

휴먼케어 로봇과 소셜 상호작용 기술 동향

Recent Trends in Human-Care Robot and Social Interaction Technology

고우리 (Woori Ko, wrko@etri.re.kr)	인간로봇상호작용연구실 선임연구원
조미영 (Miyoung Cho, mycho@etri.re.kr)	인간로봇상호작용연구실 선임연구원
김도형 (Dohyung Kim, dhkim008@etri.re.kr)	인간로봇상호작용연구실 책임연구원
장민수 (Minsu Jang, minsu@etri.re.kr)	인간로봇상호작용연구실 책임연구원
이재연 (Jaeyeon Lee, leejy@etri.re.kr)	인간로봇상호작용연구실 책임연구원
김재홍 (Jaehong Kim, jhkim504@etri.re.kr)	인간로봇상호작용연구실 책임연구원/실장

ABSTRACT

This paper examines the trends of recently developed human-care robots and social interaction technologies. As one of the solutions to the problems of an aging society, human-care robots have gained considerable attention from the public and the market. However, commercialized human-care robots do not meet user expectations for the role as companions. Current robot services based on short-term interaction and fragmentary pieces of intelligence have encountered difficulty in eliciting natural communication with humans. This results in the failure of human-robot social bonding. Social interaction is being actively investigated as a technique for improving robots' natural communication skills. Robots can form a natural bond with humans through social interaction, which consequently increases their effectiveness. In this paper, we introduce recent human-care robot-related issues and subsequently describe technical challenges and implications for the success of human-care robots. In addition, we review recent trends on social interaction technologies and the datasets required.

KEYWORDS 휴먼케어 로봇, 소셜 상호작용, 대인관계 지능, 인간로봇상호작용

1. 서론

고령층과 1인 가구의 급증으로 인해 다양한 사회 문제와 비용이 발생할 것으로 예상되고 있다.

이에 대한 해결책 중의 하나로 휴먼케어 로봇이 꾸준히 관심을 받아오고 있다. 휴먼케어 로봇은 인간과 함께 생활하면서 인간을 이해하고 정서적으로 교류하면서 상황에 맞는 생활, 정서, 인지, 건강

* DOI: <https://doi.org/10.22648/ETRI.2020.J.350304>

* 본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구개발 사업의 일환으로 수행하였음[2017-0-00162, 고령 사회에 대응하기 위한 실환경 휴먼케어 로봇 기술 개발].



등의 서비스를 제공하는 개인용 서비스 로봇의 일종이다. 로봇 기술의 발전과 사회적 필요성에 힘입어 최근까지 다양한 기능을 가진 휴먼케어 로봇이 소개되고 있다.

하지만 현재까지 상용화된 휴먼케어 로봇은 동반자(Companion)로서의 역할을 바라는 우리의 기대를 충족시키지 못하고 있다. 현재의 로봇은 인간의 다양한 특성을 제대로 이해하지 못하기 때문에 개인별 특성에 따라 대응하는 맞춤형 서비스를 제공하기에 한계가 있다. 또한, 간단한 명령에 대해 단편적으로만 응답할 수 있는 수준이므로 다양한 교류 상황에 대하여 맥락에 맞게 대응할 수 있는 능력이 매우 부족하다. 로봇의 단발적인 반응과 사전 제작된 반복 동작의 수행으로는 자연스러운 소통과 교감을 끌어낼 수 없으며, 결과적으로 유대관계 형성의 실패로 이어진다[1].

최근 몇 년간 로봇의 소통과 교감 능력을 향상시키기 위한 방법으로 소셜 지능(Social intelligence)이 주목받고 있다. 소셜 지능은 로봇이 인지 능력과 사회적 교감 능력을 바탕으로 인간과 교류함으로써 사회적 기능을 수행하도록 하는 기술이다. 소셜 지능을 가진 로봇은 인간의 능력을 모방한 인식, 판단, 표현 능력을 갖추고 있다. 소셜 조작(Social manipulation), 소셜 주행(Social navigation), 소셜 상호작용(Social interaction) 등의 기술들이 현재 활발하게 연구되고 있는 분야이다.

이 중에서 소셜 상호작용은 인간과 로봇이 정서적으로 교감하는 데 필요한 휴먼케어 로봇의 핵심 지능이다. 소셜 상호작용 지능을 가진 로봇은 주인을 이해하고 관심을 표현할 수 있으며, 상대방의 정서를 인지하고 사람처럼 반응할 수 있다. 로봇의 정서적 교감 능력을 통해 인간과 자연스럽게 유대관계가 형성될 수 있고, 이는 로봇의 효용성을 증가시킨다. 소셜 상호작용은 로봇 과학에서 중점적

으로 연구되어야 하는 유망한 분야 중의 하나로 떠오르고 있다[2].

본 고에서는 최근의 휴먼케어 로봇과 로봇 시스템 평가 동향을 소개하고 휴먼케어 로봇의 발전을 위한 기술적 난제와 시사점을 도출한다. 또한, 소셜 상호작용 기술의 최신 동향과 이들 기술의 개발에 필요한 데이터셋 구축 사례를 살펴보고자 한다.

II. 휴먼케어 로봇 연구 동향

1. 연구 동향 및 시사점

가. 연구 동향

휴먼케어 로봇은 어린이, 일반 성인, 고령자 등 모든 사람을 서비스 대상으로 할 수 있으나 현재까지의 연구는 주로 고령자를 대상으로 하고 있어 휴먼케어 로봇은 고령자케어 로봇과 동일하게 인식되고 있는 것이 일반적이며 관련 연구 동향은 다음과 같이 요약될 수 있다.

