소 특 집

최근의 음성인식 응용서비스

구 명 완
한국통신 멀티미디어 연구소 음성언어연구팀장

I. 서 론

음성인식 기술이 국내에 소개된 것이 1980년대였으나 이 때는 주로 논문연구 및 단순한 시험시스템 개발에 머물렀다. 1980년대 말부터 세계적으로 음성인식 기술의 상용화가 시작되었고, 신경망 및 퍼지 기술의 개발로 인해 국내에서도 본격적인 연구가 시도되었다. 그러나 국내에서는 일부 연구소 및 기업에서 시제품 및 시범서비스를 개통하였으나 큰 호응을 얻지 못하였다. 1990년 말부터는 음성인식 기술을 전문으로 하는 중소업체가 생겨나고 해외의 유명 회사가 국내에 진출 하면서 음성인식 기술에 대한 일반인의 관심이 증대되기 시작하였다. 특히 2000년도에서는 벤처업계의 활성화에 힘입어 음성처리 관련회사가 100여 개가 되고 다양한 제품 및 서비스가 제공됨으로써 국내에서 음성처리 관련산업이 하나의 영역을 차지하기 시작하였다[1][2].

음성인식 기술의 비전은 "2001 : A Space Odyssey"라는 영화에서 나오는 HAL9000 컴퓨터의 음성인식 능력을 최종 목표로 한다. 즉 "사람과 자유롭게 대화하고 장기 및 노래 뿐만 아니라 문제까지 풀어 줄 수 있는 컴퓨터"이다. 혹은 최근의 영화로 스타 워즈(Star Wars)에 나오는 C3PO라는 로봇의 음성인식 능력이다. 그는 수많은 언어를 이해하고 통역할 수 있는 능력을 가졌다. 그러나 현재의 음성인식 기술은 좁은 영역에서만 인식이 가능하고 음성통역도 특정 영역에서만 가능하다[3][4].

한편 미국 MIT 대학에서는 음성인식 기술을 21세기를 이끌어 갈 10대 기술의 하나로 간주 하였으며, 2001년 1월 스위스 다보스에서 열린 제31차 경제 포럼에서는 과학 기술자들이 2010년까지 우리 일상생활을 바꿔 놓을 7대 신기술로 음성인식을 선정 하였다.

본 고에서는 최근의 음식인식 기술을 이용한 대표적인 음용 서비스에 대해서 알아본다. 먼저 II장에서는 최근에 관심이 집중되고 있는 음성 포탈서비스에 대해서 국내외 동향 및 서비스 시스템에 대하여 알아보고 III장에서는 최근의 전화망 응용 서비스에 대해서 기술한다. 그리고 IV장에서는 컴퓨터 응용서비스로 이용되고 있는 음성인식 워드프로세서 및 멀티미디어 응용기술에 대해서도 설명한다. 마지막으로 V장에서 결론을 맺는다.

II. 음성 포탈 서비스

1. 배경

음성포탈 서비스의 최근의 관심은 마이크 매큐라는 사람이 만들어 냈다고 생각하여도 과언이 아니다. 그는 1998년까지 네스케이프의 부사장을 지내다 음성 포탈의 가능성을 예지하고 동료기술자와 함께 텔미닷컴(Tellme.com) 회사를 만들었다. 그는 전화를 이용하여 인터넷에서 정보를 검색하여 주는 음성 포탈이 새로운 서비스로 부상 할 것이라는 전제 하에 2000년 4월 음성 포탈 서비스를 시작하였다. 서비스가 시작한 첫날에는 시스템이 다운 될 정도로 호가 많았으며 성공

적인 서비스로 평가를 받았다. 이후 큐액닷컴(Quack.com), 비보컬닷컴(Bevocal.com), 및 텔서프(Tellsurf.com) 회사들이 경쟁적으로 서비스를 시작하였다.

음성 포탈 서비스 회사의 목표는 기존의 웹 포탈 사이트와 유사한 음성 사이트를 만드는 것을 목표로 한다. 이들의 대상은 현재 전화를 사용하고 있는 사람들이며, 특히 컴퓨터에 익숙하지 않은 어린이 노약자들도 주 대상이다. 또한 이동 중에도 사용 할 수 있는 장점이 있기 때문에 서비스가 활성화 될 수 있다는 장점이 있다. 그리고 모든 웹사이트를 음성으로 연결시키는 것이 아니라 주식, 스포츠, 일기예보, 뉴스, 영화, 음식점등 일상생활에 가장 가까운 분야의 정보 만을 취급한다.

