

음성인식을 이용한 URL Navigator 개발

전한길, 홍인숙, 김윤중
 한밭대학교 컴퓨터공학과
 e-mail:kirikko@naver.com

A Development of an URL Navigator using Speech Recognition

Han-Gil Jeon, In-Suk Hong, Yoon-Joong Kim
 **Dept of Computer Engineering, Hanbat University

요 약

기존의 인터넷 익스플로러는 높은 보급률의 이점을 지니고 있지만 인터넷을 이용하기 위한 기본 지식의 필요성과 키보드, 마우스와 같은 장치를 이용한 입력방식은 장애인 및 고령층의 사용자들이 이용하기에는 불편하다.

이와 같은 문제점을 해결하고자 본 논문에서는 웹 서비스 기술을 이용하여 기존의 음성인식 시스템을 웹 서비스로 재구성하고 클라이언트가 음성인식 웹 서비스를 요청하고 음성명령이 가능하도록 인터넷 익스플로러에 음성명령이 가능한 Navigator를 Toolbar 형태의 프로그램을 구현하였다.

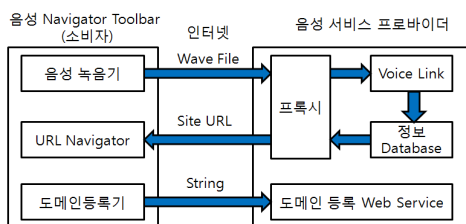
본 시스템은 클라이언트가 사용하는 음성 Navigator Toolbar와 웹 서비스를 통해서 음성 서비스를 제공하는 음성 서비스 프로바이더로 구성된다. 음성 Navigator Toolbar는 인터넷 환경에서 음성데이터를 바이너리 포맷형식인 DIME을 이용하여 음성 서비스 프로바이더로 전송하고 반환 결과를 가지고 URL Navigator를 통해 인터넷 페이지를 이동시킨다. 음성 서비스 프로바이더는 사용자가 전송한 음성데이터를 인식기를 통해 인식하고 결과를 가지고 정보 Database를 검색하여 실제 인터넷 주소를 사용자에게 반환한다.

1. 서론

전 세계적으로 유비쿼터스 기반 환경으로 변화되고 있으며, 그 기반에 자리 잡고 있는 인터넷의 사용자 수는 꾸준히 증가하고 있다. 현재 인터넷은 넓은 보급률로 인해 이용에 편리함을 제공하는 반면 이용을 위한 학습이나 입력방식의 제한으로 인해 다양한 사용자 환경에 관계없이 서비스를 제공하지 못하는 문제점이 있다.

이에 대한 해결 방안으로 기존에 이용하던 인터넷 환경을 유지시켜면서 다양한 사용자 환경에 관계없이 서비스를 제공해하기 위한 방안의 하나로 음성 인터페이스에 대한 연구가 진행되고 있다.

본 논문에서는 음성인터페이스의 이용에 대한 일원으로 기존의 음성인식 시스템을 웹 서비스로 제공하고 클라이언트가 사용하는 Navigator Toolbar에서 음성명령이 가능하도록 기능을 구현하기 위해 다음과 같은 시스템구성도를 가진다.



(그림 1 전체 시스템 구성도)

본 시스템은 클라이언트가 사용하는 음성 Navigator Toolbar와 음성 인식에 관련된 서비스를 제공하는 음성 서비스 프로바이더로 구성된다. 음성 Navigator Toolbar는 음성 녹음기, URL Navigator, 도메인 등록기로 구성된다. 음성 서비스 프로바이더는 프록시와 VoiceLick, 정보 Database, 도메인 등록 Web Service로 구성된다.

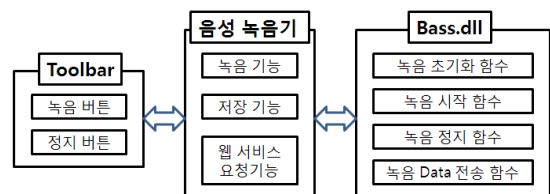
2. 음성 Navigator Toolbar 설계 및 구현

본 장에서는 사용자 환경에서 음성명령 처리가 가능한 음성 Navigator Toolbar를 설계 구현하였다.

