

2011년 11월

문화기술 (CT) 심층리포트

이달의이슈

- 음성인식 기술의 동향과 전망

CT동향

- QR코드와 회화 예술의 만남 '앰비언트 미디어 포털'
- 옷과 종이에도 적용할 수 있는 만능 터치 센서
- 사진의 미적 수준을 판단하는 이미지 인식분류 기술 외 7건

kocca



한국콘텐츠진흥원

목 차

이달의 이슈: 음성인식기술의 동향과 전망

1. 들어가며: IT 업계의 'Next Big Thing'으로 부상하고 있는 음성인식 기술 ..	1
2. 음성인식 기술의 역사와 특징	2
가. 음성인식 기술 연구의 역사 및 발전과정	2
나. 음성인식 기술의 특징	4
3. 음성인식 기술 분석	5
가. 음성인식 기술의 원리	5
나. 음성인식 기술의 구분	6
다. 최신 음성인식 기술 동향	7
(1) 언어모델링 기술	7
(2) 대화 처리 기술	8
라. 음성인식 SW 개발 동향	9
마. 음성인식 기술 선도 연구자 사례	10
4. 음성인식 활용 사례	11
가. 모바일 음성인식	11
(1) 애플의 음성인식 앱 '시리'	11

(2) 구글의 음성인식 앱	13
(3) 내비게이션	15
나. 개인화 서비스	17
(1) 보안 및 금융 분야	17
(2) 교육 분야	18
다. 기타	19
(1) TV 프로그램 검색	19
(2) 콜센터 음성처리	19
(3) 환경 제어	20
 5. 향후전망 및 시사점	 21
가. 가장 직관적인 인터페이스로 다양한 IT 기기와 융합	21
나. IT 소외계층으로 사용자 저변 확대	22
 참고문헌	 23

문화기술(CT) 동향

1. 인터랙티브 예술 ‘앰비언트 미디어 포털’, QR 코드와 회화 예술의 만남 24
2. 옷과 종이에도 적용할 수 있는 만능 터치 센서 26
3. 디지털 음악 시장의 미래, 자동차·수익모델·소셜 등 3대 변수로 지각 변동 예고 28
4. 사진의 미(美)적 수준을 판단하는 이미지 인식·분류 기술 30
5. 美 도서관 e-북 대여 확산 열기, 아마존의 킨들 이용자 대여 도서관 서비스가 기폭제 ... 32
6. 신생 소셜 TV 서비스 ‘커넥 TV’, 美 메이저 방송사 진영 동참으로 관심 집중 34
7. 선명한 컬러 동영상을 저전력으로 구현하는 디스플레이 개발 36
8. 美 학교 교육 현장에서 태블릿 PC의 교과서 대체 분위기 확산...e-텍스트북 대중화 임박 38
9. 예술, 게임 산업의 핵심 요소로 부상 中...게임 아티스트에 대한 관심도 증가세 40
10. TV 프로그램을 인식하는 아이패드 애플리케이션 ‘인투나우’ 42

이달의 이슈: 음성인식 기술의 동향과 전망

1. 들어가며: IT 업계의 'Next Big Thing'으로 부상하고 있는 음성인식 기술

- 인간의 말을 인식해 텍스트로 바꿔주거나 해당 명령을 수행하는 '음성인식(speech recognition)' 기술이 IT 업계의 'Next Big Thing'으로 부상하고 있음
 - 음성인식은 사람이 일상생활 속에서 마우스나 키보드 등을 사용하지 않고 목소리를 통해 원하는 기기 및 정보 서비스의 이용을 제어할 수 있는 기술로, 1950년대 등장해 지속적인 연구가 진행되어 왔지만 2000년대 중반까지 낮은 음성 인식률로 대중화되지 못함
 - 지금까지 특수한 용도에만 제한적으로 사용되어왔던 음성인식 관련 기술들은 최근 스마트폰으로 대표되는 휴대용 컴퓨팅 단말의 확산과 이를 지원해줄 클라우드 인프라의 확충이 맞물려 급속도로 발전하고 있음
 - 특히 얼마 전 애플(Apple)사가 공개한 '아이폰 4S(iPhone 4S)'에 탑재된 음성인식 기능 '시리(Siri)'가 공개되면서 음성인식 서비스에 대한 관심이 고조되고 있음
- 음성인식 기술 관련 세계 시장 규모는 2005년 11억 달러에서 2010년 30억 달러로 성장한 데 이어 2013년에는 54억 달러까지 성장할 것으로 전망됨
 - 국내 음성인식 시장도 2010년 1,800억 원에서 2012년 3,900억 원으로 매년 47%씩 성장할 것으로 기대를 모으고 있음¹⁾

