

인간과 로봇의 Backchanneling Interaction 에 기반한 면접 보조 로봇 설계

Design of an Interview Assisting Robot based on Backchanneling Interaction between Human and Robot

○이혁인¹, 이체은², 광영혜³, 이원형^{4*}

¹⁾ 한동대학교 전산전자공학부(TEL: 010-6604-8628; E-mail: 21700034@handong.edu)

²⁾ 한동대학교 전산전자공학부(TEL: 010-4757-6561; E-mail: 21700591@handong.edu)

³⁾ 한동대학교 전산전자공학부(TEL: 010-3343-4140; E-mail: 21500582@handong.edu)

⁴⁾ 한동대학교 전산전자공학부(TEL: 054-260-1388; E-mail: whlee@handong.edu)

Abstract During the interview, interviewees tend to focus on visual and auditory perception which is influenced by verbal and nonverbal expressions.[1] This paper propose robots which communicate more naturally with humans by focusing on backchanneling interaction which is a communication that focuses on implicit nonverbal cues of communications such as chiming in or nodding. Interaction through backchanneling reduce unnaturalness caused by robots. It also enriches simple verbal conversations and creates environments where interviewees/users are immersed in interviews by improving concentration. Therefore, in certain interview situations in which robots are interviewers, our research team designed an interview assistant robot service that accepts users' backchanneling expressions and provides expressions to people. Designed robots enable users to understand semi-verbal communications and help prepare for interview systematically and realistically.

Keywords Backchanneling, Interview Assistance, Nonverbal Expression, Human Robot Interaction, Communication Engineering

1. 서 론

면접의 중요성이 증가하는 취업 시장에서, 이를 반증하듯 면접 보조 서비스 및 로봇 개발 등 다양한 서비스들이 등장했다. 가장 널리 알려져 있는 면접 관련 서비스는 모의 AI 면접 서비스가 있고, 하드웨어 와 결합된 형태로는 Sweden HR-tech 사의 ‘Tengai’ 가 있다. 두 서비스 모두 사용자의 얼굴과 음성을 인식하고 평가를 진행하는 방식에서 동일하지만, 사람과 조화롭게 의사소통하며 훈련하는 부분에서는 다소 미흡하기 때문에 현장감을 느끼기가 어렵다. 이에 본 연구에서는 비언어적 표현과 함께 대화 상황을 더욱 실감나고 풍부하게 만드는 Backchanneling Interaction 을 도입했다.

Backchanneling 이란, 직접적이지 않거나 드러나지 않은 의사소통 방법을 뜻한다[2]. 따라서, 본 연구는 로봇이 Backchanneling 표현을 효과적으로 사용자에게 제공하기 위해, 사용자의 입력에 따른 로봇의 목과 팔 움직임, Unity 로 구현한 아바타 얼굴[3]을 통한 Feedback 을 출력하여 사용자가 실

※ 본 논문의 연구는 과학 기술 정보 통신부와 정보 통신 기술진흥 센터(IITP)의 소프트웨어중심대학 지원사업 (2017-0-00130)의 지원을 받아 수행하였음

제 대화 상황인 것처럼 몰입하여 다양한 반응을 도출할 수 있는 면접 보조 로봇을 설계했다.

2. 본 론

2.1 Overall Framework

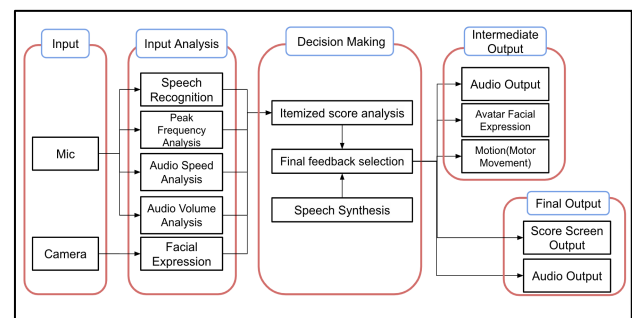


그림 1. Entire Diagram

먼저, 사용자의 언어(발화 내용의 키워드)와 Backchanneling 을 위한 비언어적 표현 인식을 위해 사용자의 얼굴 표정, 시선 방향, 목소리 톤을 인식한다. 이러한 입력을 분석하여 로봇이 면접관으로서 합당한 피드백을 결정하고, 모션(목 움직임, 팔

제스처)과 스크린 상의 3D 아바타 얼굴 표정을 통해 Backchanneling 이 포함된 대화적 상호작용을 하는 형태이다.