Navigator Toolbar는 음성녹음기, URL Navigator, 도메인 등록기로 구성된다.

2.1 음성 녹음기

본 절에서는 음성 Navigator Toolbar에서 사용자의 음성 입력을 녹음하는 음성 녹음기를 구현하였다.



(그림 2 음성 녹음기의 구성도)

음성 녹음기는 Bass.dll 라이브러리의 네 가지 기능을 기반으로 녹음버튼, 정지버튼 두 가지 Visual 인터페이스에 녹음기능, 저장기능, 웹 서비스 요청 기능을 구현하였다.

녹음 기능은 사용자가 입력한 음성을 wav 파일 포맷으로 녹음하는 역할을 담당한다. 이때 음성을 wave 형태의 파일로 초기화 하기 위한 포맷의 기본 값은 샘플링 값 8khz, 비트(bit) 샘플링 수 16비트, 모노(Mono)로 기본설정된다.

저장 기능은 사용자가 녹음한 wav 파일을 사용자의 컴퓨터에 저장하는 역할을 담당하며 사용자가 녹음을 완료한 후에 Toolbar의 정지버튼을 클릭하면 Bass.dll 라이브러리에서 제공하는 녹음 정지 함수를 호출한다.[1] 녹음 함수가 종료되면 녹음 기능을 통해서 입력된 wave 형태의 데이터를 사용자 컴퓨터의 C드라이브에 Toolbar.wav 라는 파일로 저장한다.

웹 서비스 요청기능은 녹음된 음성파일을 웹 서비스를 통해 음성 서비스 프로바이더에 전송하기 위해 음성파일을 DIME으로 패키징하고 웹 서비스를 호출한다.[1]

2.2 URL Navigator

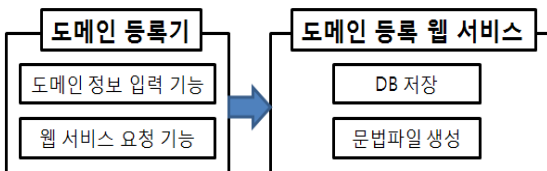
본 절에서는 음성 Navigator Toolbar에서 사용자로부터 입력받은 음성 데이터를 음성 서비스 프로바이더에서 인식한 후 반환되는 결과에 따라 해당하는 URL로 이동 시켜주는 기능을 담당하는 URL Navigator를 구현하였다.

URL Navigator는 일반적인 윈도우처럼 다뤄지는 컨테이너를 가진 COM 객체인 BandObject를 기반으로 하여 Toolbar 형태의 인터페이스에 Navigator기능을 구현하였다.

2.3 도메인 등록기

본 절에서는 음성 Navigator Toolbar에서 사용자가 직접 음성 서비스 프로바이더 내의 음성 인식기에서 사용하고 자 하는 도메인명과 URL을 직접 등록할 수 있도록 구현하였다.

도메인 등록기는 도메인 정보 입력 기능과 웹 서비스 요청 기능으로 구현되어 있다.



(그림 3 도메인 등록기 구성도)

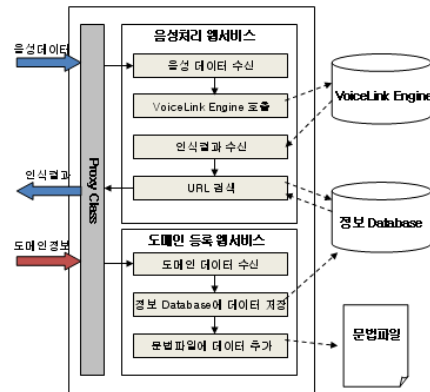
도메인 정보 입력 기능은 사용자가 도메인명과 URL을 입력하고 웹 서비스를 요청할 수 있도록 윈도우 폼으로 구현하였다.

웹 서비스 요청 기능은 사용자가 신규 도메인 정보를 입력한 후 요청 버튼을 클릭하면 도메인 값과 URL값을 음성 서비스 프로바이더의 정보 Database와 도메인 등록 웹

서비스에 전송할 수 있도록 구현하였다.

