

음성인식 시스템을 활용한 단어별 한글 단어평가에 관한 연구

고인옥, 김기돈*, 강애띠*, 김경배

서원대학교, 인포씨드*

dlsdnr64@naver.com, kdkim@infoseed.kr, dlsdnr64@naver.com, gbkim@seowon.ac.kr

A Study on the Evaluation of Korean Words Using Speech Recognition

Ko In Wook, Kim Ki Don, Kang Ae Tti, Kim Gyoung Bae

요약

본 논문은 모든 산업에 영향을 미칠 수 있는 거대한 기술인 기존 음성인식 시스템의 원리를 활용한 안드로이드 한글 단어 평가를 진행하고 단어별 음성 인식률을 점수로 환산하는 음성인식 시스템에 대하여 분석하였다. 또한, 음성인식의 정확성 향상이 미래산업에 끼치는 효과에 관하여 고찰하였다.

I. 서론

반도체 및 컴퓨터 응용기술 등의 급속한 발전과 더불어 인간의 가장 자연스러운 의사전달수단인 음성을 인간과 기계 사이의 의사소통이 매개체로 사용하기 위한 음성인식 기술에 관한 연구가 활발히 진행되고 있으며, 일부 상품화된 음성인식 시스템들이 다양한 응용 분야에 걸쳐 등장하고 있다[1].

이러한 상황에서 정확한 음성인식을 설정하도록 음성인식 시스템을 활용하여 단어 평가를 진행하고 단어별 음성인식에 대한 점수를 부여함으로써 음성인식 시 단어별 발음 특징 등을 파악하고 다양한 방안으로 보완 및 해결할 수 있다.

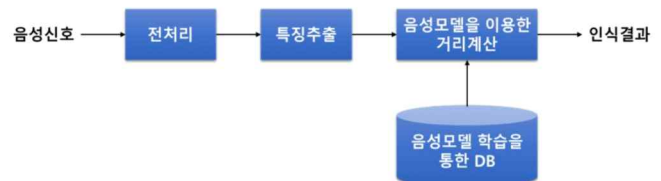
본 논문에서는 음성인식 시스템과 단어별 음성 인식률을 점수로 환산하는 기술을 소개하고 이를 활용한 안드로이드 단어 평가에 대하여 분석하였다. 또한, 단어 평가를 통한 단어별 점수를 측정함으로써 음성인식의 정밀성 향상에 대한 고찰과 정확성이 향상된 음성인식 시스템이 미래산업에 끼치는 영향에 관하여 고찰하였다.

II. 본론

“유비쿼터스”라 일컬어지는 시대가 도래하면서 컴퓨터뿐만 아니라 다양한 응용 시스템들이 인간과 기계 사이의 편리한 상호 작용을 가능케 하는 음성인식 기능을 지원하고 있으며, 앞으로도 그 수요가 점점 증가 될 것으로 보인다.

음성인식 플랫폼 이용 증가	교육, 비즈니스 번역을 위해 음성인식이 활용되어 국가별 소통에 기여
산업 분야에서의 음성인식 이용 증가	음성 입출력을 활용한 홈 공장 자동화 시스템의 이용 증가
음성인식 플랫폼 애플리케이션 성장	자동차, 가전제품 등에 음성인식을 결합하여 사용자에게 편리함을 제공하는 서비스들이 인기를 얻고 있음
스마트폰 음성 시스템 증가	안드로이드 나 IOS 등 OS의 발전으로 이용자들은 스마트폰 음성인력으로 날씨 정보, 음성 검색 등 애플리케이션을 받음

[표 1] 음성 관련 플랫폼 시장 동향



[그림 1] 음성인식 기술의 원리

대표적인 음성인식 기술의 원리는 [그림 1]과 같다. 음성신호를 사용자의 발성으로 입력시킨 후 전처리에서 발성 인식 구간을 추출하고 잡음처리를 한다. 그 후, 발음의 높낮이 등 특징을 추출하고 음성모델 학습을 통해 저장되어있는 원 단어의 데이터베이스와 비교를 통해 인식결과가 도출된다.

4차 산업 혁명 시대가 도래하면서 기존의 클라우드 컴퓨팅, 자율주행 자동차, 사물인터넷 기술, 빅데이터, 인공지능 등 다양한 기술과 결합하는 형태로 음성인식이 적극적으로 활용되고 있다.

음성인식의 궁극적인 목표는 자연스러운 발성에 의한 음성을 인식하여 실행 명령어로서 받아들이거나 자료로서 문서에 입력하는 완전한 음성/텍스트 변환(full speech-to-text conversion)의 실현이다[2].

