

디지털 교육 환경에서의 인공 지능 마이크 활용 모델

남기복⁰, 박구락*, 김재웅*, 이윤열¹, 김동현**

^{0**}공주대학교 컴퓨터공학과

*공주대학교 컴퓨터공학부

e-mail: mtgood@naver.com⁰, {ecgrpark, jykim, alphaone}@kongju.ac.kr*, dhkim@naver.com**

Artificial Intelligence Microphone Utilization Model in Digital Education Environment

Ki-Bok Nam⁰, Koo-Rack Park*, Jae-Woong Kim*, Jun-Yeol Lee¹, Dong-Hyun Kim**

^{0**}Dept. of Computer Engineering, Kongju National University

*Dept. of Computer Science & Engineering, Kongju National University

● 요약 ●

최근 4차 산업혁명의 핵심 분야 중 하나인 인공지능에 대한 많은 연구가 이루어지고 있다. 많은 기업들이 인공지능 스피커와 같은 제품을 출시하고 있으나 대부분 비서 역할만을 할 수 있도록 구성된 제품이 대부분이다. 그러나 학교와 같이 많은 사람이 존재하는 경우 시끄러운 환경에서 사용되고 있는 인공지능 스피커는 명령 인식이 제대로 되지 않아 실용도가 저하되는 단점을 가지고 있으며, 현재 인공지능 스피커는 단순한 질의응답 수준의 응대만 가능하여 다소 부족한 부분이 있다. 또한 인공지능의 급속한 발전으로 인공지능 스피커가 아닌 전자제품에 인공지능 비서 기능이 탑재된 제품도 새롭게 출시되어 인공지능 스피커가 필요 없을 수도 있기에, 본 논문에서는 학교와 같은 주변의 소음이 많이 발생하는 교육 환경에서도 소통이 가능한 인공지능 마이크를 활용할 수 있는 모델을 제안한다.

키워드: 마이크(MIC), AI 스피커(AI speaker), STT(Speech To Text)

I. Introduction

최근 IT기술의 발전으로 인공지능을 탑재한 제품이 출시되고 있다. 그중 많은 사람들에게 인기가 있는 인공지능 스피커는 가정에서 비서의 역할을 하고 있다[1]. 하지만 가정이 아닌 많은 사람들이 생활하는 학교와 같은 현장에서는 주변의 소음으로 인하여 명령을 인식하기에는 많은 어려움이 존재한다. 특히 초등학교 같은 경우에는 미디어를 이용한 교육이 활발하게 진행되고 있는 상황이나, 교사가 인공지능 스피커에 명령을 하여 필요한 답변을 듣기에는 어려움이 많다[2]. 또한 교사들이 학생들의 집중을 위해 한곳에서 수업을 진행하는 것이 아닌 교실 내에서 전, 후로 이동하면서 교육을 진행하고 있는 경우가 대부분이다. 또한 파워포인트를 통한 수업의 경우에는 컴퓨터와 조금 떨어진 곳에서도 리모콘 기능을 활용하여 이전과 이후 이동은 가능하다. 그러나 파워포인트를 이용한 수업이 아닌 웹을 이용한 수업을 하고 있다면 화면 제어를 위하여 다시 컴퓨터가 존재하는 곳으로 이동하여 제어를 해야 하는 불편함이 존재하고 있다. 본 논문에서는 이러한 불편을 해소하기 위하여 학교와 같은 주변의 소음이 많이 발생하는 교육 환경에서도 교사와 학생들 간의 소통이 가능한 인공지능 마이크를 활용할 수 있는 모델을 제안한다.

II. Preliminaries

1. Related works

1.1 TTS(Text to Speech)

TTS는 음성합성 시스템이라고 부르며 프로그램을 통해 사람의 목소리를 구현해내는 것으로 모든 단어와 문장을 쉽게 구할 수 있다. 하지만 사전 녹음된 목소리 자료의 기반으로 사용하는 프로그램은 억양이 자연스럽게 못한 단점도 있다[3].

