

소셜 로봇의 지능 지수 평가 방법

조미영 한국전자통신연구원 선임연구원

조영조 한국전자통신연구원 책임연구원

지수영 지능형 로봇(PG413) 의장, 한국전자통신연구원 책임연구원

1. 머리말

소셜 로봇이란 부여된 역할에 부합하는 사회적 행동과 규범에 따라서 인간, 로봇, 환경과 상호작용하고 소통하는 자율형 서비스 로봇을 말한다. 인공지능 기술의 발전과 함께 다양한 형태의 소셜 로봇이 등장함에 따라 소셜 로봇에 대한 사용자의 기대치가 증가하고, 로봇 지능에 대한 관심 또한 증가하고 있다. 하지만 로봇마다 사용 환경이 다르고 사용자가 다르므로 필요한 소셜 로봇의 지능도 다르고 이에 대한 평가 방법도 달라질 것이다.

로봇에게 지능이 있다면 과연 어떻게 측정할 수 있을까? 사람들이 하는 IQ 검사처럼 정량화된 방법은 없는 것일까? 대부분 관련 표준은 로봇에 포함된 단위 기술에 대한 성능평가로, 이를 지능으로 아우르는 시험 표준에 대한 연구는 미미한 실정이다. 본고에서는 소셜 로봇의 지능 지수 평가를 위한 요소지능 지수를 정의하고, 핵심

요소가 되는 사람인식 요소지능지수 평가 방법에 대해 자세히 다루고자 한다.

2. 소셜 로봇

2.1 소셜 로봇 분류

Fong은 소셜 로봇 상호작용 기술을 몇 가지 주요 범주로 분류하고 설문조사를 통해 중요한 특징을 분석하였다[1]. Bartnect은 HRI의 5가지 주요 개념(의인화, 호감도, 활동성, 지능, 안전)을 측정하기 위한 설문을 제안하였다[2]. 이러한 기존 연구를 참조하여 소셜 로봇을 형태에 따라 <표 1>에서와 같이 인간형, 동물형, 희화형, 기능형으로 나누었다.

2.2 소셜 로봇 지능지수 평가

소셜 로봇의 지능 지수는 5개의 필수 요소 지능에 대한 평가결과를 종합하여 평가한다[3]. 5개 지능 요소의 식별자를 $id \in \{HR, UM, CO,$

<표 1> 소셜 로봇의 형태별 분류

형태	인간형 소셜 로봇 (anthromorphic social robot)	동물형 소셜 로봇 (zoomorphic social robot)	희화형 소셜 로봇 (caricatured social robot)	기능형 소셜 로봇 (functional social robot)
	Class = A	Class = Z	Class = C	Class = F
설명	얼굴을 포함한 머리와 두 팔을 포함한 상반신을 기본적으로 갖고 있고 필요에 따라 이동을 위해 바퀴나 다리를 구비하여 인간의 형태와 유사한 소셜 로봇	장난감이나 애완용으로 동물의 형태를 본떠서 만든 소셜 로봇	만화에서와 같이 희망하는 상호작용 성향이나 외형 특징을 집중적으로 과장하여 희화한 형태의 소셜 로봇	주어진 작업이나 기능을 수행하기에 적합한 형태로 만들어진 소셜 로봇

<표 2> 요소 지능의 선택 가중치 설정 예

사람 인식		사용자 모델링		의사소통		몸체 움직임		자기표현	
기량요인 유/무		기량요인 유/무		기량요인 유/무		기량요인 유/무		기량요인 유/무	
얼굴 인식	○	의상 속성 인식	○	언어 대화	○	사람 회피/동반 주행	○	표현 동작	○
성별/나이 인식		머리카락 속성 인식	○	화자 전환 시점 인식	○	사용자 추종	○	감정 제어 모델	
행위/제스처 인식	○	사용자 소유물 인식		다가가기 기술	○	상호작용 행위 생성	○	음성 합성	○
사람추적	○			시선 맞춤 기술	○	제스처 생성	○	개성	
중요도: 보통		중요도: 덜 중요		중요도: 중요		중요도: 중요		중요도: 덜 중요	
W(HR)=0.20		W(UM)=0.15		W(CO)=0.25		W(BM)=0.25		W(SE)=0.15	

BM, SE}, 즉 HR(사람 인식), UM(사용자 모델링), CO(의사소통), BM(몸체 움직임), SE(자기표현)로 나타내고, 로 표시한다. 이때 소셜 로봇의 종류에 따라 요소지능별 가중치를 달리 준다. 소셜 로봇 지능 지수(SRIQ, Social Robot Intelligence Quotient)의 평점은 5개의 요소 지능 지수의 가중치 합으로 부여되는데, 이를 수식으로 표현하면 다음과 같다.

