

스마트폰을 이용한 시각장애인을 위한 음성인식 홈 리모트 컨트롤 시스템

이세훈*, 최승준*

*인하공업전문대학 컴퓨터시스템과

e-mail : seihoon@inhac.ac.kr, fotmdwns49@naver.com

Voice Recognition Home Remote Control System for the Visually Handicapped using a Smartphone

Se-Hoon Lee*, Seung-Jun Choi^o

*Dept. of Computer Systems&Engineering, Inha Technical College

● Abstract ●

시각장애인은 주변 인지 장애로 정보 습득이 어렵고, 공간 정보 습득에 있어서 제약이 있는 신체조건으로 이동에 불편을 겪는다. 또한 특정 상황에서 적절한 분위기를 인지하고 그에 맞는 대응을 하기가 어렵다. IoT 시대에 맞춰 홈 네트워크 시스템이 널리 보급되어지는 추세이지만, 여전히 시각장애인이 사용하기에는 많은 어려움이 있고, 불편을 호소하는 게 현실이다. 본 논문에서는 이러한 시각장애인을 위해 스마트폰을 이용해 음성으로 집안의 사물을 원격 제어하는 구글 음성인식 홈 리모트 컨트롤 시스템을 제안한다.

키워드: 음성인식(Voice Recognition), 시각장애인(Visually Handicapped), 원격제어(Remote Control), 홈 네트워크(Home Network)

개발한다.

I. Introduction

시각장애인은 주위 환경과 상호작용을 통해 공간 정보를 습득하는 데에 어려움이 있는 감각인지 장애이다. 장애인도 같이 살아가야 하는 지금의 스마트 세상에서는, 이러한 시각장애인을 위한 기본적인 편의성조차 보장되고 있지 않다. 장애인 정책발전 5개년 계획부터가 현재 스마트 세상에서 장애인들에게 무장애공간을 제공하고자 하는 서비스의 계획도 비미할뿐더러, 실제 대응 자체도 잘 이루어지지 않고 있다[1].

사례로, 생명과 직관적으로 연관되어 있는 횡단보도에서 시각장애인을 위한 음향신호기 시설물 설치 현황 조차도 약 50%정도에 불과하며 오작동이 많고 형식적으로 설치돼 전혀 기능을 하지 못하고 있다[2]. 또한, 은행 창구 업무 시 주변의 도움을 필요로 했던 시각장애인들의 불편함을 해소하기 위해 음성기능을 탑재한 OPT카드를 근래 출시하였으나 오류등의 어려움을 겪고 있다[3].

따라서 본 논문은 기본적으로 사람이 가장 오래 생활하는 주거 공간에 시각장애인이 불편없이 사용할 수 있는 음성이라는 감각과, 거동에 불편함이 있는 장애인의 움직임을 줄여 편의를 주고자 스마트폰을 결합하여 가정의 사물들을 제어하는 음성인식 제어 시스템을

II. 관련연구고찰

Power Voice 사의 Speech Recognition Switch는 회사 기술인 음성인식 엔진 VOIS# C&C가 탑재되어 있는 음성 인식 전등스위치이다. VOIS# C&C는

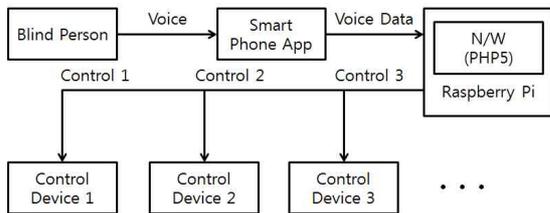
Trigger-word인식엔진을 사용하여 잡음이 심한 여건에서 오인식과 오작동에 대한 음성인식기능을 가능하게 하였고, 단어를 기반으로 하는 것이 아닌 Tri-phone기반의 HMM(Hidden Markov Model) 엔진을 사용하여 사용자는 사물과 대화를 하는 것처럼 자연스러운 발성이 가능해[4]. 표 1과 같이 VOIS# C&C는 주로 잡음에 대한 오류 및 오작동 감출에 초점을 둔 음성인식 기술이라면 본 연구에서 사용한 구글 음성인식 API[5]는 영어와 한글이 섞여 있을 때, 사용자가 올바른 결과 도출하는지 여부를 추정해 처리해준다. 또한 상시 메시지를 대기 받는 VOIS# C&C와는 달리, 사용자가 필요할 때에만 작동을 시켜 값을 받아오는 구글 음성인식 API를 사용하여 효율적인 제어를 목적으로 개발을 할 수 있다.

