

# 유비쿼터스 환경에서의 음성언어기술 정책방향

정보통신부 김병수  
한국전자통신연구원 이영직

## 1. 서론

정보통신부 1층에 설치된 유비쿼터스 드림전시관에 가보면 말로만 듣던 유비쿼터스 세상이 어떤 것인가를 느낄 수 있다. 이 전시관을 관람하다 보면 음성언어기술이 유비쿼터스 세상의 핵심기술이 될 수 밖에 없다는 생각이 절로 든다.

음성언어기술이 미래기술로서 매우 중요하다는 지적은 여러 모로 보고된 바 있다. 다보스 포럼 및 MIT대학은 광기술, 유전자기술 등과 더불어 음성정보기술을 21세기 10대 유망 신기술로 지정하고 있으며, 뉴욕타임지도 2002년도 미국 정보통신을 이끄는 10대 기술로 선정하였다. Gartner Group 2001년도 보고서도 언어정보처리 기술을 향후 10년 12대 핵심기술로 선정하고 있으며, 특히 미국, 유럽 등 선진국은 언어음성정보기술을 21세기 정보화 사회의 핵심 기술로 규정하고 관련 기술개발 등에 대규모 투자를 하고 있다.

대한민국은 IT 인프라가 세계에서 제일 잘 된 나라이다. 이러한 인프라를 생각해볼 때 우리나라가 유비쿼터스 사회가 가장 빨리 구현될 나라라 생각해도 무리가 없을 것이다. 지금 유비쿼터스 환경에 적합한 음성언어기술을 개발한다면 2012년에 우리나라의 SW산업이 국제경쟁력을 유지하는데 큰 기반이 될 것이다.

2012년의 유비쿼터스 사회는 지금과 외형상으로는 유사해 보일 것으로 생각된다. 사람들은 직장에 나가 노트북을 이용하여 업무를 할 것이고, 차를 타고 있는 중에는 텔레매틱스 단말기를 사용할 것이다. 집에 들어가면 지능형로봇이 반겨줄 것이고, 이동 중에는 휴대폰을 이용하게 될 것이다. 지금과 달라지는 점이 있다면, 음성언어를 중심으로 한 편리한 사용방법이 채용될 것이며, 여러 종류의 정보단말을 사용함에 있어 마치 하나의 단말기를 사용하는 느낌이 들게 될 것이라는 점이다. 예를 들어 TV로 영화를 보다가 밤이 늦어 중간에 잤다면, 다음날 점심때 휴대폰에 몇 마디 음성명령을 함으로써 어젯밤의 장면부터 이어서 볼 수 있게 될

것이다.

유비쿼터스 환경에서의 사용자 인터페이스에 대한 정책을 수립함에 있어 주의해야 할 점이 있다. 사람을 중심으로 기술을 바라보되, 미래의 유비쿼터스 환경, 여러 종류의 미래형 단말 간의 상호작용을 같이 고려해야 할 뿐만 아니라, 이의 보급 및 인식 확산, 기술개발, 인력양성, 산업적 파급효과 등까지 전반적으로 고려해야 한다. 정보통신부에서는 이를 위해 산학연의 음성, 언어, 영상, 인지과학, 산업공학 분야의 전문가를 모아 위원회를 구성하여 종합개발계획을 작성하고 있다.

본고에서는 유비쿼터스 환경에서의 사용자 인터페이스 중 음성언어에 해당하는 부분에 국한하여, 그 시장 및 산업동향을 기술하고, 국가경쟁력 제고를 위한 음성언어기술 정책방향을 제시한다.

## 2. 유비쿼터스 환경에서의 음성언어기술

21C의 새로운 패러다임으로 등장한 유비쿼터스 컴퓨팅에 필요한 음성언어기술은 유비쿼터스 환경에서 다양한 유비쿼터스 단말을 통해 유틸즌(유비쿼터스 환경에서 생활하는 사람)에게 보다 편리하고 지능적인 인터페이스를 제공하는 기술로, 음성 인식 및 합성 기