3. 음성 서비스 프로바이더 구현

본 장에서는 사용자의 음성을 인식하고 반환하는 서비스를 제공하기 위한 음성 서비스 프로바이더를 구현하였다. [그림 4]와 같이 음성 서비스 프로바이더는 음성처리 웹서비스와 도메인 등록 웹서비스로 구성되어 있으며 VoiceLink Engine과 정보 Database, 문법파일을 포함한다.



(그림 4 음성 서비스 프로바이더 구조도)

3.1 음성처리 웹서비스

음성처리 웹서비스는 사용자의 음성을 인식하고 인식결과에 해당하는 URL을 반환하기 위해 음성데이터 수신, VoiceLink Engine 호출, 인식결과 수신, URL 검색의 4개의 기능으로 구성되어 있다.

1) 음성데이터 수신

음성 데이터 수신은 음성 Navigator Toolbar는 음성처리 웹서비스에게 DIME 메시지 형태로 패키징하여 전달한 음성 데이터를 Microsoft.Web.Services.Dime.Dime Attachment 객체를 이용하여 녹음한 음성데이터를 수신 받는다.[2]

2) VoiceLink Engine 호출

음성처리 웹서비스의 두번째 기능은 VoiceLink Engine을 호출하고 wav파일을 전송한다.

3) 인식결과 수신

인식결과 수신 기능은 VoiceLink Engine으로부터 인식된 결과를 수신 받는다.

4) URL 검색

마지막 기능으로 정보 Database로부터 인식 결과에 해당하는 URL을 검색하여 음성 Navigator Toolbar에게 전달한다.

3.2 VoiceLink Engine

VoiceLink Engine은 음성 서비스 프로바이더에서 사용하는 음성인식 엔진으로 사용자가 입력한 음성로 인식을 수행한 후 인식결과를 반환한다.[3]

3.3 정보 Database

정보 Database는 음성 서비스 프로바이더 내의 VoiceLink 엔진과 도메인 등록 웹 서비스에서 사용하는 도메인 정보를 저장하고 있으며 다음의 두 가지 기능을 수행한다.

첫 번째 기능은 정보 Database에서는 VoiceLink Engine의 인식결과를 도메인 필드와 비교하여 해당 URL을 Procy Class로 전송하여 준다.

두 번째 기능은 도메인 등록 웹 서비스에서 신규 도메인을 등록을 위해 도메인과 URL을 입력하면 정보 Database에 신규 정보를 등록한다.

3.4 도메인 등록 웹 서비스

도메인 등록 웹 서비스는 사용자가 등록한 도메인명과 URL을 정보 Database와 문법파일에 추가하는 기능을 수행한다. VoiceLink Engine은 사용자가 발음한 음성을 인식하기 위해서 문법파일을 참조한다. 따라서 인식하고자 하는 도메인명과 URL이 정보Database와 문법파일에 등록되어야 한다.

문법파일은 VoiceLink Engine에서 인식을 수행하기 위해 필요한 단어들과 단어들의 전이 상태를 표현하는 WordNetwork로 구성된다. 다음의 과정을 통해서 문법파일에 등록된다.[4]

- 1) Word Network 파일 로드 : Open() 메서드를 이용하여 Word Network 파일을 로드한다.
- 2) 정보 추출 : ReadVersion() 메서드에서 버전정보를 추출하고 ReadCount() 메서드에서 단어와 전이 상태를 개수를 추출한다. 그리고 ReadData() 메서드와 ReadLink() 메서드에서 단어와 전이 상태의 리스트를 추출한다.
- 3) 단어 추가 : AddItem() 메서드를 이용하여 전 단계에서 추출한 단어 리스트에 해당 단어를 추가한다.
- 4) Word Network 재구성 : 새로 추가한 단어를 반영하도록 SetItem() 메서드와 SetLink() 메서드를 이용하여 Word Network를 재구성한다.
- 5) Word Network 파일 저장 : Save() 메서드를 이용하여 새로 추가한 단어를 반영한 Word Network 파일을 저장한다.

