

고령자 맞춤 능동적 상호작용의 AI스피커 개발

정재현* · 장지훈** · 문미경***

Development of AI Speaker with Active Interaction Customized for the Elderly

Jae-Heon Jeong* · Ji-Hoon Jang** · Mikyeong Moon***

요 약

인구의 고령화로 인해 노인 인구가 늘어나고, 핵가족화가 빠르게 진행되고 있다. 현재 AI스피커는 일상적으로 사람과 사람 간에 이루어지는 대화보다는 사용자의 명령어에 반응하는 형태이다. 독거노인이 먼저 얘기를 하지 않으면 AI스피커의 사용성은 떨어지게 된다. 본 논문에서는 고령자 맞춤 능동적 상호작용의 AI스피커 개발 내용에 대해 기술한다. 본 스피커는 독거노인의 움직임과 주변 상황을 파악하여 능동적으로 말을 걸어줄 수 있으며, 대화의 내용에 맞는 감정표현을 디스플레이 시킬 수 있다. 이를 통해 사용자는 AI스피커를 의인화하게 되어 친근감을 느낄 수 있고 감성 대화가 가능해짐에 따라 외로움을 달랠 수 있는 긍정적인 역할을 할 것으로 기대한다.

ABSTRACT

Due to the aging of the population, the number of the elderly is increasing, and the nuclear family is rapidly progressing. Today's AI speakers respond to user's commands rather than conversations that occur on a daily basis. If the elderly living alone do not talk first, the usability of the AI speaker will decrease. In this paper, it describes the development of AI speakers for active interaction tailored to the aged. This speaker can identify the movements of the elderly who live alone and their surroundings, actively speak to them, and display emotional expressions appropriate to the content of the conversation. Through this, users will be able to anthropomorphize AI speakers, so they can feel familiarity and emotional conversation is expected to play a positive role in easing their loneliness.

키워드

Artificial Intelligence, AI Speaker, Interaction, Emotional Expression, Hologram, Senior
인공 지능, AI스피커, 상호작용, 감정 표현, 홀로그램, 고령자

1. 서 론

한국은 인구의 고령화로 노인 인구가 늘어나고, 동

시에 핵가족화가 빠르게 진행되면서 전체 독거노인 숫자가 증가하고 있다. 2017년 기준 독거노인 수는 약 138만 명에 달하며, 1인 가구의 44%가 60대 이상 노인

* 동서대학교 학생연구원(z7942s@naver.com)

** 동서대학교 학생연구원(xxwlgns@gmail.com)

*** 교신저자 : 동서대학교 소프트웨어학과

• 접수일 : 2020. 09. 21

• 수정완료일 : 2020. 11. 03

• 게재확정일 : 2020. 12. 15

• Received : Sep. 21, 2020, Revised : Nov. 03, 2020, Accepted : Dec. 15, 2020

• Corresponding Author : Mi-Kyeong Moon

Dept. Software, Dongseo University,

Email : mkmooon@dongseo.ac.kr

가구이다. 혼자 사는 1인 노인가구인 경우 외로움으로 인해 우울증 등 정신적인 고통을 겪을 위험이 높다[1].

지역사회 통합돌봄 선도사업(커뮤니티 케어)으로 대상 어르신 중 독거노인들은 AI(인공지능스피커)와 대화하며 무료함을 달래는 동시에 생활편의 서비스를 받을 수 있게 되었다. 독거노인의 서비스 사용 비중은 음원 스트리밍 서비스 '플로'가 63.6%로 가장 높았고, 감성대화 서비스(13.4%), 날씨(9.9%), 운세(5.0%) 순이었다. 현재 AI스피커는 일상적으로 사람과 사람 간에 이루어지는 대화보다는 사용자의 명령어에 반응하는 단계인데[2], 운둔형 독거노인, 자살충동, 우울증의 독거노인이 AI(인공지능스피커)에게 먼저 얘기하고 답변하기를 기대하기 어렵다. 사용자인 독거노인의 행동패턴, 성향 등을 분석하여 사용자와의 관계에서 능동적으로 대화, 행동을 유도하는 단계로 진화될 필요가 있다. 또한 AI 스피커를 사용하는 사용자들이 AI 스피커를 단순히 정보를 전달하는 일방향적인 매체로 인지하는 것이 아니라, 일련의 상호작용이 가능한 가상의 객체(agent)로 확장해 인식한다. 그러므로 단지 음성만을 수단으로 하는 입력-수행만이 아니라 감정적인 인터랙션이 되어진다면 기기를 지속해서 사용할 가능성이 높아 질 것이기 때문에 실제 사람과 대화하는 듯한 인상을 주어 좀 더 사용자와 친밀해 질 수 있도록 할 필요가 있다.