음성인식 API 종류	각 회차별 틀린 개수 (각 회차: 14개)										계 (오인식률)
	1회	2회	3회	4회	5회	6회	7회	8회	9회	10회	
Naver	5	3	3	3	6	4	3	1	2	4	34(24.28%)
카카오	4	4	6	4	5	4	5	4	4	6	46(32.85%)

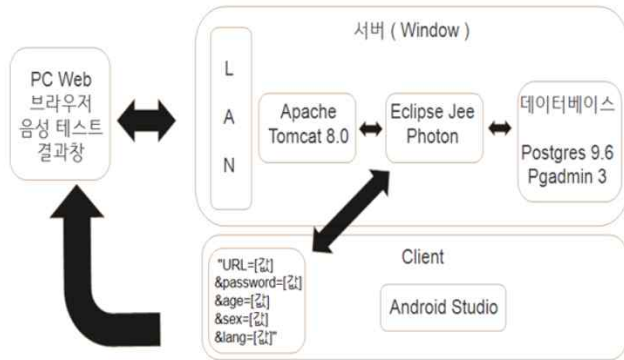
[그림 2] 국내 대표적인 음성인식 오픈 API(네이버, 카카오)의 종류별 14개 한글 음성인식 결과[3]

가~까지 14개의 음절을 음성인식에 입력시킨 결과이다. 이렇게 다양한 방법의 연구로 음성 인식률이 호전되고 있지만, 아직 음성인식의 보완점으로 지적되는 것이 음성인식의 정확성이다. 위의 이유로 일부 한글 단어를 선별한 후 수집한 단어들을 사용자에게 단어별 발음의



난이도 등에 대하여 질의하고 음성 인식률을 점수로 환산하여 결과값으로 도출하는 안드로이드 단어 평가를 시행하였다.

음성인식 평가에 사용된 개발 도구는 Tomcat 8.0, Eclipse Jee Photon, Postgres 9.6, Pgadmin 3, Android Studio 3.5.0, SDK 19~28, JDK - 1.8으로 구성되었으며, 시행한 음성인식 시스템은 [그림 1]과 같은 원리로 작동한다.



[그림 2] 한글단어 평가 어플 작동 원리

개발 툴	
명칭	용도
Tomcat 8.0	웹 서버 소프트웨어
Eclipse Jee Photon	Java 웹 프로젝트 프로그래밍
Postgres 9.6	데이터베이스
Pgadmin 3	PostgreSQL 관리
Android Studio 3.5.0	안드로이드 프로그래밍 소프트웨어
SDK 19 ~ 28, JDK - 1.8	웹, 안드로이드 개발 필요 툴

[그림 3] 음성평가 상세 개발 도구

데이터베이스에서 단어 내용 순서에 해당하는 단어를 사용자의 음성으로 해당 단어를 인식시켜 결과를 서버에 저장하는 방식을 사용하였다. 음성인식 기술은 Google API를 사용하였다.

[그림 4] 어플 작동

로그인 후 Posgres 9.6 데이터베이스에 저장되어있는 순서의 한글 단어와 단어에 대한 질의에 답을 한다. 질의에 대한 내용은 표준어 여부, 단어 사용빈도, 발음의 난이도, 기타 비속어 여부 등 16개로 이루어져 있다.

질의에 답변 후 작은 소리, 중간 소리, 큰 소리 총 3번의 음성인식을 실행한다. 이때 사용자의 발성을 통한 단어의 음성인식 발음을 입력하고 화면에 출력함과 동시에 Apache Tomcat 8.0 서버에 저장한다.

Apache Tomcat 8.0 서버에 저장된 사용자의 발음과 데이터베이스에 저장된 원 단어와 프로그래밍을 통해 비교한 후, 단어들의 질의응답 점수와 인식률에 대한 점수를 통합하여 환산한다. 통계 결과는 [그림 5]와 같다.

[그림5] 평가 통계

기존의 음성인식의 원리를 활용한 단어 평가를 통하여 단어별 질의응답과 인식률을 종합하고 통계로 나타냄으로써 인식률 결과를 기반으로 한 단어별 인식률에 대한 고찰과 인식률이 낮은 단어의 발음인식 개선 등 음성인식 정확성 향상에 대하여 고찰하였다,

또한, 음성인식 시장의 동향으로 보아 정확성이 향상된 음성인식 시스템이 다른 서비스들과 결합한다면 미래산업에 끼치는 파급력이 클 것이다.

III. 결론

본 논문에서는 인간과 기계 사이의 편리한 상호 작용을 가능케 하고 모든 산업에 영향을 미칠 수 있는 거대한 기술인 음성인식 시스템에 대한 분석을 위해 기존 음성인식 시스템의 원리를 활용한 한글 단어평가를 시행하여 단어별 인식률을 점수로 환산하고 결과를 도출함으로써 음성인식 시스템의 정확성 향상에 대한 고찰과 음성인식 시스템의 정확성 향상이 음성인식을 활용한 공장 자동화 등 미래산업과 다양한 기술에 결합하여 더욱 정확한 음성인식 시스템으로 응용될 수 있는 효과를 기대한다.

ACKNOWLEDGMENT

본 연구는 2019년도 중소벤처기업부의 기술개발사업 지원에 의한 연구임 [S2768125]

참 고 문 헌

- [1] 김형준. (1994). 음성인식. 대한음성학회 1994년도 2월 학술대회지, pp. 1~1
- [2],[3] 최승주, 김종배. (2017). 음성인식 오픈 API의 음성인식 정확도 비교분석. Asia-Pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology. pp. 412~417