1.2 STT(Speech to Text)

STT는 사람이 말하는 음성 언어를 컴퓨터가 해석해 그 내용을 문자 데이터로 전환 하는 것을 말한다. 그러나 정확한 발음이 되지 않았거나 주변의 선풍기 같은 바람소리 그리고 시끄러운 상황에서는 변환이 잘 되지 않는 단점이 있다. 또한 여러 사람이 동시에 이야기 했을 경우 잘 인식 되지 않는 문제점도 있다. [4]

III. The Proposed Scheme

본 논문은 블루투스 방식의 무선 마이크를 이용하여 컴퓨터의 웹 화면 및 각종 응용 프로그램을 제어하는 모델이다. 다음의 <Fig. 1.>은 명령 인식 및 제어처리를 위한 시스템 구성도이다.

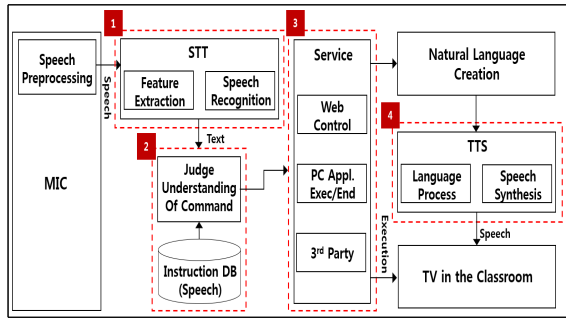


Fig. 1. System Architecture

학교 현장에서 사용하고 있는 컴퓨터를 활용하여 음성명령 처리를 위한 구성은 다음과 크게 4단계로 이루어져 있다. 첫째, 블루투스 방식의 무선 마이크로부터 인식된 음성 데이터의 특징 및 음성인식을 통해 컴퓨터가 이해할 수 있는 텍스트로 변환한다. 둘째, 변환된 텍스트를 이미 구성되어 있는 명령 데이터베이스를 이용한 검색을 통하여 명령어의 이해 및 판단을 한다. 셋째, 판단된 명령어로 컴퓨터의 프로그램 서비스를 제어한다. 서비스에는 웹페이지를 제어하는 기능과 문서 편집기 같은 일반 프로그램 또는 날씨 정보와 같은 외부 연동 프로그램을 통하여 명령을 송신하게 되고 명령어에 따라 필요한 기능이 작동된다. 넷째, 서비스 중 스피커 음성으로 통보해야 하는 경우가 발생하면, 언어처리를 통하여 음성 합성하여 컴퓨터와 연결된 스피커로 송신한다. 그리고 음성이 아닌 화면으로 표현해야 할 경우 컴퓨터와 연결된 TV모니터로 출력한다.

IV. Conclusions

인공지능 스피커는 대중화 되어 있지만 주변의 소음이 대량으로 발생하는 학교와 같은 교육 환경에서는 음성 인식에 많은 문제가 있다. 이러한 단점을 보완하고자 본 논문에서는 인공지능 마이크를 이용한 모델을 제안한다. 특히 멀티미디어 수업을 진행하는 학교 현장에서 교사들이 컴퓨터까지 가지 않고 무선 마이크를 통해 컴퓨터에게 명령을 할 수 있으면 수업을 조금 더 슬기롭게 진행할 수 있고, 교사와 학생들의 의사소통이 원활하게 될 것이며, 교사들의 수업의 질이 향상될 것으로 기대된다. 향후 연구로는 제안 모델에 대한 구현과 음성 인식 향상을 위한 연구가 계속되어야 할 것이다.

of Quality of Artificial Intelligence Speakers on Use Intentions," Korean Academic Society Of Business Administration, pp. 460-463, Aug. 2018.

- [2] Joongang, <https://news.joins.com/article/22503268>
- [3] Choi, Yeunju-Jung and Youngmon-Kim and Youngwan-Suh and Youngjo-Kim and Hoim, "An end-to-end synthesis method for Korean text-to-speech systems," *Phonetics and Speech Sciences*, Vol. 10, No. 1, pp. 39-48, 2018.
- [4] Seongjoon Chu and MinGyu Lee and JinGyeong Jeong and Chiwan Song and HanJong Ko and HanSung Kim and Won-Gil Hong, "A Study on the Similarity Measure between Speech to Text Result and Real Text Using Smith-Waterman Algorithm," *Korea Institute Of Communication Sciences*, pp. 1323-1324, Jun, 2017.

REFERENCES

- [1] Il Woo Park and Gyoo Gun Lim, "A Study on the Effects