

Alexa를 이용한 대학안내 시스템

김태진 · 김동현*

The University Guidance System using the Alexa

Tae Jin Kim · Dong Hyun Kim*

Division of Computer Engineering, Dongseo University, Busan 47011, Korea

요 약

음성인식 기술은 사용자의 음성을 인식하여 명령을 실행하는 기술로 현재는 자연어 처리 기법과 접목되어 인공지능 음성인식 서비스로 제공되어 지고 있다. 이러한 인공지능 음성인식 서비스는 IoT기기를 제어하거나 뉴스 또는 날씨와 같은 정보 제공 분야에서 활용되고 있다. 정보 제공 분야의 하나인 대학 정보는 주로 웹을 통하여 제공되나 너무 많은 데이터를 제공하기 때문에 사용자가 신속하게 원하는 정보를 검색할 수 없는 문제가 있다. 이 논문에서는 사용자가 찾고자 하는 정보를 음성으로 인식하고 제공하기 위한 대학 안내 시스템을 설계하고 구현한다. 대학정보를 각 주제별로 분류하고 이를 제공하기 위한 람다 함수를 설계한다. 그리고 알렉사 스킴 키트를 이용하여 노드제이에스로 구현된 모듈을 아마존웹서비스에 업로드하여 음성인식을 이용하여 대학 정보를 제공한다.

ABSTRACT

The voice recognition technology is to recognize the voice of an user and execute the command. Recently, the voice recognition is evolving to the artificial intelligence voice recognition by adding the scheme of the natural language processing. The AI voice recognition is exploited to control the IoT devices or provide the information, such as the news or the wether. The University Information which is one of fields serviced by the information provider is mainly presented on the web. However, since too much information are presented on the web, it is difficult for an user to find efficiently the specific information which the user want to know. In this paper, we design and implement the university guidance system to recognize the user voice searching the information and provide the result using the voice. To do this, we classify the university data and design the lambda function to provide the data.

키워드 : 알렉사, 안내시스템, 대학시스템, 음성명령시스템, 아마존웹서비스

Key word : Alexa, Guidance System, University System, Voice Commanding System, Amazon Web Service

Received 21 October 2017, Revised 30 October 2017, Accepted 02 November 2017

* Corresponding Author Dong Hyun Kim(E-mail:pusrover@dongseo.ac.kr, Tel:+82-51-320-1801)

Division of Computer Engineering, Dongseo University, Busan, 47011, Korea

Open Access <https://doi.org/10.6109/jkiice.2017.21.11.2061>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

음성은 인간이 가지고 있는 가장 기본적인 기능중의 하나로 다른 사람 또는 사물에게 의사를 전달할 수 있는 중요한 수단 중 하나이다[1]. 컴퓨터에게 의사를 전달하기 위하여 음성을 이용하면 매우 편이하며 키보드로 명령을 입력하는 것보다 매우 빠르게 입력할 수 있다. 특히, 마이크이외의 부착 공간이 필요하지 않기 때문에 자동응답장치, 자동항법장치, 통제시스템 등 다양한 분야에서 활용되고 있다.

음성 인식 기술은 음절, 모음 등의 단어 단순 인식 수준에서 출발하였으나 음성 인식율을 높이기 위하여 다양한 음성 모델 및 최소분류에러(Minimum Classification Error) 등의 기법이 개발되었다. 현재는 자연어 처리 기술이 접목되어 인공지능 음성인식 서비스가 일반인들에게 제공된다[2]. 해외에서는 아마존의 에코가 서비스되고 있으며 국내에서는 누구(NUGU)와 지니(Genie) 등이 서비스되고 있다[2]. 이 서비스들은 음성으로 주어진 명령을 인식하여 뉴스 또는 날씨 등의 정보를 사용자에게 제공하거나 또는 멀티미디어 데이터를 검색하여 제공한다. 또한 IoT 기능을 가진 기기들과 연결하여 음성으로 제어할 수 있게 한다.

