	
<ul style="list-style-type: none"> • 샤오미 지원 LLM 액세스 • AI오디오 안경, 11개국 번역 • 11시간 음악재생, 음성내비, AI채팅 	
중국/솔로/AirGo	
<ul style="list-style-type: none"> • 카메라 없음(탈착식 카메라) • OpenAI ChatGPT음성으로 액세스 • GPT-4o 멀티모달/10개국 번역 	
중국/라이웨이케/Meta Lens S3	
<ul style="list-style-type: none"> • 광각120도 4K 고화질 스포츠 카메라 • 4K u_LED • AI챗봇, 음성대화형 번역 기능 	
중국/이븐 리얼리티/G1	
<ul style="list-style-type: none"> • u_LED(단색-녹색) 	
중국/샤지/OptoX AI	
<ul style="list-style-type: none"> • 카메라, 스피커, 마이크 • GPT-4o 멀티모달 AI 액세스 • 최종 30분 영상만 저장 후 시청 	
중국/TCL RayNeo/RayNeo	
<ul style="list-style-type: none"> • SLAM 내비/제스처 인식 • u_LED Waveguide • 멀티모달 LLM GPT-4/음성인식 • RayNeo Ring 3DoF Beam • 실시간 다국어 번역 	

한국/시어스랩/AINoon	
<ul style="list-style-type: none"> • 멀티모달 LLM GPT 기반의 음성안내/AI Agent 기반 길안내 서비스 	
한국/ETRI/책읽기 AI 글라스	
<ul style="list-style-type: none"> • 구글 AI 및 최대 8개 카메라 탑재 • 멀티모달 LLM 기반 책읽기 서비스 • 실시간 다국어(6개국) 동시 번역 	
한국/엑스퍼트아이엔씨/게더X	
<ul style="list-style-type: none"> • ChatGPT 기반 음성인식 • CES 2025 전시 	

출처 참고자료 활용하여 저자 작성[2-7].

생성형 AI 기술은 XR 디바이스에서 'AI 글라스'라는 XR 디바이스의 시장 정체기를 탈피할 수 있는 새로운 인터페이스 기술을 만들어 가고 있다.

표 3은 국내·외 기업 혹은 기관에서 개발 중인 AI 글라스 현황을 나타내었다. 주로 미국 혹은 중국 기업 중심으로 기술개발이 이루어지고 있다. 미국 기업의 경우 메타, 구글, 쉐콤 등 글로벌 ICT 기업 중심으로 기술개발 중이며, 중국의 경우 벤처기업 중심으로 기술개발이 이루어지고 있다[2-7].

미국의 메타는 안경형 폼 팩터 AI 글라스인 Orion을 출시[2], 마이크로 LED 프로젝터로 70도 시야각, 13PPD를 구현하고 있으며 Llama 3 모델을 사용하여 Meta AI 서비스를 제공하고 있다.

미국의 메타는 AI 글라스인 Ray-Ban Meta 시리즈를 출시[3], 카메라, 마이크, 스피커로 구성된 인터페이스로 Meta AI를 적용하여 음성인식을 통한 질의응답 혹은 번역 기능을, Orion과 동일하게 Llama 3 모델을 사용하여 Meta AI 서비스를 제공하고 있다.

중국의 솔로(Solos)는 AirGoTM을 출시하였다[4]. OpenAI ChatGPT 음성으로 액세스하며 GPT-4o mini로 10개국에 대해 음성 번역 기능을 제공, 사용자 맞춤형 디자인을 제공하기 위해 안경렌즈 부분과 다리 템플 부분이 이탈착되는 구조를 제공하고

있다.

중국의 TCL RayNeo는 RayNeo 시리즈를 출시하였다[5]. AR 표현을 위해 u-LED waveguide를 장착하고 있으며 멀티모달 LLM GPT-4를 적용하여 음성 Whisper 및 음성 내비게이션 기능을 제공하고 있으며 얼굴 인식 기능과 함께 문장 번역 기능을 제공하고 있다. 번역에 제공되는 언어로는 영어, 중국어, 한국어 등 8개국의 통/번역 기능을 제공하고 있다.

한국의 시어스랩은 “AINoon”을 시연하였다[6]. 멀티모달 거대언어모델 기반 GPT를 탑재하고 있으며 특히 AI 에이전트 모델을 이용한 AI 글라스 출시도 예정하고 있다.

