

# 모바일 하이네임을 위한 음성인식 웹 서비스 환경 구축

고유정\*, 홍인숙\*, 김윤중\*, 송은숙\*\*  
 \*한밭대학교 컴퓨터공학과,  
 \*\*한국인식기술  
 e-mail:youlony@hanbat.ac.kr

## A Development of Speech Recognition Web Services Environment for Mobile Hi-Name

You-Jung Ko\*, In-Suk Hong\*, Yoon-Joong Kim\*, Eun-Suk Song\*\*  
 \*Dept of Computer Engineering, Hanbat University  
 \*\*Korea Recognition Technology

### 요 약

한국인식기술에서는 명함정보를 원격에서 관리하는 모바일 하이네임(Hi-Name)을 출시하였다. 하지만 모바일 기기에서는 명함 정보를 이용하기 위해 작은 화면에 검색어를 펜으로 입력해야하는 불편함과 공간상의 제약성으로 음성인식엔진의 탑재가 어려우므로 분산인식환경이 필요하다. 따라서 본 연구에서는 모바일 환경에서 음성인식기술을 이용하여 명함정보를 이용할 수 있는 하이네임의 웹 서비스 환경을 구축하였다.

### 1. 서론

한국인식기술에서 개발한 개인용 명함관리시스템인 하이네임(Hi-Name)[1]은 명함의 자동입력, 그룹관리, 검색, 입력정보의 수정, 지도검색, 주소록 인쇄, 문자메시지 전송(SMS), 전자우편 발송, 휴대폰 동기화, Outlook 동기화 등 다양한 기능을 탑재한 제품을 출시하여 명함관리의 대중화를 선도하고 있다. 하이네임은 사용자 개인이 컴퓨터와 연결하여 명함을 인식하고 관리하는 제품이다. 현재 명함관리 프로그램은 단순히 PC에 정보를 저장하여 사용하던 방식에서 벗어나, 장소나 시간에 구애받지 않고 원격지에서 노트북, PDA, 휴대폰을 이용하여 접속이 가능한 서비스로 발전하고 있는 추세이다.

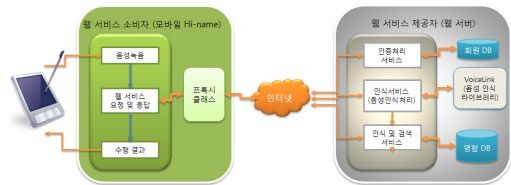
한국인식기술에서는 이와 같은 추세에 따라 명함정보를 원격에서 관리하는 모바일 하이네임을 출시하였다. 모바일 하이네임은 모바일 기기를 이용하여 명함을 인식하고 검색할 수 있는 제품이다. 모바일 하이네임은 기존의 오프라인 위주로 수행하던 명함관리를 원격에서 간편하게 관리하고 검색할 수 있지만 모바일 기기의 특징상 화면이 작아 펜이나 스타일러스로 검색어를 클릭하는 불편함이 발생하였다. 이러한 불편함을 해결하기 위한 방법으로 모바일 기기에서 명함 정보를 펜이 아닌 음성으로 검색어를 입력하는 방법이 필요하게 되었다.

따라서 본 연구에서는 음성 명령으로 명함을 검색할 수 있는 모바일 하이네임을 위한 음성인식 웹 서비스 환경을 구축하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 설계한 모바일 하이네임을 위한 음성인식 웹 서비스 환경의 구성도를 살펴보고, 3장에서는 모바일 클라이언트인 웹 서비스 소비자 및 4장에서는 웹 서비스 제공자의 기능을 알아본다. 5장에서는 실험 및 결과를 통해 시스템을 성능을 알아보며 6장에서는 결론 및 향후 연구 방향에 대해 설명하도록 한다.

### 2. 모바일 하이네임의 음성인식 웹 서비스 구성

본 연구에서 구축하고자 하는 모바일 하이네임을 위한 음성인식 웹 서비스 구성도는 (그림 1)과 같다. 모바일 기기는 웹 서비스 소비자로서 음성녹음, 음성인식 요청 및 응답, 수신결과처리 기능이 있다. 서버는 웹 서비스 제공자로 인증처리 서비스, 인식 서비스, 검색 서비스기능이 있다.



