VoiceXML을 이용한 자동차 정보 안내 시스템 구현

양정수, 김동규, 김정현, 노용완, 홍광석
성균관대학교 정보통신공학부

An Implementation of Automobile Information System using VoiceXML

Jung-Su Yang, Dong-Gyu Kim, Jung-Hyun Kim, Yong-Wan Roh, Kwang-Seok Hong
School of Information and Communication Engineering, Sungkyunkwan University

요 약

음성 인식 기술이 발달함에 따라 음성 인식 기술을 이용한 응용의 개발이 중요한 문제로 떠오르고 있다. VoiceXML은 전화기를 통한 음성 인터페이스를 위한 XML 언어로서 손쉬운 방법으로서 음성 인터페이스를 설계, 구현할 수 있도록 만들어진 언어이다. 본 논문에서는 이를 이용해 전화를 통하여 음성으로 자동차 정보 안내 시스템을 사용할 수 있는 사용자 인터페이스를 구현한다. 구현된 시스템 및 서비스는 VoiceXML의 장점을 활용하여 원거리에서 편리하게 사용자가 자동차의 정보를 안내받고 제어할 수 있는 인터페이스가 제공되어 지고 사용자의 보안을 위하여 로그인 기능을 제공한다. 본 논문에서는 자동차의 단말기와의 인터페이스 자체보다는 음성 인터페이스의 설계 및 구현에 중점을 두었다. 10인의 피실험자가 각 10회씩 총 100회를 실험한 결과 99.3%의 인식률을 보였다. 추후 차세대 자동차 텔레메틱스 서비스와 연동하면 구현되어진 시스템의 활용이 증대될 것이라 판단된다.

I. 서 론

XML(eXtensible Markup Language)은 W3C(World Wide Web Consortium)가 HTML의 한계를 극복하기 위해 만든 마크업 언어이다. VoiceXML은 이러한 XML을 이용한 특화된 마크업 언어의 일부이다. AT&T, IBM, Lucent Technologies, Motorola 등이 설립한 VoiceXML 포럼에서는 1999년 8월 음성 응용 프로그램 개발을 위해 VoiceXML(Voice eXtensible Markup Language) 0.9 버전을 발표하였고, 2000년 3월 이를 보완한 1.0 버전을 정식으로 제안하였다.

W3C에서는 이 제안을 받아들여 2000년 5월 웹의 대화형 마크업 언어로서 VoiceXML을 표준으로 공인하였다. 현재 1.0 버전의 문법 부분을 보완한 VoiceXML 2.0버전이 나와 있다[1].

VoiceXML 기술의 주된 목적은 음성이나, 간단한 버튼 조작으로 컨텐츠를 이용할 수 있기 하기 위함이다. 우리가 컴퓨터나 기타 장치를 사용할 때 불편을 느끼는 것 중 하나는 키보드나 마우스 등의 제한된 방법을 이용해 명령을 수행해야 하는 것이다. 이 부분의 해결을 위해 음성 인식 기술들이 개발되었으나, 음성인식을 실용 제품에 사용하기 위해서는 음성의 95%이상의 인식률이 요구되지만, 현재 사람이 자유로이 말하는 것을 95%이상 인식하는 것은 거의 불가능하다. 이러한 기술적인 문제 때문에 아직 VoiceXML은 그 가능성의 범위 보다 널리 사용되지 못하고 있다[2].

이러한 인식률을 향상 시킬 기술들이 꾸준히 연구되고 있으며, 인식률이 향상 될수록 VoiceXML의 활용 범위는 넓어질 것이라고 생각된다. 본 논문에서는 이러한 VoiceXML을 이용하여 자동차 정보 안내 및 제어를 편리하게 이용하도록 도와주는 "VoiceXML을 이용한 자동차 정보 안내 시스템 "을 구현하였다. 이 시스템은 사용자가 음성을 통해 어디서나 간단하게 자동차의 상태와 제어를 할 수 있기 때문에 사용자에게 편의를 제공한다. 이러한 장점으로 실질적인 응용서비스 구현시 VoiceXML을 활용하는 사례가 증가 하고 있다[3].