한국의 ETRI는 AR 디바이스와 함께 “책읽기” AI 글라스를 개발하여 전시하였다. 탑재된 카메라, 스피커, 마이크를 활용하여 멀티모달 LLM 기반의 책 읽기를 실시간 다국어(6개국) 동시번역 가능한 서비스 제공이 가능하다.

표 3에서 조사한 바와 같이 개발된 AI 글라스 기술의 구성요소기술들을 XR 디바이스의 일종인 AR 디바이스 구성요소기술들과 상호 비교 분석하면 다음과 같은 특징들이 존재한다.

먼저, AR 디바이스는 실세계와 가상세계의 구성과 상호연결을 통해 사용자에게 시점 기반 소통의 동적 환경을 제공하기 위한 기술개발이 주로 이루어지고 있으며, 주요한 기술로는 AR 디바이스 자체와 실·가상 객체와 인공지능, 엣지 및 시점 기반 네트워크 구조를 포함하고 있다[7].

반면, 앞서 표 3에서 보여준 AI 글라스 기술 동향에 따르면 AI 글라스는 터치 인터페이스를 완전히 탈피한 구조이며, 일반인이 손쉽게 착용이나 사용이 가능한 안경형 폼 팩터에 카메라(시각), 마이크(음성), 스피커(청각), 기타 센서 기반의 경량화 가능한 인터랙션 인터페이스 요소와 공간컴퓨팅 및 생성형 인공지능을 기반으로 서비스를 제공하는 구조를 포

함하고 있다.

특히, AI 글라스는 안경형 렌즈를 통한 증강현실(AR)을 표현할 때 디스플레이는 대다수의 기업이 5mm 이하 크기의 디스플레이가 가능한 마이크로 LED 광모듈을 채택하여 극소형화를 이루고 있다. 결국 AI 글라스는 극소형의 마이크로 LED와 함께 마이크로 LED 웨이브가이드 형태의 디스플레이 렌즈 구조를 적용하여 안경형 소형화의 기술적 진일보를 완성하고 있다.

증강현실 표현이 주요한 기능의 하나인 AR 디바이스 또한 이러한 극소형 마이크로 LED 웨이브가이드 구조를 적용하고 있으며, 다만 AR 디바이스에서는 더욱 실감 있는 콘텐츠를 구현하기 위해 더 큰 분해능의 디스플레이 디바이스 기술개발은 지속적으로 이루어질 것으로 예상된다.

그림 1은 지금까지 언급한 AR 디바이스와 AI 글라스의 차이점을 개념도로 표현한 것이다.

AR 디바이스와 AI 글라스의 차이점을 다시 한번 언급하면 다음과 같은 서로 다른 특징의 차이점이 존재한다.

즉, AR 디바이스는 실세계와 가상세계의 구성과 상호연결을 통해 사용자에게 시점 기반 소통의 동적 환경을 제공하기 위한 실·가상 객체와 인공지능, 엣지 및 시점 기반 네트워크 기술을 포함하고 있

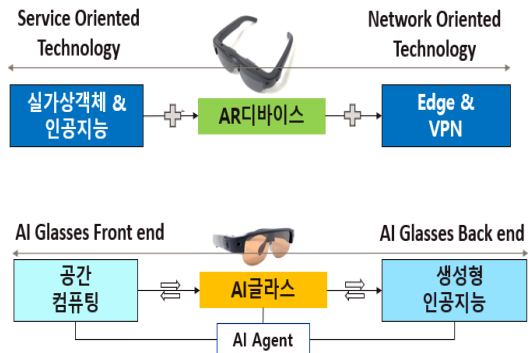


그림 1 AR 디바이스와 AI 글라스의 서로 다른 특징

표 4 AI 글라스와 AR 디바이스의 공통 기술

공통 기술	개념
인터랙션 트래킹	<ul style="list-style-type: none"> • 콘텐츠를 지각, 인지, 조작, 입력할 수 있도록 돕는 기술과 사용자의 생체 데이터 및 공간 데이터를 실시간 추적하는 기술 • 카메라, 스피커, 마이크 등의 물리 요소와 제스처, SLAM, 6DoF 등 멀티모달 트래킹 등의 요소를 포함
디스플레이	<ul style="list-style-type: none"> • AR 콘텐츠를 표시하는 기술 • 실물과 같은 수준의 AR 콘텐츠 표현을 목표 • 마이크로 디스플레이 및 렌즈 등 디스플레이 광학 요소를 포함
영상합성 렌더링	<ul style="list-style-type: none"> • 콘텐츠를 생성하고 생성된 콘텐츠를 실물과 정합하는 HW 및 SW기술
통신네트 워크	<ul style="list-style-type: none"> • 사용자 관련 데이터, 데이터 세트 및 AR 콘텐츠 등을 실시간으로 저지연으로 통신하는 통신인프라 포함의 Connectivity 기술

는 반면, AI 글라스는 터치 인터페이스에서 완전히 탈피한 안경형 폼 팩터에 카메라(시각), 마이크(음성), 스피커(청각), 기타 센서 기반의 경량화 가능한 인터랙션 인터페이스 요소와 공간컴퓨팅 및 생성형 인공지능을 기반으로 서비스를 제공하는 간편한 구조의 명확한 특징의 차이가 있다.