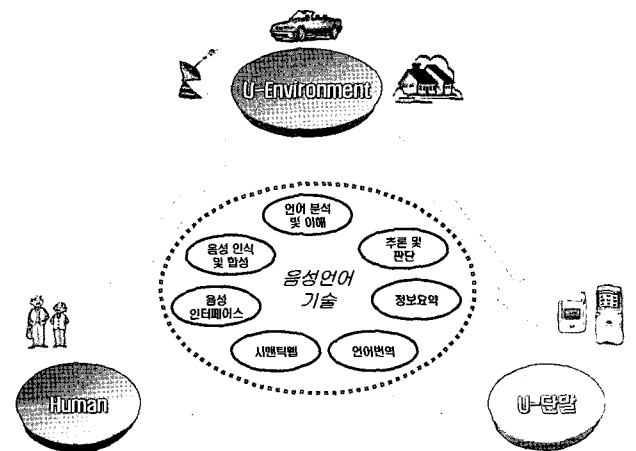


그림 1 유비쿼터스 환경의 음성언어기술

술, 언어 분석 및 이해 기술, 지식 처리 및 추론/판단 기술, 시맨틱웹 기술 등을 포함하고 있다.

유비쿼터스 환경의 음성언어기술은 텔레매틱스, 홈네트워크, 차세대PC 등의 음성 인터페이스, 지능형 로봇의 음성질의응답 서비스, 시맨틱웹을 통한 의미기반 콘텐츠 검색 서비스 등의 구현에 핵심 요소 기술로 적용되고 있다.

### 3. 산업 전망

#### 3.1 시장 전망

음성언어 산업분야는 음성인식 및 합성, 정보검색, 기계번역으로 나뉘지며, 세계 시장은 2005년 35억 9천만 달러에서 2010년 310억 4천만 달러로 연평균 40% 이상의 성장이 예상되고, 국내 시장은 2005년 1억 5천만 달러에서 2010년 6억 4천만 달러로 연평균 30% 이상의 고속 성장을 할 것으로 예상된다.

(단위: 억 달러)

구분	2006	2007	2008	2009	2010
세계	83.7	111.6	157.0	220.7	310.4
국내	2.2	3.0	3.8	5.0	6.4

출처: IDC 2004, ABI, VIA 2002, 언어정보산업협회의회 2004

#### 3.2 기술 전망

유비쿼터스 환경에서의 음성언어처리기술 중 음성처리, 정보검색, 언어번역 분야의 기술전망을 집중적으로 살펴보고자 한다.

음성처리기술은 1960년대부터 연구가 태동되어, 1980년대에는 고립단어 인식, 1990년대에는 연속음성 및 대화체 인식을 연구하였으며, 이 과정에서 많은 발전을 거쳤으나 일부 단어인식 기술을 제외하고는 아직 실상용화에는 미치지 못하고 있는 형편이다.

음성처리기술은 그 난이도가 높아 주로 국가주도로 연구가 수행되었다. 미국은 1990년대에 DARPA 프로젝트를 통해 항공편 문의시스템, 전화망 대화 연속음성 인식, 방송뉴스 인식, 지능형 대화 인터페이스 등의 기술을 개발하였으며, 유럽에서는 독일 Verbmobil 프로젝트를 통해 휴대형 자동통역기를 개발하였으나, 상용화되지 않았다. 일본 ATR에서는 음성번역통신연구소를 설립하여 음성처리분야 연구를 수행하였다.

이 기간 동안 미국, 유럽, 일본, 한국의 음성처리분야 연구기관들은 C-STAR 국제공동연구를 통해 다국간 자동통역 연구를 수행하였으며, 한국에서는 ETRI

가 참여하였다.

2004년부터 유럽에서는 European 6th Frame 프로젝트에 음성인식을 활용한 휴먼인터페이스 분야 연구를 시작했으며, 3년에 걸쳐 4000만 유로를 투입하고 있다.

정보검색기술 분야는 1990년대 이전에 저자, 제목, 출판사 등 서지정보를 이용한 문헌정보검색기술이 개발되었으며, 1990년대 후반에 야후, 구글 등에서 키워드를 이용한 웹페이지 검색기술이 개발되어 활용되고 있다. 그러나 아직도 사용자들이 불편함을 느끼고 있으며, 이를 개선하기 위해 고급 자연어처리 및 텍스트마닝 기술을 이용하여 의미기반 정보 검색을 개발하거나, 질문을 자연스럽게 입력하면 정답만을 알려주는 질의응답형 정보검색 기술, 도메인별 특화 검색기술 등을 개발하는 추세이다. 이외에도 주제 탐지 및 추적 기술 등도 개발되고 있다.