compare Google voice Recognition API and VOIS# C&C

	Google 음성인식 API	VOIS# C&C
특징	<ul style="list-style-type: none"> 녹음버튼이 있어 원하는 순간의 원하는 단어만 인식기능 수행 	<ul style="list-style-type: none"> 녹음버튼없이 상시 음성대기모드에서 음성인식기능 수행
장점	<ul style="list-style-type: none"> 영어와 한글이 함께 들어간 음성 검색가능 흔치 않은 낱말 인식 가능 개인별 발음차이에 따른 인식을 강화 훈련 	<ul style="list-style-type: none"> 잡음환경에서 오인식이나 오작동없는 음성인식가능 대화하듯이 자연스럽게 발성 가능 특이한 발성패턴이나 사투리에 적용
단점	<ul style="list-style-type: none"> 잡음이나 소음에 대한 오작동 	<ul style="list-style-type: none"> 혼합단어 인식 미흡

III. 음성인식 홈 네트워킹 시스템 설계

음성인식 홈 원격제어 시스템의 구성도는 그림 1과 같다. 시각장애인이 스마트폰 앱에 음성 값을 전송하면 해당 스마트폰은 N/W에 음성데이터로 전송을 하고, 값을 받은 N/W는 Raspberry Pi에 프로그래밍 되어있는 제어 프로그램에 의해 각각 연결되어 있는 장치들을 음성 값에 맞게 제어한다.



system configuration

개발 환경은 어플리케이션을 개발을 위해 Java Android SDK를 사용하였고 서버를 Raspberry Pi에 Raspbian OS를 설치하여 APM(apache + PHP5 + MySQL)을 구축하였다. 제어할 전기 제품은 GPIO에 연결하였으며, 디바이스 드라이버는 C언어로 구현하였다. 사용자 앱에서 음성을 통해 명령을 하면, 서버에 전달되어 해당 디바이스 드라이버가 수행된다.

앱의 실행 흐름은, MainActivity가 실행되고 MainActivity에서는 앱의 실행 여부에 대해서 확인을 한다, 후에 앱을 사용하겠다는 설정을 하고 스마트폰 화면을 활성화 시키면 LockScreenActivity가 실행되고 LockScreenActivity에서는 사용자의 터치를 기다렸다가 터치가 있으면 음성 API를 이용하여 음성 값을 받는다. 음성값이 지정해둔 값과 일치하면 해당 제어에 관한 내용을 라즈베리파이에서 구축하였던 APM으로 값을 보내고, 서버가 받은 제어 값을 판단하여 라즈베리파이에서 LED를 제어하는 프로그램을 실행시킨다 해당 작업은 반복적으로 이루어지다가, 사용자가 스마트폰 화면을 비활성화 하면 ScreenReceiver, ScreenService, RestartReceiver에 의해 화면이 비활성화 되어있는 동안에도 서비스가 죽지 않고 사용자가 스마트폰 화면을 다시 활성화 할 때까지 대기하고, 화면이 활성화 되면 LockScreenActivity를 실행시켜서 다시 음성을 받아들 준비를 한다.

IV. Conclusions

본 연구에서는 구글 음성인식 API를 이용해 시각장애인을 위한 홈 네트워킹 음성 제어 시스템을 구축하였다. 시스템은 오픈소스 하드웨어와 소프트웨어를 활용하였으며, 버튼 제어 방식 대신 음성 인식 방식을 사용하였으며, 앱 자동 실행을 통해 지속성을 확보하였고 가정 내 전기 낭비를 줄일 수 있으며, 앱 사용의 편의성을 통해 실제 활용 가능성을 극대화하였다. 향후 연구 결과를 실제 시각 장애인 가정에 적용시키는 현장 연구가 필요하다.

References

- [1] Korea Blind Union, <http://www.kbuwel.or.kr/>
- [2] Gi-Hwan Kwon, etc., Development of Alarms sound for the visually impaired using GPS, The Road Traffic Authority, 2014.12
- [3] Korea Welfare Foundation for the Visually Handicapped, http://www.hsb.or.kr/client/notice/viw.asp?p_idx=553&cpa ge=2&cate=K01
- [4] Korea Power Voice - VOIS# C&C, <http://www.kpvoice.com/>
- [2] Google Voice recognize API Sample code, http://www.androi dside.com/plugin/mobile/board.php?bo_table=B56&wr_id=262061430964989498826044