본 논문은 2장에서 VoiceXML의 구성에 대해 언급하였으며, 3장에서는 본 논문이 구현한 "VoiceXML을 이용한 자동차 정보 안내 시스템"의 구성환경과 순서도를 제시하였으며, 4장에서는 실험 및 결과를 5장에서는 결론 대해 기술하였다.

II. VoiceXML

2.1 VoiceXML 게이트웨이

이종 망간의 연동을 위해서는 망간 연동을 가능하게 해주는 게이트웨이 장비가 필수적이다. VoiceXML 또한, 전화망과 인터넷망의 연동, 웹 서버로의 URL 전송, 웹 서버로부터 수신된 VoiceXML 문서를 분석하고 렌더링 기능을 수행하는 VoiceXML Gateway를 필요로 한다[4].

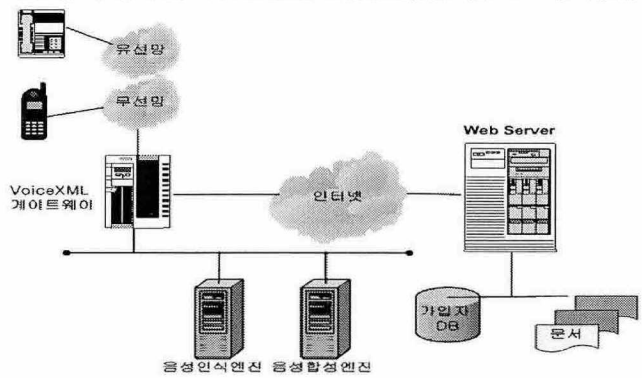

그림 1. VoiceXML 문서와 VoiceXML 게이트웨이

그림 1의 VoiceXML 문서와 VoiceXML 게이트웨이의 구성을 보면 사용자가 유.무선 전화로 콜을 하면 전화망을 타고 콜이 VoiceXML 시스템으로 전해지고 시스템내부에서 전화의 콜 신호를 받아들여 인터프리터를 동기화시켜 인터넷망을 통해서 웹 서버로부터 VoiceXML 문서를 요청한다. 웹서버에서는 해당 VoiceXML 문서를 찾아서 VoiceXML 시스템으로 문서를 전송하면 해당 문서를 인터프리터에서 파싱하고, 시나리오를 진행시켜서 사용자에게 TTS를 통해 만들어진 합성음성을 전달한다[5].

2.2 VoiceXML의 구조

그림 2에서 Document Server는 웹서버를 말하며 HTTP 클라이언트 응용 프로그램에 해당하는 VoiceXML Interpreter가 VoiceXML Interpreter Context를 통하여 요청하는 것을 처리한다. Document Server는 VoiceXML Interpreter의 요청에 따라 적절한 처리 과정을 거쳐서 (VoiceXML) Document를 만들고 그것을 HTTP 요청에 의한 응답으로 VoiceXML Interpreter에 전달한다.

VoiceXML Interpreter는 서버가 보낸 Document를 해석하고 사용자와의 상호작용(Interaction)을 수행하기 위해 Implementation Platform을 제어하는 컴퓨터 프로그램을 말한다.

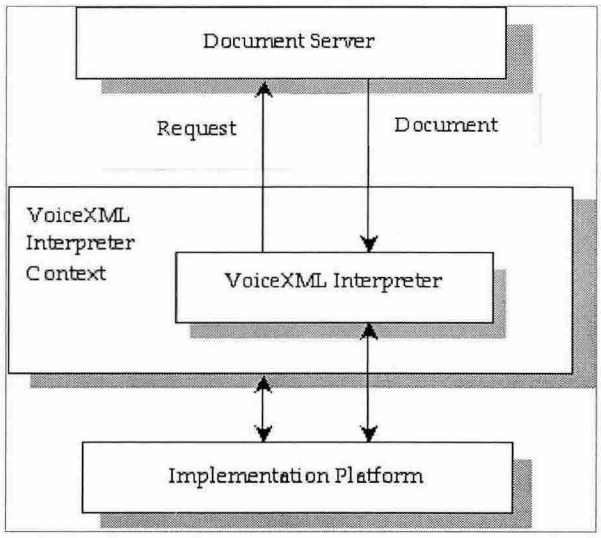

그림 2. VoiceXML 구조 모델

VoiceXML Interpreter Context는 웹서버(document server)와 HTTP를 통해 데이터를 주고받는 HTTP통신을 담당하며, VoiceXML interpreter가 VoiceXML문서를 해석하게 하고, VoiceXML Interpreter와는 독립적으로 Implementation Platform과 상호작용을 한다.

