

휴머노이드 로봇과 교감/공감 대화를 위한 Affective AI 기술 동향

김상훈 한국전자통신연구원 체화복합지능연구실 책임연구원

1. 머리말

최근 AI는 단순한 인지·판단 중심 기술을 넘어, 인간과 실제 물리 환경에서 상호작용하는 피지컬 AI(Physical AI) 영역으로 빠르게 확장되고 있다. 이에 자율주행 시스템, 협동 로봇, 서비스 로봇, 의료·재활 로봇 등 다양한 형태의 피지컬 AI가 개발·보급되고 있으나, 인간의 정서 상태와 심리적 반응, 사회적 맥락을 충분히 반영하지 못한다는 한계가 여전히 존재한다. 특히 피지컬 AI는 인간과 동일한 공간에서 직접 물리적 행동을 수행하기 때문에, 기능적 안전성뿐 아니라 정서적 안정감, 신뢰성, 사용자 수용성이 기술 확산의 핵심 요소로 작용한다. 따라서 사용자가 느끼는 불안감이나 긴장감, 위협 인식은 실제 사고 발생 여부와 관계없이 피지컬 AI의 활용을 제한하는 주요 요인으로 지적된다.

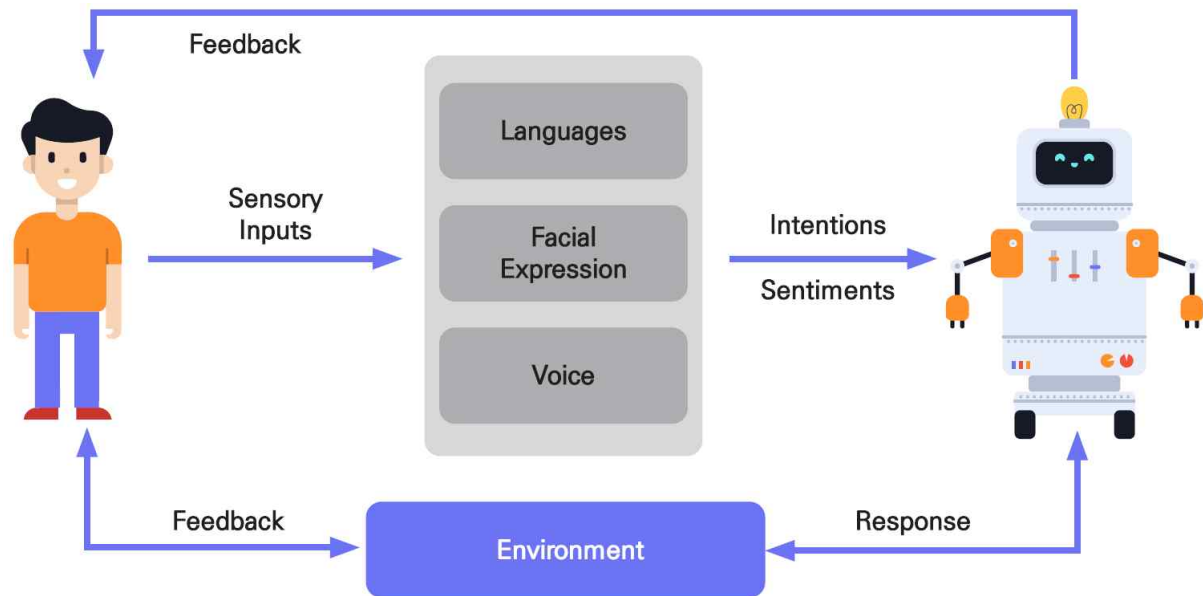
이러한 문제를 해결하기 위한 핵심 기술로 Affective AI가 주목받고 있다[1]. Affective AI는 인간의 감정을 인식·해석하고, 상황에 적절하게 반응하거나 감정을 표현할 수 있도록 설계된 AI 기술을 의미한다. 최근에는 피지컬 AI를 좀 더 인간 중심적으로 동작하게 만드는 핵심 요소로 부상하고 있으나, 현재까지의 연구는 주로 사용자 경험 개선이나 HMI(인간-기계 인터페이스, Human-Machine Interface) 수준에 머물러 있으며, 피지컬 AI의 제어 및 의사결정 계층과 본질적으로 통합된 연구는 아직 제한적인 실정이다.