미국은 대학을 중심으로 고령자의 삶의 질(Quality of life) 향상을 위한 연구가 대규모로 진행된 바 있으며, 유럽연합은 기초 기술 경쟁력을 바탕으로 휴먼케어 로봇에 대한 연구개발을 지속 중인데, 중증 환자를 위한 시설 중심의 간병 로봇보다는 가정에서 고령자의 독립적인 생활을 보조하기 위한 기술 위주로 연구가 진행되어 왔다. 국내도 미국/유럽연합과 비슷한 방향으로 연구가 진행되어 왔으며, 2019년부터 물리적 지원(이송, 욕창 예방, 배설, 식사)이 가능한 돌봄 로봇 기술 개발을 시작하였다.

일본은 전 세계에서 고령화가 가장 빨리 진행되어 미국/유럽연합/한국 등과 달리 고령자 대상 로봇 개발 초기부터 물리적 지원을 중심으로 하는 간병 로봇을 활용한 사회 문제 해결에 집중해 왔으나 최근 다른 국가들과 같이 물리적 지원 이외의 서비

표 1 휴먼케어(고령자케어) 로봇 서비스 사례

분야	서비스 예시
건강 관리	건강(혈압/맥박/당뇨) 모니터링/코칭, 건강상태 파악 후 음식섭취 코칭 및 먼지/습도/온도 등 제어, 특수질환 유병자 생활관리(코칭), 원격화상 진료/상담, 병원 예약, 노인 병원에서 환자안내/사전문진, 건강이상 징후 감지(인지/언어/행동장애)
생활 지원	일상생활 모니터링을 통한 코칭, 위험상황 인식/대처/예방, IoT 연동을 통한 편의기능(조명/에어컨/TV 제어), 부족한 식재료 주문, 주방 보조(요리 레시피 제공), 물건/식사 배달, 전화걸기/받기, 고령자 전용 금융/여행예약 인터페이스
정서 지원	가족/친구들과의 소셜 네트워크 연결, 대화/감정 상호작용을 통한 반려기능, 심리 치료(스트레스/우울증), 간호시설의 레크리에이션(음악/체조/노래/퍼즐 등)
인지 지원	기억/판단 지원, 약복용 알림, 치매 예방 콘텐츠, 물건 찾아주기, 기억 더듬기(가족 간 이벤트, 본인 기억), 인지장애 고령자 대상 개인화된 케어, 물 마실 시간 알림
정보·안전	정보제공(날씨/교통/여행/관광/금융정보), 일정알림(이벤트/기념일/TV/기상/취침/식사), 보안 시스템 연동(침입자 감지 등)

스 제공을 목표로 연구하고 있다.

물리적 서비스를 주로 제공하는 간병 로봇을 제외하면 고령자케어 로봇은 표 1과 같이 건강관리, 생활 지원, 정서 지원, 인지 지원, 정보·안전 서비스를 제공하는 것을 목표로 주로 연구가 진행되어 왔다.

물리적으로 지원하는 돌봄 로봇을 제외한 휴먼케어 로봇 연구의 주요 특징은 다음과 같이 요약될 수 있다[3].

- 실용적 측면의 실 서비스 제공이 강조되어 원격제어(Tele-presence) 기술 적극 활용
- 타 분야 대비 서비스와 콘텐츠 비중이 높음
- 다학제 간 연구가 활성화되어 있고, 실 환경에서의 테스트 및 사용자 연구를 포함
- 가정에서의 독립생활 지원이 주요 목표, 요양기관에서의 지원도 목표에 포함

상기 주요 특징에서 유추할 수 있듯이 현재까지의 연구는 주로 로봇으로 고령자케어를 위한 서비스를 제공할 때의 기술적 이슈 해결 및 사용자의 수용성 분석에 집중되었다. 기술적 관점에서 볼 때 현재까지의 연구는 로봇에 탑재된 화면에서 고령자가 직접 메뉴를 선택하여 서비스를 제공받거나 로봇이 음성, 영상 등을 입력받아 인식을 수행하여 일회성 서비스를 제공하는 형태가 대부분이었다. 하지만 이러한 수준의 기술은 사람의 다양한 특성을 제대로 이해하지 못하고 다양한 교류 상황에 대한 대응 능력이 부족하여 고객의 기대를 충족시킬 수 없다.

나. 기술적 난제와 시사점

휴먼케어 로봇을 위한 소셜 상호작용 기술의 현재 수준 및 앞으로 해결해야 할 기술적 난제를 살펴보기 위한 참조 자료로 표 2와 같이 유럽연합 로보틱스 기술 로드맵의 소셜 상호작용 기술 레벨 정의가 있다[4].

표 2의 소셜 상호작용 기술 레벨을 구체적으로 설명하면 다음과 같다.

- 레벨 0: 사람과의 소셜 상호작용이 없음
- 레벨 1: 사전 저장된 스크립트에 따라 상호작용. 인식 기술을 활용하지 않는 단방향
- 레벨 2: 사전 입력한 정보를 활용하여 로봇

표 2 인간과의 소셜 상호작용 기술 레벨

레벨	정의
0	사람과의 소셜 상호작용 없음
1	시간적으로 제한된 상호작용
2	시간이 확장된 상호작용
3	사람의 행위에 의해 조절되는 상호작용
4	장기 상호작용
5	상호작용을 통한 작업의 조절
6	상호작용을 통한 개인에 대한 지식의 축적
7	다수 사용자와의 장기 상호작용

상호작용 행위 개인화(사람마다 다른 반응)

