

## 특 집

# 유/무선 통신망에서의 음성인식 상용서비스 개발 동향

구 명 완

KT

## I. 서 론

음성인식 기술의 중요한 응용서비스 중의 하나가 유·무선 통신망을 이용한 음성인식 상용서비스이다<sup>[1][2]</sup>. 유/무선 통신망에서의 음성인식 상용서비스는 1990년 대부터 본격적으로 도입되었으며<sup>[3]</sup>, 최근 동향은 대용량 단어를 기반으로 하는 서비스에 음성인식기술을 도입하는 방안과 기존에 운용되고 있는 서비스에 음성인식 기술을 접목 하여 업무의 효율성을 높이는 방안 및 음성인식 신규 서비스를 개발하는 방안이 있다. 대용량 기반 음성인식 서비스의 대표적인 사례가 음성인식 전화번호 안내 서비스 개발이다. 현재 미국, 유럽의 많은 국가가 2002년부터 시험 및 상용서비스를 도입하고 있다<sup>[4]</sup>. 기존에 운용되고 있는 서비스에 음성인식을 적용하는 대표적인 사례로는 콜 센터를 자동화 하는 방안이다. 현재 콜 센터는 인터넷과 접목되어 컨택트 센터(contact center)로 진화되고 있으며 컨택트 센터를 위한 음성인식 기술 응용서비스가 개발되고 있다. 마지막으로 음성인식 신규 상용서비스로는 단말기, 혹은 서버에서 음성다이얼 서비스 및 전자비서 서비스 등이 있다. 앞에서 기술한 대부분의 서비스는 전용 시스템에서 VXML(voice extensible markup language) 기반 시스템으로 변경되고 있다<sup>[5]</sup>. VXML기반 시스템에 대해서는 본 특집호의 다른 곳에서 자세히 기술되어 있으므로 참조해 주시기 바란다.

본 고에서는 통신망에서의 음성인식 상용서비스 개발을 서비스 관점에서 조명하여 제II장에서

는 최근에 도입되고 있는 음성인식 전화번호 안내 서비스에 대해서 상세히 기술하며, 제 III장에서는 음성인식 다이얼 서비스 그리고 제 IV장에서는 음성인식 콜 센터 상용 서비스 개발 동향에 대해서 미국에서 현재 서비스 중에 있는 “How may I help you?”서비스 중심으로 기술한다. 또한 제 V장에서는 KT에서 최근에 개발한 서비스를 소개하며 마지막으로 VI장에서 결론을 맺는다.

## II. 음성인식 전화번호 안내 서비스

### 1. Bell South의 511

511 서비스는 도로 및 날씨 상태에 대한 실시간 정보, 사고, 도로 폐쇄, 공사 구역, 대중교통수단의 일정 및 관광 정보 등을 알려주는 서비스이다. 이 서비스에 음성인식을 도입하기 위하여 BeVocal은 Nuance의 음성 인식 엔진을 사용하여 VXML 기반 음성인식 소프트웨어 플랫폼의 디자인과 구현을 개발하였다. 특히 BeVocal은 Castle Rock Consultants과 같이 작업을 하며 CARS(Castle Rock's Condition Acquisition and Reporting System)라는 데이터 수집 및 관리 툴에 자사의 VXML 기반 응용 프로그램을 결합시켰다. BellSouth는 호스팅 및 전체 개발을 관리하였다. 현재 8개 주(알래스카, 아이오와, 켄터키, 메인, 미네소타, 뉴햄프셔, 뉴멕시코, 버몬트) 컨소시엄으로 참여하고 있으며 511로서는 미국 내에서 가장 큰 시스템이다.

이 서비스는 2002. 7월부터 미네소타 주에서 시작하였으며 2002년 말에는 나머지 7개 주에서도 서비스를 제공하기 시작하였다.