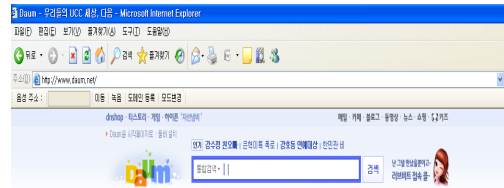
4. 실험 및 결과

본 연구에서 구현한 시스템의 성능을 분석하기 위해 음성인식 기능과 도메인 등록과정 두 가지 실험을 하였다.

4.1 음성 인식 기능 실험

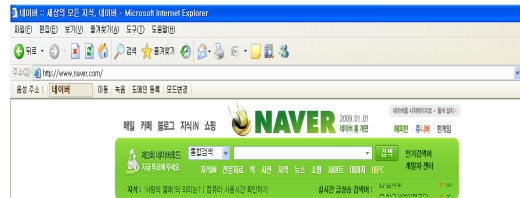
음성 인식 기능 실험은 사용자가 입력한 음성정보를 시스템에서 인식하고 그 반환 결과에 대한 실험으로 다음과 같은 실험 과정으로 진행하였다.

- 1) 클라이언트의 컴퓨터에 음성 URL Navigator Toolbar를 설치한다.
- 2) 설치가 끝나면 인터넷 익스플로러를 실행한다.
- 3) 사용자는 모드변경 버튼 클릭을 통해 음성 모드로 변환한다.



(그림 6 URL Navigator 설치 후 인터넷 익스플로러)

- 4) 녹음 버튼을 클릭한 후 음성을 녹음한다.
 - 마이크를 통해 “네이버” 이라고 음성 입력
- 5) 음성 녹음이 종료되면 종료 버튼을 클릭한다.
 - 종료버튼을 클릭하면 “네이버”가 녹음된 wav 파일을 DIME으로 패키징한 후 SOAP 프로토콜로 음성 서비스 프로바이더로 전송



(그림 7 음성 인식 결과)

- 6) 인식결과를 반환하고 페이지를 이동시켜 준다.
 - 음성 서비스 프로바이더는 “네이버”를 인식한 후 정보 Database에서 “네이버”의 URL 필드 값인 <http://www.naver.com>을 URL Navigator에 반환한다.
 - URL Navigator는 [그림 7]과 같이 <http://www.naver.com>으로 페이지를 이동시켜주는 것을 확인 할 수 있다.

이와 같은 실험 결과 클라이언트의 음성 녹음기와 Navigator 음성 서비스 프로바이더의 음성인식기 웹 서비스와 정보 Database가 정상적으로 동작되고 있음을 확인할 수 있었다.

5.2 도메인 등록 실험

도메인 등록 실험은 사용자가 입력한 도메인과 URL을 음성 서비스 프로바이더의 정보 Database와 도메인 등록기 웹 서비스에 등록하는 실험이다. 실험에 앞서 정보 Database와 문법파일에 “다음”을 포함해 50개의 정보를 등록한 상태에서 다음의 실험 과정으로 진행하였다.

- 1) Navigator Toolbar의 도메인 등록 버튼을 클릭하여 도메인 등록 폼을 연다.
- 2) 도메인과 URL을 입력하고 요청버튼을 클릭한다.

- 도메인에 네이버, URL에 <http://www.naver.com>을 입력한다.
- 요청 버튼을 클릭하여 음성 서비스 프로바이더의 정보 Database와 도메인 등록기 웹 서비스를 호출한다.

(그림 8 도메인 등록 화면)

다음 [그림 9]와 같이 도메인 정보가 정보 Database에 등록된 것을 확인할 수 있다.

v_idx	v_domain	v_url	v_fixdate
48	다음	http://www.daum.net	2008-12-07 오후
49	파란	http://www.paran.com	2008-12-07 오후
50	한밭대학교	http://www.hanbat.ac.kr	2008-12-07 오후
51	야후	http://kr.yahoo.com	2008-12-08 오후
53	삼성전자	www.samsung.com/sec	2008-12-08 오후
54	싸이월드	http://www.cyworld.com	2008-12-08 오후
55	네이버	http://www.naver.com	2008-12-08 오후

(그림 9 신규 도메인을 등록 후 정보 Database)

실험 결과 “네이버”라는 신규 도메인 정보가 정보 Database에 정상적으로 등록되었음을 확인할 수 있다.