Table 1. 세계 음성인식 시장규모 성장 추이 (단위: 10억 달러)

구분	2005년	2010년	2011년	2012년	2013년	2005~2013년 CAGR(%)
시장규모	1.1	3.0	3.7	4.5	5.4	22

자료: 디오텍(Company Analysis), HMC 투자증권

- 아직 인간의 모든 언어 표현을 이해하는 음성인식 기술은 개발되지 못했지만, 비교적 정형화된 문장이나 일정 범위의 어휘로 한정될 경우 현존 기술로도 높은 정확도를 담보할 수 있다는 측면에서 상당한 수준의 기술적 진보를 달성함
 - 네트워크와 컴퓨팅 기술의 발달로 음성 인식률이 개선되었을 뿐만 아니라 정보기기가 소형화되고 이동성이 중요시되면서 음성으로 간편하게 제어할 수 있는 음성인식에 대한 수요는 더욱 증가할 것으로 전망됨

1) 출처: 디오텍, HMC 투자증권

- 본 보고서에서는 IT 수요의 기폭제로서 빠르게 성장하고 있는 음성인식 기술의 특징 및 시장 동향을 분석하고, 향후 다양한 분야에서의 활용 가능성과 향후 전망을 모색하고자 함

2. 음성인식 기술의 역사와 특징

가. 음성인식 기술 연구의 역사 및 발전과정

- 음성인식 기술 연구는 1952년 미국 통신업체 에이티엔티(AT&T) 벨연구소(Bell Laboratories)의 단일 음성으로 말하는 숫자 인식 시스템 ‘오드레이(Audrey)’ 개발을 통해 시작됨
- 1963년 IBM은 세계 최초로 음성을 통해 16개의 영어단어를 인식할 수 있고 간단한 숫자 계산이 가능한 ‘슈박스(Shoebbox)’라는 장비를 공개함

Figure 1. IBM의 음성인식 시스템 ‘슈박스’



자료: IBM

- 이후 미국, 영국, 일본, 소련 등의 국가 연구소들이 인간의 발화를 인식하는 전용 하드웨어를 개발해 4개의 모음과 9개의 자음을 지원하는 수준까지 음성인식 기술을 확장시킴
- 1971년부터 76년까지 美 국방부 산하 국방첨단연구사업국(DARPA)은 음성인식 연구 역사상 가장 큰 프로젝트 중 음성이해연구(Speech Understanding Research) 프로그램을 진행함²⁾

2) 당시 국방첨단연구사업국(DARPA)은 거대한 컴퓨터 자원을 투입해 무려 1,000단어의 연속음성인식과 체계적인 음성 인식 데이터베이스를 구축하는 데 성공함

- 민간에서 연구된 많은 음성인식 프로그램들이 공개되지 않은 반면, 국방과 관련한 음성인식 프로젝트들은 다수가 개방적으로 논의됨
- 1980년대 초 음성인식 기술은 인식할 수 있는 단어가 1,000단어에서 1만 단어까지 늘어나면서 군사용, 로봇, 건강 분야 등에 널리 활용됨
 - IBM은 통계적 기법을 이용한 대규모 음성인식 시스템 ‘은닉 마르코프 모델(Hidden Markov Model, HMM)³⁾’을 개발하였는데, 이 방식을 통해 단순히 소리의 패턴을 찾기 보다는 음성을 모델링하고 알려지지 않은 소리도 단어가 될 수 있는 가능성을 고려하기 시작함
- 1990년대는 음성인식 기술이 최초로 상용화된 시기로, 드래곤(Dragon)社は 낱말단위의 인식과 표현이 가능한 소비자용 음성인식 제품 ‘드래곤 딕테이트(Dragon Dictate)’를 출시함
 - 1996년에는 벨사우스(BellSouth)社가 최초의 전화번호 안내 서비스를 시작, 다이얼인(Dial-In) 방식으로 소비자가 수화기에 대고 말한 내용을 기반으로 정보를 제공함
 - 1997년에는 연속적인 발화를 인식할 수 있는 음성인식 애플리케이션 ‘드래곤 내추럴리스피킹(Dragon NaturallySpeaking)’이 출시되기도 함
- 도쿄 공업대학의 사다오키 후루이(Sadaoki Furu) 박사는 음성인식 기술의 발전 동향을 1세대에서 5세대로 구분하고 세대별 특징을 아래와 같이 기술함