- 레벨 3: 사람의 행위에 대한 인식을 활용하여 로봇 상호작용 행위 개인화
- 레벨 4: 반복 상호작용을 통해 사용자 행위/외형으로부터 개인에 대한 신원인식, 모델(프로파일) 자동생성·추론을 기반으로 로봇 상호작용 행위 개인화 및 상호작용 이력 기반 행위 적응(adaptation)
- 레벨 5: 상호작용하는 동안 로봇이 관찰을 통해 사용자의 단기 목표, 기본 의도, 욕구를 추론하여 수행하던 작업의 조정에 활용
- 레벨 6: 상호작용하는 동안 관찰을 통해 선호도, 습관, 목표, 믿음 등 사용자 장기 특성에 대한 지식을 축적하여 사용자와의 상호작용 및 수행하던 작업의 조정에 활용
- 레벨 7: 사람-사람 상호작용/관계 모델 이해, 사람-로봇 관계 모델 이해 및 이를 통한 다자 상호작용(multi-party interaction)

상기 총 8단계의 기술 레벨에서 기술 레벨 1에서 3까지는 단기 상호작용, 레벨 4 이후는 장기 상호작용과 관련된 기술이다. 현재 개발된 로봇 제품은 대부분 레벨 3에 해당하며, 수행 중인 연구는 대부분이 레벨 3에 해당하나 일부는 레벨 4에 해당한다.

인간과 로봇의 자연스러운 상호작용을 위해서는 레벨 3까지의 단순 반응형 상호작용이 아니라 레벨 4에서 레벨 7까지의 장기 상호작용 기술 개발이 필수적이며, 이때 로봇은 사람과의 교류행위(상호작용)에 대한 학습을 통해 상호작용의 수준을 스스로 높일 수 있어야 한다. 이때 선결 조건이 장기 상호작용이 가능할 정도로 레벨 3까지 개발된 기술이 로봇에서 안정적으로 동작할 수 있는 수준이 되어야 한다는 것인데, 현재의 기술은 그 수준에 이르지 못한 상태이다.

이러한 기술 수준의 문제와 가격 경쟁력 등의 복합적인 문제로 Jibo, Kuri 등의 로봇이 사업을 중단하게 되어 관련 기업이나 연구자들이 일부 위기감을 느끼고 있다.

로봇 소셜 상호작용에 대해 15년 이상 연구를 해 온 Cornell 대학 Guy Hoffman 교수는 2019년 IEEE 스펙트럼 기고에서 현재 로봇 기술 수준이 기업이 약속한 수준의 사용자 경험을 제공하지 못하고 있고, 킬러앱(killer app)의 부재, AI 스피커와의 경쟁의 어려움 등 다수 전문가들의 진단을 소개하면서 현재 상황이 실패가 아니라 기회라는 의견을 피력하였다[5]. 다만 로봇의 성공을 위해서는 단기적인 단순 반응형 상호작용에 머물러 있는 현재 기술의 한계를 극복하고 장기 상호작용이 가능한 수준의 기술 개발의 필요성을 강조하였다. 이를 위해 자연스러운 로봇 제스처 생성, 정교한 대화, 장기/감성 상호작용 등을 가능하게 하기 위해 로봇 스토리텔링 전문가들과의 협업을 강조하였다.

이와 같이 휴먼케어 로봇이 일상에서 사람에게서 서비스를 제공할 수 있도록 하기 위해서는 레벨 3 수준의 단순 반응형 상호작용이 아니라 레벨 4 이상의 장기 상호작용 기술 개발이 필수적이며, 이때 로봇은 사람과의 교류행위(상호작용)에 대한 학습을 통해 상호작용의 수준을 스스로 높일 수 있어야 한다.

2. 로봇 챌린지 및 평가동향

로봇은 인공지능 기술의 집합체로 조작, 주행, 인간로봇상호작용 등 다양한 지능을 필요로 한다. 휴먼케어 로봇의 핵심 지능인 소셜 상호작용 지능은 음성·시각 정보를 이용한 지각과 주변 상황을 고려한 상황 인지, 인간과 소통할 수 있는 표현 능력을 포함한다. 휴먼케어 로봇의 발전과 더불어 이러한 소셜 상호작용 지능 수준에 대한 관심 또한

높아졌다. 본 절에서는 로봇 챌린지 등 로봇의 지능평가 동향에 대해 살펴보고자 한다.

인공지능 시스템을 대상으로 한 대표적인 평가 방법은 Alan Turing이 1950년 제안한 튜링테스트이다. 하지만 튜링테스트는 사람의 질문에 묵비권을 행사할 경우 인공지능의 기술적인 결함 때문인지 아니면 의도된 침묵인지 구별할 수 없다는 맹점을 가지고 있다[6]. 이에 따라 판정자 마음의 불확실성으로 인공지능 시스템이 튜링테스트를 통과할 수 있으므로 인공지능 시스템의 지능 수준을 평가할 수 있는가는 여전히 논란의 여지가 있다.

MIT에서는 어휘 지식 베이스인 ConceptNet 4를 대상으로 유아용 웹슬러 지능검사를 이용해 지능을 평가하고자 하였다[7]. 그 결과 ConceptNet 4는 4세 유아 수준의 지능으로 단어, 공통점 찾기, 정보 부분에서는 우수하지만, 추론이나 이해는 부족하다는 결론을 내렸다. 이 지능평가 방법은 인간을 대상으로 한 평가 방법을 시스템에 적용해 인간의 지능과 비교했다는 측면에서는 의의가 있지만, 언어지능에 국한되어 있으며, 정보처리 능력을 다루는 인간지능 평가 방법으로 로봇의 사회적 상호작용 평가를 하기에는 적합하지 않다는 한계를 가지고 있다.