2.2 음성 분석

마이크를 통한 Google 음성인식[4]과 사전에 학습된 머신러닝을 이용해 발화 키워드를 구별(시작 인사, 좋음, 공감, 나쁨, 면접종료)하였다[4]. 또한, 목소리 톤 분석을 위해 음성정보를 44.1kHz로 Sampling 되고, FFT를 적용시켜 평균 주파수 값을 구했다. 이렇게 얻어진 사용자의 음성 분석 값과 스피치 분석에 따른 감정 값을 기준으로 로봇은 어떤 Backchanneling Service를 표현할지 결정하고 수행한다.

2.3 영상 분석

사용자의 표정, 얼굴 방향 등을 인식하기 위해 얼굴 신호를 입력으로 받아온다. 먼저, 카메라 센서를 통해 사용자의 얼굴이 인식될 경우, 얼굴 표정을 구현하는 신체 특징점들의 34개 좌표를 설정하고 이를 기반으로 사용자의 20가지 얼굴 요소의 값들을 결정한다. 이후, 실시간으로 받아오는 값들로 7가지 감정을 생성 및 분석한다.

2.4 Decision Making

면접상황에서 긍정적인 표현, 부정적인 표현의 범주를 나눈 기준은 구체적이며 솔직한 답변인가 혹은 애매모호하고 추상적인 답변인가이다. 예를 들어 피면접자가 “열심히 하겠습니다”라고 발언하는 것은 추상적인 표현이므로 Bad Category에 들어간다. 앞서 계산되어 얻어진 사용자의 음성과 영상 분석에 따른 값을 기준으로 로봇은 어떤 Backchanneling Service를 표현할지 결정하고 수행한다. 로봇의 표현을 결정하는 원리는 음성과 영상 분석을 통해 도출되는 분석 값들 중 우선시되는 가치를 중점으로 두어 다양한 경우에 따른 적절한 로봇의 감정을 설정한다. 본 연구는 면접이라는 상황 속에서 피면접자의 언어적 문맥이 면접관에게 자신의 의견을 전달하는 데 있어 가장 큰 비중을 차지한다고 판단하였기에 언어적 내용을 주(Primary) 가치로 설정하였다.

최종적으로 사용자의 Response로 출력하고자 하는 로봇의 감정이 결정되어지면, 이에 대한 표정 좌표 설정 값들을 아바타 측으로 전달하여 아바타의 얼굴이 모니터로 출력되고, 최종 출력 감정에

따른 모션 분류 값을 Zigbee 통신을 통해 로봇의 메인 보드로 전달하여 보드에 내장되어 있는 모션들에 매칭될 수 있도록 한다.

[관련 Youtube 링크-<https://youtu.be/lpbMVm9swr0>]

2.5 로봇의 Backchanneling 표현

로봇이 사용자에게 Backchanneling 표현을 효과적으로 제공하기 위해 한정된 모션만 출력하는 것이 아니라 대화 상황에서 등장하는 긍정, 부정, 아리송, 돌발, 공감 등 다양한 표현들을 출력하도록 하였다. 대화 상황에서 목과 팔 움직임이 큰 영향을 끼치기 때문에[5] 해당 부위에 모터를 추가하여 상응하는 모션을 취하도록 설계하였다. 또한, 사용자의 답변에 대한 아바타의 얼굴 표정을 로봇 머리에 부착한 모니터로 출력하여 사용자에게 더욱 사실적으로 피드백을 제공한다.

3.결과 및 검토

결과적으로, Socket 통신을 기반으로 Python Speech 분석결과와 Robotics 모션 출력 값, 그리고 아바타 표정이 제대로 동작하였다. 표현 모듈간의 미세한 동기화 문제가 있으나, 향후 연구에서 비동기 방식으로 수정 후 이에 대한 실증이 필요할 것으로 보인다.

참고문헌

- [1] Mehrabian, Albert. *Nonverbal Communication*, Transaction Publishers, 1972.
- [2] Park, H. W., Gelsomini, M., Lee, J. J., Breazeal, C., *Telling Stories to Robots: The Effect of Backchanneling on a Child's Storytelling*, MIT Media Lab, March, 2017.
- [3] Ramanarayanan, Vikram., Pautler, David., Lange, Patrick., Suendermann-Oeft, David., *Interview with an Avatar: A Real-Time Cloud-Based Virtual Dialog Agent for Educational and Job Training Applications*, ETS Research Memorandum, March, 2018.
- [4] Choi, S.J., Kim, J.B., *Comparative Analysis of Speech Recognition Accuracy in Speech Recognition Open API*, August, 2017.
- [5] Seongmi Jeong, Dong-Hee Shin, Jihyang Gu, *A Case Study on the Nonverbal Immediacy of the Robot*, Department of Interaction Science at Sungkyunkwan University, July, 2015.