Table 2. 음성인식 기술의 발전과정

구분	특징
1세대 (1952년~1968년)	- 개별숫자, 음절, 모음, 음소 인식시스템 개발을 시도
2세대 (1968년~1980년)	- 동적시간 정합(Dynamic Time Warping, DTW) ⁴⁾ 기술의 제안 및 사용 - 고립단어 인식(짧은 음성명령어나 간단한 음성제어 등에 사용)의 가시화 - IBM 연구소: 대어휘 음성인식 ⁵⁾ 기술 개발 시도 - 벨연구소: 화자 독립 인식 시스템 개발 - 카네기멜론대학(CMU): 음소동적추적 기반의 연속 음성인식 기술 개발
3세대 (1980년~1990년)	- 2세대의 템플릿 기반 음성인식에서 은닉 마르코프 모델과 같은 통계적 기반의 음성인식 기술로 전환 - 신경회로망 기반 인식 기술, n-gram 기반의 언어모델 기술 제안 - 카네기멜론대학의 ‘스핑크스(SPHINX)’ 시스템(자연언어 + 불특정화자 연속 음성인식)이 개발되면서 대어휘 음성인식 시스템 개발의 전기를 맞이함

3) 음성이 발생되는 과정을 모델링하여 각 대상 단어 또는 음소마다 고유의 모델을 할당하여 입력되는 음성이 어떤 음성 모델로부터 발생되었을 확률이 가장 높은지를 측정하여 인식하는 방법으로, 높은 인식률과 편리한 학습성으로 널리 활용되고 있음

4) 1970년대 음성인식 분야에서 적은 수의 패턴으로 짧은 패턴을 인식하기 위해 제안되었으며 1990년대 말부터 가속도

3.5세대 (1990년~2007년)	- 음성인식 오류를 최소화하기 위해 MCE(Minimum Classification Error), MMI(Multi Media Interface) 같은 변별학습 기법이 사용됨 - 배경 잡음, 개별 화자의 음성 특성, 마이크로폰, 전송 채널, 반향 등에도 강인한 음성인식 가능한 기법이 사용됨 - 대어휘 방송 뉴스 인식 시스템 및 전화망에서의 음성인식 시스템 개발
4세대 (2007년~)	- 정보 서비스를 액세스하기 위한 대화시스템과 유비쿼터스 환경에서 음성으로 정보를 전사, 이해, 요약해줄 수 있는 방송뉴스, 미팅, 강의, 발표, 회의기록, 재판 기록, 음성 메일 등에 적용하기 위한 연구가 계속 진행

자료: 사다오키 후루이, 스트라베이스 재구성

나. 음성인식 기술의 특징

- 음성인식 기술은 △친숙하고 편리하며 △이동 및 작업 중에도 상시 입력이 가능하고 △화자의 고유 정보를 전달할 수 있으며 △입력속도가 빠르다는 장점이 있음

Table 3. 음성인식 기술의 특징

특징	활용 유형	활용 분야
편리함	음성인식은 인간에게 친숙한 정보 전달 방법이기 때문에 별도의 학습이나 훈련 없이도 기기를 손쉽게 사용할 수 있음	TV 등 가전제품 이용 및 게임 플레이
상시 입력 가능	손과 발이 자유롭지 못한 상황에서도 정보를 입력할 수 있기 때문에 이동시에도 기기 사용이 가능하며 안전성과 생산성을 제고할 수 있음	모바일 기기 사용 및 물류 등의 작업 수행
화자의 고유 정보 전달	화자의 음성을 통해 신원, 심리, 건강상태, 언어능력 등을 파악할 수 있어 개인별 서비스 제공이 가능함	보안, 금융, 의료, 교육 등
고속 입력 가능	입력 속도가 타자에 비해 빠르기 때문에 고속 또는 실시간으로 정보를 처리 가능	콜센터, 방송, 통번역

자료: 스트라베이스

- 음성은 인간에게 가장 친숙한 정보 전달 방식으로, 별도의 학습이나 훈련 없이도 일상제품의 복잡한 기능을 제어하는 데 용이함
 - 특히 게임분야에서 현실감과 몰입도 향상을 위해 키보드와 마우스 조작을 음성인식으로 대체하는 데 활용될 수 있음
- 음성인식 기술은 이동 시 정보 입력에 유리한데, 모바일 단말 이용이나 운전 중 기기 조작 시 안전성과 효율성 향상에 기여할 수 있음
- 화자의 음성을 통해 신원, 심리, 건강상태, 언어 능력 등을 파악할 수 있어 개인별 서비스 제공에도 적합함

센서 및 모바일 장치의 보급과 성능 향상에 기반해 제스처 및 필기 인식 분야에서 다시 활발히 연구되고 있음
5) 수만에서 수십만 개의 인식어휘를 갖는 음성인식시스템