이러한 음성인식 서비스가 개발이 되면 주정부 입장에서는 전화 회선 풀링(pooling)과 콜 센터의 아웃 소싱으로 비용을 절감할 수 있고 주민들은 8개 주 어느 곳에서 어떤 전화로도 실시간 여행 정보를 이용할 수 있다는 장점이 있게 된다.

Bell South에 의해서 제공되는 서비스 외에 미국 내 다른 지역의 다른 기관에 의해서도 음성인식 서비스가 제공되기 시작했다. 샌프란시스코만(Bay) 지역에서는 2003년 1월부터 MTC(Metropolitan Transportation Commission)에 의한 음성인식 서비스가 제공되고 있다. 이 서비스는 PB Farradyne에 의해 개발되었는데 Nuance의 음성 인식 엔진 및 VXML 기반 플랫폼을 사용하였다. 워싱턴 주에서도 2003년 2월부터 DOT(Department of Transportation)에 의해 Avaya IVR(Interactive Voice Response) 플랫폼 기반 음성인식 서비스가 제공되기 시작되었다. 이 서비스는 Nuance의 음성 인식 엔진, TTS 및 VXML 기반 플랫폼을 사용하여 개발되었다.

## 2. 일본의 #9887 서비스

#9887 서비스는 휴대전화 사용자가 안내에 따라 주소와 이름을 입력하면 전국 900만 건의 기업과 점포의 최신 전화번호를 음성으로 알려준다. 정보 이용료는 부과되지 않고 통화료만 부과된다. 음성인식이 성공을 못하더라도 통화료는 부과되며 금액은 전화를 거는 장소에 따라 다르다. 현재 이 서비스는 치바현에서 제공되고 있으며 만약 동경에서 이 서비스를 이용한다면 동경에서부터 치바현까지의 통화요금만 부과된다. 또한 3000여 개의 주요업체를 등록하여 주소 없이 이름만으로 검색할 수 있도록 하였다. 2002. 7. 1부터 J-phone 휴대전화 가입자 대상으로 서비스가 시작되었으며 2002. 10. 1부터는 DDI PHS, TU-KA 휴대전화가입자 대상으로 확장되었다.

이 서비스는 JMS(Japan Multimedia Ser-

vice)사에 의해 제공되고 있으며 필립스 음성인식 엔진을 사용하였으며 Infocom 사가 전체 시스템 개발을 담당하였다.

## 3. 핀란드의 020200 서비스

- 기존의 전화번호 안내원(Personal-assistance) 서비스를 필요로 하지 않거나, 원하지 않는 사람들을 대상으로 낮은 가격에 터치 톤 버튼 혹은 음성인식 기술을 사용하여 개인의 유선 전화번호와 무선 전화번호, 기업전화번호, 이름정보 등을 알려 주는 서비스이다. 현재 서비스 번호는 020200이며 SMS(short message service)를 이용하여 검색한 전화번호를 알려 주거나 검색된 전화번호로 자동으로 전화를 걸어주는 기능도 제공하고 있다. 또한 필요할 경우 안내원에게 전화를 연결시켜주는 기능도 있다.

이 서비스는 핀란드의 Fonecta Ltd와 이스라엘의 포네틱 시스템(Phonetic System)사가 공동으로 개발하였으며 포네틱 시스템사의 음성인식 엔진과 응용시스템을 사용하여 2001년 11월부터 핀란드 최초로 서비스를 시작하였다. 현재 약 5백만 정도의 유/무선, 기업전화 번호의 검색이 가능하다.

## 4. 덴마크의 음성인식 전화번호 안내 서비스

덴마크의 TDC(Tele Danmark)에서는 3백5십만 가입자 번호와 40만 기업 전화번호 대상으로 음성인식 자동 전화번호 안내 서비스를 2002년 6월부터 제공하고 있다. 이 서비스는 이용 요금 등을 이유로 기존의 안내원 기반 전화번호 안내를 사용하지 않는 사람들을 대상으로 신규 고객 창출을 목적으로 시작되었다. 이를 위하여 2001년 12월부터 2002년 4월까지 3개 도시의 230,000 일반 가입자, 47,000 기업 전화번호 대상으로 시험서비스를 수행하였다.