한편, Affective AI의 산업적 활용 범위도 빠르게 확대되고 있다. 헬스케어 기반 AI 상담 서비스, 차량 내 HMI, 감성형 서비스 로봇 등 다양한 분야에서 적용 사례가 증가하고 있다. 국내의 경우 로봇·제조·모빌리티 분야 피지컬 AI 하드웨어 및 시스템 기술 경쟁력은 세계적 수준에 근접해 있으나, 인간 정서와 상호작용하는 지능화 기술 측면에서는 글로벌 선도국과 여전히 격차가 존재한다. 이에 따라 피지컬 AI와 Affective AI를 융합해 인간의 상태를 이해하고, 상황에 적응적으로 반응할 수 있는 인간 공존형 AI 시스템을 구현하기 위한 연구의 필요성이 더욱 커지고 있다.

이번 원고 2장에서는 Affective AI의 개념과 필요성을 설명하고, 3장에서는 관련 국내외 기술 동향을 살펴본다. 마지막으로 4장에서는 주요 응용 분야와 향후 활용 가능성을 논의하고자 한다.

2. Affective AI 기술 개념 및 필요성

Affective AI는 인간의 감정(Emotion), 기분(Mood), 정서 상태(Affective State)를 다양한 데이터 기반으로 인식하고, 이를 계산적 모델(Computational Model)로 해석·표현하며, 상황과 목적에 적합한 반응을 생성하는 AI 기술을 의미한다[그림 1].



출처: Chenxi Li, AIVA: An AI-based Virtual Companion for Emotion-aware Interaction, 2025

[그림 1] Affective AI 시스템 구성

즉, Affective AI는 단순한 감정 분류 기술을 넘어, 감정이 발생한 원인과 사회적·상황적 맥락을 이해하고 시간에 따른 정서 변화를 지속적으로 추적함으로써, 궁극적으로 인간과 유사한 수준의 정서적 상호작용 능력을 구현하는 것을 목표로 한다. 초기 연구는 얼굴 표정이나 음성 억양과 같은 단일 모달 기반의 기본 감정 분류에 집중했으나, 최근에는 멀티모달 데이터 융합, 대규모 학습 모델, 개인화 기반 정서 모델링 등으로 연구 범위가 확장되고 있다[2].

최근 AI 기술은 인식, 추론, 생성 분야 전반에서 급속한 발전을 이루고 있으며, 특히 LLM(거대언어모델, Large Language Model)과 멀티모달 AI의 등장은 인간과 기계 간 상호작용 방식에 근본적인 변화를 가져오고 있다. 그러나 기존 AI 시스템은 주로 논리적 판단과 정보 처리 능력 중심으로 발전해 왔기 때문에, 인간의 감정, 기분, 정서 상태와 같은 비정형적이고 맥락 의존적인 요소를 충분히 반영하지 못한다는 한계를 지니고 있었다.

반면 최근에는 음성·영상·텍스트·생체신호 등을 통합적으로 분석하는 멀티모달 AI 기술과 LLM의 발전을 통해, 기계가 인간의 감정을 좀 더 정교하게 이해하고 이에 적절히 대응할 수 있는 가능성이 확대되고 있다. 이에 따라 Affective AI는 단순한 기술적 흐름을 넘어 의료, 교육, 돌봄, 국방, 공공 서비스, 메타버스 등 사회 전반에 구조적 변화를 유발할 핵심 기술로 부상하고 있다. 따라서 Affective AI를 차세대 전략 기술로 정의하고, 중장기적 관점에서 체계적인 연구개발 방향과 기술 확보 전략을 수립할 필요성이 점차 커지고 있다.