아마존 에코는 마이크와 스피커/모니터 일체형의 에코기기를 활용하여 서비스를 제공하며 특히 개발용 인터페이스인 알렉사 스킬 키트(Alexa Skill Kit, ASK)를 개발자에게 제공한다. 개발자는 알렉사 스킬 키트를 이용하여 노드제이(Node.js) 또는 파이썬(Python) 언어로 작성된 스킬 코드를 아마존 웹 서비스에 업로드하여 공개할 수 있다. 그리고 일반 사용자는 공개된 스킬 코드를 이용하여 IoT 기기를 제어하거나 또는 맞춤형 정보를 서비스 받을 수 있다. 알렉사 스킬 키트를 이용하여 개발된 스킬은 130여개의 음악, 뉴스, 날씨 그리고 기업 안내 등의 다양한 분야에서 정보 서비스를 제공하고 있다.

정보 서비스의 대표적인 응용 중 하나인 대학 정보는 주로 웹을 이용하여 제공되어 진다. 특정 대학에 대한 정보를 파악하기 위하여 사용자는 대학 홈페이지 또는 대학정보 알리미 서비스에 컴퓨터를 이용하여 접속한다. 그리고 키워드 검색 기능을 이용하거나 또는 페이지 브라우징을 통하여 원하는 정보를 검색한다.

그러나 웹을 이용한 대학정보는 사용자가 신속하게

원하는 정보를 검색할 수 없는 문제가 있다. 키워드 검색 기능을 이용할 때 정확한 키워드를 사용하지 않으면 많은 수의 데이터가 검색된다. 메인 페이지에서 제공하는 정보의 양이 많을 경우에 찾고자 하는 정보의 위치를 정확히 모르면 많은 수의 페이지 클릭이 필요하다. 예를 들어, 부산 D 대학의 메인페이지에서 1차 메뉴로 제공되는 메뉴의 수는 100여개 정도이며 사용자가 원하는 정보를 찾기 위하여 페이지를 최대 100여 번 방문해야 한다.

이 논문에서는 이러한 문제를 해결하기 위하여 사용자가 찾고자 하는 대학 정보를 음성으로 인식하고 검색된 결과를 음성으로 제공하는 대학 안내 시스템을 설계하고 구현한다. 제공하는 대학 정보는 대학소개, 전공, 대학활동, 입학정보 등 주제별로 분류하고 세부 데이터를 작성한다. 그리고 서비스를 위한 람다함수를 설계하고 알렉사 스킬 키트와 노드제이에스를 이용하여 구현한다. 구현된 람다 함수는 아마존웹서비스에 업로드 하여 사용자에게 대학 정보를 제공한다.

이 논문은 다음과 같이 구성된다. 2장에서는 음성안내시스템의 관련 연구를 기술한다. 3장에서는 대학 정보 데이터의 분류 기법과 람다 함수를 설계하고 4장에서는 알렉사 스킬 키트로 구현된 람다함수의 구현 결과를 제시한다. 마지막으로 5장에서 결론 및 향후 연구를 기술한다.

II. 관련연구

2.1. 알렉사 스킬 키트

에코는 2015년에 미국 아마존에서 시작한 서비스로 에코 마이크를 통하여 입력된 음성 명령을 인식하여 처리한다. 인식된 음성 명령을 처리하는 각각의 서비스는 스킬(Skill)로 아마존웹서비스에서 수행되며 알렉사 스킬 키트는 사용자 스킬을 개발하기 위한 개발, 인터페이스다.

알렉사 스킬 키트는 크게 4가지로 분류된다. 첫 번째는 커스텀 스킬(custom skill) API이다. 커스텀 스킬은 임의의 사용자 요구를 처리하기 위한 스킬로 웹으로부터 정보를 획득하거나 우버 또는 피자헛 등의 타 서비스와 통합하기 위하여 사용할 수 있다. 두 번째는 스마트 홈 스킬(smart home skill) API다. 스마트 홈 스킬은

전등, 도어락 그리고 카메라 등 클라우드와 접속될 수 있는 IoT 기기들을 사용자가 제어할 수 있게 지원하는 스킬이다.

세 번째는 비디오 스킬(videl skill) API다. 비디오 스킬은 클라우드 기반의 비디오 서비스를 사용자가 제어할 수 있도록 지원하는 스킬이다. 마지막으로 플래쉬 브리핑 스킬(flash briefing skill) API이다. 이 스킬은 뉴스 또는 공공기관의 특보 또는 속보 서비스를 통합하여 지원하기 위한 스킬이다.