본 논문에서는 AI스피커에 IoT기술을 접목하여 고령자 맞춤형 능동적 상호작용이 가능한 AI스피커의 개발 내용에 대하여 기술한다. 본 스피커는 독거노인의 움직임을 파악하여 능동적으로 대화를 걸어줄 수 있으며, 대화의 내용을 분석하여 감정을 표현할 수 있는 캐릭터를 디스플레이 시킬 수 있다. 이를 통해 사용자는 AI스피커를 의인화하게 되어 친근감을 느낄 수 있고 감성 대화가 가능해짐에 따라 외로움을 달랠 수 있는 긍정적인 역할을 할 것으로 기대한다.

II. 관련연구

AI스피커는 단순하게 소리를 전달하는 도구에서 인공지능 알고리즘을 이용해 사용자와 음성으로 의사소통을 할 수 있는 스피커가 되었다. 대표적인 제품으로 아마존 에코와 구글 홈, SK텔레콤의 NUGU, 카카

오의 카카오미니, KT의 기가지니 등이 있다. 이들은 대부분 음악 재생, 뉴스, 스포츠, 날씨 등 필요한 정보 제공, 그리고 기타 기능으로 알람 또는 일정 메모 기능 등을 가지고 있다 [3-5]. 최근에는 목소리로 제어하는 AI스피커가 시각적 기능까지 지원하는 '디스플레이형 디바이스'로 진화하고 있다. SK텔레콤과 LG 유플러스가 디스플레이형 AI 스피커를 출시한 데 이어 KT도 음성 인식을 지원하는 디스플레이형 디바이스 '기가지니 테이블TV'를 선보였다.

AI스피커가 점차 보편화됨에 따라 AI스피커의 사용성에 대한 연구들이 많이 이루어지고 있다 [6-8]. 연구 [9-10]에서는, 사용자들은 AI 스피커의 일상 대화에 대한 기대감은 높으나 충족되지 못하고 있음을 보여주면서, AI스피커가 먼저 말을 거는 것과 같은 능동적인 행동을 하는 것을 긍정적으로 기대하는 것으로 연구결과를 보여주었다. 연구[11]에는 정서별 감정발화에 따른 사용성을 평가하였다. 이 연구에 따르면, 음성을 기반으로 하는 AI스피커의 경우, 인간과 같은 감정표현을 할 수 있는 능력을 갖추게 하는 것이 중요함을 설명하였고, 부정적인 감정을 배제하고 긍정적인 감정을 포함하는 표현에 선호를 보인다고 설명하고 있다. 이러한 AI스피커의 사용성에 대한 연구들을 바탕으로 본 연구에서는 능동적으로 상호작용을 하면서 감정표현이 나타날 수 있는 AI스피커의 개발에 초점을 두었다.

III. 고령자 맞춤형 능동적 상호작용의 AI스피커 개발

본 논문에서 기술하는 고령자 맞춤형 능동적 상호작용의 AI스피커에 대한 전체 내용을 도식화하면 그림 1과 같다.