## 5. 스웨덴의 음성인식 전화번호 안내 서비스

덴마크의 TDC 제공 서비스와 비슷하게 비싼 요금으로 인해 전화번호 안내 서비스를 받지 않

는 고객들을 대상으로 수요를 창출하기 위한 서비스로서 스웨덴 텔레콤인 Telia의 자회사인 Response AB가 2001년 중반부터 음성인식기반의 저가형 자동 안내서비스를 제공하기 시작하였다. 현재 이 서비스는 기존의 안내원에 의한 고가형 서비스와 병행하여 저가형 서비스로 제공되고 있다.

### 6. 포르투갈의 118 서비스

포르투갈 텔레콤에서는 2001년부터 기존의 전화번호 안내 서비스인 118 서비스에 음성인식 기술을 이용하여 자동화 하였다. 이 서비스는 전화 가입자와 주요 회사의 전화번호를 안내해주며 음성인식이 불가능한 회사의 전화번호 문의나 다른 종류의 문의에 대해서는 안내원으로 연결해준다.

### 7. 아르헨티나의 음성인식 전화번호 안내 서비스

아르헨티나 텔레콤에서는 두 종류의 음성 인식 서비스를 제공하고 있다. 하나는 도시나 지역에 대한 지역 번호를 알려주는 것이고, 다른 것은 전화번호가 바뀐 사람이나 회사에 대해 구 전화번호를 말하면 새로 바뀐 전화번호를 알려주는 것이다.

## III. 음성다이얼 서비스

### 1. 스프린트사의 음성다이얼 서비스

스프린트사가 제공하는 음성다이얼 서비스는 단말기 기반 서비스와 네트워크 기반 서비스로 나뉘어진다. 단말기 기반 서비스는 단말기에 기본적인 음성인식 기능을 수행하도록 하는 단말기 기반 서비스이다. 예를 들면 삼성전자의 A500 PCS 전화기에 보이스 시그널(voice signal)사의 음성인식 기술을 적용하여 전화번호 숫자 음, 간단한 음성 명령이 가능하도록 한다. 이 단말기는 네트워크 기반 서비스와 병행하여 사용될 수 있다. 네트워크 기반 서비스는 음성인식 기능이 서버에서 수행되는 서비스로서 서비스 가입자는 웹상

에서 음성다이얼 기능이 필요한 전화번호와 대표 이름을 개인별로 등록하여 사용한다. 요금은 서비스 개시 몇 개월의 무료 서비스에 이어 월 \$5 정도이다. 개인별로 500개 정도의 이름을 등록할 수 있으며 각 이름 당 5대의 유선전화, 무선전화, 사무실 전화 등의 번호를 입력할 수 있다. 헤이아니타(HeyAnita)에 의해 개발되었으며 음성인식 엔진으로 스피치 워크(SpeechWorks)사의 제품을 사용한다. 또한 음성다이얼 서비스 이외에도 증권, 날씨, 뉴스, 운세 등의 서비스도 제공하고 있다. 또한 기업체 내의 직원용으로 사용할 수 있는 기업용 솔루션도 제공하고 있다. 2001년 6월부터 음성인식 기반 정보 및 다이얼 서비스를 시작 하였으며 2001년 10월부터는 현 위치에 가까운 식당을 안내해 줄 수 있는 식당 안내 서비스도 개시하였다. 이 서비스는 미국 내에 있는 12,000개 이상의 도시에 있는 260,000개 이상의 식당에 대한 정보를 제공하여 주며 사용하는 방법은 \*TALK을 누른 다음 "Call Restaurants"라고 말하면 된다.