음성 포탈 서비스의 성공 여부는 소비자들이 가장 필요로 하는 것이 무엇이고 또 손 쉽게 자료를 추적해 음성으로 바꿀 수 있는 분야를 찾는 것이다. 현재 음성 사이트에 가장 적합한 정보로는 주식 시세처럼 수시로 바뀌고 게임 정보처럼 비교적 간단한 것들이다.

음성 포탈 서비스를 제공 하여 주는 방법은 세 가지가 있다. 먼저 음성 포탈 전문사업자가 직접 제공 하여 주는 방법이다. 현재 비보컬닷컴이 대표적인 회사 이다. 이러한 방식은 사업자가 음성 포탈 사이트를 독립적으로 구축하고 사용자도 이 사이트만을 대상으로 따로 등록하여야 한다. 두 번째 방식은 통신 서비스 업체가 제공하는 방식이다. 즉 전화 통신 서비스를 제공 하는 업체가 음성 포탈서비스 제공자와 제휴하여 자사의 네트워크를 사용하도록 하는 것이다. 그리고 전화 사용자들에게 인터넷을 접할 수 있게 하여 주어 고객 서비스를 강화 하자는 것이다. 실제로 AT&T사는 텔미닷컴에 많은 투자를 하였으며 2001년 말까지 자사의 1500만 고객에게 음성포탈 서비스를 해 줄 것이라는 보도도 있었다. 세 번째 방식은 인터넷 서비스 제공자가 운영 하는 방법이다. 즉 자사의 인터넷 서비스를 사용하고 있는 사용자들에게 전화를 통한 인터넷 서비스를 제공하여 준다는 것이다. 현재 AOL(America On-Line) 사가 큐액닷컴사를 병합하여 서비스를 제공하여 주고 있다.

2000년도 초반에는 음성포탈 서비스가 독립적인 서비스가 될 수 있다는 가능성이 있었으나 시간이 감에 따라 사용자들은 음성 포탈 서비스를 받기 위하여 자기가 사용하는 인터넷 포탈 외에 또 다른 계정을 만들어야 한다는 사실에 실망을 하였으며 또한 음성 포탈 업체도 독립적으로 컨텐트를 제작하는 것보다는 유명 사이트의 컨텐트를 이용하는 것이 유리하다는 판단 하에 독립적인 비즈니스 모델에서 다른 사업과 결합하여 시너지 효과를 야기하여 주는 방식으로 선회하고 있다. 이러한 방향전환 이유로는 신생 음성 포탈 업체가 자금력 및 네트워크가 부족하고 기존의 통신, 인터넷 업체가 기존의 사업모델에 한계를 느낀 나머지 새로운 돌파구를 찾아야 한다는 사실에 서로 공감을 이루었기 때문이다[5].

2. 해외의 음성 포탈 서비스

미국의 대표적인 세가지 음성 포탈 서비스에 대해서 알아 보자

1) 텔미닷컴(http://www.tellme.com)

음성 포탈 서비스를 2000년 4월 10일에 시작하였으며 200년 7월 24일에 정식으로 서비스를 개시하였다. 전화번호는 1-800-555-TELL로 무료 서비스되고 있다. 제공되는 서비스는 주식정보, 날씨, 스포츠, 영화정보 등이다. 또한 주와 도시 이름 혹은 우편번호를 말하면 근처의 날씨, 교통정보, 택시회사와 연결하여 줄 수 있다. 그리고 웹사이트에 주로 사용하고 있는 정보를 입력한 후 "내 주식", "내 회사" 혹은 "나의 집" 등만 말하더라도 소유하고 있는 주식 및 집 혹은 회사의 정보가 자동으로 입력되는 개인화 서비스도 가능하다. 현재 AT&T로부터 약 6천만 달러의 투자를 받고 있으며 AT&T사의 네트워크를 이용하여 서비스하고 있다. 앞으로 사업 확장을 위하여 AT&T사와 적극 협력할 예정으로 되어 있다. 사용하고 있는 음성인식 엔진은 뉴앙스(Nuance)엔진이다.

2) 비보컬닷컴
(http://www.bevocal.com)

전화번호는 1-800-4-BVOCAL이며 음성 포탈 사이트로 승부를 걸고 있다. 다른 사이트와의 차별 점은 비즈니스 찾기(Business Finder)라는 서비스 이다. 이 서비스는 사용자가 찾고자 하는 편의점, 식당, 주유소, 혹은 백화점의 전화 번호를 알려 주는 서비스로서, 현재 있는 위치를 컴퓨터와 대화를 통하여 말하면 자세히 알려 준다. 예를 들면 햄버거 체인 이름을 말하면 컴퓨터는 사용자가 알고자 하는 위치를 계속 물어서 전화 번호를 알려 주거나 전화로 연결 하여 준다. 특히 출발지와 도착지를 말하면 주행하여야 할 도로 이름 및 시간을 전화로 알려 주거나 전자 메일로 알려 줄 수 있게 하였다. 현재 백 만개 이상의 지역과 2000개 정도의 상호 명을 제공하여 준다.