- 따라서 보안 및 금융 분야에서 보안인증 및 신용평가 등에 활용될 수 있으며, 언어 교정 및 외국어 교육에도 사용될 수 있음
- 음성을 통한 입력 속도가 빠르기 때문에(컴퓨터 타자의 2~3배) 긴급 상황에서 신속하게 원하는 정보 입력과 검색이 가능함
- 반면 음성인식 기술은 사용자에게 따른 인식률 차이, 주변 잡음, 인식대상 어휘 제한 등 보편적 활용에 여러 한계를 노출하고 있음
- 따라서 음성인식 관련 연구자들은 모든 사용자를 대상으로, 최대한 많은 어휘, 자연스러운 발화(대화체)를 인식해 인식률을 높이는 방향으로 연구를 진행 중
- 성별, 연령, 사투리 등에 대한 방대한 음성모델 DB를 확보해야 향후 개인차에 따른 음성인식률 향상이 가능할 것으로 전망됨

3. 음성인식 기술 분석

가. 음성인식 기술의 원리

- 음성인식 기술은 입력받은 음성을 컴퓨터가 분석하고 특징을 추출한 다음, 미리 수집된 음성모델 데이터베이스와의 유사도를 측정해 가장 유사한 것을 문자 혹은 명령어로 변환함
- 음성인식 기술은 일종의 패턴 인식 과정으로, 사람마다 목소리와 발음, 억양 등이 다르기 때문에 최대한 많은 사람들로부터 음성 데이터를 수집해 이로부터 공통된 특성을 추출, 기준 패턴을 생성함
- 음성인식 과정은 크게 전처리부와 인식부로 구분됨

Figure 2. 음성인식 기술의 원리



자료: Digieco(2010.8)

- 전처리부에서는 사용자가 발성한 음성으로부터 인식 대상이 되는 구간을 찾아 잡음 성분을 제거하고 인식 과정을 위한 특징을 추출함
- 인식부에서는 입력된 음성을 음성 데이터베이스와의 비교를 통해 가장 가능성 있는

단어를 인식결과로 출력하게 되며, 단순 명령어가 아닌 문장을 인식할 때는 언어모델을 이용해 비교 단어를 제한하여 인식 성능을 높임

- 이러한 과정은 사람이 태어나서 말을 배우는 인지 과정을 모방한 것임
- 사람은 수많은 반복훈련을 통해 여러 가지 단어나 문법을 머릿속에 정형화하고, 새로 들은 음성을 이와 비교해 알아내는 인지과정을 거침

나. 음성인식 기술의 구분

- 음성인식 기술은 주로 인식할 수 있는 발성의 형태, 인식 대상 화자 및 단어에 따라 분류됨
 - 발성 형태별로는 가장 초보적 단계의 고립단어 인식에서 연결단어(connected word recognition) 인식, 연속어 인식(continuous speech recognition), 핵심어 인식 등으로 분류됨
 - 인식 대상 화자별로는 특정 화자 또는 사용자가 자신의 음성으로 미리 인식기를 훈련시키는 화자 종속 인식(speaker dependent recognition), 임의의 화자 발성을 인식할 수 있는 화자 독립 인식(speaker independent recognition), 사용자가 자신의 목소리에 대한 인식률을 높이기 위해 화자 독립 인식기를 자신의 목소리에 적응시키는 화자 적응(speaker adaptation) 인식으로 분류됨
 - 인식 대상 단어별로는 인식할 수 있는 대상단어가 고정되어 있는 고정단어 인식(fixed vocabulary recognition)과 인식 대상 단위를 수시로 갱신할 수 있는 가변단어 인식(flexible vocabulary recognition)으로 구분됨

Table 4. 음성인식 기술의 구분 및 특징

구분		특징
발성의 형태	고립단어 인식	- 고립된 형태로 발성된 음성만을 인식할 수 있음 - 음성인식의 가장 초보적 단계로, 현재 가장 많이 상용화되어 있음
	연결단어 인식	- 여러 개의 단어를 연결시켜 발성해도 인식 가능하며, 제한된 대상 단어의 조합으로 여러 형태의 음성인식이 가능함 - 고립단어 인식에 비해 난이도가 높으며, 인식률이 낮음
	연속어 인식	- 자연스럽게 발성한 연속된 음성을 인식할 수 있음 - 현재까지 인식률이 그다지 높지 못하며 특히 자연스러운 대화 형태의 발성인 경우 인식률이 매우 낮음
	핵심어 인식	- 자연스럽게 발성한 연속된 음성 중에서 인식 대상 단어만을 추출해 인식함 - 예를 들어, 열차, 비행기 자동예약 시스템에서 사용자가 발성한 여러 가지 정보 중 지명에 해당하는 것만 알고 싶을 경우 이 방식을 이용함
인식 대상 화자	화자종속 인식	- 특정 화자 또는 사용자가 자신의 음성으로 미리 인식기를 훈련시키는 과정이 요구되며, 이 경우 인식기는 훈련된 음성만을 인식할 수 있음 - 비교적 구현이 간단해 단말기 등에 탑재돼 응용되고 있으나 사용자가 훈련 과정을 거쳐야 하는 불편함이 있음