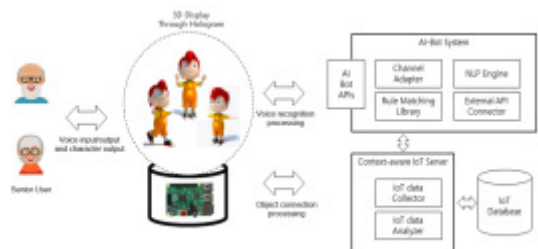


그림 1. 시스템 개념도
Fig. 1 System conceptual diagram

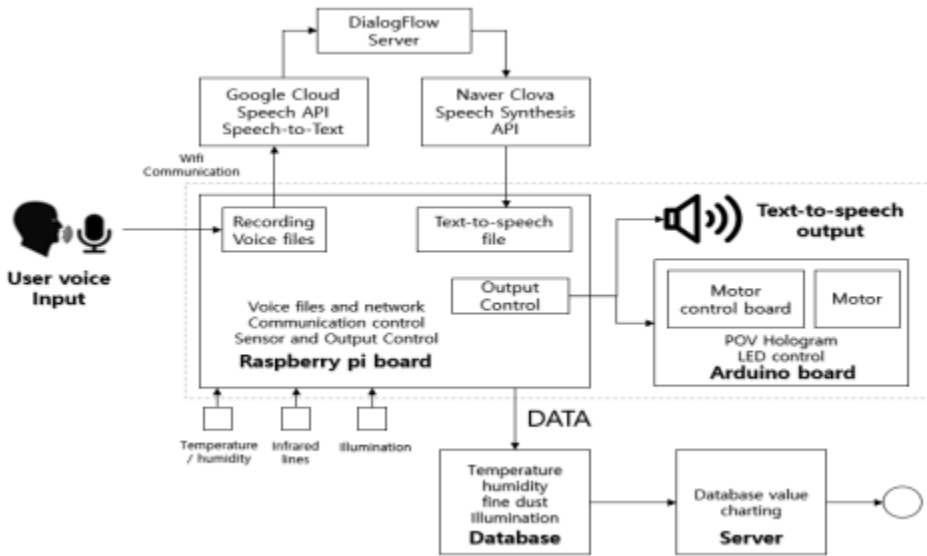


그림 2. 라즈베리파이 보드 음성인식 스피커의 시스템 구성도
 Fig. 2 System Configuration of raspberry pi board speech recognition speaker

본 연구에서는 기존의 AI스피커에 IoT기술을 접목하여 사용자가 있는 공간의 날씨 정보, 온도, 습도, 움직임 정보 등을 파악하여 능동적으로 사용자 요구에 맞춤형 서비스를 제공해 줄 수 있다. 또한 AI스피커에 홀로그램 디스플레이를 장착하여 감정을 나타내는 이모티콘을 3D로 나타냄으로써 친근함과 시각효과를 줄 수 있도록 하였다.

3.1 구성도

그림 2는 라즈베리파이 보드에 각종 센서와 디스플레이를 장착한 AI스피커의 시스템 구성도이다. 확장성과 범용성이 뛰어난 라즈베리파이를 기반으로 하여 음성인식 및 서버와의 통신 기능, 각종 센서값을 처리한다. 라즈베리파이에 장착된 센서들로부터 감지된 온도/습도 데이터를 입력받아서 관제서버에 전달 후 사용자의 거주지 상황을 모니터링 할 수 있는 장치를 포함하고 있다.

3.2 음성인식 및 자연어처리 개발방법

사용자의 음성을 라즈베리파이에서 입력을 받아 녹음 파일로 생성하여 Google Cloud Speech API를 통하여 Text로 변환된 후, 변환된 Text를 DialogFlow

Server로 전송한다. DialogFlow Server에 전송된 Text에서 사용자의 요구를 확인하고 온도 및 미세먼지 확인 등 여러 가지의 기능을 수행하며 상황에 맞는 캐릭터와 표정이 반환된다. 반환된 Data를 라즈베리파이에서 다시 수신하여 Naver Clova Speech Synthesis API를 이용하여 반환된 Text 내용을 음성 파일로 변환하고 출력제어를 통해 음성과 영상이 출력되어 사용자에게 정보를 전달한다. 그림 3은 그 과정을 도식화한 것이다.