이외에도 비행기의 도착 및 출발시간을 알려주거나 전자메일을 읽어주거나 간단한 음성다이알 서비스까지도 해주고 있다.

3) 큐액닷컴(http://www.quack.com)

전화번호는 1-800-73-QUACK이며 처음에는 독립적인 음성포탈사이트로 성장 하고자 하였으나 2000년 8월 31일 AOL에 합병 되면서 AOL 가입자들에게 음성포탈 서비스까지도 제공하여 주는 방식으로 전략을 바꿨다. 즉 AOL의 컨텐트를 이용하는 것이 자체적으로 제작하는 것보다 사업에 유리하다는 결과이다. 합병이 된 후에는 1-800-AOL-1234 전화번호로 AOLbyPhone이라는 서비스를 2천 7백만 AOL 가입자에게 제공하고 있다. 현재는 이 서비스 이용료로 매달 4.95달러를 받고 있으며 그 대신 광고는 하지 않고 있다. 그러나 스포츠 게임 등의 메뉴에서 정보를 제공한 후에 게임에 관련된 할인 쿠폰 제공 등과 같은 광고는 제공하고 있다.

현재 이 서비스를 사용하고 있는 사람은 20만 가입자(2001년 1월 기준)가 있으며 백 만호 이상의 호가 발생하였다[6]. 현재까지 운영한 결과를 분석하여 본 결과 사용자들의 80%이상이 유선 전화를 사용하고 있었으면 오직 20%정도만 핸드폰을 사용한 호였다는 것이다. 이것은 많은 AOL가입자들이 직장에서 인터넷을 사용하여 개인적인 정보를 검색하지 않거나 아직까지도 직장에서 인터넷을 사용하지 않는 다는 것을 나타낸 것이다. 또한 무선 인터넷 시장과 음성인식 전화 서비스는 다른 시장을 갖고 있다는 것도 의미한다. AOLbyPhone을 사용하여 "나의 스포츠"라고 말을 하면 AOL가입자가 좋아하는 팀 및 관련 정보도 제공해 준다.

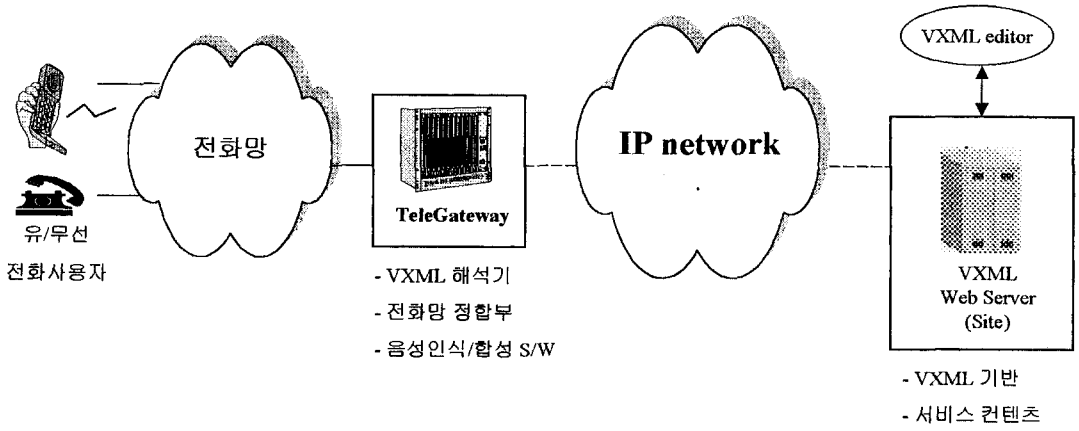
최근에는 노텔(Nortel)회사와 전략적 제휴를 맺어 음성 포탈 서비스 프랫폼을 노텔에도 제공하여 준다. 사용하는 음성인식 기술은 스피치 워크(SpeechWorks) 제품이다.