다만, AR 디바이스와 AI 글라스는 표 4와 같은 하드웨어의 기능적 공통점은 포함하고 있다.

AI 글라스의 또 다른 특징은 생성형 인공지능과 공간컴퓨팅 기술을 기반으로 서비스를 제공하고 있다.

생성형 AI 기반의 서비스로는 카메라, 마이크, 스피커 기반의 음성인식, 음성생성, 문장생성 등을 거대언어모델 멀티모달 ChatGPT, Llama, Whisper 등을 적용하여 서비스를 제공하는 기술을 개발 중이다.

AI 글라스에서 적용되고 있는 또 다른 기술은 공간컴퓨팅 관련 기술이다. 공간컴퓨팅 기술의 적용은 아직은 초기 단계이지만 직접적으로 적용하고자 하는 기술로는 AI 에이전트(Agent)[11] 기술로 판단된다.

AI 글라스에 탑재된 AI 에이전트는 카메라가 공간의 시각 정보를 분석하고 공간의 이벤트 상황을 인식하여 자동으로 필요한 AI들을 찾아내어 실시간

정보를 안내하는, 서비스에 제공하는 에이전트의 기능을 구현하고 있다.

즉, AI 글라스에 탑재된 AI 에이전트는 물리적이거나 소프트웨어적인 인터페이스를 통하여 공간을 감지하고 자신이 인지한 내용과 데이터를 기반으로 합리적인 결정을 내려 최적의 성과와 결과를 도출하는 데 편리하게 사용할 수 있는 기술이다. 이는 멀티모달 거대언어모델과 함께 가격 대비 효율성을 제공하는, 터치 인터페이스를 탈피하는 새로운 인터페이스가 될 것으로 판단된다.

구체적인 가능성의 예로써 미국 구글에서 최근 발표한 AI 에이전트의 초기 기술인 “프로젝트 아스트라”에 따르면 구글 픽셀 폰에 탑재된 AI 에이전트를 통하여, 카메라를 통해 공간에 존재하는 사물을 파악하고 해당되는 사물에 대해 설명해 주고 주변 경관을 비추면 위치를 파악해 주는 모습을 보여주고 있다.

또한, 아직은 AI 에이전트가 탑재된 안경형 폼 팩터의 모습을 정확하게 보여주고 있지는 않지만, 영상 속에서 AI 글라스(예전의 구글 글라스와는 다른 형태의 안경형 글라스에 더 접근한 것으로 판단)에 탑재된 AI 에이전트를 소개하고 있다.

또 다른 예로써 한국의 시어스랩은 “AINoon” AI 글라스에 AI 에이전트를 탑재하고 있다. 알려진 예로 AI 글라스의 카메라에서 양파, 당근, 파 등의 재료를 보여주면 해당되는 식재료를 이용한 요리법 등을 보여주는 것으로 알려져 있다[6].

V장은 AI 글라스 관점에서 멀티모달 거대언어모델 기반 생성형 인공지능 기술과 AI 글라스 탑재가 가능한 공간컴퓨팅 적용 기술을 개념적으로 대분류한 결과를 소개한다.

이러한 분류의 목적은 AI 글라스에서 제시하는 초기의 특징 기술들을 파악하고 재분류하여 현시점 이후 향후 AI 글라스가 기술적으로 추구하여야 하

는 방향성을 제시할 수 있는 역할을 할 것으로 판단된다.

V. AI 글라스와 공간컴퓨팅, 생성형AI, AI 에이전트 기술의 연관성

그림 2에서는 그림 1에서 개념적으로 보여준 AI 글라스의 전후방 기술인 공간컴퓨팅과 생성형 인공지능 기술 및 AI 글라스 요소기술에 대해서 각 기술이 가지는 핵심 요소기술을 나열하고 있다.