다음은 도메인 등록기 웹 서비스에서 새로운 도메인이 등록되는 과정이다.

- 1) 음성 인식 문법파일을 열고 정보를 복사한다.
- 2) 기존에 등록된 정보들의 패턴을 확인하고 새로운 도메인이 들어갈 위치를 정한다.
 - 기존에 입력된 정보들의 Network를 파악하여 “네이버”가 들어갈 위치를 정한다.
- 3) 문법 파일을 재구성하고 저장한 후 인식기에 포팅 시킨다.

이와 같은 실험 결과를 통해 클라이언트의 도메인 등록 요청과 도메인 등록기 웹 서비스가 정상적으로 동작되고 있음을 확인할 수 있었다.

6. 결론

사용자들이 인터넷을 이용함에 있어서 입력장치의 불편함을 개선하기 위해 본 연구에서는 사용자가 음성을 입력방식으로 이용할 수 있도록 하기 위해서 음성 Navigator Toolbar와 음성 서비스 프로바이더를 구현하였다.

음성 Navigator Toolbar는 세 가지 기능으로 구성되어 있다.

첫 번째 음성 녹음기는 클라이언트가 음성으로 명령이 가능하도록 음성 녹음 기능, 저장 기능과 웹 서비스 요청 기능으로 구성되어 있다. 음성 녹음 기능은 클라이언트의

음성 입력 녹음한다. 저장 기능은 녹음된 데이터를 Wav 파일로 저장한다. 웹 서비스 요청 기능은 녹음한 Wav 파일을 DIME으로 패키징하여 SOAP 프로토콜로 음성 서비스 프로바이더에 전송해주는 역할을 담당한다.

두 번째 URL Navigator는 음성 인식 결과인 URL을 반환 받아 클라이언트의 인터넷 페이지를 이동시켜주는 역할을 담당한다.

세 번째, 도메인 등록기는 클라이언트가 도메인 정보를 직접 등록할 수 있도록 웹 서비스를 요청한다. 클라이언트는 도메인과 URL을 입력한 후 요청 버튼을 클릭함으로써 도메인을 등록할 수 있다.

음성 서비스 프로바이더는 네 가지 기능으로 구성되어 있다.

첫 번째 프록시 클래스는 사용자의 음성 파일을 수신 받아 Voice Link 엔진으로 전송해주거나 정보 Database로부터 결과를 반환받아서 음성 Navigator Toolbar에 반환한다.

두 번째 VoiceLink Engine은 사용자가 입력한 음성파일을 인식하고 그 결과를 정보 Database로 전송한다

세 번째 정보 Database는 VoiceLink Engine의 음성 인식 결과를 도메인 필드와 매칭 시켜서 URL필드의 값을 프록시 클래스로 전송한다.

네 번째 도메인 등록 웹 서비스는 클라이언트가 요청한 도메인 정보를 정보 Database에 추가하고 문법파일을 생성하고 인식기에 포팅 하는 역할을 수행함으로써 사용자가 실제로 이용할 수 있도록 도와준다.

7. 참고 문헌

1. 백종현, 놀라운 새로운 세계 웹 서비스, 마이크로소프트웨어, 2003,
2. 고유정, 모바일 환경에서 인증과 음성인식을 위한 웹 서비스 구현, 석사학위 논문, 2004
3. VoiceLink, <http://www.vlink.co.kr/>
4. WordNetwork, <http://labrosa.ee.columbia.edu/doc/HTKBook21/node128.html>
5. 오신영, 웹 서비스를 위한 음성인식 시스템 구현, 한밭대학교, 석사학위 논문, 2002
6. 오지영, XML 웹 서비스와 인식기술을 이용한 웹 브라우저 구현, 한밭대학교, 석사학위논문, 2002