3. 국내의 음성 포탈 서비스

국내에서의 음성 포탈 서비스는 2000년 1월 20일 제너웨이가에 의해서 시작되었다. 전화번호는 1588-0852이며 웹 사이트(http://www.tellmetellme.com)에서 사용자 등록을 하면 사용 할 수 있다. 현재 제공하는 서비스는 일정관리, 이 메일, 주소록 및 뉴스를 안내 해 준다. 2000년 9월 22일에는 보이시안닷컴(voician.com)에서 음성 포탈 서비스를 시작하였다. 전화번호는 1588-0888이며 웹 사이트(http://www.voician.com)에서 등록을 한 후 사용할 수 있다. 이 사이트에서는 전자메일, 일정 관리 등 개인정보 관리 서비스 뿐만 아니라 증권, 교통, 날씨 등에 관한 서비스도 해 주고 있다. 마지막으로 헤이아니타(http://www.heyanita.com)에서는 2000년 12월 6일부터 전화번호 030311로 음성 포탈을 서비스 해 주고 있다. 기존의 서비스와의 차이점은 음식점 소개 및 항공기 예약까지 가능하다는 것이다[7].

4. 표준언어 음성 포탈 서비스

표준언어를 사용하는 음성 포탈 서비스란 음성 입출력 및 시나리오를 표현하는 언어를 표준으로 정하여 음성 포탈 서비스를 제공하여 주는 것이다. 현재 대표적인 표준언어로 VXML (voice extensible markup language)이 있