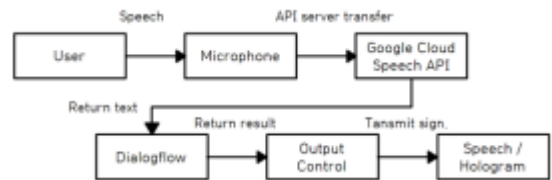


그림 3. 스피커 음성 및 캐릭터 처리
 Fig. 3. Speaker voice and character processing

3.3 능동적 상호작용을 위한 개발 방법

사용자가 AI스피커를 통해 먼저 말을 걸고 서비스를 요청하는 경우의 일반적 질문 유형과 질문 예시는 표 1과 같다.

표 1. 일반적 질문 유형과 예시

Table 1. Common types and examples of questions

Description	Example
When a Senior user asks questions about the weather	Tell me the weather today. Is it raining today? Tell me the current weather in Seoul.
When a Senior user asks a question about fine dust.	Tell me about the fine dust "to day". Tell me about "Seoul" fine dust.
When a Senior user asks a question about the date, day of the week, or time.	Tell me today's date. Tell me the day of the week. ○Month, ○. What time is "America"?
When a Senior user requests a song	Play me some "exciting" music. Play me a song by "○○".
When a Senior user requests the end of the song.	Turn off the music.
When a Senior user requests a schedule or alarm.	Tell me today's schedule. Please register your alarm at 7 p.m.
When a Senior user requests news information.	Read me the main news today. Tell me the soccer schedule.

AI스피커가 능동적으로 먼저 대화를 시도하기 위해서는 사용자의 주변상태 또는 거주지의 상황을 확인해야 한다. 이를 위해 라즈베리 파이 보드에는 온도센서, 습도센서, 적외선센서와 조도센서가 부착한다. 이러한 센서와 API를 통해 데이터를 입력받아서 서버에 전달되고 데이터 분석을 통해 상황인식을 할 수 있다. 확인할 수 있는 상황으로는 사용자의 움직임이 장시간 인지되지 않는 경우, 또는 사용자의 움직임은 감지되었지만 말을 걸지 않는 경우가 있다.

- ① 사용자의 움직임이 장시간 인지되지 않은 경우 : 적외선센서를 이용하여 사용자가 움직임이 있는지 5분 간격으로 확인한다. 사용자가 1시간 이상 움직임이 없다면 AI스피커는 사용자에게 “무기력하신가요? 음악을 재생할까요?”식으로 먼저 질문을 하게 된다.
 - ② 사용자의 움직임은 감지되었지만 말을 걸지 않는 경우 : 온도센서, 습도센서, 조도센서와 미세먼지를 이용하여 사용자에게 말을 걸게 된다.
- 온·습도 센서 : 실내의 온도와 습도를 10분 단위로 측정된 뒤 표 2의 계절별 실내 적정 온도와 비교한다. 온도가 낮을 경우 “실내의 온도가 낮습니다. 보일러를 켜는 것이 어떨까요?”식으로

AI 스피커가 질문을 한다.

- 조도센서 : 수면시간을 제외한 시간의 실내 밝기가 일정 수치 밑으로 내려가거나 일어날 시간, 잠잘 시간이 되었을 때 말을 걸게 된다. 실내 밝기가 낮을 경우 “실내가 많이 어둡지 않으신가요? 전등을 켜는 것이 좋을 것 같아요”식으로 AI 스피커가 질문을 한다.
- 미세먼지 : AirKorea의 OpenAPI를 이용하여 미세먼지의 값을 XML형식으로 받아들이고 뒤 Dictionary형식으로 변환을 한다. 나오는 값으로는 미세먼지(PM10)와 초미세먼지(PM2.5)에 대한 값이 나오며, 각각에 대해 미세먼지의 농도(pm10Value)와 24시간 예측농도(pm10Value24), 미세먼지의 24시간 등급(pm10Grade)과 1시간 등급(pm10Grade1h)에 대해서 나오게 된다. 등급으로는 1~4등급으로 표시되며, 순서대로 좋음, 보통, 나쁨, 매우 나쁨으로 구분된다. 등급이 좋을 경우 “오늘의 미세먼지는 좋습니다. 산책하기 좋은 것 같아요”식으로 AI 스피커가 질문을 한다.