AI 글라스에 탑재하고 있거나 탑재될 것으로 예측되는 AI 글라스 관점의 멀티모달 거대언어모델[8]의 생성형 인공지능을 기술 분류하면 표 5와 같이 분류할 수 있다.

AI 글라스 관점의 공간컴퓨팅 기술은 현재 AI 에이전트에서 일부 기술을 적용하여 AI 글라스에 탑재하여 일부 기술을 적용하고 있다.

표 6에서 제시하고 있는 공간컴퓨팅 기술은 XR 디바이스 관점의 공간컴퓨팅 기술은 고려하지 않고 있으며, AI 글라스 탑재가 예상되는 공간컴퓨팅 기술 분야로만 한정하여 분류하였다[12].

AI 에이전트는 기능성과 복잡성에 따라 단순반사

표 5 AI 글라스 관점의 멀티모달 거대언어모델 기반 생성형 인공지능 기술

대분류	기술 분야
지각능력	<ul style="list-style-type: none"> AI 글라스의 인터랙션 트래킹 기술 요소인 카메라, 스피커, 마이크, 기타 센서와 관련된 입력 데이터 요소 문자인식, 음성인식, 이미지인식, 뇌파인식 등이 있음
학습능력	<ul style="list-style-type: none"> 지각능력에서 입력되는 다양한 형태의 데이터 세트를 학습하는 기능 학습에 활용되는 데이터 세트는 텍스트, 이미지, 비디오, 오디오 등이 있음 문자학습, 문장학습, 음성학습, 영상학습 등이 있음
생성능력	<ul style="list-style-type: none"> 지각능력에서 입력되는 다양한 형태의 데이터 세트를 학습하여 새로운 데이터를 생성하는 능력 문자생성, 문장생성, 음성생성, 영상생성, 생각생성 등이 있음
행동능력	<ul style="list-style-type: none"> 학습한 대량의 데이터 세트에서 데이터 세트 안의 구조분석, 패턴, 스타일 등으로 이해하며 의사결정을 예측하고 학습 및 추론된 결과를 결정하고 설명으로 행동하는 능력 목표결정, 정보획득, 작업구현 등의 실행 기능들은 AI 에이전트에서 연관되어 기능을 제공

반응, 모델기반 반응, 목표기반, 유틸리티 기반, 학습기반 등으로 분류하고 있으나 아직은 기술 수준이 AI Stage 2를 만족스럽게 달성하지 못하고 있는 실정이다.