(그림 1) 하이네임을 위한 음성인식 웹 서비스 환경 구성

시스템 동작 과정은 모바일 기기를 이용하여 명함정보

를 검색하고자 하는 사용자명을 녹음하여 음성인식 웹 서비스로 전송한다. 음성인식 웹 서비스 제공자는 인증처리 후 음성파일을 인식하여 전화번호 등의 검색정보를 모바일 기기로 응답한다.

### 3. 웹 서비스 소비자

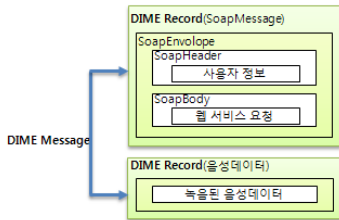
웹 서비스 소비자는 모바일 기기에 탑재된 명함 관리 프로그램이다. 웹 서비스 소비자는 동작과정은 (그림 2)와 같이 사용자로부터 음성을 녹음 받고 녹음된 음성은 웹 서비스 제공자에게 전달 후 인식 결과를 수신 받아 결과에 대한 명령을 수행한다.



(그림 2) 모바일 클라이언트에서 녹음, 요청 및 수신절차

음성녹음 기능은 Pocket PC 2002에서 제공하는 VoiceRecord.dll파일을 이용하여 녹음기능을 구현하였다. 녹음 wave 파일의 포맷은 오디오 샘플크기 8비트, 채널 1(모노), 오디오 샘플속도 16kHz, 오디오 형식 PCM방식으로 저장하였다.

녹음된 웨이브파일을 웹 서비스 호출시 SOAP 메시지에 파일을 첨부하기 위해 DIME(Direct Internet Message Encapsulation)[2]기술을 이용하였다. 본 연구에서 구현한 DIME 메시지 구조는 (그림 3)과 같다. DIME 메시지 구조는 SOAP 메시지 형식으로 구성된 DIME 레코드와 음성데이터를 첨부하는 DIME 레코드로 구성된다.



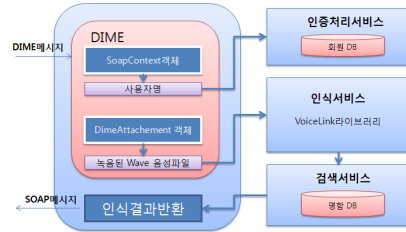
(그림 3) 음성데이터를 첨부하는 DIME 메시지 구조

첫 번째 SOAP 메시지 형식으로 구성된 DIME 레코드는 SOAP Envelope안에 SOAP 헤더와 SOAP 바디를 포함한다. SOAP 바디에서는 웹 서비스 제공자의 메서드를 호출하였다. 이와 같이 구성된 SOAP 메시지는

DimeAttachment 객체를 이용하여 DIME 첫 번째 레코드에 첨부하고 두 번째 DIME 레코드에는 녹음한 음성데이터를 첨부한다.

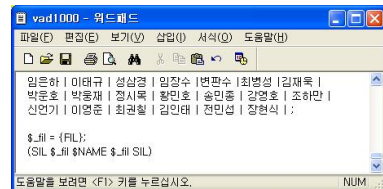
### 4. 웹 서비스 제공자

웹 서비스 제공자는 (그림 4)와 같이 웹 서비스 소비자에게 인증처리 서비스, 음성 인식 서비스, 그리고 검색 서비스를 제공한다. 먼저 웹 서비스 제공자는 웹 서비스 소비자로부터 DIME메시지를 전달받아 사용자 정보와 음성파일 분리한다. 사용자 정보는 인증처리 서비스를 이용하여 음성 인식 웹 서비스 권한을 부여받는다. 인증된 사용자는 음성파일을 인식 서비스를 이용하여 인식한 다음 인식한 결과에 따라 검색 서비스의 명함 DB에서 일치하는 정보를 찾아 웹 서비스 소비자에게 응답한다.