### 2. 야후사의 바이폰(By phone) 서비스

야후사의 야후바이폰("Yahoo! by Phone")은 음성 다이얼 서비스를 제공하여 주는 것 이외에도 전자 메일이나 음성 메일을 듣고, 주식 정보, 날씨, 스포츠 스코어, 뉴스를 듣게 해주는 서비스다. 이 모든 서비스는 1-800-MY-YAHOO(1-800-699-2466)로 전화를 걸고 주소록에 있는 사람의 이름을 말하면 그 사람에 대한 정보를 들려주며, 직장이나 휴대폰, 집 전화번호, 팩스 및 다른 전화번호를 비롯하여 전자메일 주소까지 알 수 있다. 현재 한 달에 \$4.95 서비스료를 받고 있으며 웹 사이트(<http://phone.yahoo.com>)를 통하여 전화번호를 등록할 수 있다.

## IV. 음성인식 콜 센터 서비스

음성인식 콜 센터 서비스란 콜 센터로 전화를

거는 사람들의 다양한 요구를 자동으로 인식하여 미리 설정한 업무로 자동 분류하여 담당자로 연결해 주는 서비스이다. 일반적으로 콜 센터로 연결되는 전화는 매우 다양한 문의를 요구하는데 기존에는 자동 응답 시스템을 사용하여 메뉴로 업무를 나누거나 안내원이 직접 응답하였다.

전형적인 자동 응답 시스템은 메뉴 리스트를 제공하고 사용자가 여러 경로를 거치면서 선택하게 한다. 메뉴 시스템은 MFC 시스템(“x를 원하시면 1번을 누르시고, y를 원하시면 2번을 누르시고,...”)이나 음성 라벨 (“~이면 ‘예’라고 말씀하시고, 아니면 ‘아니오’라고 말씀하십시오.”) 또는 둘을 합한 방법(“x를 원하시면 1번을 누르시거나 또는 ‘일번’이라고 말씀하십시오.”)을 이용하여 구현할 수 있다. 이와 같은 방법들은 선택할 목록이 적거나 사용자가 쉽게 이용할 수 있다면 유용하지만, 어떤 태스크의 경우에는 메뉴가 너무 계층적이고 길어서 사용하기가 불편하다. 한편, 메뉴가 너무 단순할 경우에는 사용자가 원하는 것이 어느 카테고리에 속하는지를 판단하는데 어려움을 느낄 수도 있다. 이러한 경우에 종종 사용자들은 안내원과 연결하기 위하여 ‘0’을 누르거나 아무 것도 하지 않게 된다. 그러므로 선택 사항의 개수와 메뉴에 담은 정보를 어느 정도 자세하게 할 것인지를 결정하는 것은 기술과 경험을 필요로 한다.

대조적으로 안내원이 직접 서비스를 할 경우에는 사용자에게 “무엇을 도와 드릴까요?”라고 묻고 난 후 사용자가 자신의 요구나 문제점을 기술하게 함으로써 간단하게 처리할 수가 있다. 안내원은 당연히 업무에 대해 충분히 파악하고 있어서 전화를 담당자나 해당되는 자동 안내 모듈로 연결할 수 있다.

### 1. AT&T사의 HMIMY 서비스

HMIHY(“How May I Help You?”) 서비스는 사용자가 마치 사람에게 말하는 것처럼 인사를 받고 요구를 하는 것처럼 느끼도록 하기 위하여 만든 시스템이다<sup>[6]</sup>. 시스템이 하는 일은 사용자가 원하는 바를 인식하고 이해하는 것으로

써 모든 것을 이해하는 것이 아니라 호를 적절한 곳으로 연결할 수 있을 만큼만 알면 되는 것이다. 다음은 사용자와 HMIHY와의 대표적인 대화를 보여주는 예이다.

예제 대화)

Machine: AT&T. How may I help you?  
(AT&T입니다. 무엇을 도와드릴까요?)