대표적인 로봇 챌린지인 RoboCup의 리그 중 RoboCup@Home은 2009년부터 현재까지 서비스 로봇의 지능 수준을 비교 평가하기 위해 매년 개최되고 있다[8]. 각 출전팀은 오픈 플랫폼을 이용해 정해진 시나리오에 따라 작업을 수행하며 가장 높은 점수를 얻은 팀이 승리하는 방식이다. 로봇은 불특정 다수를 상대로 복잡하고 다양한 환경에서 주행, 인식, 조작 지능을 조합해 주어진 시나리오를 수행한다(그림 1). RoboCup@Home 2018부터는 오픈 플랫폼에 일본 도요타의 HSR(Human Support Robot)과 소프트뱅크 로보틱스의 Pepper 로봇 플랫



그림 1 RoboCup@Home 챌린지

폼을 추가하여 경기를 펼치고 있으며, 평가 시나리오는 지속적으로 업데이트되고 있다.

유럽 국가를 중심으로 한 또 다른 챌린지인 RoCKin@Home은 노인을 대상으로 한 독립된 실내 환경을 가정하고 과제(Task) 기반과 기능성(Functionality) 기반 방식으로 나누어 로봇을 평가했다[9]. 평가 시스템 없이 심판에 의해 평가되는 RoboCup@Home과 달리 RoCKin@Home은 객관적 시스템에 의한 평가로 재현성과 반복성에 초점을 맞추어 평가를 진행하였다.

그 외에도 HSR 로봇을 이용한 WRC 챌린지[10]의 'Tidy up the room' 시나리오에서는 미등록 물체 인식, 조작, 경로 계획 등의 지능을 평가하고자 했다. 하지만 이러한 챌린지 기반의 평가 방식은 지능보다 성능에 초점을 두고 있으며, 무엇보다 문제의 난이도에 따른 지능의 수준을 측정할 수 없다는 문제점이 있다.

앞서 살펴본 챌린지와 달리 로봇의 지능을 세부적으로 분류하고 정의하는 지능 체계에 대한 연구도 진행되었다. 그 중 하나인 ALFUS(Autonomy Levels for Unmanned Systems)에서는 임무 복잡성, 환경적 어려움, 인간 상호작용을 기준으로 지능 수준을 레벨 0(원격조정)에서 레벨 10(완전하고 지능적인 자율성)으로 정의하고 있다[11].

Horizon 2020에서는 로봇의 능력을 이동, 조작,

인지 등으로 분류하고 능력별 지능 수준을 세부적으로 정의하고 있다[12]. 특히, 상호작용 지능은 인간과 로봇의 협업을 위한 안전성 등에 대한 물리적 상호작용과 자연스러운 인간과 로봇의 상호작용에 대한 소셜 상호작용으로 지능을 분류하고 수준을 정의하였다. Horizon 2020은 로봇의 지능을 세부 단위 지능으로 나누어 체계적으로 정의하고 있지만 이를 평가하기 위한 구체적인 방법은 제시하지 못하고 있다.

휴먼케어 로봇이 인간에게 동반자로서 좀 더 다가가기 위해서는 자연스러운 사회적 상호작용이 중요하다. 이를 위해 사회 인지적 관점에서 소셜 지능에 대한 이해가 필요하다. 한국전자통신연구원에서는 그 답을 인간을 대상으로 한 소셜 지능 평가에서 찾고자 ESI(Evaluation of Social Interaction) [13] 방법을 로봇에 적용하였다. ESI는 표준화된 관찰 기반 목표지향 절대평가 도구로서 27개의 사회적 상호작용 기술로 개인의 사회적 상호작용 능력을 평가한다. 특히, 상대방과의 거리에 관한 ‘위치시키기’, 몸과 얼굴이 상대방의 정면을 향하고 있는지 평가하는 ‘향하기’ 등 신체적 지지와 관련된 평가항목은 로봇의 소셜 지능 평가에 대한 방향을 제시하고 있다(표 3).

ESI를 적용한 소셜 지능 평가를 통해 사회 인지적 관점에서 휴먼케어 로봇의 사회적 상호작용 수준을 객관적이고 정량적으로 평가할 수 있다. 뿐만

아니라 휴먼케어 로봇의 강점과 약점을 분석해 추가 기능을 구현하고 상호작용의 기능 개선을 위한 기초 자료로 활용될 수 있을 것이다.

III. 소셜 상호작용 기술 동향

1. 대인관계 지능

Howard Gardner에 따르면 소셜 지능은 대인관계 지능(Interpersonal intelligence)과 자기이해 지능(Intrapersonal intelligence)으로 나뉜다[14]. 대인관계 지능은 다른 사람들과 교류하고, 이해하며, 그들의 행동을 해석하는 능력을 말하며, 자기이해 지능은 자기 자신을 이해하고, 느낄 수 있는 인지적 능력을 말한다. 휴먼케어 로봇의 경우 사람과의 유대관계 형성이 중요하므로, 대인관계 지능이 필수적이라 할 수 있다. 따라서 본 절에서는 로봇의 대인관계 지능 관련 연구들을 소개한다.

2. Sense-Think-Act 패러다임

감지(Sense)-사고(Think)-행동(Act)은 많은 로봇 문제를 해결하는 데 사용할 수 있는 의사결정(Decision making) 루프이다. 예를 들어 로봇의 주행 문제를 해결하기 위해서는 먼저 전방에 장애물이 있는지 감지하고, 그런 다음 앞으로 나아갈 수 있는지, 회전해야 할지, 후진해야 할지 생각(사고)한다. 그리고 그 결정에 따라 행동한다.