2.2. 음성 인식을 이용한 기존 응용

[3]에서는 음성인식시스템과 다목적 시각 언어를 연동한 실시간 쇼핑 시스템을 제안한다. 하지만 홈쇼핑 시스템의 음성인식 리모컨이 아직 일반화 되지 않아 많은 어려움이 있다. 자동차 환경에서 음성인식을 수행하는 장치 중 AVN 보드 내에 소프트웨어 형태로 설치되어 주 프로그램에서 송출되는 음성인식기 제어 신호를 수신하는 시스템을 [4]에서 제시하였다. 그러나 자동차 주행속도에 따라 소음량 때문에 음성인식 성공률이 낮다.

[5]에서는 AVR에 입력된 음성을 확인한 후 처리된 값에 대한 검증을 통하여 문을 여는 디지털 도어락을 제안하였다. 디지털 도어락을 음성으로 한다면 비밀번호를 말하게 되어 보안상의 문제가 되어 보안 문제 해결방안에 연구가 필요하다. VoiceXML의 레코드 태그 기능에 기반하여 음성인식 시스템을 이용하는 철도에 약서비스를 [6]에서 제시하였다. 조용한 사무실이 아닌 잡음이 있는 기차역, 거리 환경, 터널 통과 할 때 음성인식이 어려운 문제가 있다.

라즈베리파이, 마이크, 거울, 인터넷을 이용한 스마트 미러를 [7]에서 제시하였다. 일반적인 가정 소음에 무리가 없지만 조금 더 자극적인 소리가 발생하면 사용자의 명령보다 자극적인 소리에 반응하는 문제가 있다.

III. 대학 안내 시스템 설계

3.1. 시스템 구조

본 논문에서 제안하는 아마존 에코를 이용한 대학안내시스템은 WiFi가 있는 대학건물장소에서 대학정보

를 아마존 에코를 이용하여 음성으로 정보를 사람들에게 제공하는 것을 제안한다. 본 시스템은 그림 1과 같이 구성되며, 아마존 웹 서비스에 Node.js를 이용하여 작성된 람다 함수를 스킬로 업로드한다. 그리고 사용자가 음성으로 대학 정보를 요구하면 아마존 웹 서비스에 업로드된 스킬이 검색하고 결과를 사용자에게 음성으로 제공한다.

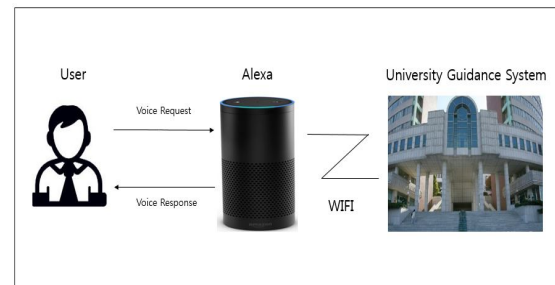


Fig. 1 system concept

본 시스템의 핵심은 사용자가 대학내에서 대학에 관한 정보를 찾고자 할 때 스마트폰, 태블릿 등을 이용하여 대학홈페이지를 접속하지 않는다. 그리고 에코가 구비되어 있는 장소에서 음성으로 정확한 정보를 빠른 시간 내에 찾을 수 있도록 한다.

3.2. 대학 정보

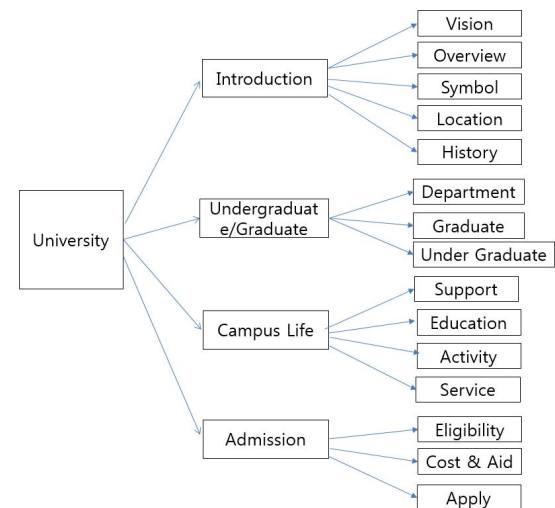


Fig. 2 university information classification

그림 2는 대학안내시스템에서 제공하는 대학정보의 분류를 보여준다. 대학정보는 크게 대학소개, 대학/대학원, 대학생활 그리고 입학인 5가지 영역으로 구성한다.

