



음성 인식 및 시각장애인용 UI를 통한 시각장애인 접근성 어플리케이션

Visually impaired accessibility application through speech recognition and UI for persons with Visual Impairments

**저자
(Authors)** 고한슬, 한성준, 임헌준, 안이삭, 장희진, 이기영, Sung-Jun Park
Han-Seul Ko, 한성준, heon-Jun Lim, Lee-Sak Ahn, Hee-Jin Jang, Ki Young Lee, Sung-Jun Park

**출처
(Source)** [대한전자공학회 학술대회](#) , 2018.6, 1558-1560(3 pages)

**발행처
(Publisher)** [대한전자공학회](#)
The Institute of Electronics and Information Engineers

URL <http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07516114>

APA Style 고한슬, 한성준, 임헌준, 안이삭, 장희진, 이기영, Sung-Jun Park (2018). 음성 인식 및 시각장애인용 UI를 통한 시각장애인 접근성 어플리케이션. 대한전자공학회 학술대회, 1558-1560

**이용정보
(Accessed)** 경기대학교
203.249.3.***
2021/03/14 16:35 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

음성 인식 및 시각장애인을 위한 UI를 통한 시각장애인 접근성 어플리케이션

*고한솔, 한성준, 임현준, 안이삭, 장희진, 이기영
인천대학교 정보통신공학과

e-mail : jjhkhs1214@naver.com, han_sj@naver.com, jumong@naver.com
dltkrskfk@hanmail.net, isk04100@naver.com

Visually impaired accessibility application through speech recognition and
UI for persons with Visual Impairments

*Han-Seul Ko, Sung-Jun Park, heon-Jun Lim,
Lee-Sak Ahn, Hee-Jin Jang, and Ki Young Lee
Department of Information & Telecommunication Engineering
Incheon National University

Abstract

An application to improve the accessibility of mobile apps for the visually impaired is proposed in this paper. The proposed application consists of speech recognition technology, noise cancellation technology, and customized service providing technology based on visual impairment classification. Designing an intuitive UI for the visually impaired and implementing frequently used functions in smart phones is different from existing applications.

I. 서론

최근 스마트폰의 역할이 단순 정보 제공을 넘어 쇼핑, 엔터테인먼트, 안전, 교육 등과 같은 영역으로 확대되어 생활에 필수적인 요소로 자리 잡았지만, 시각장애인의 스마트폰 활용 수준은 현저히 낮은 상태이다. 공공기관 뿐만 아니라 모든 법인이 제공하는 웹사이트에 대한 접근성 제공이 의무화 되었음에도 불구하고 아직도 이를 준수하지 않은 웹사이트들이 대다수이며

실제로 시각장애인을 위한 정보접근권이 제대로 보장되지 않아 Web, App 사용에 많은 불편함을 겪고 있다. 이에 본 논문에서는 시각장애인을 위한 App 접근성 향상을 위해 음성인식 및 noise cancelling 기술을 이용한 다양한 기능 서비스 제공이 가능하게 하고, 기존 키보드 형식 외에 점자 키보드와 음성 인식을 통한 회원가입 진행 및 직관적인 소통의 수단이 되도록 할 수 있는 어플리케이션을 제안한다. 또한 시각 장애 분류에 따른 사용자 각 개인에 맞춘 유니버설 디자인을 통해 쉽고 편리한 서비스를 제공할 수 있을 것이다.

본문에서 위에서 제안한 시스템의 구성과 적용기술에 대해 기술하고 결론에서 시스템의 기대효과 및 발전방향 등을 기술하여 마무리한다.

II. 시각장애인 접근성 어플리케이션

1. 시스템 구성

사용자는 App을 통해 음성으로 서비스를 요청하고, 제공 받을 수 있다. App은 서비스 제공을 위해 DB서버에서 데이터를 제공받는다. DB서버에는 프로그램 데

이더 외에 사용자의 개인정보나 사용자의 음성 등이 저장되어, 사용자 맞춤 서비스에 사용된다. 그림 1은 개발하려는 시스템의 구성도이다.



그림 1. 시스템 구성도

2. 적용기술

2.1 이미지 이진화를 통한 글자 edge 추출

이미지 이진화로 글자의 edge를 검출하여, 상품 이미지에 있는 글자를 추출한다. 이미지에 글자나 로고 등이 과하게 사용된 경우, 콘텐츠의 품질이 저하되고 본래의 정보가 왜곡 될 가능성이 높다. 하지만 쇼핑물 이미지의 특성상, 글자나 로고가 과하게 사용되는 경우가 빈번하다. 따라서 이와 같은 문제를 해결하기 위해, 이미지 이진화 작업을 수행한다. 또한 효율성을 위해, 문자가 있는 것으로 추정되는 영역만 분석하는 방법을 사용한다. 색깔과 윤곽선을 기반으로 텍스트 추정 영역을 찾는다. 이 때, 이미지를 흑백조로 변환하면 글자를 탐지하는데 훨씬 용이해진다.