User : I have a question about my June bill.  
(6월 요금에 대해 문의할 게 있습니다.)

Machine: Okay, what is your question?  
(예, 무슨 내용인가요?)

User : There's a call to Tallahassee, but I don't know anyone there.  
(Tallahassee로 전화 걸었다고 나왔는데요, 그 곳에 아는 사람이 없습니다.)

Machine: I can help you with that. Are you calling from your home phone?  
(도와드리겠습니다. 지금 집에서 전화하시는 건가요?)

User : I'm calling from my house.  
(집에서 전화 거는 겁니다.)

Machine: Please hold on while I transfer your call.  
(잠시만 기다려 주십시오. 전화를 다른 곳으로 연결해 드리겠습니다.)

상기의 예제는 사용자의 대화로부터 기계가 어떻게 대응하는지에 대한 내용과 요금관련 안내원으로 연결해 주는 상황을 나타낸 것이다. 이러한 서비스가 되기 위해서는 사용자와 콜 센터 안내원과의 수 많은 대화를 녹취하여 사용자가 원하는 내용에 가장 적당한 안내원과의 연결에 도움이 되는 핵심단어를 찾아 내야 한다. 이때 핵심단어는 여러 단어가 될 수 있고 또 이 단어에 해당되는 담당 안내원도 여럿이 될 수 있으므로 상관관계를 최대한 나타내어 주는 관련 확률을 녹취한 대화로부터 찾아 내어야 한다<sup>[6]</sup>. HMIHY는 대화중의 내용을 62가지로 분류하여 담당자에게 호 전환하거나 자동안내 서비스를 제공하고 있다.

## V. KT의 최근 개발 서비스

### 1. 폰북 서비스

KT의 폰북(phone-book) 서비스는 KT가 자체 개발한 음성인식, 합성 및 VXML 기반 솔루션인 휴보이스(HUVOIS)를 사용한 개인 맞춤형 음성 다이얼링 시험 서비스이다. 이 서비스는 VXML기반 서비스이며 휴보이스 폰북 서비스 전용 웹사이트(<http://www.huvois.com>)에서 이름과 전화번호를 등록한 후, 서비스 대표 전화번호로 전화를 하면 음성으로 전화를 걸 수가 있다. 현재 이 서비스는 연구소 자체 내에서 음성인식 무인자동 교환 서비스의 후속으로 시험 운용 중에 있다<sup>[7]</sup>. 이 서비스를 이용하면 전화번호를 외우거나 수첩을 찾는 불편을 없앨 수 있고 발신번호 확인기능 번호를 등록해 놓으면 대표 전화번호로 전화를 걸었을 때, 비밀번호를 입력할 필요도 없다. 이름의 추가, 수정, 삭제는 HTML을 기반으로 한 웹에서 이루어지며, 음성 인식 및 합성은 VXML 기반으로 구현된 시스템에서 이루어진다.

### 2. 1588-0114 서비스

고객이 전화를 걸면 고객과 가장 가까운 곳의 지점이나 기관으로 자동 연결되는 지능망 서비스로서 전국대표번호 1588이 있다. 1588-0114는 1588에 가입한 기업체 명칭을 대상으로 기업체 이름 혹은 상품명 등을 음성으로 말하면 자동 연결해 주는 서비스다. 과거에 기업체 음성다이얼 서비스로 제공하던 시험서비스(1577)의 후속으로<sup>[8]</sup> 1588 가입고객의 일부를 대상으로 서비스를 하고 있다. 현재 약 1000여 개의 회사명이 등재되어 있으며, 웹사이트(<http://www.15880114.com>)에서 확인할 수 있다.