로봇의 대인관계 지능도 마찬가지다. 대표적인 연구로 참고문헌 [15]는 감지-사고-행동 패러다임을 적용하여 자폐아와 놀아주는 로봇을 개발하였다(그림 2). 먼저 감지 단에서는 자폐아의 활동/음성/얼굴 표정, 장난감의 위치/자세, 로봇의 현재 제스처/머리 자세 등을 인지하여 메타데이터(Meta-data)로 저장하고, 사고 단에서는 이를 기반

표 3 사회적 상호작용 신체적 지지 평가 항목

항목	설명
향하기	얼굴이나 몸이 상대방을 향함
바라보기	상대방과 편안한 눈 맞춤(적절한 눈 맞춤 빈도와 시간)
위치시키기	상대방과 적당한 거리를 두고 자신을 위치시키기(너무 가깝거나 멀지 않도록 한다)
터치	상대방과 신체적으로 접촉
조절	의미 없는 충동적 행위 제어



출처 L. J. Wood et al., "Developing Kaspar: A Humanoid Robot for Children with Autism," *Int. J. Social Robotics*, <https://doi.org/10.1007/s12369-019-00563-6>, CC BY 4.0.

그림 2 자폐아-로봇 간 상호작용 예시

으로 관심 영역(Region of interest)을 알아낸 후 규칙 기반(Rule-based) 메커니즘에 따라 의사결정을 한다. 행동 단은 결정된 의사를 수행하기 위한 제어 신호를 생성한다. 이러한 과정을 통해 로봇은 자연스러운 시선 처리(Eye gaze)와 더불어 아이와 다양한 놀이를 하며 교감할 수 있다.

3. 기계학습 기법

그러나 위와 같은 규칙 기반 방법은 개발자가 로봇에게 일어날 수 있는 다양한 상황을 미리 예측하고 그에 따른 의사결정 규칙을 정의해야 한다는 어려움이 있다. 따라서 최근에는 로봇 스스로 학습을 통해 문제 해결 방법을 도출하는 기계학습(Machine learning) 방법이 활발히 활용되고 있다. 기계학습 알고리즘은 크게 확률 기반, 기하 기반, 회귀기법으로 분류된다. 확률 기반 기계학습 알고리즘은 대부분 베이즈(Bayes) 정리에 기반하며, "X가 주어졌을 때 Y가 발생할 확률"을 계산한다고 보면 된다. NBC(Naive Bayes Classifier)나 HMM(Hidden Markov Model)이 이에 해당한다. 참고문헌 [16]은 coupled HMM으로 발화 문장과 바디 제스처 간의 관계를 학습하여 로봇이 사용자의 현재 감정 상태에 맞는

행위를 생성한다.

기하 기반 기계학습 알고리즘은 주어진 입력특징을 벡터로 만들어 특정 벡터끼리의 기하학적인 관계를 기반으로 추론을 진행하는 시스템을 말한다. k-NN(k-Nearest Neighbors), SVM(Support Vector Machine), k-means clustering이 이에 해당한다. 참고문헌 [17]은 로봇이 고령자에게 운동 동작을 알려주고 이를 끝까지 완수하도록 도와주는 시스템을 제안하였다. k-NN을 사용하여 사용자의 현재 행위를 인식하였으며, FSM(Finite State Machine)을 이용하여 로봇의 행위(인사하기, 운동 동작 보여주기, 박수치기 등)를 선택하였다.

참고문헌 [18]에서는 전시장 안내 로봇이 방문자들에게 다른 이들의 경험과 감흥을 공유하고 설명하였다. Photochar이라는 소프트웨어를 통해 방문자들이 찍은 사진과 설명을 수집하였으며, SVM을 통해 각 사진을 세 가지 카테고리(간단한 설명, 새로운 통찰, 전시물 간 관계를 설명)로 분류하였다. 분류 결과에 따라 로봇은 인기 있었던 전시물의 사진을 보여주며 설명과 함께 관계된 다른 전시물을 추천해 주기도 한다.

회귀기법 기계학습 알고리즘은 주로 지도학습(Supervised learning) 문제를 푸는 데 쓰이며, 인공지능 경망이 여기에 해당한다. 인공지능 경망은 구조에 따라 CNN(Convolutional Neural Network), RNN(Recurrent Neural Network), DQN(Deep Q-Network), AE(AutoEncoder) 등으로 나뉘며, 얼마나 깊느냐에 따라 딥러닝이라고도 한다. 참고문헌 [19]에서는 로봇이 CNN을 통해 사용자 영상 정보를 이해하고, multimodal DQN을 이용하여 네 개의 소셜 행위(기다리기, 쳐다보기, 손 흔들기, 악수하기) 중 하나를 선택적으로 수행하였다. 이러한 행위 생성 능력을 학습하기 위해 Pepper 로봇이 식당, 로비 등의 공용 공간에서 14일간 사람과 상호작용하며 데

이터를 수집하였다.

참고문헌 [20]에서는 RNN의 일종인 LSTM(Long Short-Term Memory)을 이용하여 사람이 먼저 질문을 하지 않아도 로봇 주도적인 발화와 행위를 생성하였다. 이를 위해 카메라 판매점에서 점원과 손님이 상호작용하는 것을 약 11시간 동안 기록하였으며, 로봇은 LSTM을 이용해 비슷한 과거의 기록을 찾아 적절한 대답을 하거나 새로운 정보를 제공하였다.

참고문헌 [21]은 variational AE를 이용하여 로봇의 발화 음성에 적합한 발화 제스처를 함께 생성하는 방법을 제안하였다. Variational AE는 encoder와 decoder로 구성되는데, encoder는 발화 음성을 암호화하며, decoder는 이 결과로부터 로봇 움직임을 생성한다.

4. 종단 간(End-to-End) 학습

앞서 소개한 기계학습 기법 중 최근 대두되고 있는 기술인 딥러닝은, 입력부터 출력까지 사람의 개입이 없다는 의미에서 종단 간(end-to-end) 학습이라고도 한다. 그러나 앞서 소개한 대인관계 지능을 구현한 딥러닝 기술들은 완전한 종단 간 방식이라고 보기 어렵다. 참고문헌 [19]의 소셜 행위 생성 방법을 예로 들면, 입력으로는 미가공 영상 데이터를 그대로 사용하지만, 출력의 경우 사람이 미리 정의해 둔 네 개의 행위 중 하나를 선택하여 재생하는 방식이다. 따라서 로봇은 대상과의 관계나 상황적 맥락과 관계없이 기계적인 행위만을 반복하게 될 수 있다.