대학소개는 대학이념, 대학 조직도 등 대학을 구성하는 기본적인 정보를 알려주는 주제이며 대학의비전, 대학개요, 대학상징, 대학위치 그리고 대학역사로 구성한다. 대학/대학원은 4개의 대학원, 2개 단과대학 그리고 15개 학부의 관한 정보를 알려주는 주제이며 학부, 대학, 대학원으로 세분한다.

대학생활은 학생들이 대학생활에 필요한 학사일정, 장학, 기숙사 정보 등의 주제를 다루며 학사지원, 대학교육, 대학활동 그리고 대학 질문으로 구성한다. 대학입학은 대학을 입학에 필요한 자격, 비용 등의 정보를 알려주는 역할을 하는 주제이며 대학입학 자격, 비용 그리고 지원으로 구성한다.

3.3. 램다구성

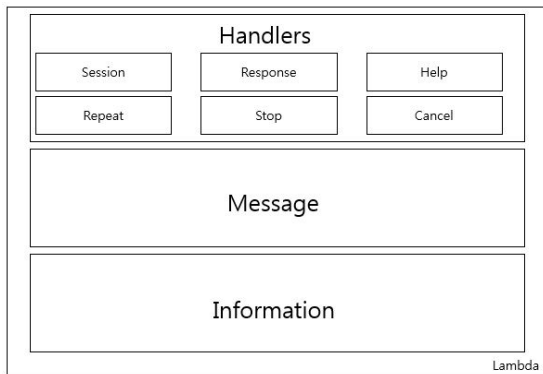


Fig. 3 lambda design structure

그림 3은 아마존 웹 서비스의 램다의 구성도이며 핸들러, 메시지 그리고 정보인 3개의 램다 함수로 구성된다. 핸들러 램다는 아마존 스킬 키트에서 제공하는 API를 활용하여 알렉사 스킬 사용, 메시지 처리, 대학정보를 음성으로 처리해 준다. 이 램다 함수는 세션, 응답, 도움, 반복, 멈춤 그리고 취소의 기능으로 구성된다.

세션은 사용자가 찾고자하는 정보를 응답기능으로부터 가져와 알렉사가 사용자에게 음성으로 제공하는 역할을 한다. 응답은 사용자가 찾고자하는 정보의 내용

을 정보람다와 비교하여 단어를 찾으며 일치하는 단어가 있을 경우 사용자에게 정보를 제공 하고 단어가 없을 경우 정보가 없음을 사용자에게 알리는 역할을 한다. 도움, 반복, 멈춤 그리고 취소의 기능은 메시지 램다의 내용을 처리하는 역할을 한다.

메시지 램다는 사용자에게 알렉사의 도움의 메시지, 반복메시지, 중단 메시지 그리고 취소 메시지의 내용을 저장하고 있는 역할을 하는 램다이다. 정보 램다는 사용자가 찾고자 하는 정보를 대학정보를 주제별로 나누어 주제에 맞게 세부사항을 작성하여 정보를 가지고 있는 역할을 하는 램다이다.

IV. ASK를 이용한 구현

아마존을 이용한 대학안내 시스템은 앞서 보여드린 대학정보 분류와 아마존 웹 서비스의 램다를 이용하여 사용자가 찾고자 하는 정보를 음성으로 제공하는 것을 구현하였다. 아마존 에코를 이용한 대학안내시스템은 WiFi가 있는 장소에서 구동이 가능하며 구현결과는 그림 4와 같이 나타내도록 한다.