## VI. 결 론

본 고에서는 유/무선 통신망에서의 상용서비스의 개발 동향에 대해서 알아 보았다. 개발환경은 VXML기반의 음성인식, 합성기술을 사용하여 인터넷에 개인 정보를 저장하고 유/무선 전화망을 통해 음성인식 기능이 가능하도록 하여 사용자들이 유비쿼터스한 환경에서 사용할 수 있도록 하고 있다. 상용서비스는 음성인식 안내를 중심으로 하는 서비스와 개인 정보를 이용하는 음성다이얼 서비스 마지막으로 콜 센터의 자동화에 음성인식 기능을 제공하는 서비스로 크게 나누어 국외 및 KT의 동향을 기술하였다.

음성인식 안내 서비스의 최종 목표는 음성인식 114 서비스를 완성하는 것이나 현재는 음성 인식 성능의 한계로 저가형 서비스로 제공하고 있으며 유럽 많은 나라에서는 고가의 기존 안내원 기반 서비스와 병행하여 서비스 중에 있다. 음성다이얼 서비스는 웹과 연동되어 웹상에서 이름을 등록 삭제가 가능할 뿐 아니라 전화기를 통해서도 등록할 수 있는 방향으로 가고 있으며 주로 통신회사 중심의 서비스가 진행되고 있으나 최근에는 인터넷 포털 서비스 제공자들도 상용서비스를 제공하고 있다. 음성인식 콜 센터는 안내원이 주로 응대하던 콜 센터에 음성인식 기능을 제공하여 고객의 요구를 처리해 줄 수 있는 담당자에게로 전화를 연결해주는 서비스이다. 최근에는 대화의 중요성을 나타내는 핵심 의미 기반 단어와 담당 업무와의 연결을 위한 기술이 발전함에 따라 관련 서비스의 활성화가 기대된다.

앞으로는 다양한 주변 환경에 관계없이 성능이 보장되는 기술이 발전하게 되고 대화중에 사용자의 의도를 보다 정확히 파악할 수 있는 알고리즘이 개발됨에 따라 보다 다양하고 편리한 상용 서비스가 국내외에 제공될 것이다.

## 참 고 문 헌

- (1) 구명완, “음성인식기술의 현황과 전망”, 대한 전자공학회 학회지 제20권 제5호, pp.548-557 1993년 5월.
- (2) 구명완, “최근의 음성인식 응용서비스,” 대한 전자공학회지, pp. 587-595, 2001년 5월
- (3) 양재우, “음성정보처리 기술동향 및 전망,” 음성정보처리 산업협의회 창립총회 및 세미나, pp.27-53, 2001년 2월.
- (4) W. S. Meisel, *Speech recognition update*, 2001년-2003년 6월호
- (5) 구명완, “유/무선 전화망을 통한 음성 인터넷 서비스 기술 연구,” 제9회 한국음성과학회 학술발표대회 논문집, pp.69-74, 2000, 10월.
- (6) G. D. Fabbriozio, et. al., “AT&T HELP DESK,” ICASSP 2002, pp.2681-2684
- (7) 김희경, 김문식, “음성인식 무인 자동교환서비스의 시범운용,” 한국음향학회 제17회 음성통신 및 신호처리 학술대회, 2000년
- (8) 박용기, “한국통신의 음성언어 연구현황,” 제9회 한국음성과학회 학술발표대회 논문집, pp.3-14, 2000, 10월.

## 저 자 소 개



## 구 명 완

1982년 2월 연세대학교 전자공학과 졸업, 1985년 2월 한국과학기술원 전기 및 전자공학과 졸업(석사), 1991년 8월 한국과학기술원 전기 및 전자공학과 졸업(박사), 1996년 11월~1997년 12월 : 미국 벨 연구소 방문연구원, 1985년 4월~2003년 3월 16일 현재 : KT 서비스 개발연구소 음성인식서비스개발 팀장(경영직), 2003년 3월 17일~현재 : KT 인력 개발실 총무팀(서울대 KT-MBA 과정 전일제 이수 중), <주관심 분야: 음성서비스시스템 개발 및 응용, 음성 인식, 음성합성, 자동통역>