고령자케어 로봇을 개발하고 있는 한국전자통신 연구원은 완전한 종단 간 학습 기반의 대인관계 지능 연구들을 수행하고 있다. 참고문헌 [22]에서는 다수의 유튜브 TED 강연 영상을 학습데이터로 사

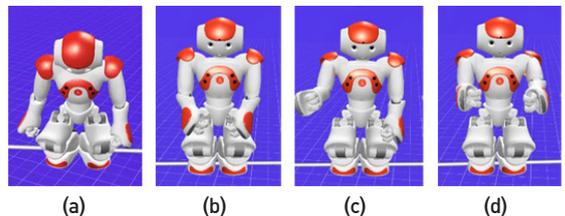


출처 Y. Yoon et al., "Robots learn social skills: End-to-end learning of co-speech gesture generation for humanoid robots," Int. Conf. Robotics Autom. (ICRA), 2019.

그림 3 발화 문장에 따른 제스처 생성 모습. 입력 문장은 "What is your body language communicating to me? What's mine communicating to you?"이며, 밑줄 친 단어에 대한 제스처를 보여준다.

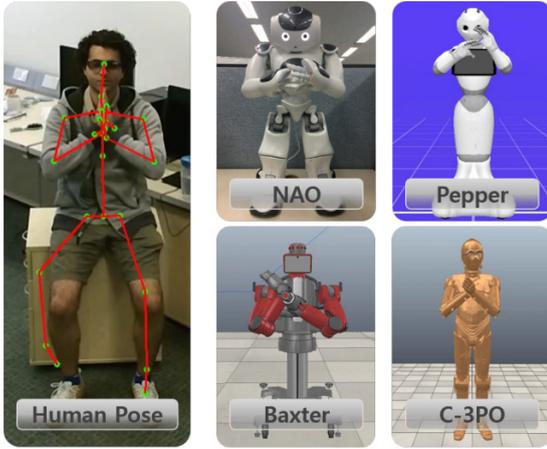
용하여 RNN기반 발화 제스처 생성 모델을 개발하였다. 입력은 발화 문장이며 출력은 바디 제스처이다. 이때 제스처는 일련의 포즈로 표현하였기 때문에 사전에 정의하지 않은 다양한 제스처가 생성되었다. 그림 3은 생성된 로봇 제스처의 예시를 보여준다.

참고문헌 [23]에서는 RNN의 일종인 LSTM을 이용하여 사용자의 행동에 반응하는 로봇의 소셜 행위를 생성하였다. 예를 들어 사용자가 문을 열고 들어오면 로봇은 고개를 숙여 인사를 하고, 손으로 얼굴을 가리고 흐느껴 울면 안아주려 한다. 이때, 입력은 일련의 사용자 포즈이며 출력은 일련의



출처 W. R. Ko et al., "End-to-End Learning-based Interaction Behavior Generation for Social Robots," Workshop on Social Human-Robot Interaction of Service Robots in International Conference on Social Robotics (ICSR), 2018.

그림 4 LSTM 기반 소셜 행위 생성 결과. 사용자가 (a) 문 열고 들어올 때 (b) 쳐다보지 않을 때 (c) 악수하려 손 내밀 때 (d) 얼굴 가리고 울 때



출처 Reprinted with Author's Permission from T. Kim and J. H. Lee, "C-3PO: Cyclic-Three-Phase Optimization for human-robot motion re-targeting based on reinforcement learning," arXiv:1909.11303, 2019.

그림 5 강화학습 기반 로봇 포즈 변환 결과

로봇 포즈이다. 이를 학습하기 위해 사람과 사람이 상호작용하는 데이터를 5,000세트 수집하였으며 샘플 데이터와 이를 확인할 수 있는 코드는 참고문헌 [24]에 공개되었다. 그림 4는 학습된 결과를 보여준다.

위의 연구들은 모두 제스처를 일련의 포즈로 표현하였다는 공통점이 있다. 이를 로봇의 움직임으로 변환(retargeting)하기 위해서는 로봇의 형태(configuration)를 고려해 역기구학(inverse kinematics)을 풀어야 한다는 번거로움이 있다. 참고문헌 [25]에서는 강화학습(reinforcement learning)과 AE를 이용하여 사람의 포즈를 다양한 휴머노이드 로봇 포즈로 변환하였다(그림 5).

IV. 로봇용 데이터셋 구축 사례

앞서 살펴본 바와 같이 소셜 상호작용 연구 분야도 다른 분야와 마찬가지로 딥러닝 기반의 연구들이 주류를 이룬다. 다양한 특성을 지니는 대용량의 학습 데이터셋들은 관련 연구를 발전시킬 수 있는

주요 원동력 중의 하나이다. 로봇 분야에서는 로봇의 이동, 주행 및 조작 연구에 필요한 환경과 물체 대상의 데이터셋들이 가장 풍부하다. 반면에 인간을 대상으로 하는 데이터셋은 종류가 많지 않으며, 데이터의 양도 충분하지 않다. 이는 인간과 로봇이 직접 상호작용하는 데이터를 장기적으로 수집하기가 쉽지 않기 때문이다. 또한, 수집 대상자의 확보에도 한계가 있으며 개인정보 보호 문제도 데이터셋의 공개를 어렵게 하는 요인이 된다. 본 장에서는 인간과 로봇의 소셜 상호작용 연구에 활용될 수 있는 주요 공개 데이터셋들을 간단히 살펴본다.