앞서 언급한 바와 같이 AI 글라스의 구성요소기술인 인터랙션-트래킹, 디스플레이, 영상표현 렌더링, 통신네트워크 기술과 관련이 있을 것으로 판단

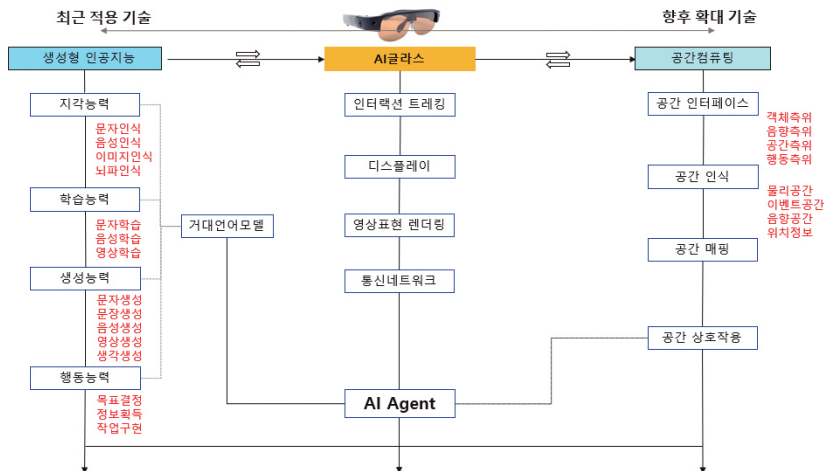


그림 2 AI 글라스 전후방 기술의 적용 기술 분류

표 6 AI 글라스 탑재 공간컴퓨팅 적용기술 분야

대분류	기술 분야
공간 인터페이스	<ul style="list-style-type: none"> • AI 글라스 착용자의 위치에 따라 다르게 보이는 공간의 객체를 착용자의 움직임을 감지하고 직접 만지거나 변형이 가능하게 하는 기술 요소로 공간과 AI 글라스 경계면에서 처리하는 기술 요소 • 객체측위, 공간측위, 행동측위, 음향측위 등이 있음
공간인식	<ul style="list-style-type: none"> • AI 글라스 착용자의 위치에 따라 다르게 나타나는 공간의 객체(오감객체)에 대해 위치, 모양, 크기, 움직임 등을 이해하는 기술 • AI 글라스에서는 카메라, 스피커 마이크 등의 시청각 공간의 중요성으로 오감 관련 공간인식 기술이 중요할 것으로 판단 • 물리공간, 이벤트공간, 음향공간, 위치정보 등이 있음
공간매핑	<ul style="list-style-type: none"> • AI 글라스 착용자의 위치에 따라 다르게 보이는 공간의 객체를 위치, 모양, 크기, 움직임 등을 공간의 위치에 정하고 위치정보와 함께 특징 정보를 매핑하고 추적하는 기술 • 실제 객체의 공간위치와 가상공간의 매핑하는 기술 등이 있음
공간 상호작용	<ul style="list-style-type: none"> • 공감매핑에 의해 생성된 공간 디지털 정보와 AI 글라스의 인터랙션 트래킹 기술에 의해 콘텐츠와 상호작용하는 기술 • AI 글라스의 경우 오감 상황인식 기반의 공간상호작용 필요성으로 AI 에이전트 기술 등이 있음

되는 거대언어모델 기반 생성형 인공지능기술과 공간컴퓨팅 기술 간 관계를 그림 2에서 제시하였다.

AI 글라스-생성형 AI-AI 에이전트-공간컴퓨팅으로 연결되는 기술 간 관계는 현재까지도 개발되고 있는 기술임에 따라 간단하게 결론에서 언급하고자 한다.

VI. 결론

지금까지 공간컴퓨팅과 거대언어모델 기반의 생성형 AI 및 AI 에이전트를 실질적으로 결합하여 사용자에게 직접적이고 실용적인 가치를 제공하며 대중화에 성공할 것으로 판단되는 새로운 인터페이스인 AI 글라스 기술과 전후방 기술의 특징들에 대해 분석하였다.

다만, 아직은 AI 글라스 기술이 스마트폰과 같은 누구나 하나쯤은 보유하여 사용할 수 있는 모바일

컴퓨팅 기기가 되기 위해서는 다음과 같은 예시의 기술들을 개발할 필요성을 언급하며 본고를 끝맺음하고자 한다. 이 장에서 제안하는 기술들은 AI 글라스에 탑재하여도 사용자가 부담스럽게 생각하는 가격 상승 요인이 되지 않는 것으로 판단된다.

기술적 해결 사항으로는 첫 번째로 AI 글라스 하드웨어 요소기술 중 디스플레이 개발 및 하드웨어 신호 처리 인터페이스 요소이다. 사용자 착용이 편리한 안경형 폼 팩터를 구현하기 위해서는 극소형 디스플레이 즉, 디스플레이 렌즈를 포함하여 예상으로는 0.1인치 이하의 크기가 필수적이며(물론 저가형 디스플레이) ① 극소형 고화질 디스플레이 (~30PPD 이상) ② 극소형 디스플레이(10만 화소 내외) 등이 필요할 것으로 판단된다. 또한, 신호처리 인터페이스 요소로는 MIPI, SPI, QSPI, I2C, USB 등과 같이 각각의 특징들이 존재하고, 국제표준 활동과 밀접한 관련이 있지만 AI 글라스 구조에 합당한 신호처리 인터페이스 요소에 대한 논의가 필요할 것으로 판단된다.

두 번째로 AI 글라스 탑재 공간컴퓨팅 관련 기술에서는, 현재 AI 글라스에서는 사용자가 센싱하는 공간정보는 카메라 기반의 시각 정보가 주를 이루고 있어 이를 오감 기반 상황인식의 이벤트 공간까지 확대할 필요가 있을 것으로 판단된다. 하나의 예로, 공간음향 기반의 AI 글라스 기술은 가능할 것으로 판단된다.

세 번째는 멀티모달 거대언어모델의 확장이다. 각종 센서를 입력으로 하는 멀티모달[9]과 함께 뇌파 기반의 멀티모달 거대언어모델에 대한 연구는 궁극적으로 최고의 난이도를 가지는 Neuro-GPT[10]의 방법이 될 것으로 예측된다. 현재는 멀티모달 거대언어모델에서 문장생성, 영상생성(특히 LVM) 등에 대한 기술개발이 활발하게 이루어지고 있으나 장기적으로 인간의 언어체계와는 다른 신호 구조를

가지고 있는 뇌파에 대한 기본 모델 제시와 함께 고난이도인 뇌파 기반의 멀티모달 거대언어모델에 대한 연구개발이 필요할 것으로 판단된다.