인식 대상 단어	화자독립 인식	- 임의의 화자 발성을 인식할 수 있는 기술로, 미리 수백 또는 수천 명의 음성에 관한 정보를 추출하여 데이터베이스화함으로써 별도의 훈련 과정 없이 사용가능함 - 현재 대부분의 상용화 시스템이 채택하고 있음
	화자적응 인식	- 화자종속 및 화자독립의 절충안으로, 사용자가 자신의 목소리에 대한 인식률을 높이기 위해 화자 독립 인식기를 자신의 목소리에 적용시키는 방식
	고정 단어 인식	- 인식할 수 있는 대상단어가 고정되어 있음 - 대상단어를 교체할 경우, 여러 사람의 음성 데이터를 녹취, 분석해 음성 모델을 구축하는 과정을 거쳐야 하므로 시간과 비용이 많이 소모됨
	가변 단어 인식	- 인식 대상단어를 수시로 갱신할 수 있음 - 음성인식기가 모든 음소에 대한 정보를 갖고 있으면서도 대상 단어가 갱신될 경우 음소의 조합으로 인식 대상 단어의 모델을 생성함

자료: 스트라베이스

- 음성인식 기술은 모든 사용자를 대상으로 보다 많은 어휘, 자연스러운 대화체를 인식하여 인식률을 높이는 방향으로 발전해 가고 있음

Figure 3. 음성인식 기술의 발전방향



자료: 정보통신산업진흥원(2011.10)

다. 최신 음성인식 기술 동향

(1) 언어모델링 기술

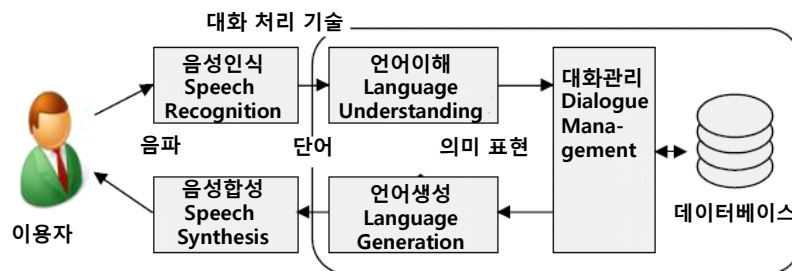
- 언어모델링 기술은 실제 사용되는 언어적 패턴의 규칙성을 수집해 음성인식 시스템이 보다 나은 성능을 내는 데 활용됨
 - 최근 음성검색 시스템이 상용화되면서 어휘 수가 기하급수적으로 증가하고, 웹 검색의 경우 특정 도메인으로 대상 영역을 한정할 수 없게 되면서 언어모델의 대용량화와 지속적 확장이 요구되고 있음

- 언어모델링에서는 일반적으로 N-gram 차수를 이용한 통계적 모델이 대중화되어 있으며, 최근에는 어휘 수를 제한하지 않는 대용량 분산 언어모델링 기술이 널리 활용되고 있음
 - N-gram 언어모델은 1개의 단어나 음절이 서로 다른 단어 또는 음절들과 함께 발현될 확률을 구하는 방식으로, 복잡한 자연언어를 표현하기에 적합한 구조를 가지고 있지만, 대용량 학습 데이터를 먼저 갖춰야 한다는 단점을 갖고 있음
- 구글 역시 자체적인 클러스터링 기술을 이용하여 무제한의 어휘수와 n-gram 개수를 기반으로 한 언어모델을 제공하고 있음
 - 구글의 자체 클러스터링 기술은 맵리듀스(MapReduce)라는 프로그래밍 모델과 라이브러리로 대용량 데이터를 처리할 목적으로 개발되었으며, 병렬/분산 처리 프로그래밍 경험이 없는 개발자도 분산 리소스를 활용할 수 있게 설계됨

(2) 대화 처리 기술

- 음성인식, 음성합성과 함께 음성 대화 시스템을 구성하는 대화 처리 기술은 언어이해, 대화관리, 언어생성 모듈로 구성되어 있음
 - 언어이해 모듈은 말한 사람의 발화를 언어처리하고 분석해 발화자의 의도를 표현하는 의미표현(meaning representation)을 생성함
 - 대화관리 모듈은 대화 흐름과 의도를 감안해 최선의 대화 전략을 계산해, 응답에 필요한 의미표현을 생성함
 - 언어생성 모듈은 시스템이 응답하게 될 단어나 문장을 입력된 의미 표현으로부터 생성함