표 3은 AI스피커가 사용자에게 능동적으로 말을 거는 경우에 대해서 나타낸 것이다.

표 2. 계절에 따른 실내 적정 온·습도

Table 2. Seasonal optimal indoor temperature and humidity

Season	Optimal temperature	Optimal humidity
Spring-Fall	19~23℃	50%
Summer	24~27℃	60%
Winter	18~21℃	40%

표 3. AI스피커가 사용자에게 능동적으로 말을 거는 경우

Table 3. The AI speaker is actively speaking to the user.





Situation	AI speaker
If the room temperature is lower than the proper temperature.	Aren't you cold? Turning on the boiler looks good.
If the room temperature is higher than the proper temperature.	Isn't it hot? It looks good to turn on the fan.
Higher than humidity suitable for room temperature.	I think it's humid inside. Turning on the dehumidifier looks good.

Less than humidity suitable for room temperature.	I think the room is dry. It looks good to turn on the humidifier.
If the outdoor fine dust level is level 3 or 4.	Fine dust is bad today. It looks good to wear a mask.
If the outdoor fine dust level is level 1 or 2.	Fine dust is good today. (If the weather is good) It looks nice for a walk.
If it's mealtime.	It's time for breakfast / lunch / dinner.
If the room brightness is low.	I think you should turn on the lights.
When it's time to wake up / go to bed.	It's time to get up. (If it's past bedtime) It's time to go to bed.
If the user does not move for a certain amount of time.	Are you sick somewhere? Are you lethargic? Shall I play a song?

3.4 대화내용 분석 감정표현 방법

신체언어는 언어를 매개로 하지 않는 얼굴표정, 동작, 자세, 눈의 움직임 등 신체 움직임에 의한 비언어적 커뮤니케이션 (Nonverbal Communication) 행위를 말한다. 출력되는 음성의 의미에 따라 캐릭터의 표정이 각기 다르게 표현되도록 한다. 인간의 감정은 희(기쁨), 노(분노), 애(슬픔), 락(즐거움), 애(좋아함), 오(싫어함), 욕(욕망) 7가지로 구분할 수 있는데, 이에 대한 AI스피커의 반응은 표 4와 같이 4가지로 구분할 수 있다. AI스피커 기사용자들은 스피커의 발화에서 부정적인 감정을 적절히 배제하고, 긍정적인 감정을 충분히 포함하는 표현에 선호를 보인다는 연구 결과가 있다 [2]. 이를 참고하여 일반적인 질문에 대한 답을 하는 평소 표정도 즐거운 감정이 더 표현되도록 하였으며, 사용자의 주변 환경에 따라 좋은 상황일 때는 기쁨으로, 적절하지 못한 상황일 때는 슬픔으로 표정을 나타내었다. 또한 AI스피커가 사용자의 질문을 이해하지 못하는 경우에는 당황스러운 표정을 나타내도록 하였다. 표정은 홀로그램 디스플레이에 나타내기 위해 계란모양의 이모티콘 형태로 3D 애니메이션 구현을 하였다.

표 4. 대화내용에 따른 AI스피커 감정표현 유형
Table 4. Cases of the AI speaker emotion expression according to conversation content

Emotion	Situation and expression
Joy	 <ul style="list-style-type: none"> -The weather is clear, fine dust level is good. - Meal time reminder, music play
	 <ul style="list-style-type: none"> - Cloudy weather, poor fine dust level - Residential conditions (temperature, humidity) are not appropriate.
Embarrassment	 <ul style="list-style-type: none"> - The AI speaker does not understand the user's question.
	 <ul style="list-style-type: none"> - The AI speakers answer common questions from users.

IV. 결과 및 실험

그림 3은 라즈베리파이 보드와 각종 센서를 사용하여 만든 AI스피커의 시제품이다. AI스피커와 연동되는 디스플레이는 (주)엠플즈에서 홀로그램 시제품으로 만든 회전형 LED 디스플레이 장치이다. 이는 1mm간격의 LED를 사용하여 소형 팬형식의 홀로그램 디스플레이를 개발한 것이다 (그림 4). 3D영상은 두뇌 활성화 효과 측면에서 2D보다 뛰어나며, 인지능력 또한 상승함에 대해 연구 결과가 있었다 [12-13]. 이로인해 본 연구에서도 2D디스플레이와 차별되게 3D 영상이 나올 수 있는 홀로그램을 사용한 것이다.