Implementation Platform은 H/W와 S/W를 모두 포함하며, 전화 수신기능, 전화상호전환(call transfer)기능, 음성인식기능, 음성합성기능, 음성과 오디오 재생기능, 음성과 오디오 녹음 기능 등을 수행한다. 그림 1에서 보이듯이 VoiceXML Interpreter와 VoiceXML Interpreter Context가 모두 Implementation Platform을 제어할 수 있다.

III. 자동차 정보 안내 시스템

3.1 기존 자동차 정보 안내 시스템

현재 자동차의 상태를 알고자 할 때 직접 가지 않고 알아내는 것은 어려움이 있다. 또한, 자동차의 상태를 알기 위해 직접 가기가 힘든 상황일 때 자동차의 현재 상태를 알기에는 많은 어려움이 따른다.

차량 밖에 사용자가 있을 때 사용자가 차량의 상태 및 기타 정보를 확인하고 제어하는 서비스의 초기 단계이며 이 서비스를 이용하기 위해서는 중앙센터의 안내원과 연결을 통하여 사용자가 차량의 상태 안내 및 제어를 할 수 있는 시스템이다.

3.2 구현한 자동차 정보 안내 시스템

기존 시스템의 안내원을 통해서 자동차 정보 안내 및 제어를 하지 않고 자동차의 상태 정보를 얻기 위하여 자동차 내의 정보단말기(PDA)가 센서와 연결된 ECU와 통신을 연결하여 정보를 얻고 얻어진 상태 정보를 DB에 저장하게 되고 상태 정보가 저장된 DB에 구현된 시스템을 이용하여 저장된 정보를 TTS를 통해 음성으로 안내하거

나 제어 명령을 전달하게 된다.

본 논문에서 구현된 시스템을 이용하면 언제 어디서나 음성 몇 마디로 간단하게 자동차의 정보를 획득하고 제어할 수 있기 때문에 사용자의 편의를 도모할 수 있다.

3.3 로그인

처음 자동차 정보 안내 서비스의 VoiceXML 인터페이스 시스템에 접속하게 되면 사용자를 승인하기 위하여 회원번호와 비밀번호를 이용해 로그인하는 방식을 사용하였다.

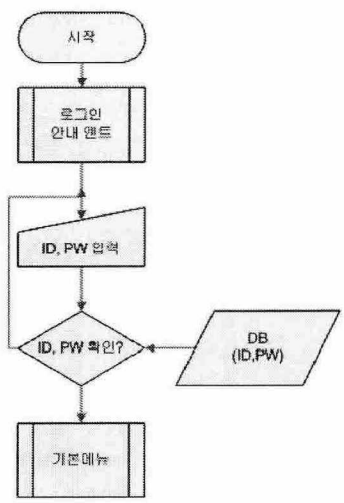

그림 3. 로그인 과정 순서도

로그인은 DTMF톤을 사용하거나(숫자) 음성 입력을 사용하여 이용할 수 있도록 제공된다. 하지만 음성은 주위의 잡음이나 음질상의 문제로 인해 인식률이 낮을 수 있으므로 DTMF톤을 사용하여 입력시키는 방법으로 한다.

그림 3에서 회원의 정보가 들어 있는 DB에 접속하여 회원번호와 비밀번호를 비교하여 로그인을 승인하는 방식을 사용한다.

3.4 자동차 정보 안내 서비스

자동차 정보 안내 서비스는 앞에서 언급한 내용처럼 사용자가 직접 자동차에 가지 않고도 현재 상태를 알 수 있다.

현재 자동차 정보 안내 서비스에서 가능한 부분은 문 상태, 등 상태, 시동 상태, 침입 상태, 손상 상태, 연료 상태, 배터리 상태, 현재 위치와 같은 부분이다.