Figure 4. 음성 대화 시스템 구성도



자료: 정보통신산업진흥원(2011.10)

- 대화 처리 기술에서는 인간과 기계의 대화를 진행시키는 대화관리 방법이 핵심요소이

며, 이에 따라 언어이해와 언어생성 방법론이 정해짐

- 대화관리 방법은 특정 목적에 적합한 대화 전략을 전문가가 직접 구축하는 지식기반 방법론(knowledge based approach)과 특정 도메인의 코퍼스 데이터로부터 학습하거나 예제화해 대화전략을 구축하는 데이터 주도 방법론(data-driven approach)으로 구분됨
- 일반적인 대화관리 방법들은 버스정보 안내, 비행기 예약 정보 안내 및 여행 정보 안내 등과 같은 정보 서비스를 위한 대화형 인터페이스에 맞도록 설계됨
- 그러나 이러한 정보 서비스를 위한 대화형 시스템은 목적을 빨리 달성하는 대화 흐름으로만 대화 전략이 구축되거나 학습되어 있어, 사용자의 발화에 대해 같은 대답과 흐름으로만 일관한다는 한계가 있음
- 최근에는 외국인과의 대화 상대나 역할 놀이 상대로 음성 대화 처리 기술을 외국어 교육에 활용하고자 하는 시도가 계속되고 있음
- 스웨덴 왕립공대 KTH는 영어 학습자를 위한 역할 놀이 대화시스템 'DEAL⁶⁾'을 개발했으며, MIT에서는 몰입식 중국어 학습을 위한 대화형 음성인식 시스템 'Family Island⁷⁾'를 소개함

라. 음성인식 SW 개발 동향

- 세계 최대 오픈소스 소프트웨어 개발 및 배포 사이트 소스포지(Sourceforge)에는 총 52건의 음성인식 관련 SW가 공개되고 있고, 이 중 Medscribber Community, Modular Audio Recognition Framework 등의 조회 건수가 높게 나타남

Table 5. 소스포지에 공개된 음성인식 관련 SW

구분	특성
Medscribber Community	- 태블릿 PC에서 음성인식 기능을 최적화하는 전자의료기록 SW
Modular Audio Recognition Framework(MARF)	- 오디오 알고리즘(음성, 사운드)과 자연언어 텍스트 분석 및 자바(Java)에서 구현 가능한 음성인식 기술을 통합한 크로스 플랫폼 프레임 워크
Talking Calculator	- 계산기의 숫자와 계산 결과를 음성으로 낭독하는 소프트웨어
Voice Conference Manager	- 컨퍼런스 콜 이용 시 음성인식, 생체인식, 텍스트를 음성으로 변환 시켜주는 소프트웨어
GoMad	- 윈도우 기반의 애플리케이션 이용 시 마우스나 키보드 대신 음성을 통해 제어할 수 있는 시스템

자료: 소스포지(Sourceforge)

6) GUI(Graphic User Interface) 인터페이스를 통해 학습자와 대화를 진행하도록 설계됨

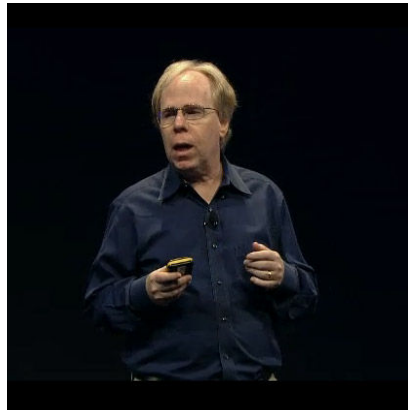
7) 가족관계도를 학습자에게 제시해 4단계의 수준별로 진행하는 교육 프로그램으로, 음성인식 오류나 대화처리 오류를 최소화하기 위해 아주 제한적 도메인과 시나리오를 차용함

- 한편, 음성인식 SW는 이동 중 휴대 단말기에 글자를 입력할 때도 활용되는데, 최근 스마트폰의 보급으로 그 수요가 급증하고 있음⁸⁾
 - 보행 중 음성으로 검색하거나 문자를 입력하는 실용적인 앱이 특히 인기를 얻고 있는데, '드라이브세이프리 프로(DriveSafe.ly Pro)'는 운전 중 화면을 터치하지 않고도 메시지를 작성할 수 있으며 음성으로 문자 메시지 송수신이 가능함
 - 이외에도 특화된 검색 서비스를 제공하는 앱들이 주목받고 있는데, 지식검색 서비스를 제공하는 '차차앤서(ChaCha Answer)', 음악 검색이 가능한 '사운드하운드(Soundhound)', 전화번호를 찾을 수 있는 '보이스 114' 등이 대표적임