이를 ‘뇌파 AI 글라스’에 탑재하면 기존의 RNN으로 처리하던 방법을 혁신할 수 있는 뇌파 해석에 대한 새로운 방법의 제안이 될 수도 있을 것으로 판단된다.

마지막으로, AI 글라스 사용자 근처에서 서비스가 가능한 AI 에이전트 확장의 구조 개발이다. 즉, 앞서 언급한 바와 같이 AI 글라스에 탑재된 AI 에이전트는 궁극적으로 공간을 감지하고 공간의 이벤트 상황을 인식하여 자신이 인지한 내용이나 데이터를 기반으로 자동으로 필요한 AI들을 찾아내어 합리적인 결과를 도출하는 서비스를 제공하는 기술을 개발하고자 하고 있다.

따라서 이러한 기술은 소형언어모델(SLM) 등으로 시도가 가능하겠지만, AR 착용자인 사용자 근처에서 AI 에이전트 기능이 구현되는 에지(혹은 모바일)와 같은 네트워크 구조에 ‘에지 AI 에이전트 모델 및 아키텍처’가 좀 더 심도 있게 개발되어야 할 것으로 판단된다.

결국, 공간에 산재해 있는 대규모 특정 앱, 웹 등이나 데이터 세트를 학습하고 새로운 데이터 및 콘텐츠를 생성하기에 거대언어모델 생성형 인공지능(특히 생성능력 분야), AI 콘텐츠 에이전트 등에 대한 관점의 기술개발은 사용자가 더욱 편리하고 안전하게 사용하는 도구가 될 것으로 판단된다.

지금까지 공간컴퓨팅과 거대언어모델 기반의 생성형 AI를 실질적으로 결합하여 사용자에게 직접적이고 실용적인 가치를 제공하여 시장의 한계상황을 돌파할 것으로 기대되고 있는 AI 글라스 기술에 대해 알아보았다.

결론적으로, AI 글라스는 공간컴퓨팅과 거대언어모델 기반의 생성형 AI, AI 에이전트를 실질적으로

결합 구현하면, 사용자에게 직접적이고 실용적인 가치를 제공하며 대중화에 성공하는 새로운 인터페이스가 될 것으로 기대한다.

약어 정리

GPT	Generative Pre-trained Transformer
LVM	Large Vision Model
MIPI	Mobile Industry Processor Interface
PPD	Pixel Per Degree
RNN	Recurrent Neural Network
SLM	Small Language Model
SPI	Serial Peripheral Interface
VPN	ViewPoint Network

참고문헌

- [1] <https://aws.amazon.com>
- [2] <https://about.meta.com/ko/realitylabs/orion/>
- [3] <https://www.ray-ban.com>
- [4] <https://solosglasses.com>
- [5] <https://www.rayneo.com>
- [6] 김동원, “시어스랩, AI 에이전트가 활동할 ‘스마트글라스’ 개발,” 더에이아이, 2024. 12. 12. <https://www.newstheai.com/news/articleView.html?idxno=6767>
- [7] 송기봉 외, “Trusted Reality(신뢰현실) 기술, 포스트-스마트폰 관점에서,” 전자통신동향분석, 제33권 제6호, 2018. pp. 154-164. doi: 10.22648/ETRI.2018.J.330615
- [8] H. Chen et al., “Multi-Modal Generative AI: Multi-modal LLM, Diffusion and Beyond,” arXiv preprint, 2024. doi: 10.48550/arXiv.2409.14993
- [9] S. Wu et al., “NExT-GPT: Any-to-Any Multimodal LLM”, in Proc. Int. Conf. Mach. Learn., (Vienna, Austria), Jul. 2024.
- [10] W. Cui et al., “Neuro-GPT: Developing A Foundation Model for EEG,” arXiv preprint, 2023. doi: 10.48550/arXiv.2311.03764
- [11] Z. Deng et al., “AI Agents under Threat: A Survey of Key Security Challenges and Future Pathways,” arXiv preprint, 2024. doi: 10.48550/arXiv.2406.02630
- [12] 이승환, “공간컴퓨팅과 AI시대, 플랫폼 진화와 정책이슈 전망,” 국회미래예제 23-10호, 국회미래연구원, 2023. 12. 31. <https://www.nafi.re.kr/new/research.do?mode=view&articleNo=6910&article.offset=20&articleLimit=10>