또한 본 논문에서 구현된 시스템에서는 총 8종류의 정보 안내 서비스를 제공하고 있는데, 더 다양한 정보 안내를 하고자 할 경우에는 VoiceXML을 사용하였기에 쉽게 서비스 메뉴를 추가할 수 있는 장점이 있다.

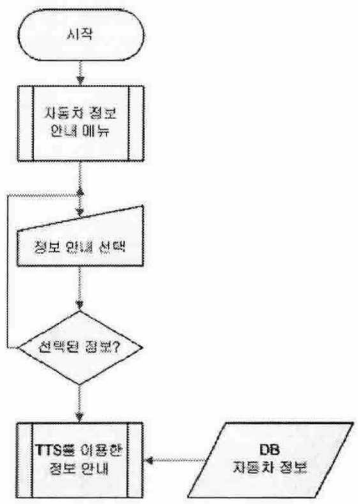

그림 4. 자동차 정보 안내 서비스 순서도

3.5 자동차 제어 서비스

자동차 제어 서비스는 문 제어, 등 제어, 시동 제어, 경보시스템 제어를 할 수 있도록 제공되는 서비스이다.

본 서비스는 사용자가 전화를 통하여 자동차의 상태를 확인하고 문이 잠겨있지 않다면 문을 잠금 상태로 손쉽게 제어할 수 있는 장점이 있다.

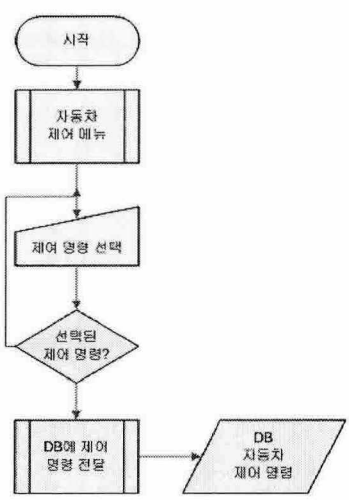

그림 5. 자동차 제어 서비스 순서도

Ⅳ. 실험 및 결과

VoiceXML을 이용한 자동차 정보 안내 및 제어 명령 인식률을 실험 하였다. 실험에 참가한 인원은 모두 10명이고 각 10회씩 총 100회의 실험을 하였다.(정보 안내 : 문 상태, 등 상태, 시동 상태, 침입 상태, 손상 상태, 연료 상태, 배터리 상태, 현재 위치)(제어 : 문 제어, 등 제어, 시동 제어, 경보시스템 제어)

4.1 실험 환경

VoiceXML서버는 연구실에 설치하였고, 자동차 정보단

말기는 PDA(HP-HX4700)을 사용하였으며 DB서버는 MS-SQL을 사용하였다. VoiceXML서버에 설치된 미들웨어는 ODBC를 이용하여 DB에 원격 접속하여 DB의 정보를 실시간으로 업데이트하고 자동차 정보단말기(PDA)와 TCP/IP 소켓 통신을 통하여 DB의 정보를 주고 받을 수 있도록 구성하였다. 그리고 자동차 ECU는 ATmega128을 사용하여 구현하였다.

4.2 실험 과정

사용자가 전화를 이용하여 VoiceXML 서버에 접속하여 안내 멘트에 따라 로그인, 자동차 정보 안내, 자동차 제어, 여러 하위 메뉴를 선택하여 서비스 되는 과정을 살펴본다. 그림 6에서 VoiceXML 사용의 실제 예를 나타내었다.