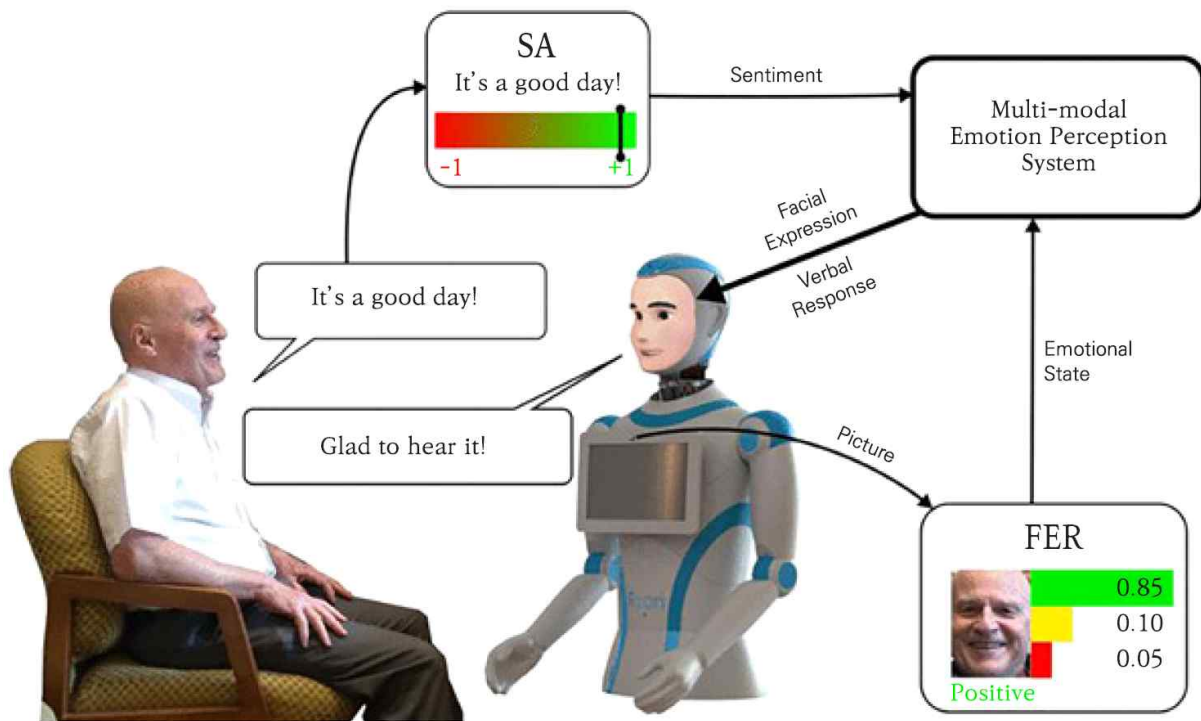
3. Affective AI 기술 동향

Affective AI 시스템은 인간의 정서 상태를 다층적·단계적으로 처리하는 구조를 가지며, 일반적으로 입력(Input)-해석(Interpretation)-추론(Reasoning)-반응(Response)의 흐름으로 구성된다.

최근 글로벌 연구 동향을 살펴보면, Affective AI는 단일 모듈 중심 구조에서 벗어나 멀티모달 인지, 정서 추론, 행동 생성 기능이 통합된 형태의 아키텍처로 진화하고 있다. 특히 최근 연구의 가

장 두드러진 특징은 멀티모달 감정 인식 기술의 고도화다. 기존에는 단일 입력 채널에 의존해 감정을 인식하는 방식이 주를 이루었으나, 최근에는 영상, 음성, 텍스트, 생체신호 등을 동시에 분석함으로써 감정 인식의 정확성과 안정성을 향상시키기 위한 연구가 활발히 수행되고 있다[3]. 예를 들어 얼굴 표정 인식, 음성 톤 분석, 발화 내용의 정서적 의미 분석을 결합해, 외현적으로 표현되는 감정과 실제 내재된 정서 상태 간의 불일치를 보완하려는 시도가 증가하고 있다.