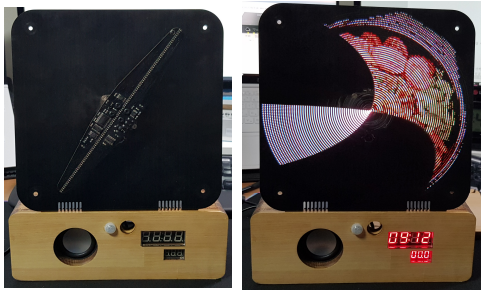


그림 4. AI스피커 시제품
Fig. 4 AI speaker prototype

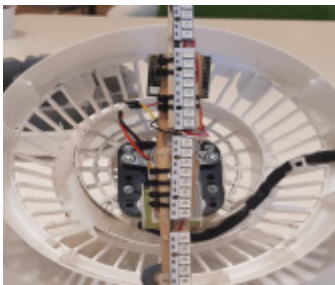


그림 5. 1mm 소형 홀로그램
Fig. 5 1mm small hologram

시제품으로 개발된 AI스피커는 기존에 AI스피커를 사용해본 경험이 있는 사람 중 6명을 대상으로 3일 동안 사용해보도록 하였다. 사용자로부터 받은 결과는 표 5와 같으며, 각 항목당 10점 만점을 기준으로 하였다.

표 5. AI스피커 사용성 조사
Table 5. Usability survey of AI speakers

Type	Survey Items	Score
Accuracy	AI speaker understands my words well.	9.33
	AI speaker answers my questions correctly.	8.50
The appropriateness of expressing emotions	The AI speaker display is useful.	8.33
	The emoticon expression on the AI speaker display is appropriate for dialogue.	8.83
	Emoticons on AI speaker display enhance realism.	8.50
The appropriateness of active conversation	The AI speaker speaks first.	8.67
	It is good for the AI speaker to speak first.	8.50
	It's not good for the AI speaker to speak first.	4.83
	The frequency with which AI speakers speak first is reasonable.	4.67

AI스피커 사용성 조사에 따르면 AI스피커가 감정표현을 하는 것에 대해서는 아주 긍정적인 결과가 나왔다. 그러나 AI스피커가 먼저 말을 거는 것에 대해서는 긍정적인 반응도 높았지만, 부정적인 반응도 다소 있는 것으로 결과가 나왔다. 이는 낮은 AI스피커가 먼저 말을 걸면, 이에 대해 대답을 해야 하는 부담감이 작용된 것으로 보인다. 능동적 대화 여부에 대해 설정을 할 수 있으며, 말을 거는 시간간격을 조정할 수 있도록 함으로써 이를 극복할 수 있을 것으로 판단된다.

V. 결론 및 향후개선 방향

본 논문은 AI스피커에 IoT기술을 접목하여 고령자 맞춤형의 능동적 상호작용이 가능한 AI스피커의 개발 내용에 대하여 기술하였다. 본 AI스피커는 사용자가 요청하는 명령에 답을 하는 것뿐만 아니라 사용자의 주변상황을 파악하여 능동적으로 말을 걸어준다. 또한 대화의 내용과 적절한 감정을 이모티콘을 이용하여 표현함으로써 사용자가 좀 더 친근감을 느낄 수 있도록 하였다. 본 스피커를 통해 독거노인들이 혼자 무료하게 시간을 보내는 것을 AI스피커와 대화를 함으로써 생활에 활력을 줄 수 있게 됨으로써, 우울증이 많이 개선될 것으로 기대한다. 실험대상자가 6명으로 진행되었던 점이 연구결과의 통계적 한계가 있는 것으로 보인다. 추후 연구에서 실험대상자의 인원을 늘려 이를 극복하고자 한다.