<수식 1>

$$SRIQ = \sum_{id} W(id) \times EIQ(id)$$

예를 들어, 인간형 소셜 로봇에 대하여 <표 2>과 같이 선택 가중치 W(HR)=0.20, W(UM)=0.15, W(BM)=0.25, W(SE)=0.15, W(CO)=0.25를 부여하고, EIQ(HR) = 80, EIQ(UM) = 60, EIQ(BM) = 70, EIQ(SE) = 50, EIQ(CO) = 70로 평가되면, 소셜 로봇 지능

지수는 SRIQ = 67.5가 된다.

기량요인(skill factor)란 소셜 로봇의 지능을 구성하고 요소 지능을 유발하는 요인이 되는 세부 기량을 의미하며, 각 요소 지능은 3~4개의 기량요인으로 구성되어 있다. 소셜 로봇의 사용 환경 및 응용 분야에 따라 기량요인의 유/무가 결정되면, 이를 고려해 선택 가중치를 부여하게 된다. 요소 지능별 기량요인에 대한 자세한 설명은 3장에서 다루도록 한다.

3. 요소지능지수 정의

3.1 사람인식(human recognition)

소셜 로봇이 서비스 대상이 되는 사람의 정보를 카메라 등 입력 센서를 이용하여 얼마나 지능적으로 인식하느냐를 판단하는 사람인식 요소 지능은 <표 3>에서와 같이 얼굴 인식, 성별/나이 인식, 행위/제스처 인식, 사람 추적과 같은 4

<표 3> 사람 인식 요소 지능의 기량 요인

코드	기량요인	설명
HR ₁	얼굴 인식	등록된 사람에 대한 얼굴 인식
HR ₂	성별/나이 인식	사용자의 성별과 나이 인식
HR ₃	행위/제스처 인식	일정 거리에서 사용자의 제스처 인식
HR ₄	사람 추적	이동하는 사용자의 검출 및 추적

<표 4> 사용자 모델링 요소 지능의 기량 요인

코드	기량요인	설명
UM ₁	의상속성 인식	의상의 색상, 패턴, 타입, 소매길이 등 인식
UM ₂	머리카락 속성 인식	머리카락 길이, 색상, 스타일 등 인식
UM ₃	사용자 소유물 인식	안경, 휴대폰 등 사용자 소유물 인식

<표 5> 의사소통 요소 지능의 기량 요인

코드	기량요인	설명
CO ₁	언어 대화	서비스 영역, 명령어 단위(문장/단어) 등을 고려한 지능
CO ₂	화자 전환 시점 인식	비언어 대화 요소 중 화자 전환 관련 지능
CO ₃	다가가기 기술	상호작용 자세 중 대화 상대와의 거리와 향하기(몸통)를 고려한 지능
CO ₄	시선 맞추기 기술	상호작용 자세 중 대화 상대의 얼굴 바라보기(고개), 시선 맞춤을 고려한 지능

<표 6> 몸체 움직임 요소 지능의 기량 요인

코드	기량요인	설명
BM ₁	사람 회피/동반 주행	사람을 피하거나 사용자와 일정 거리를 유지하며 동반해 지정 경로 지점을 거쳐 이동
BM ₂	사용자 추종	사용자와 일정 거리를 유지하며 사용자를 따라서 이동
BM ₃	상호작용 행위 생성	사용자의 교류 행동에 대응하는 상호작용 행위 생성
BM ₄	제스처 생성	사람과 자연스러운 상호작용을 위해 제스처 생성

<표 7> 자기표현 요소 지능의 기량 요인

코드	기량요인	설명
SE ₁	표현 동작	얼굴표정(눈, 눈썹, 코, 입), 감성 제스처 등
SE ₂	감정제어 모델	정해진 감정패턴, 정서루프 생성, 장기적 감정제어 모델 등
SE ₃	음성 합성	녹음된 음성, TTS 발화, TTS에서의 피치/톤 제어 등
SE ₄	개성 표현	외향성, 동의성, 성실성, 신경성, 개방성

개의 기량 요인으로 구성된다.

3.2 사용자 모델링(user modeling)

소셜 로봇이 사용자 이해를 위해 외적 정보를

얼마나 지능적으로 인식하거나 판단할 수 있느냐에 대한 사용자 모델링 요소지능은 <표 4>에 서와 같이 의상 속성, 머리카락 속성 인식, 사용자 소유물 인식과 같은 3개의 기량 요인으로 구

성된다.

3.3 의사소통(communication)

소셜 로봇이 사용자와 의사소통할 때 필요한 지능으로 <표 5>에서와 같이 언어 대화, 상호작용 자세, 비언어 대화, 행위 학습과 같은 4개의 기량 요인으로 구성된다.