또한 LLM을 중심으로 한 정서 이해 기술의 발전은 Affective AI 연구 패러다임에도 중요한 변화를 가져오고 있다. [그림 2]와 같이 LLM은 감정 표현이 포함된 언어적 맥락과 담화 구조를 이해하는 데 강점을 가지며, 단기적인 감정 인식을 넘어 장기 대화 과정에서의 감정 변화 추적과 감정 발생 원인에 대한 추론까지 가능하게 한다. 이는 Affective AI가 단순한 센서 기반 감정인식 기술을 넘어, 인간의 인지와 정서를 함께 이해하는 인지 기반 정서모델(Cognitive-Affective Model)로 발전하고 있음을 보여준다.



출처: Hojjat Abdollahi et al, Artificial Emotional Intelligence in Socially Assistive Robots for Older Adults: A Pilot Study, IEEE Transactions on Affective Computing, 2022

[그림 2] Affective AI 기반 인간-로봇 간 교감 대화 상황

최근에는 사용자의 현재 감정 상태를 사후적으로 인식하는 수준을 넘어, 감정 변화의 패턴과 경로를 학습해 미래의 정서 상태를 예측하는 연구도 활발히 진행되고 있다. 이러한 감정 예측 기술은 스트레스, 우울, 번아웃과 같은 정서적 위험 상태를 조기에 탐지하고 선제적으로 개입할 수 있다는 점에서 의료·복지 분야를 중심으로 높은 관심을 받고 있다.

3.1 국내 동향

ETRI(한국전자통신연구원, Electronics and Telecommunications Research Institute)는 최근 AI가

인간과 자연스럽게 대화하고 감정을 이해하며 정서적으로 교감할 수 있도록 하는 공감 상호작용 AI(empathic Interaction AI)를 핵심 연구 분야로 설정하고, 음성, 언어, 시선, 표정, 제스처 등 다양한 신호를 통합적으로 처리하는 멀티모달 상호작용 기술 개발을 본격적으로 추진하고 있다. 특히 소음 환경에서도 안정적인 음성 인식이 가능하도록 고도화하는 한편, 단순한 발화 내용 분석을 넘어 억양, 발화 맥락, 대화 흐름 등을 종합적으로 고려해 사용자의 감정 상태를 추론하고 공감형 응답을 생성하는 기술에 중점을 두고 있다. 이러한 기술은 인간과의 대화를 좀 더 자연스럽게 인간친화적으로 구현하기 위한 기반 기술로 활용되고 있다.

또한 공감 상호작용 AI를 실제 물리적 객체와 결합하기 위해 휴머노이드 로봇 및 교감형 AI 에이전트 연구도 병행하고 있다. 대화 이해와 감정 추론 결과를 로봇의 표정, 몸짓, 행동 제어와 연계함으로써, 인간과 정서적으로 상호작용할 수 있는 기술 구현을 목표로 하고 있다. 이를 통해 돌봄, 교육, 안내 등 다양한 사회 서비스 분야에서의 활용 가능성을 확대하고 있으며, 인간의 행동과 감정을 좀 더 정밀하게 이해하기 위한 대규모 멀티모달 데이터 구축과 인간 이해 AI 연구도 지속적으로 추진 중이다. 특히 라이프로그 및 감정 반응 데이터를 기반으로 공감형 대화 모델의 학습과 성능 고도화를 진행하고 있으며, 연구자와 산업계의 활용을 지원하기 위해 관련 데이터와 기술의 개방도 확대하고 있다.

KT는 자사의 AICC(AI Contact Center) 솔루션을 기반으로 상담 음성을 자동 분석하고, 상담 내용과 함께 고객의 정서적 반응에 대한 인사이트를 제공하는 통화 분석 서비스를 상용화하고 있다. 해당 서비스는 음성 인식 기술과 음성 신호 분석 기술을 결합해 상담과정에서 나타나는 고객의 감정 변화 패턴을 분석하고, 불만이나 분노 가능성이 높은 통화를 식별하는 데 활용된다. 이를 통해 기업은 민원 대응 품질을 향상시키고, 문제 가능성이 높은 통화를 사후 분석 대상으로 선별할 수 있다. 특히 이 사례는 감정 인식 기술이 독립적인 기능이 아니라, 콜 QA(품질 관리, Quality Assurance) 및 VOC(고객의 소리, Voice of Customer) 분석 기능과 통합된 형태로 상용화됐다는 점에서 의미가 있다.