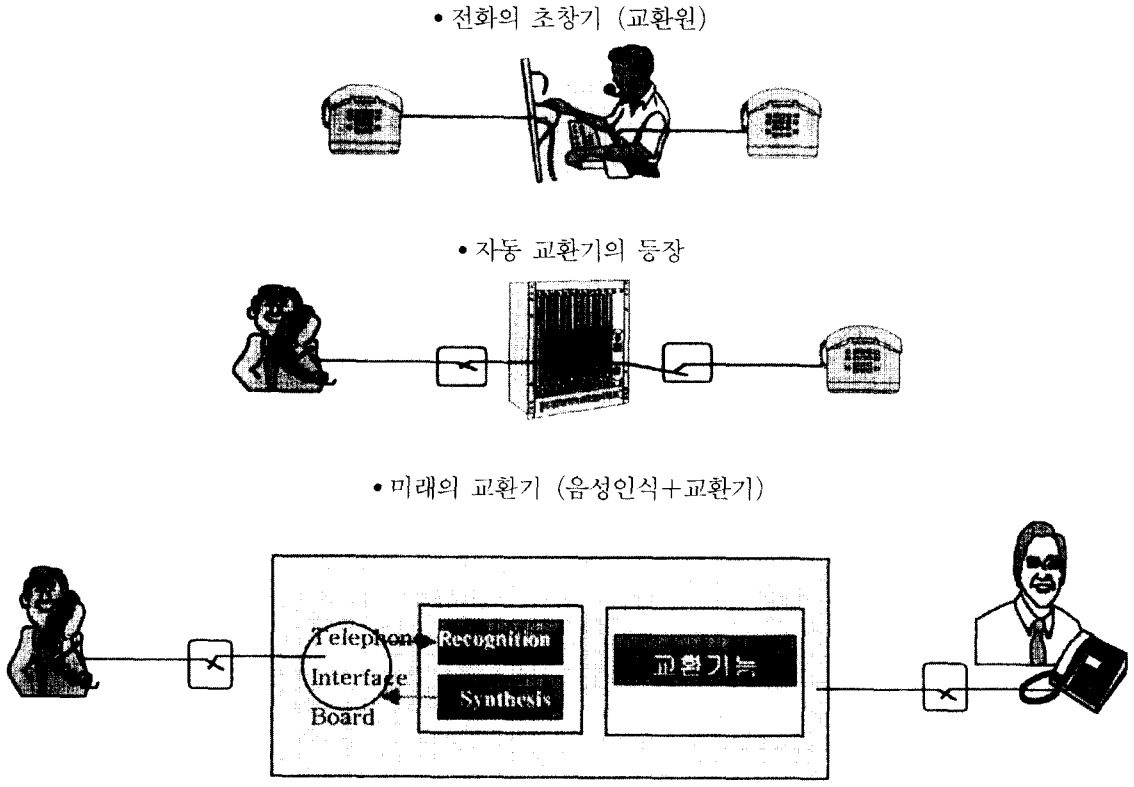
〈그림 1〉 VXML 기반 음성 포탈 서비스 구성도

다. VXML은 1999년 AT&T, 루슨트, 모토로라 및 IBM이 모여서 VXML 포럼을 결성하면서 본격적으로 연구 되기 시작하였다. VXML의 목적은 서비스 제공자가 아니라 컨텐트 제공자가 시나리오를 직접 작성 할 수 있도록 하는 것이다. 그렇게 함으로써 인터넷의 정보를 유. 무선 가입자가 쉽고 빠르게 검색 할 수 있도록 하는 것이다. 1999년 8월에 VXML 스펙 0.9가 제시 되었으며 2000년 3월에는 1.0이 제안 되었고 5월에는 W3C(World Wide Web Consortium)에 음성 브라우져(browser)의 표준방식으로 제안 되었다. 7월에는 이 방식이 워킹 드래프트(Wordking Draft)로 인정되어 현재까지 국제표준으로 연구 되어지고 있다. 최근에는 W3C와 VXML포럼이 공동으로 VXML 2.0 버전을 연구 중에 있으며, 1.0과의 차이점은 언어처리에 초점을 맞추어 합성기용 언어와 XML 기반 문법이 추가 될 것이며 또한 1.0의 내용을 좀 더 분명히 하는 것이라고 한다[8].

〈그림 1〉에는 VXML기반 음성 포탈 서비스의 개념도가 나타나 있다[9]. 텔리게이터웨이(TeleGateway)는 VXML해석기(VXML interpreter)와 전화망 정합부 및 음성인식, 합성기로 구성되어 있다. VXML 웹 서버에서는 VXML 에디터(editor)를 사용하여 음성 입, 출력이 가능한 대화형 시나리오를 작성하여 VXML 문서로 저장한다. 이 문서는 테리게이터웨이로 전송되어 VXML해석기로 입력이 되고, VXML 해석기는 문서를 해석하여 사용자에게 음성으로 질문을 하거나 음성을 인식하는 시나리오로 변경시켜 준다. 그리고 전화망 정합부는 유/무선 전화를 컴퓨터와 연결하는 역할을 한다.

Ⅲ. 전화망 응용서비스

전화망을 통한 음성인식 응용서비스의 기본 아이디어는 초창기 교환기에서 기인한다. 〈그림 2〉에는 전화교환기의 발전과정을 나타내었다. 초창기에는 교환원이 있어서 사용자가 누구를 바꾸라고 할 경우 교환원이 전화를 연결 해 주었다. 즉 사용자의 음성을 교환양이 자동으로 인식하여 전화 연결을 해 준 것이다. 자동 교환기가 개발이 된 후에는 상대방의 전화번호만 있으면 자동으로 전화가 연결이 되었다. 그러나 접촉하는 사람들이 늘어남에 따라 모든 전화번호를 외울 수가 없으며 또한 음성인식 기술의 향상으로 사람의 전화번호 혹은 이름 및 상호를 말하면 자동으로 전화가 걸리게 하는 교환기가 개발되게 되었다. 미래에는 이렇게 음성인식이 기본적으로 장착된 교환기가 선 보일 것이다. 음성인식 전화망응용서



〈그림 2〉 교환기의 발전방향

비스는 통신 사업자들이 주도적으로 제공하여 왔다.

1. 국외 동향

외국의 통신 사업자들이 음성언어연구를 하는 목적은 기존에 안내양에 의해서 운용되고 있는 서비스를 음성인식 및 합성기술을 이용하여 자동화 하는 것이다. 대표적인 응용 사례들로 회사 내에서 사람이름을 말하면 자동으로 연결해주는 음성인식 무인자동 교환시스템이 있다. 현재 미국 내에서 가장 많이 사용되고 있는 응용서비스로 AABS(Altenative automatic Billing Service)가 있다[10]. 이 서비스는 안내양이 과금의 방식(착신자 요금 부담, 콜링 카드 등)을 묻는 과정을 음성인식 기술을 적용하여 자동화 한 서비스이다. 현재 하루에 백만 호 이상 뜬다고 한다. 또한 최근에 AT&T에서는 "How may I help You?"라는 고객 문의를 자동화하는 서비스가 시험운용 중에 있다고 한다[11]. 이 분야의 최종 운용 목표는 114 안내를 자동화 하는 것이다. 이를 위해선 10만 단어 이상을 인식할 수 있는 기술이 필요한데 현재는 1만 5천 단어 정도를 인식할 수 있는 서비스가 운용 중에 있다.