미국에서는 국가차원의 개발 프로젝트로 START 프로젝트, AQUAINT 프로그램, TREC 질의응답 연구개발 로드맵 등이 진행되고 있다.

현재의 웹은 사람이 읽어보아 그 뜻을 이해할 수 있으나, 컴퓨터는 그 의미를 이해하지 못한다. 따라서 현재의 웹은 향후 컴퓨터가 의미를 이해할 수 있는 시맨틱웹 환경으로 전환될 예정이며, 이에 따른 정보검색기술이 필요하게 될 것이다. 이를 위해 미국에서는 W3C 등을 통해 시맨틱웹을 표준화하고 있으며, EU에서는 시맨틱웹의 기반인 온톨로지를 구축하는 OntoWeb project를 진행 중이다.

언어번역 분야는 정보화사회가 발전해 갈수록 그 기술에 대한 사회적 요구가 커져온 분야이다. 초기에는 규칙 기반 방식의 언어번역 기술이 개발되었으나 상용화에 실패하고, 지금은 통계 및 데이터 기반 방식으로 전환하여 동종언어간에는 상용화에 성공하였다. 현재 동종언어간 체감번역률은 90% 이상이지만, 이종언어간 체감번역률은 60% 정도로 상용화 수준에 미치지 못하고 있다.

미국 Systran사는 유럽어를 바탕으로 17개 언어번역기를 개발하였고, 일본 ATR에서는 영어, 중국어, 인도네시아어 간의 번역기술을 개발하고 있다. 언어번역기술이 활용되는 분야는 웹 문서 번역, 기술 매뉴얼 번역, 특허문서번역 등 고품질 특화 번역 등이다. 상용화 예로는 일본 특허청의 특허 번역 서비스인 PAT Transer, 전문번역가 기술 매뉴얼 번역 지원 도구인 TRADOS 등을 들 수 있다. 한국에서는 2005년 12월 7일에 한영 특허문서 자동번역 상용서비스를 개시하였다.

## 4. 주요 정책방향

정보통신부에서는 IT839 중 3대 인프라를 BCN, USN과 함께 소프트웨어 인프라로 정하고, 이 소프트웨어 인프라의 중요한 한 분야로서 HCI분야를 핵심전략 산업으로 육성할 계획이다.

주요 정책목표는 2012년에 유비쿼터스 환경에서의 HCI 분야 소프트웨어 세계시장 선점이며, 이를 위해 음성언어기술 보급 및 인식확산 추진, 세계시장 선점을 위한 음성언어분야 전략기술 집중육성, 유비쿼터스 환경에서의 HCI 관련 법·제도 개선에 힘을 쏟을 예정이다.

### 4.1 음성언어기술 보급 및 인식확산 추진

음성언어기술을 사용한 사용자인터페이스 기술은 키보드, 마우스, 키패드, 버튼 등 지금까지의 사용자 인터페이스 방식에 비해 좀더 사람에 가까운 방법이다. 일반인들이 기대하는 자연스러운 사용자 인터페이스는 말만 하면 모든 것이 저절로 이루어지는 것이다. 하지만 현재의 음성인식 기술은 그 요구수준에 미치지 못하며, 5년이 지난다 하더라도 이 수준을 만족시키기 어려운 것이 사실이다. 유비쿼터스 환경에서의 사용자 인터페이스기술이 잘 활용되기 위해서는 음성인식기술이 잘 동작되도록 사용자의 발성방식을 부지불식간에 적응시킬 필요가 있다. 음성인식 기술을 처음 접한 사람의 경우 음성인식률이 50% 이하이지만 익숙해지면 90%를 넘는 인식률을 보이는 것이 그 증거이다. 우리나라가 유비쿼터스 환경에서의 사용자 인터페이스 기술분야에서 국제경쟁력을 가지기 위해서는 일반 사용자가 이러한 사실을 인식하고 기술을 활용하는 것이 매우 중요하다.