국내 중소기업인 마음AI(maum.ai)는 FAST AICC를 중심으로 음성봇, 챗봇, 통화 분석, VOC 분석 기능을 포함한 통합형 AICC 솔루션을 상용 제공하고 있다. 금융, 보험, 제조 분야 기업을 대상으로 실제 구축 사례를 확보하고 있으며, 상담 내용과 고객 반응을 분석해 상담 품질 평가 및 클레임 탐지에 활용하고 있다. 특히 VOC 분석 과정에서 고객의 강한 불만 표현, 반복 민원, 부정적 정서가 포함된 통화를 자동 분류함으로써, 감정 신호를 운영 지표로 활용하는 체계를 구축하고 있다. 이는 Affective AI가 단순한 감정 인식 수준을 넘어, 운영 효율성 향상과 리스크 관리 지원을 위한 실질적 상용 기술로 적용된 대표적인 국내 사례로 평가된다.

3.2 국외 동향

해외에서는 공감 능력을 갖춘 대화형 AI 기술 개발이 심리 지원, 사회적 상호작용, 교육 등 다양한 분야에서 활발히 이뤄지고 있다.

대표적으로 와이사(Wysa)와 같은 AI 기반 정신건강 지원 챗봇은 CBT(인지행동치료, Cognitive Behavioral Therapy) 기법을 적용해 사용자의 감정 상태를 분석하고 정서적 지지를 제공함으로써, 실제 임상적 효능과 높은 사용자 참여율을 보고한 바 있다. 이러한 시스템은 공감적 언어 표현과

대화 전략을 활용해 이용자와의 치료적 연대감(Therapeutic Alliance)을 형성하며, 정신건강 분야에서 AI가 실질적인 지원 도구로 활용될 수 있음을 보여준다[4].

사회적 상호작용 분야에서는 스웨덴의 퍼햇 로보틱스(Furhat Robotics)가 개발한 퍼햇(Furhat) 로봇이 대표적인 사례로 주목받고 있다. 퍼햇은 음성 기반 대화 기능뿐 아니라 얼굴 표정, 시선 처리, 고개 움직임 등 다양한 비언어적 행동을 결합해 좀 더 자연스럽게 인간친화적인 상호작용을 구현한다. 해당 로봇은 교육, 의료, 공공 서비스 등 다양한 사회적 환경에서 활용되고 있으며, 인간-로봇 상호작용 연구를 위한 주요 플랫폼으로도 활용되고 있다[5].

한편, 미국 베린트(Verint)의 코기토(Cogito)는 상담 과정에서 고객과 상담원의 음성을 실시간으로 분석해 감정 상태를 추정하고, 상담원에게 말의 속도, 음성 톤, 공감 표현 사용 여부 등에 대한 코칭 신호를 제공하는 서비스를 상용화했다. 이 시스템은 감정 인식 결과를 기반으로 상담원의 행동을 실시간으로 유도하는 구조를 가지며, Affective AI가 단순한 감정 분석을 넘어 업무 수행 가이드 및 행동 개입 도구로 활용되는 사례로 평가된다[6].

또 다른 사례로 영국 옥스퍼드대(University of Oxford) 창업업체인 리얼아이즈(Realeyes)는 웹캠 기반 얼굴 분석 기술을 활용해 광고 시청자의 감정 반응과 주의(Attention) 수준을 측정하는 상용 플랫폼을 제공하고 있다. 기업은 광고 콘텐츠를 사전에 테스트하고, 시청자의 감정 반응 데이터를 기반으로 광고의 효과를 예측·개선할 수 있다. 이는 Affective AI가 개인 맞춤형 서비스에 국한되지 않고, 집단 수준의 감정 데이터 분석을 기반으로 의사결정을 지원하는 기술로 활용되고 있음을 보여주는 대표적 사례라 할 수 있다[7].