이미지를 흑백으로 변환하여 윤곽선을 추출한 후 잡음을 제거한다. 잡영이 제거된 윤곽선영역에서, 텍스트로 판단되는 사례와 텍스트가 아닌 것으로 판단되는 사례를 나누어 학습데이터를 구축한다. 이 데이터는 CNN(Convolution Neural Network)를 학습시킬 데이터로 사용이 된다.

2.1.2 딥러닝을 통한 글자 인식

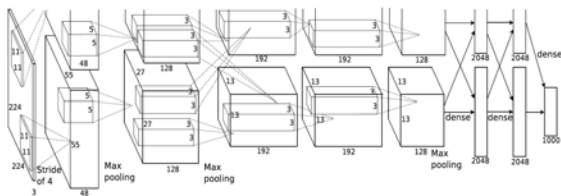


그림 2. AlexNet의 구조

이미지 이진화 작업을 통하여, 추출된 글자 이미지 데이터를 ILSVR에서 이미 검증이 된 AlexNet의 구조

와 weight를 사용한 CNN을 구성하여 전이학습을 통한 글자를 추출하는 것을 넘어 글자를 인식 할 수 있도록 하였다.



그림 3. openCV를 이용한 이미지 이진화 결과

2.2 음성인식

2.2.1 TTS (Text To Speech)

TTS는 음성을 문자로 변화하는 기술로, AWS Amazon polly API를 활용하여 위와 같이 추출된 글자를 TTS 기술을 사용하여 시각장애인들에게 제공한다.

2.2.2 STT (Speech To Text)

시각장애인들이 서비스를 요청할 경우, 인간의 언어를 기계적으로 분석하고 컴퓨터가 이해할 수 있는 형태로 변환하는 STT기술과 자연어 처리 기술을 이용한다. 이 기술 또한 Amazon polly를 활용하여 구현한다.

2.2.3 음성인식의 정확도 향상

음성인식의 정확도 향상을 위하여 사용자의 성문을 저장한다. Noise Cancelling을 통한 음성인식의 정확도를 향상시킨다.

2.3 사용자 환경 맞춤 서비스 제공

App을 처음 사용하게 될 때, 사용자는 본인의 시각장애 등급 분류를 입력하고 DB서버에 저장한다. 그 후 음성을 통한 본인 확인으로 App에 로그인 하면, 본인의 시각장애 분류에 적합한 환경과 서비스를 제공한다. 음성인식 외에 간단한 제스처를 제공하여 편리성을 증가시키고 필요에 따라 사용자는 직접 제스처를 만들어 등록할 수 있다.

III. 결론 및 발전방향

본 논문에서는 시각장애인들의 모바일 접근성을 향상시키기 위한 이미지 이진화를 통해 글자 edge 추출, 음성인식 기술과 사용자 환경에 따른 맞춤 서비스를 제공하는 기술들을 제시하였다. 또한 음성등록과 Noise cancelling을 통한 음성인식의 정확도를 향상시켜 음성인식기술의 불확실한 정확성을 개선시켰다.

향후 발전방향으로는 시각장애인들의 스마트폰 활용 및 접근성을 향상시켜 장애인들의 정보화 수준을 상승시키고 모바일 Web, App 접근성에 대한 고찰 및 장애인들의 접근성 향상 부분에 많은 관심을 갖게 되기를 기대한다. 시각장애인이나 노인들에서 확대되어 외국인 여행객들에게 맞춤 서비스를 제공할 수 있는 음성 어플리케이션이나 교육용 어플리케이션으로 연계하여 극대화된 활용분야에 대해 생각해볼 수 있다.

ACKNOWLEDGMENT

이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 한국연구재단-현장맞춤형 이공계 인재양성 지원사업의 지원을 받아 수행된 연구임.

(NRF2017H1D8A1029391)

참고문헌

- [1] D. Cireşan, U. Meier, and J. Schmidhuber.
Multi-column deep neural networks for image
classification. Arxiv preprint arXiv:1202.2745,
2012.
- [2] Convolutional Neural Network Benchmarks:
<https://github.com/jcjohnson/cnn-benchmarks>
- [3] 이장래, SQL Server, 루비페이퍼, 2017.
- [4] 이명희, A study on web accessibility of people
with visual impairment, 조선대학교 정책대학원
석사학위논문, 2006.