VXML : 안녕하세요. 자동차 정보 안내 시스템입니다.
VXML : 회원번호 7자리를 눌러주세요.
사용자 : (회원번호 7자리 입력 - 1234567)
VXML : 비밀번호 4자리를 눌러주세요.
사용자 : (비밀번호 4자리 입력 - 1234)
VXML : 회원 승인이 완료되었습니다. OOO회원님 안녕하세요. 자동차 메뉴로 이동합니다.
VXML : 자동차 메뉴입니다. 자동차 정보 안내, 자동차 제어 중 하나를 선택해 주세요.
사용자 : 정보 안내.
VXML : 자동차 정보 안내를 선택하셨습니다. 문 상태, 등 상태, 시동 상태, 연료 상태, 배터리 상태, 침입 상태, 손상 상태, 현재 위치 중 하나를 선택해 주세요.
사용자 : 문 상태.
VXML : 문 상태를 선택하셧습니다. 현재 차량의 문이 열려 있는 상태입니다. 자동차 메뉴로 이동합니다.
VXML : 자동차 메뉴입니다. 자동차...
사용자 : 제어. (멘트 도중 입력 가능)
VXML : 자동차 제어를 선택하셨습니다. 문 제어, 등 제어, 시동 제어, 경보 시스템 제어 중 하나를 선택해 주세요.
사용자 : 문 제어.
VXML : 문 제어를 선택하셨습니다. 현재 차량의 문이 열려 있는 상태입니다. 문을 잠그시겠습니까? 예, 아니오로 말씀해 주세요.
사용자 : 예.
VXML : 차량에 명령을 전달하였습니다. 명령 전달이 완료되었습니다.(명령 전달 완료 후 리턴 값을 확인) 차량의 문이 잠금 상태로 변경 되었습니다.

그림 6. 구현된 시스템 사용의 실제 예

4.3 실험 결과

회원 번호와 비밀번호의 경우 입력을 DTMF로 받으므로 100%의 인식률을 보인다. 하지만 사용자의 음성을 인식할 때 외부의 잡음이나 전화기상의 잡음, 부정확한 발음 등으로 인하여 인식기가 인식을 하지 못하는 경우가 간혹 발생한다.

메뉴	인식률	메뉴	인식률
문 상태	98%	배터리 상태	99%
등 상태	98%	현재 위치	100%
시동 상태	99%	문 제어	100%
침입 상태	100%	등 제어	99%
손상 상태	100%	시동 제어	100%
연료 상태	100%	경보 시스템 제어	99%
-	-	평균	99.3%

표 1. 인식률 결과

V. 결 론

본 논문에서는 사용자가 전화 또는 휴대전화를 이용해 손쉽게 음성으로 VoiceXML을 이용하여 자동차 정보 안내와 제어 인터페이스 시스템을 구현하였다.

음성은 인간의 가장 기본적인 의사소통 수단으로서, 편리함과 경제성의 측면에서 다른 인터페이스 방법에 비해서 우수한 특성을 가진다. 최근 음성 인식, 합성 기술의 발달과 정보 통신 기술의 발전으로 유무선 전화망을 통하여 컴퓨터와 인터넷을 연결한 음성서비스를 제공하는 업체와 이를 이용한 서비스의 개발 및 연구도 늘어나는 추세이다.

이러한 서비스 중 하나인 VoiceXML을 이용하여 구현된 자동차 정보 안내 시스템의 사용자 인터페이스는 추후 개발되어지는 텔레메틱스 서비스와의 연동을 통하여 보다 복잡해지고 다양해지는 텔레메틱스 서비스의 사용에 불편함이 해소될 것이고 향후 텔레메틱스 시장의 발전과 함께 텔레메틱스 서비스에서 음성을 활용한 사용자 인터페이스를 구현 할 수 있는 VoiceXML의 역할도 역시 증대될 것이다.

감 사 의 글

본 연구는 엔지비(주)의 차세대 자동차선행연구와 정보통신부 및 정보통신연구진흥원의 대학 IT 연구센터 지원사업의 연구 결과로 수행되었음(IITA-2005-C1090-0501-0019).

참 고 문 헌

[1] http://www.voicexml.org/specs/VoiceXML-100.pdf
[2] 김한수, 황인준, "VoiceXML 기반 EPG 검색 시스템", 한국정보과학회 논문지 C - 컴퓨팅의 실제 Vol.10 No.4 , pp 351~363, 2004
[3] 장민석, 방초균, "웹 환경에서 VoiceXML을 이용한 음성인터페이스 활용방안", 한국정보과학회 봄 학술 발표 논문집 Vol. 29. No. 1, pp.451, 2002
[4] VoiceXML Forum : "VoiceXML 2.0 candidate Draft": http://www.voicexml.ogr
[5] W3C, "Speech Recognition Grammar Specification version 1.0": http://www.w3.ogr/TR/speech-grammar