4. 맺음말

최근 AI는 단순히 화면 기반 환경에서 질문에 응답하는 수준을 넘어, 로봇과 자율주행차처럼 실제 물리 공간에서 이동하고 행동하는 피지컬 AI 기술로 빠르게 발전하고 있다. 그러나 현재의 피지컬 AI는 인간과 동일한 공간에서 상호작용함에도 불구하고, 사용자가 느끼는 감정 상태나 심리적 반응을 충분히 고려하지 못하는 한계를 지닌다. 예를 들어 기술적으로는 안전성이 확보된 시스템이라 하더라도, 사용자는 '로봇이 너무 가까이 다가와 불편하다'거나 '자율주행차의 움직임이 불안하다'고 느낄 수 있다. 이러한 정서적 불안과 위협 인식은 실제 사고 발생 여부와 관계없이 피지컬 AI에 대한 거부감과 불신으로 이어질 가능성이 크다.

이와 같은 문제를 해결하기 위한 핵심 기술로 Affective AI가 주목받고 있다. Affective AI는 얼굴 표정, 음성, 언어 표현, 행동 패턴 등 다양한 신호를 분석해 인간의 감정 상태를 추론하는 기술이다. 지금까지는 주로 상담 챗봇이나 고객 서비스와 같은 비물리적 서비스 환경에서 활용돼 왔으나, 향후에는 피지컬 AI의 '눈과 귀' 역할을 수행하며 인간의 상태를 이해하고 이에 맞추어 행동을 조절하는 핵심 기술로 발전할 것으로 예상된다. 특히 로봇과 자율 시스템이 인간의 생활 공간으로 더욱 깊숙이 진입하는 상황에서, 인간의 감정과 심리적 맥락을 이해하지 못하는 피지컬 AI는 사회적 수용성을 확보하기 어려울 가능성이 높다.

우리나라는 로봇, 제조, 자율 시스템 분야에서 높은 수준의 하드웨어 및 시스템 기술 경쟁력을 보유하고 있는 만큼, 여기에 Affective AI를 융합한 인간 중심 피지컬 AI 기술을 선도적으로 연구·개발할 필요가 있다. 이는 단순한 기술 고도화를 넘어, 인간과 AI가 안전하고 자연스럽게 공존할 수

있는 새로운 지능형 사회 환경을 구축하는 기반이 될 수 있다.



[그림 3] Affective AI 기반 휴머노이드 로봇 (출처: ETRI)

Affective AI는 인간과 AI의 관계를 단순한 기능적 도구 수준에서 벗어나 사회적 상호작용 주체로 확장시키는 핵심 기술이라는 점에서 중요성이 크다. 그러나 기술 경쟁력 확보에만 집중하고 윤리적·제도적 기반 마련이 병행되지 않을 경우, 개인정보 침해, 감정 조작, 알고리즘 편향성과 같은 문제로 인해 향후 강한 사회적 규제와 신뢰 위기에 직면할 가능성이 존재한다. 따라서 향후에는 국가 차원의 전략적 연구개발 계획과 단계적 실증, 그리고 윤리·제도체계 정비를 함께 추진함으로써, 기술 혁신과 사회적 수용성을 동시에 확보할 필요가 있다.

[참고문헌]

- [1] 이해수, 유재홍, 안성원, "피지컬 AI의 현황과 시사점", SPRI, ISSUE REPORT, IS-202, 2025.
- [2] Chenxi Li, "AIVA: An AI-based Virtual Companion for Emotion-aware Interaction", 2025.
- [3] Hojjat Abdollahi et al, "Artificial Emotional Intelligence in Socially Assistive Robots for Older Adults: A Pilot Study", IEEE Transactions on Affective Computing, 2022.
- [4] Inkster, B., Sarda, S., & Subramanian, V., "An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation", JMIR mHealth and uHealth, 6(11), e12106, 2018.
- [5] Furhat Robotics, <https://www.furhatrobotics.com/>, 2025.
- [6] Today Digital, "Verint Acquires Cogito, Develops a Bot That Scores Live Customer Conversations", <https://www.cxtoday.com/contact-center/verint-acquires-cogito-develops-a-bot-that-scores-live-customer-conversations/>, 2024.
- [7] Realeyes, "The Global Standard for Human Verification," <https://realeyes.ai/company/>, 2026.

※ 출처: TTA 저널 제225호