두 번째 목적은 새로운 서비스를 제공하는 것이다. 예를 들면 영화안내, 교통안내 서비스 등이 있다. 또한 최근에는 인터넷을 음성으로 검색할 수 있는 인터넷 음성 포탈 서비스가 있다[12]. 현재는 날씨, 증권 및 교통 등 단순한 서비스를 제공하지만 향후 서비스의 종류가 다양해 질 전망이다. 그리고 이탈리아에서는 RAILTEL(Railway telephone information service)라는 서비스를 시험운용하고 있다. 이 서비스는 이탈리아의 도시 사이의 열차 발착 및 도착시간 뿐 아니라 요금 및 서비스 내역을 알려주는 서비스 시

스템이다[13]. 일본 KDD에서는 장난전화 퇴치 서비스가 있다. 일본 사람이 국외여행 중 일본으로 전화를 할 경우 일본 교환원이 응대하는 서비스가 있는데 장난전화가 심하여 일본어로 특정 단어를 말하도록 하면 일본어를 모르는 사람은 입력을 못하도록 하는 서비스이다[14].

2. 국내 동향

국내에서 가장 많이 사용되고 있는 전화망 응용서비스는 음성인식 증권 서비스이다. 현재 많은 증권정보 안내 사이트에 음성인식 기술이 적용되어 있다. 최근에는 미국 스피치워크제품을 사용하여 증권 매매까지도 가능한 서비스가 출현하였다. 또한 항공기, 기차표 예매에도 음성인식 기술이 사용되고 있다.

현재 전화망 응용서비스를 제공하는 방식은 크게 두가지가 있다. 첫 번째는 앞장에서 소개 한 바와 같이 음성 포탈 서비스를 통하는 방식이다. 두 번째는 전화망 응용 서비스 시스템을 독립적으로 개발하는 방식이다. 현재 대부분의 개발회사가 이러한 방식을 택하고 있다.

대표적인 통신사업자인 한국통신의 전화망 응용서비스는 다음과 같다(http://voice.kotel.co.kr). 현재 한국통신에서 자체 개발한 음성인식 엔진을 이용하여 기업체음성 다이얼링(1577), 음성인식 증권정보 서비스(700-3399), 그리고 음성인식 열차예매 서비스(700-1188)를 해 주고 있다[15][16][17] 그리고 최근에는 Windows NT상에서 구동되는 음성인식 무인 자동교환서비스도 개발한 바 있다[18]. 이 시스템은 연구개발 본부내에서 1400명의 이름을 대상으로 시범 서비스 중에 있다[19].

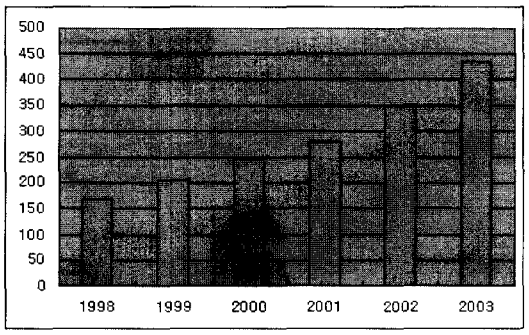
Ⅳ. 컴퓨터 응용서비스

1. 음성 워드프로세서

음성을 이용하여 컴퓨터에게 명령을 내리는 음성명령기가 상용화되어 있으며 음성으로 워드프로세서를 작동할 수 있는 음성 워드프로세서 소프트웨어도 개발되어 있다. 현재 미국에서는 L&H가 Dragon 회사를 인수하여 이 분야의 시장을 80% 점유하고 있다고 한다. L&H의 음성워드프로세서인 "Voice Xpress"는 64,000단어를 실시간 인식할 수 있다. IBM에서는 "Via-Voice"라는 소프트웨어를 출시하여 판매하고 있다. 현재 미국에서는 4종류 이상의 음성 워드프로세서가 150불 정도의 가격으로 판매 되고 있다고 한다[20]. 2001년도 11월에는 MS가 음성 명령기가 운영체제에 기본적으로 내장된 휘슬러를 출시할 예정에 있어 이 분야의 시장을 놓고 각축이 예상된다. 그러나 한국어는 빠져 있다고 한다. 〈표 1〉은 Dataquest 99에 나와있는 자료로서 음성인식 소프트웨어의 시장을 보여주고 있다.

〈표 1〉 음성인식 워드프로세서 시장

(단위 : $M)



Source : Dataquest 99

2. 멀티미디어 응용서비스

최근에는 음성인식 기술이 향상되고 멀티미디어 서비스가 제공됨에 따라 음성인식 기술을 이용한 멀티미디어 서비스가 선보이고 있다. 대표적인 서비스로는 음성인식 기술을 이용하여 영상을 검색하는 기술이다[21]. 예를 들면 뉴스의 동화상이 인터넷에 있을 경우 뉴스 내용 중 특정단어가 있는 부분의 뉴스만 검색할 경우에 음성인식 기술이 사용된다. 즉 뉴스 아나운서의 음성을 인식하여 특정 단어가 있을 경우 그 부분을 검색하여 동영상을 보여 준다. 또한 MPEG7에