(그림 4) 음성 인식 및 검색 서비스와 처리절차

음성 인식을 위한 인식기는 음성정보기술업체인 에스엘투(SL2)에서 구축한 화자독립방식의 가변어휘인식을 지원하는 VoiceLink 3.4[3]를 이용하여 웹 서비스 환경으로 구축하였다. 또한 명함정보가 명함DB에 등록될 때 새로운 인식대상 단어로 Grammar파일(vad1000.ABNF)에 자동으로 추가되도록 기능을 구현하였다. Grammar파일은 (그림 5)와 같은 형태로 구성되었으며 표준 ABNF에 따른다.



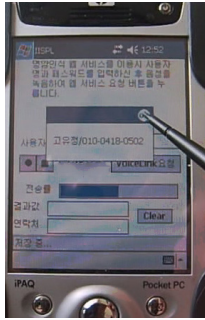
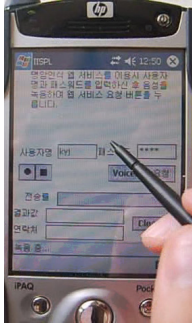
(그림 5) Grammar파일(vad1000.ABNF)의 일부

### 5. 실험 및 결과

본 시스템의 성능을 알아보기 위해 실험은 다음과 같은 환경에서 수행하였다. 음성녹음을 위해 사람이 없는 일반 강의실에서 오디오 샘플 크기 16bit, 채널1(mono) 오디오 샘플속도 16KHz, 오디오 형식 PCM 방식으로 녹음 환경을 구축하였다. 실험에 이용한 기종은 HP iPAQ H5420이다.

(그림 6)은 모바일 하이네임의 인터페이스 화면이며, (화면 7)은 모바일 하이네임 음성인식 결과를 나타내는 그림이다.

[7] 김학균, 정영준 외 2인, “지능형 홈네트워크 서비스를 위한 PDA 기반 음성처리 및 정보검색 기술”, 2006.01, 정보과학회지 제 24권 제1호



(그림 6) 인터페이스 화면 (그림 7) 음성인식 결과

발성자는 성인 여자 3명, 성인 남자 2명을 대상으로 명함 DB에 들어있는 사람이름을 대상으로 총 20개의 단어를 각각 5번씩 발음 하였으며, 인식 결과는 500회 중 482번 정확히 인식하였다. 또한 무선 환경에서 테스트한 결과 동일한 화자가 10개의 단어를 각 2회씩 발음하였을 때, 10회 중 96번을 정확히 인식하여 전화번호를 반환하였다.

## 6. 결론

본 연구에서는 모바일 하이네임을 위한 음성인식 웹 서비스 환경을 구축함으로써 입력이 불편한 모바일 하이네임 명함관리 프로그램에 음성인식 기술을 적용하여 사용자와의 인터페이스를 편리하게 제공하였다. 또한 음성인식 엔진을 모바일 기기에 탑재하지 아니함으로 인식 시 요구되는 자원이용의 제약성을 극복하였다.

향후에는 화자인식기능을 추가하여 사용자 인증을 음성으로 인식하고자 한다.

## 참고문헌

- [1] Hi-name, <http://www.hiname.net/>
- [2] Jeannine Hall Gailey, "Sending Files, Attachments, and SOAP Messages Via Direct Internet Message Encapsulation", <http://msdn.microsoft.com/en-us/magazine/cc188797.aspx>
- [3] VoiceLink, <http://www.vlink.co.kr/>
- [4] 박영주, 장광한 외 4인, "모바일환경에서의 VoiceXML을 이용한 cs 모델에 관한연구", 2002, 음성통신 및 신호처리 학술대회 논문집 19권 1호
- [5] Casey Chesnut, "WSE and the Compact Framework", 2002, <http://www.learnmobile.net/MobileClient/Tutorials/cfWSE>
- [6] Thomas Erl, "Service-Oriented Architecture: A Field Guide to Integrating XML and Web Services", 2004