감사의 글

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음(2019-0-01817)

References

- [1] Miral Welfare Foundation, "Lonely old man living alone, It's a problem before us.", URL: <http://www.miral.org>
- [2] J. H. Jang and D. Y. Ju, "Usability Test of

- Emotional Speech from AI Speaker," In *Proc. of Conf. on The HCI Society of Korea*, Jeju, Korea, Feb. 2019, pp. 705-712.
- [3] H. Yoon, "AI Speaker Trend," *The Korean Institute of Electrical Engineers*, vol. 68, no. 10, Oct. 2019, pp. 16-21.
- [4] C. H. Ban, D. H. Kim, and B. Y. Choe, "Design and Implementation of Order Settlement System Using Artificial Intelligence Speaker," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 14, no. 6, Dec. 2019, pp. 1181-1186.
- [5] H. W. Kim, K. R. Noh, O. J. Kwon, and S. J. Ahn, "Technology Convergence Map Creation and Country Profile Analysis in the Field of Artificial Intelligence," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 12, no. 1, Feb. 2017, pp. 139-146.
- [6] H. Lee, C. H. Cho, S. Y. Lee, and Y. H. Keel, "A Study on Consumers Perception of and Use Motivation of Artificial Intelligence(AI) Speaker," *J. of the Korea Contents Association*, vol. 19, no. 3, Mar. 2019, pp. 138-154.
- [7] S. Yoon and K. Hea, "A Study on the Direction of AI Speaker Conversation Interaction for Middle-aged Elderly," In *Proc. of Conf. on The Korea Contents Societys*, Seoul, Korea, May 2019, pp. 23-24.
- [8] C. W. Kim, S. B. Kim, and K. P. Lee, "Uncover adoption pattern of AI speaker via Google Home, Amazon Echo, SK NUGU," In *Proc. of Conf. on KSDS(Korean Society of Design Science)*, Seoul, Korea, June 2017, pp. 158-159.
- [9] J. Ahn, Y. Kang, and G. Kim, "A Study on User Preference according to the types of answer in Emotional Conversation with AI speaker," In *Proc. of Conf. on KSDS*, Ilsan, Korea, Nov. 2019, pp. 227-228.
- [10] S. J. Kim, J. Park, W. Jo, and M. H. Yeoun, "Proactive Dialogue Attempts and Memory of Users Personal Taste to Improve Persistent Usage of AI Speaker," In *Proc. of Conf. on KSDS*, Ilsan, Korea, May 2019, pp. 76-77.
- [11] Y. Lee and D. Cho, "Factors that Affect User Satisfaction toward Continuous Usage of AI Speakers - Focusing on The Mediation Effect of Emotional Attachment," *J. of Korea Society of Visual Design Forum*, vol. 63 no. 0, 2019, pp. 87-100.
- [12] G. Y. Noh, D. J. Park, and H. J. Jang, "An Experimental Study of User Experience for 3D Video Game: Presence, Arousal, Recognition Memory, and Brain Activity Pattern," *J. of Cybercommunication Academic Society*, vol. 31 no. 2, June 2014, pp. 45-83.
- [13] T. E. Kim, "A study on the Fabrication of Mixed Reality Content with Coloring Technology," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 14, no. 4, Aug. 2019, pp. 737-742.

저자 소개



정재헌(Jae-Heon Jeong)

2013년~ 동서대학교 컴퓨터공학부

※ 관심 분야 : BCI기술 응용, 웹 시스템개발



장지훈(Ji-Hoon Jang)

2013년~ 동서대학교 컴퓨터공학부

※ 관심 분야 : AI융합, 웹 애플리케이션 기술



문미경(Mi-Kyeong Moon)

1990년 이화여자대학교 전자계산학과 졸업(이학사)

1992년 이화여자대학교 대학원 전자계산학과 졸업(이학석사)

2005년 부산대학교 대학원 컴퓨터공학과 졸업(공학박사)

2008년 동서대학교 컴퓨터공학부 교수

※ 관심분야 : 소프트웨어공학, AI융합기술 응용