JPL First - Person Interaction[26]은 인간과 로봇이 상호작용하는 행동을 로봇 시점에서 촬영한 데이터셋이다. 상호작용하는 인간이 우호적인지 적대적인지를 로봇이 판단하기 위한 목적으로 구축되었다. 악수하기, 껴안기, 때리기 등 총 7개의 행동을 촬영한 180개의 비디오로 구성된다. 초기 소셜 상호작용 연구에 활용되었으며 딥러닝 기반의 연구에 활용되기에는 데이터의 양이 부족하다.

MHHRI(Multimodal Human-Human-Robot Interactions)[27]은 로봇과 상호작용하고 있는 인간의 성격 특성과 상호작용할 때의 집중도를 측정하기 위해 구축된 약 6시간 분량의 데이터셋이다. 비디오, 오디오, 웨어러블 센서 데이터 등의 멀티모달 데이터로 구성된 것이 특징이다.

Multimodal Dataset of Human-Robot Hugging Interaction[28]은 로봇이 안전하고 자연스럽게 인간과 포옹할 수 있도록 학습하기 위한 데이터셋이다. 가속도, 근전도, 압력 등을 측정할 수 있는 웨어러블 센서를 착용한 33명의 참가자가 휴머노이드 로봇인 백스터와 포옹을 하였으며, 이때 측정되는 멀티모달 데이터를 저장하였다. 총 353개의 에피소드로 구성되어 딥러닝 모델 연구에 활용되고 있지만, 단일 동작만으로 구성된 것이 아쉬운

점이다.

NTU RGB+D 120[29]는 인간의 일상 행동을 인식하는 연구를 위해 구축된 데이터셋으로 114,480개의 3차원 영상 데이터셋을 포함한다. 세계 최대 규모의 데이터셋으로 딥러닝 기반의 다양한 행동 인식 연구에 활용되고 있다. 인사하기, 토닥이기, 밀치기 등의 인간들 간의 상호작용을 촬영한 8천여 개의 행동 데이터들은 로봇이 인간의 행동 습성을 이해하고 학습하는 용도로 활용될 수 있다.

고령자케어 로봇을 연구하고 있는 한국전자통신 연구원은 딥러닝 기반의 로봇 인공지능 연구를 가속화시킬 수 있는 데이터셋들을 공개하였다[24]. 고령자의 일상 행동을 로봇 시점에서 촬영한 시각 데이터셋인 ETRI-Activity3D는 고령자의 행동 특성을 연구하는 다양한 분야에서 활용되고 있다. 로봇이 발화하는 문장에 적절한 제스처를 스스로 생성하기 위한 데이터셋, 고령자와의 대화 상황에서 교류행위 시점을 인식하기 위한 데이터셋, 그리고 고령자의 행동에 대응하는 로봇의 상호작용 행동을 생성하기 위한 데이터셋들은 로봇의 소셜 상호작용 연구를 위해 구축되었다. 고령자를 대상으로 획득된 상호작용 데이터셋들은 데이터의 부족으로 어려움을 겪고 있는 고령자케어 로봇의 발전에 도움이 될 것으로 기대된다.

V. 결론

본 고에서는 최근의 휴먼케어 로봇 연구와 평가 동향을 소개하고 기술적 난제와 시사점을 도출하였다. 또한, 소셜 상호작용 기술의 최신 동향과 딥러닝 기반의 기술 개발에 필요한 데이터셋 구축 사례를 살펴보았다.

인공지능, 컴퓨팅 파워, 초고속 통신 등의 기술은 하루가 다르게 발전하고 있으며, 저출산/고령

화, 삶의 질 향상 추구 등의 사회적 변화도 지속되고 있다. 이러한 기술의 발전과 사회적 변화는 휴먼케어 로봇 시장의 성장을 더욱 촉진시킬 것으로 보인다.

휴먼케어 로봇의 성공을 판가름하는 핵심 요소는 로봇의 정서적 교감 능력을 바탕으로 하는 인간과의 유대관계 형성이다. 이를 가능하게 하는 소셜 상호작용에 대한 연구는 앞으로 더욱 가속화될 것으로 전망된다. 또한, 소셜 상호작용 연구에 필요한 공개 데이터셋이 매우 부족한 상황에서 연구에 필요한 데이터를 확보하기 위한 전략적인 노력이 필요하리라 판단된다.

용어해설

인간로봇상호작용(Human-Robot Interaction, HRI) 인간과 로봇이 다양한 의사소통 채널을 통해 인지적이고 정서적으로 상호작용을 할 수 있도록 구현해 주는 기술

약어 정리

AE	AutoEncoder
ALFUS	Autonomy Levels for Unmanned Systems
CNN	Convolutional Neural Network
DQN	Deep Q-Network
ESI	Evaluation of Social Interaction
FSM	Finite State Machine
HMM	Hidden Markov Model
k-NN	k-Nearest Neighbors
LSTM	Long Short-Term Memory
MHHRI	Multimodal Human-Human-Robot Interactions
NBC	Naive Bayes Classifier
RNN	Recurrent Neural Network
SVM	Support Vector Machine