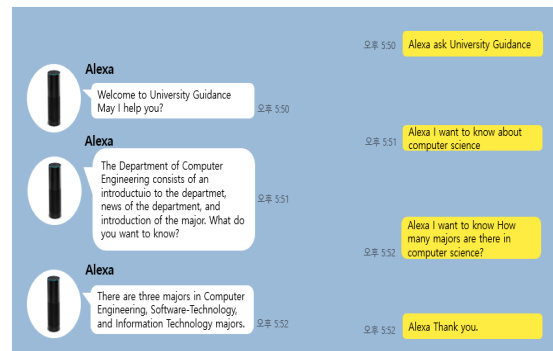


Fig. 4 finding the information of div. of computer engineering using the implemented skill of the Alexa

그림 4는 컴퓨터공학부의 전공 정보를 알기위해 알렉사를 이용하여 정보를 찾는 과정을 대화형식으로 보여준다. 그림 4에서 보듯이 Alexa로 시작하는 명령어에 대하여 아마존 웹 서비스에 구현된 대학 안내 시스템 스킬이 처리하여 검색된 결과를 음성으로 제공한다.

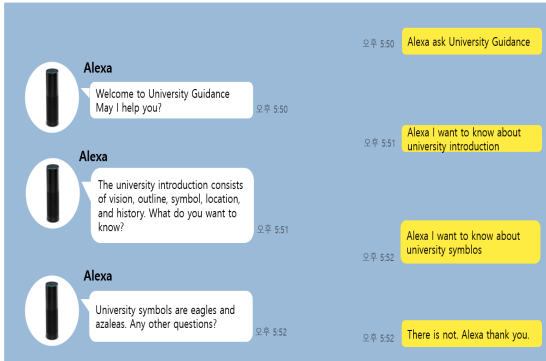


Fig. 5 find the symbol of the university using the Alexa

그림 5는 대학소개의 대학상징에 대한 정보를 알기 위해 알렉사를 이용하여 정보를 찾는 과정을 대화 형식으로 보여준다.

```
const handlers = {
  'NewSession': function () {
    this.attributes.speechOutput = this.t('WELCOME_MESSAGE', this.t('SKILL_NAME'));
    this.emit('ask', this.attributes.speechOutput, this.attributes.repromptSpeech);
  },
  'ResponseIntent': function () {
    const itemSlot = this.event.request.intent.slots.item;
    let itemName;
    if (itemSlot && itemSlot.value) {
      itemName = itemSlot.value.toLowerCase();
    }
    const dongTitle = this.t('DISPLAY_CARD_TITLE', this.t('SKILL_NAME'), itemName);
    const mydongseo = this.t('dongseo');
    const dongseo = mydongseo[itemName];
    if (dongseo) {
      this.attributes.speechOutput = dongseo;
      this.attributes.repromptSpeech = this.t('dongseo_REPEAT_MESSAGE');
      this.emit('askWithCard', dongseo, this.attributes.repromptSpeech, cardTitle, dongseo);
    } else {
      let speechOutput = this.t('dongseo_NOT_FOUND_MESSAGE');
      const repromptSpeech = this.t('dongseo_NOT_FOUND_REPROMPT');
      if (itemName) {
        speechOutput += this.t('dongseo_NOT_FOUND_WITH_ITEM_NAME', itemName);
      } else {
        speechOutput += this.t('dongseo_NOT_FOUND_WITHOUT_ITEM_NAME');
      }
      speechOutput += repromptSpeech;
      this.attributes.speechOutput = speechOutput;
      this.attributes.repromptSpeech = repromptSpeech;
      this.emit('ask', speechOutput, repromptSpeech);
    }
  }
};
```

Fig. 6 the part of handler lambda code

그림 6에서 보여주는 핸들러 람다는 대학정보를 음성으로 처리하는 역할을 하며 함수는 세션, 응답함수로 구성되어 있다.

세션 함수는 사용자가 찾고자 하는 정보를 응답세션로부터 데이터를 전달받아 사용자에게 음성으로 정보를 제공하는 역할을 한다. 응답 함수는 사용자가 찾고자 하는 정보의 내용을 정보 람다와 비교하여 단어를 찾으려 일치하는 단어가 있을 경우 사용자에게 정보를 제공하고 단어가 없을 경우 정보가 없음을 사용자에게 알리는 함수 역할을 한다.