음성언어기술의 보급 및 인식 확산을 위해서는 이 분야의 신기술을 채용하는 서비스나 단말기에 기술사용 보조금을 마일리지 방식으로 지급하는 방법이 있을 수 있다. 아울러 흔히 사용하는 음성명령어를 표준화하는 방식을 도입하는 것도 바람직하다. 이미 언어음성 정보산업협의회에서는 이를 위해 PDA 음성명령어를 2003년에 표준화한 바 있다. 2012년 유비쿼터스 환경을 염두에 두고 이러한 노력을 계속 확대해 나갈 필요가 있다.

음성언어기술의 보급 및 인식 확산을 위해서는 요소 기술 별로 표준인증 규범을 정하고, 인증된 기술만을 국가에서 구매함으로써 저가 입찰에 따른 언어 음성업체의 부실화를 방지하는 방안도 바람직하다.

### 4.2 세계시장 선점을 위한 언어음성분야 전략기술 집중육성

유비쿼터스 환경에서의 사용자인터페이스 기술은 현

재 사용되는 음성다이얼이나 음성인식 콜센터와는 달리, 음성과 키패드, 음성과 영상, 음성과 정보보호기술, 음성과 지식처리기술 등이 복합적으로 기능하게 될 것이다. 따라서 부분기술 하나의 성능향상도 중요하지만, 이보다 복합적인 기능요소들이 적절히 역할을 분담하여 사용자에게 편리한 서비스를 제공하는 것이 중요하다. 예를 들면 텔레매틱스 단말기에 행선지를 음성으로 입력하는 경우, 그 인식대상 어휘 수가 25만 개 이상으로 매우 크므로 음성인식 성능이 낮아지게 된다. 이럴 때 음성인식 결과 5개를 가능성이 높은 순서로 화면에 제시하고, 이중 하나를 터치패널 입력으로 선택하는 방식을 들 수 있다.

유비쿼터스 환경에서의 사용자인터페이스는 음성인식, 영상인식, 의미이해 등 부분기술뿐만 아니라 인지과학분야의 사용자모델링까지 고려하여 기능목표를 설정해야만 사회적으로 받아들여질 수 있을 것이다.

현재 이러한 복합적인 요소를 고려하여 유비쿼터스 환경에서의 사용자인터페이스 종합개발계획을 수립하고 있으며, 곧 그 윤곽을 드러낼 것이다.

### 4.3 유비쿼터스 환경에서의 HCI 관련 법·제도 개선

유비쿼터스 환경에서는 개인 고유의 정보가 많이 활용되므로 개인생활 침해의 가능성이 높아진다. 이러한 점을 염두에 둔 제도개선은 2012년의 유비쿼터스 사회 도래를 위해 필수적인 일이다.

아울러 4.1장에서 지적한 바와 같이 음성언어기술의 보급 및 인식확산을 위해 인증 분야, 표준화 분야, 보급지원 분야에 법·제도적인 뒷받침이 필요하다고 본다.

## 5. 맺음말

앞의 내용을 근거로 할 때, 유비쿼터스환경에서의 사용자 인터페이스기술이 국가적으로 매우 필요한 기술임은 자명하다. 이러한 분야에서 세계시장 선점을 하기 위해서는 이 기술에 대한 일반인의 인식을 확산시키는 것도 매우 중요하며, 유비쿼터스 환경에서 사용하게 될 핵심기술의 개발 역시 중요하다. 이 분야는 기존의 음성인식, 영상인식 분야뿐만 아니라 시맨틱웹, 온톨로지, 분석-추론-판단 등의 지식처리 기술이 바탕이 되어야만 한다. 이러한 기술개발을 바탕으로 일반 사용자의 기술에 대한 인식이 확산된다면, 2012년에 대한민국은 유비쿼터스 강국으로 세계 속에 우뚝 서게 될 것으로 확신한다.

---

김 병 수



성균관대학교 행정학과(학사)  
네델란드 암스텔담대학교 신문방송학과  
(석사)  
1987년~현재 정보통신부 S/W진흥팀장  
E-mail : bskim@mic.go.kr

이 영 직



서울대학교 전자과(학사)  
한국과학기술원 전자과(석사)  
미국 Polytechnic University 전자과  
(박사)  
1989년~현재 한국전자통신연구원 음성/  
언어정보연구부장  
E-mail : ylee@etri.re.kr