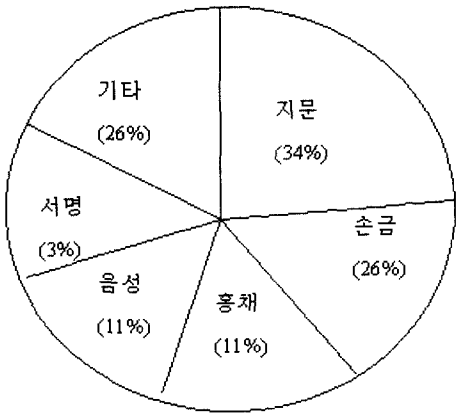
음성인식 기술이 이용되고 있다. MPEG7에서는 MPEG 그룹에서 개발중인 ISO/IEC 표준으로 멀티미디어 컨텐트 표현에 대한 인터페이스를 정의하고 있다. 다시 말해 멀티미디어 콘텐트가 갖는 내용을 해석하여 다른 컴퓨터에서 사용자가 찾고자 하는 내용을 검색할 수 있게 해 준다. 이를 위하여 멀티미디어 정보를 영상, 음성별로 내용단위로 인코우딩하고 사용자가 검색을 하면 인코우딩한 내용을 검색하여 준다. 음성인식 기술은 MPEG7의 오디오를 표현해주는데 사용되는 하나의 방식이다.

또 다른 음용 서비스로 음성 요약(gisting) 서비스가 있다. 이것은 음성을 요약하여 정리 해주는 서비스이다. 예를 들면 5분짜리 뉴스를 요약정리하여 1분 짜리 뉴스로 만들어 주는 서비스이다. 현재는 중요한 단어, 혹은 문장을 찾아 내는 방식으로 진행하지만 미래에는 요약 정리된 새로운 문장을 만들 수 있을 것이다.

3. 음성인증 서비스

음성인증 분야는 생체 인증분야의 하나로 개발되어 왔다. 생체 인증분야는 군사 및 법조계에서만 주로 사용되어 왔지만 최근에는 전자장치 및 전자 상거래등과 같이 상용목적으로 사용되기 시작하였다. 이러한 생체 인증기술의 장점은 암호키 관리가 필요없으며 생체 정보를 사용하는 것이 인증에 거부감이 없다는 것이다. 반면 온 라인 상태로 등록이 되는 것이 아니라 미리 사람이 가서 신원을 확인 한 후에 등록이 되는 단점이 있다. 생체 인증은 크게 지문, 손금, 홍체, 음성 및 사인을 이용하는 방법으로 나누어지며 현재 가장 많이 사용되고 있는 생체 인증 기술은 지문이다. 국제 생체 그룹(International Biometric Group)에 따르면 1999년의 생체시장 규모는 5천 8백만 달러 였으며 사용되고 있는 생체 정보의 우선 순위는 〈그림 3〉에 있다.

벨지움 은행인 바코브(http://www.bacob.be)는 전화를 통하여 계좌이체 등 송금을 할 경우에 음성인증 기술을 사용하고 있다[22]. 먼저 사용자는 음성으로 세 종류의 패스워드를 각각 세



〈그림 3〉 생체종류별 매출규모(%)

번 말하여 등록 한다. 그리고 송금을 할 경우에 사용자는 전화를 걸어서 기존의 방식과 같이 개인번호를 누르고 또한 미리 등록된 패스워드 중 하나를 말하면 된다. 사용하기는 쉬우나 아직까지 전화기 종류, 주변 소음 및 유, 무선의 여부에 따라 성능에 보장되지 않는다는 단점이 있다. 또한 미리 녹음을 할 경우에는 인증 문제점도 발생이 된다는 것이다.

V. 결 론

본 고에서는 최근의 음성인식 응용서비스에 대해서 알아 보았다. 먼저 현재 관심이 고조되고 있는 음성 포탈 서비스에 대해서 국내외 현황을 알아 보고 표준 언어인 VXML기반 음성 포탈 시스템에 대해서도 기술 하였다. 전통적으로 음성인식 기술의 응용분야인 전화망 응용서비스를 알아 보았으며 또한 멀티미디어 응용 서비스 및 음성 워드 프로세서에 대해서도 간략히 기술하였다.