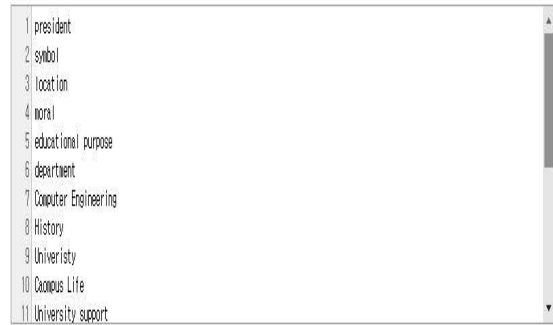


Fig. 7 the part of the word list recognized by the implemented skill

그림 7은 대학안내시스템의 사용자가 알렉사에게 질문을 할 때 알렉사가 반응하는 단어의 모음이다중 일부를 보여준다.

V. 결론

정보서비스의 대표적인 응용 중 하나인 대학정보는 주로 웹을 이용하여 제공되어 진다. 그러나 웹을 이용한 대학정보는 사용자가 신속하게 원하는 정보를 검색할 수 없는 문제가 있다. 정확한 키워드를 사용하지 않으면 많은 수의 데이터가 검색될 수 있고 메인페이지에서 제공하는 정보의 위치를 정확히 모르면 많은 수의 페이지 클릭이 필요하다.

이 논문에서는 이러한 문제를 해결하기 위하여 사용자가 찾고자 하는 대학 정보를 음성으로 인식하고 검색된 결과를 음성으로 제공하는 대학 안내 시스템을 설계하고 구현하였다. 제공하는 대학 정보는 대학소개, 전공, 대학활동, 입학정보 등 주제별로 분류하고 분류에 따라 세부 데이터를 작성하였다. 그리고 서비스를 위한 람다함수를 설계하고 스킬로 구현하였다. 구현된 스킬은 아마존 웹 서비스에 업로드 하여 사용자에게 대학 정보를 제공한다.

본 연구의 향후 과제로는 질문자의 발음과, 주위 환경에 영향을 받아 에코의 음성 인식이 떨어지는 문제를 해결하기 위한 기법을 개발하는 것이다.

ACKNOWLEDGMENTS

Following are results of a study on the “Leads Industry-university Cooperation” Project, supported by the Ministry of Education

REFERENCES

- [1] D. S. Song and D. G. Jeong, “Design of a voice-recognized elevator control system,” In *Proceedings of Korean Institute of Information Technology Summer Conference*, Kyung-ki Do, Suwon, pp. 240-243, 2010.
- [2] D. J. Kim, G. R. Lee and H. W. Kim, “Analysis of speech recognition technology and personal assistant service using AI,” In *Proceedings of The Korean Institute of Communication and Information Sciences Summer Conference*, Jeju Do, Jeju, pp. 213-214, 2017.
- [3] Y. J. Kim, “The Real-time Shopping System using Multipurpose Visual Language with Voice Recognize,” *International Journal of The Korea Academia-Industrial cooperation Society*, vol. 16, no. 6, pp. 4164-4169, Jun. 2015.
- [4] S. J. Lee and S. M. Kang, “A Study of Automatic Evaluation Platform Speech Recognition Engine in the Vehicle Environment,” *International Journal of Korea Institute of Communication Sciences*, vol. 37, no. 7, pp. 538-543, Jul. 2012.
- [5] T. M. Kwak and S. G. Moon, “Digital Doorlock with Voice Recognition,” In *Proceedings of Korean Institute of Information Technology Summer Conference*, Gwangju, pp. 345-348, 2012.
- [6] B. S. Kim and S. H. Kim, “A Study on Realization of Speech Recognition System based on VoiceXML for Railroad Resevation Service,” *International Journal of Korean Society for Railway*, vol. 14, no. 2, pp. 130-136, Apr. 2011.
- [7] J. B. Park, D. G. Jeong, J. Y. Cho, and B. G. An, “Design of IoT Smart Mirror System based on Speech Recognition,” In *Proceedings of Communication and Information Sciences Fall Conference*, Seoul, pp. 380-381, 2016.



김동현(Dong Hyun Kim)

부산대학교 컴퓨터공학과 공학박사

※관심분야 : 데이터베이스, 공간데이터베이스, GIS, AI 음성인식, 증강현실



김태진(Tae Jin Kim)

동서대학교 컴퓨터공학부

※관심분야 : 데이터베이스, AI음성인식