마. 음성인식 기술 선도 연구자 사례

- 마이클 코언(Michael Cohen) 박사는 뉘앙스 커뮤니케이션즈⁹⁾의 공동 창업자로 2004년 구글로 이직한 이후 음성인식 기술팀을 이끌고 있음
 - 지난 6월에는 휴대전화보다 성능이 떨어지는 마이크와 내부 기계 소음 때문에 정확도가 낮은 PC에서의 음성인식 서비스를 발표, 2010년 공개한 '보이스 액션' 서비스 이후 또 한 차례 업계의 주목을 받고 있음
 - 마이클 코언은 스마트폰과 PC 등 기기의 장벽을 넘어선 음성인식이 가능해지고 있으며, 이는 사람들이 있는 곳이라면 어디서든 기계가 사람의 음성에 반응해 움직이는 시대가 시작된다는 것을 의미한다고 주장함

Table . 음성인식 기술 연구자 마이클 코언 약력



- 작곡가로 7년간 활동
- 스탠포드연구소(SRI) 소속 연구원
- 1994년 뉘앙스 커뮤니케이션즈(Nuance Communications) 공동 창업
- 2004년 'Voice User Interface Design' 저서 출간
- 2004년 구글로 스카웃, 現 구글 음성검색 시스템의 개발책임자
- 2011년 6월 PC를 이용한 음성인식 기능 발표

자료: 구글

8) 구글 안드로이드 마켓과 애플 앱스토어의 음성인식 애플리케이션 개수는 약 2,503개로 집계됨

9) 뉘앙스 커뮤니케이션즈는 구글과 함께 최신 음성인식 기술을 선도하는 업체로, 공교롭게도 마이클 코언이 구글로 이직한 후 구글의 음성인식 기술을 책임지고 있고 그가 떠난 후 뉘앙스는 애플과 함께 '시리' 개발에 관여한 것으로 알려짐

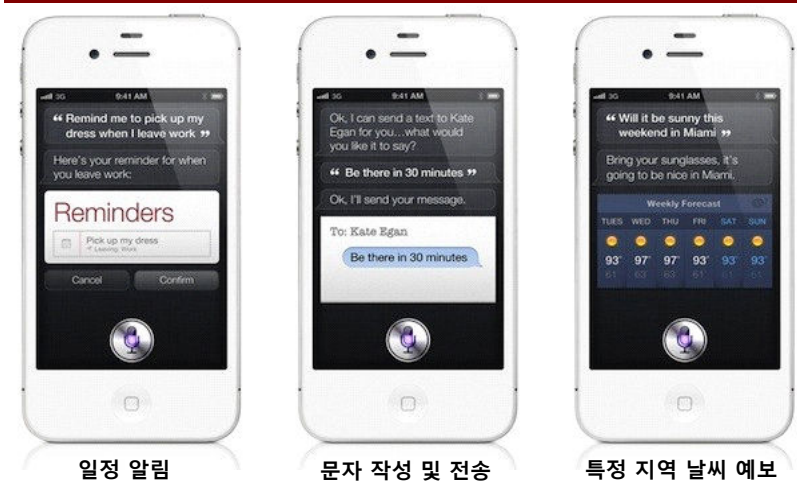
4. 음성인식 활용 사례

가. 모바일 음성인식

(1) 애플의 음성인식 앱 '시리(Siri)'

- 애플은 지난 10월 신규 스마트폰 '아이폰 4S(iPhone 4S)' 출시와 함께 사용자의 음성 명령을 인식하고 웹과 온라인 서비스를 검색해 답변을 제공하는 인공지능형 음성인식 서비스 '시리¹⁰⁾'를 발표함

Figure 5. 애플의 음성인식 앱 '시리' 실행 화면



일정 알림

문자 작성 및 전송

특정 지역 날씨 예보

자료: 애플

- '시리'는 단순히 사용자의 음성을 인식하는 것뿐 아니라 현재 위치와 일정 계획, 주소록 등 상황 맥락에 따른 답변을 제시, 사용 시간에 따라 이용자 선호도를 파악해 답변을 제시하는 기능도 갖추고 있음