3.4 몸체 움직임(body movement)

소셜 로봇이 사회적 활동을 위해 몸체를 움직이는 지능 요소로서, <표 6>에서와 같이 사람 회피/동반 주행, 사용자 추종, 상호작용 행위 생성, 제스처 생성과 같은 4개의 기량요인으로 구성된다.

3.5 자기표현(self-expression)

소셜 로봇이 자기 자신을 표현하는 지능 요소로서, <표 7>에서와 같이 표현 동작, 감정제어 모델, 음성 합성, 개성과 같은 4개의 기량 요인으로 구성된다.

4. 사람인식 요소지능지수 평가

4.1 사람인식 요소지능지수의 평점 부여방법

사람인식 요소지능을 위한 4개의 기량요인은 소셜 로봇의 활용 및 기능에 따라 W(code)로 표시되는 가중치를 갖는다. 사람인식 요소지능지수 $EIQ(HR)$ 는 기량요인의 가중치 합인 다음의 (수식 2)로 결정된다[4].

$$EIQ(HR) = \sum_i (W(HR_i) \times Score(HR_i))$$

여기서 Score는 $0 \leq Score(HR_i) \leq 100$ (for $i=1, 2, 3, 4$)의 값을 가진다.

소셜 로봇이 사람인식과 관련된 4개의 기량요인을 모두 가지고 있다면 0.25로 균등한 기준 가중치를 부여하며, 그 합은 1이 된다. 다만, 기량요인의 수에 따라 가중치 합 최대값은 달라지며, 기량요인의 중요도에 따라 선택가중치를 달리 준다.

4.2 얼굴인식 시험평가 방법

얼굴인식 기술의 실험환경 성능을 보장받기 위해서는 기존 알고리즘 평가 방법이 아닌 시스템 레벨의 평가가 필요하다. 이를 위해 실물을 대신하여 고해상도 디스플레이 장치를 통해 출력되는 얼굴 (동)영상을 이용한 KS B 7306:2016 표준[5]의 환경과 방법을 적용한다.

얼굴 인식률 $Recog(HR_i)$ 을 구하고 평균 응답시간 가중치 W_{RT} 를 고려하여 얼굴인식 기량요인 평점은 $Score(HR_i) = W_{RT} \times Recog(HR_i)$ 로 구한다. 여기서, 평균 인식 응답시간(RT)은 0.5초 이내가 되어야 하며, 1.5초가 지나서 응답하는 것은 의미가 없다고 간주하여 다음 식에 의해 결정한다.

$$W_{RT} = \begin{cases} 1 & \text{if } RT \leq 0.5 \text{ sec} \\ 1.5 - RT & \text{if } 0.5 < RT < 1.5 \text{ sec} \\ 0 & \text{if } RT \geq 1.5 \text{ sec} \end{cases}$$

예를 들어, 평균 인식 응답시간이 1초인 경우 평균 응답시간 가중치(W_{RT})는 0.5가 된다.

4.3 성별/나이 인식 시험평가 방법

성별/나이 인식 기술의 시험평가를 위해 KS B 7306:2016 표준[5]에서 요구하는 환경과 방법을 따른다. 성별 인식률은 성공과 실패로 평가하며, 나이 인식률은 인식 오차가 ± 5 세 이내이면 성공, 그 이상이면 실패로 판단한다. 성별/나

이 인식률은 (수식 4)와 같이 평균 성별 인식률과 평균 나이 인식률 합의 평균으로 구한다.

<수식 4>

$$Recog(HR_2) = (\frac{1}{N} \sum_{n=1}^N R_n^g + \frac{1}{N} \sum_{n=1}^N R_n^a) / 2$$

(수식 4)의 성별/나이 인식률 $Recog(HR_2)$ 에 (수식 3)의 평균 응답시간 가중치 W_{RT} 를 고려하여 성별/나이 인식기량요인의 평점은 $Score(HR_2) = W_{RT} \times Recog(HR_2)$ 로 구한다.

4.4 행위/제스처 인식 시험평가 방법

소셜 로봇의 행위 및 제스처 인식 성능평가를 위해 로봇의 이동에 의한 진동, 카메라의 움직임 요소가 반영된 KS B 7315:2019 표준[6]을 따른다. 평가용 행위 및 제스처는 소셜 로봇에서 주로 활용되는 인사하기, 악수하기 등 10개 이상의 제스처를 선정한다. 영상 촬영 시 개개인의 자연스러운 움직임이 반영될 수 있도록 행동 및 제스처에 대한 설명은 최소화한다. 모든 행동의 시작은 차렷 자세에서 시작해 차렷 자세로 끝나도록 하며, 데이터베이스로 저장할 경우 시퀀스별로 제스처가 속한 프레임의 위치(시작과 끝), 제스처 정보를 태깅한다.