최근에는 음성인식 기술의 중요성을 정부도 공감하면서 정보통신부, 산업자원부가 음성처리 기술에 과감한 투자를 하고 있어 음성처리 관련 연구자들이 고무 되어 있다. 반면 한국시장의 잠재성 때문에 최근에는 세계 유수의 음성인식기술

업체가 한국에 진출하고 있다. 한국인의 자존심인 우리 말에 관련된 기술을 우리 손으로 개발하고 장기적으로 외국어 음성인식기술까지 개발하기 위해선 산, 학, 연의 사심없는 공동연구가 어느 때보다 필요한 시기이다

참 고 문 헌

〔1〕 구명완, "음성인식기술의 현황과 전망", 대한전자공학회 학회지 제20권 제5호, pp. 548-557 1993년 5월.

〔2〕 양재우, "음성정보처리 기술동향 및 전망," 음성정보처리 산업협의회 창립총회 및 세미나, pp.27-53, 2001년 2월.

〔3〕 C.H. Lee, "Voice user interface(VUI) for new services over internet and wireless communications," Proc. of 9-th Korea Telecom International Symposium, pp.323-339, Oct. 2000.

〔4〕 S. Yamamoto, "Corpus-based spoken language translation technologies," Korea Telecom Journal, pp.4-12, Dec. 2000.

〔5〕 케이박, "음성포탈 혁명," 전자신문, 2000년 5월 16일 17면

〔6〕 W.S. Meisel, Speech recognition update, pp.13-14, no. 92, Feb. 2001.

〔7〕 정민화, "국내외 음성포탈 동향," 음성정보기술산업 워크샵, 2001년 1월

〔8〕 W.S. Meisel, Speech recognition update, pp.7-8, no. 92, Feb. 2001

〔9〕 구명완, "유/무선 전화망을 통한 음성 인터넷 서비스 기술 연구," 제 9회 한국음성과학회 학술발표대회 논문집, pp.69-74, 2000, 10월.

〔10〕 구명완, "음성인식기술의 현황과 실용화 전망", 한국음향학회 하계 학술발표대회 논문집 제17권 제1(s)호, pp.23-28, 1998. 7월.

〔11〕 A Gorin, et al. "How may I help you", Proc. IVTTA'96, pp.57-60, 1996

〔12〕 S Goose, M. Newman, C. Schmidt and L. Hue, "Enhancing Web accessability via the Vox Portal and a Web-hosted dynamic HTML<->VoxML converter, Computer Network, Vol.33, Issues 1-6, pp.538-592, 2000.

〔13〕 R. Billi, et al., "Field trial evaluation of two different information inquiry system", 1996 IEEE workshop IVTTA, pp.129-134, Sep. 30-Oct 1, 1996.

〔14〕 S. Kuroiwa, et al., "Prank call rejection system for home country direct service", 1996 IEEE workshop IVTTA, pp.129-134, Sep.30-Oct 1, 1996.

〔15〕 박용기, "한국통신의 음성언어 연구현황," 제9회 한국음성과학회 학술발표대회 논문집, pp.3-14, 2000, 10월.

〔16〕 류창선, 전호현, 김재인, "한국통신 음성인식 전화정보 시스템의 음성 인식기 운용 현황," COMSW99, 1999년

〔17〕 전호현, 장경애, 류창선, "대용량 음성인식 증권정보 시스템의 개발 및 상용 서비스 운용 현황," 차세대 통신소프트웨어 학술대회(NCS'99), 1999년

〔18〕 류창선, 전호현, 구명완, "음성인식 철도정보안내서비스의 시범운용," 한국음향학회 제17회 음성통신 및 신호처리 학술대회, 2000

〔19〕 김희경, 김문식, "음성인식 무인 자동교환서비스의 시범운용," 한국음향학회 제17회 음성통신 및 신호처리 학술대회, 2000년

〔20〕 PC Magazine, Sep.1999.

〔21〕 Y. Wang, et. All, "Multimedia content analysis," IEEE Signal Processing magazine, pp.12-36, Nov. 2000.

〔22〕 M. H. Shrif, "Standards for biometric

identification," IEEE Communications magazine, pp.48-50, Jan., 2001

저 자 소 개

丘 明 完

1960년 4월 26일생, 1982년 2월 연세대학교 전자공학과 졸업, 1985년 2월 한국과학기술원 전기 및 전자공학과 졸업(석사), 1991년 8월 한국과학기술원 전기 및 전자공학과 졸업(박사), 1996년 12월~1997년 12월 : 미국 벨연구소 음성연구부. 객원연구원(한국과학재단 후원 post-doc 과정), 1985년 4월~현재 : 한국통신 멀티미디어 연구소 음성연구팀장 (경영직), 〈주관심 분야 : 음성인식, 음성합성, 자동통역, 음성응용 서비스 개발〉