참고문헌

- [1] M. De Graaf, S. B. Allouch, and J.van Diik, "Why do they refuse to use my robot?: Reasons for non-use derived from a long-term home study," in Proc. ACM/IEEE Int. Conf. Human-Robot Interaction (HRI), Vienna, Austria, Mar. 2017, doi: 10.1145/2909824.3020236.
- [2] G.-Z. Yang et al., "The grand challenges of Science Robotics," *Sci. Robotics*, vol. 3, no. 14, 2018, doi: 10.1126/scirobotics.aar7650.
- [3] 김재홍, 장민수, "고령자케어 서비스 연구동향," *로봇공학회지: 로봇과 인간*, 제13권 제4호, 2016, pp. 6-12.
- [4] The Partnership for Robotics in Europe, "Robotics 2020 Multi-Annual Roadmap: For Robotics in Europe Horizon 2020 Call ICT-2017(ICT-25, ICT-27, ICT-28)," Feb 12, 2016, pp. 211-212, https://www.eu-robotics.net/cms/upload/topic_groups/H2020_Robotics_Multi-Annual_Roadmap ICT-2017B.pdf.
- [5] G. Hoffman, "Anki, Jibo, and Kuri: What We Can Learn from Social Robots That Didn't Make It," *IEEE Spectrum*, May, 2019, <https://spectrum.ieee.org/automaton/robotics/home-robots/anki-jibo-and-kuri-what-we-can-learn-from-social-robotics-failures>
- [6] 장길수, "튜링테스트 오류 발견했다," *로봇신문*, 2016. 7. 11.
- [7] S. Ohlsson et al., "Measuring an artificial intelligence system's performance on a verbal IQ test for young children," *J. Experimental Theoretical Artif. Intell.*, vol. 29, no. 4, 2017, pp. 679-693, doi: 10.1080/0952813X.2016.1213060.
- [8] <http://www.robocupathome.org/>
- [9] <http://rockinrobotchallenge.eu/home.php>
- [10] <https://worldrobotsummit.org/en/wrs2020/challenge/>
- [11] H. M. Huang and E. R. Messina, "Autonomy levels for unmanned systems (ALFUS) framework volume II: framework models initial version," No. Special Publication (NIST SP)-1011-II-1.0, 2007.
- [12] SPARC Robotics, "Robotics 2020 multi-annual roadmap for robotics in Europe," SPARC Robotics, EU-Robotics AISBL, The Hague, The Netherlands, accessed Feb 5, 2018.
- [13] <https://www.innovativeotsolutions.com/>
- [14] H. Gardner, "Frames of mind: the theory of multiple intelligences," Basic Books, New York, 1983.
- [15] A. Zarakı et al., "Toward autonomous child-robot interaction: development of an interactive architecture for the humanoid kaspar robot," in Proc. Workshop Child-Robot Interaction HRI, Vienna, Austria, Mar. 6, 2017, <http://orca.cf.ac.uk/id/eprint/129029>.
- [16] A. Aly and A. Tapus, "Prosody-based adaptive metaphoric head and arm gestures synthesis in human robot interaction," in Proc. Int. Conf. Adv. Robot. (ICAR), Montevideo, Uruguay, Nov. 2013, doi: 10.1109/ICAR.2013.6766507.
- [17] S.F.R. Alves, M. Shao, and G. Nejat, "A socially assistive robot to facilitate and assess exercise goals," in Proc. Int. Conf. Robot. Autom., Montreal, Canada, May 2019, https://morobae.github.io/papers/morobae_p10_alves.pdf.
- [18] K. Matsumura, T. Gompei, and Y. Sumi, "Robot behavior designed to encourage conversations between visitors in an exhibition space," in Proc. IEEE Int. Symp. Robot Human Interactive Commun., Edinburgh, UK, Aug. 2014, doi: 10.1109/ROMAN.2014.6926269.
- [19] A. H. Qureshi et al., "Robot gains social intelligence through multimodal deep reinforcement learning," in Proc. IEEE-RAS Int. Conf. Humanoid Robots (Humanoids), Cancun, Mexico, Nov. 2016, doi: 10.1109/HUMANOIDS.2016.7803357.
- [20] P. Liu et al., "Learning proactive behavior for interactive social robots," *Autonomous Robots*, vol. 42, 2018, pp. 1067-1085, doi: 10.1007/s10514-017-9671-8.
- [21] T. Kucherenko, "Data driven non-verbal behavior generation for humanoid robots," in Proc. ACM Int. Conf. Multimodal Interaction, Boulder, CO, USA, Oct. 2018, pp. 520-523, doi: 10.1145/3242969.3264970.
- [22] Y. Yoon et al., "Robots learn social skills: End-to-end learning of co-speech gesture generation for humanoid robots," in Proc. Int. Conf. Robot. Autom. (ICRA), Montreal, Canada, May 2019, doi: 10.1109/ICRA.2019.8793720.
- [23] W.-R. Ko et al., "End-to-End Learning-based Interaction Behavior Generation for Social Robots," in Proc. Int. Conf. Social Robotics (ICSR), Qingdao, China, 2018.
- [24] Artificial Intelligence for Robots, "Datasets," <https://ai4robot.github.io/datasets/>
- [25] T. Kim and J.-H. Lee, "C-3PO: Cyclic-Three-Phase Optimization for human-robot motion retargeting based on reinforcement learning," arXiv:1909.11303, 2019.
- [26] M. S. Ryoo et al., "Robot-centric activity prediction from first-person videos: What will they do to me?" in Proc. ACM/IEEE Int. Conf. Human-Robot Interaction (HRI), Portland, OR, USA, Mar. 2015, doi: 10.1145/2696454.2696462.
- [27] O. Celiktutan, E. Skordos, and H. Gunes, "Multimodal human- human-robot interactions (MHRI) dataset for studying personality and engagement," *IEEE Trans. Affective Comput.*, vol. 10, no. 4, 2019, pp. 484-497, doi: 10.1109/TAFFC.2017.2737019.
- [28] K. Bagewadi, J. Campbell, and H.B. Amor, "Multimodal dataset of human- robot hugging interaction," arXiv:1909.07471, 2019.
- [29] J. Liu et al., "NTU RGB+D 120: A large-scale benchmark for 3d human activity understanding," *IEEE Trans. Pattern Anal. Mach. Intell.*, Early Access, 2019, doi: 10.1109/TPAMI.2019.2916873.