행위/제스처 인식 성능은 인식 대상의 행동이 발생하는 시작 프레임부터 마지막 프레임까지 구간 정보가 있는 시퀀스가 입력으로 주어졌을 때, 로봇이 입력된 행동을 얼마나 정확하게 분류하였는가를 측정하기 위해 인식 정확도 $Recog(HR_3)$ 로 행위/제스처 인식 성능을 평가한다.

행위/제스처 인식 정확도는 입력 시퀀스 데이터의 인식 유무에 따라 결정되므로 (수식 3)의 평균 응답시간 가중치(W_{RT})가 1이 되어 행위/

제스처 인식 기량요인의 평점은 $Score(HR_3) = Recog(HR_3)$ 이 된다.

4.5 사람 추적 시험평가 방법

사람 추적 기술은 로봇의 행동인식 및 휴먼추종을 위한 요소 기술로서 추적 기술의 정확도는 로봇 서비스의 품질을 좌우한다. 특히, 로봇의 사람추적 성능평가에는 사람의 움직임과 로봇의 이동을 고려한 평가용 데이터베이스가 필요하다. 소셜 로봇의 사람추적 기량요인의 시험평가를 위해 TTA.KO-10.1138 표준[7]에서 제시된 바와 같이 추적 대상의 크기와 자세 변화 및 물체/사람에 의한 가려짐을 반영한 평가용 DB를 이용한다.

사람추적 정확도는 검출한 사람 영역과 표준치 사람 영역 간의 겹치는 정도를 평균하여 (수식 5)에 의해 계산한다.

<수식 5>

$$Recog(HR_4) = \frac{1}{N} \sum_{j=1}^N d(h_j, g_j) \times 100\%$$

여기서, N은 입력 영상에서 검출해야 하는 표준치 개수, g_j 는 j번째 표준치의 영역 정보, h_j 는 j번째 표준치와 신원 구별자가 동일한 검출 영역 정보를 가리킨다. d는 g_j 와 h_j 가 나타내는 두 사각형 간의 IntersectionOverUnion(IOUS)로 정의한다. 사람추적 정확도는 각 프레임 단위로 측정되므로 (수식 3)의 평균 응답시간 가중치(W_{RT})가 1이 되어 사람추적 기량요인의 평점은 $Score(HR_4) = Recog(HR_4)$ 이 된다.

5. 맺음말

본고에서는 소셜 로봇의 지능 지수 평가를 위

한 요소지능과 이를 구성하는 기량요인을 정의하였다. 그리고 사람인식 요소지능지수 평점 부여 방법과 각 기량요인의 평가 방법에 대해 살펴 보았다. 이를 통해 소셜 로봇의 지능이라는 개념을 이끌어 내고 개별 요소기술에 대한 성능평가 표준을 통합하여 이를 활용하고자 하였다.

향후 소셜 로봇 지능 시험 표준을 통해 소셜 로봇의 지능 수준을 검증함으로써 산업체는 로

봇 기술 및 지능에 대한 목표치를 설정하고 이를 통해 진일보한 기술을 개발하는 발판을 마련할 수 있을 것이다. 소비자는 지능 수준이 검증된 제품을 공급받을 수 있으며, 또한 지능지수를 통해 제품을 선택하는 기준을 얻을 수 있다. 더 나아가 고수준의 지능을 가진 소셜 로봇의 개발로 국가 경쟁력 향상에 기여할 수 있을 것이다. TTA



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

참고문헌

- [1] Terrence Fong, et al., "A Survey of Socially Interactive Robots", Robotics and Autonomous Systems, 42 (2003) 143-166
- [2] Lewis R. Goldberg, "An Alternative "Description of Personality: The Big-Five Factor Structure," Journal of Personality and Social Psychology, Vol. 59, No. 6, 1216-1229, 1990
- [3] TTA.KO-10.1225 소셜 로봇의 지능지수 평가 방법
- [4] TTA.KO-10.1279-Part1 소셜로봇의 요소지능지수 평가방법-제1부: 사람인식
- [5] KS B 7306:2016 서비스 로봇을 위한 시스템 레벨에서의 얼굴인식 성능평가방법
- [6] KS B 7315:2019 데이터베이스 기반 휴먼 행동 및 제스처 인식 성능평가
- [7] TTA.KO-10.1138 가정환경에서 서비스 로봇을 위한 영상 장치 기반 사람 추적 알고리즘 성능평가 방법