Table 5. 애플의 음성인식 앱 '시리'의 특성

구분	특성
특정 명령어 불필요	- 명령어의 맥락을 이해하고 위치 정보를 포함한 안내를 제공함 - 예를 들어 "맛있는 피자가 먹고 싶은데!" 라고 말하면, 주변의 피자 레스토랑 정보를 Yelp(미국 생활정보 검색업체) 평점을 이용해 제공 - 주소록에 관계를 설정해 두었다면 누가 배우자인지, 친구인지, 부모인지 구별하므로 사람의 이름을 말하지 않고 "어머니의 이메일 주소는?" 등의 질문이 가능함
학습인식 기능 탑재	- 사용자의 목소리를 학습해 인식률을 높이고 더 정확한 답변을 도출함
빠른 구동	- 직관적인 UI로 구동이 쉽고, 답변 시 딜레이가 거의 없음

10) 2010년 애플이 인수한 업체의 명칭이기도 한 '시리'는 아이폰 용 무료 앱인 '시리 어시스턴트(Siri Assistant)'를 개발한 바 있음

주요 기능 연동

- 특정 날짜에 일정을 잡도록 요구하면, 해당 일에 중복되는 일정이 없는지 체크하고, 충돌이 발생할 경우 다른 일정으로 변경하도록 권고

자료: 스트라베이스

- 구글이나 MS의 음성인식 앱과 비교해 ‘시리’의 가장 큰 차별점은 자연어 음성인식 기능을 구현하고 있다는 것으로, 딱딱한 명령어 대신에 좀 더 자연스러운 대화형 인터페이스를 구축하고 있음
- ‘시리’의 핵심 기술이 가진 차별성은 단순히 음성 인식이라기보다는 사람의 말을 해석하는 능력에 있음
 - ‘시리’는 단순한 UI가 아니라 이용자의 말을 잘 알아듣고 거기에 맞게 반응하는 인공지능 서비스라고 볼 수 있음
- ‘시리’에는 천재 물리학자 스티븐 울프람(Stephen Wolfram)이 개발한 지능형 검색엔진 울프람알파(Wolframalpha)¹¹⁾가 적용됨
 - 울프람알파는 기존 검색 서비스가 인터넷에 있는 정보들을 나열하는 것과 달리, 자체적으로 수집한 방대한 정보를 활용해 ‘의미 있는 답’을 만들어 제공함¹²⁾
 - 울프람알파는 문맥에 맞는 결과를 찾아준다는 점에서 ‘연산능력을 갖춘 지식엔진’으로 평가받고 있음
 - 울프람알파는 API 형태로 제공되며, 기업이 자사 시스템에 이를 연동해 사용하고 로열티를 지불하게 됨
 - 스티브 잡스는 평소 스티븐 울프람과 친분을 맺어 왔으며, 애플이 ‘시리’에 울프람알파 기술을 적용하는데 큰 영향을 준 것으로 알려짐
- ‘시리’의 구현 과정은 정확하게 알려져 있지 않지만, 아이폰으로 음성을 받아 웹 DB에서 텍스트로 변환하고, 이들 중 웹 검색이 필요한 일부를 울프람 검색 엔진에 넣어 결과 값을 도출한 후 다시 아이폰에 적용하는 클라우드 방식인 것으로 보임
 - 아이폰은 음성 신호와 각종 데이터를 클라우드 웹DB와 주고받을 때 Wi-fi나 3G 네트워크를 사용함
 - 3G나 Wi-fi 연결이 되어 있지 않으면 ‘시리’가 작동하지 않는다는 사실이 이러한 추정을 가능케 함

11) 홈페이지 주소는 www.wolframalpha.com임. 16세에 입자 물리학 논문을 썼고 20세에 박사학위를 받고 미국 캘리포니아공대 교수로 임용된 천재 물리학자 스티븐 울프람은 1987년 울프람 연구소(Wolfram Research Company)를 설립. 1988년 물리학과 수학 데이터를 연산하는 매스매티카를 개발했고, 이를 바탕으로 2009년 울프람알파 검색엔진을 선보였음. 울프람알파는 매스매티카라는 강력한 연산엔진을 통해 전 세계에 있는 다양하고 방대한 지식을 연산 가능한 형태로 정형화(Prototyping)하고, 이들을 컴퓨터로 연산함으로써 질문에 맞는 답을 제시함

12) 예를 들어, 구글에서 'Korea'를 검색하면 'Korea'라는 단어가 들어간 수많은 웹 사이트를 제시하지만, 울프람알파는 면적과 해안선 길이 등의 지리 정보와 인구통계, GDP 등을 정리해서 보여줌. 울프람알파는 경제나 지리 정보를 비교할 때 가장 유용한데, 예를 들어 영국과 덴마크의 인구밀도를 비교하는 질문에 즉각적으로 답을 제시해 줌

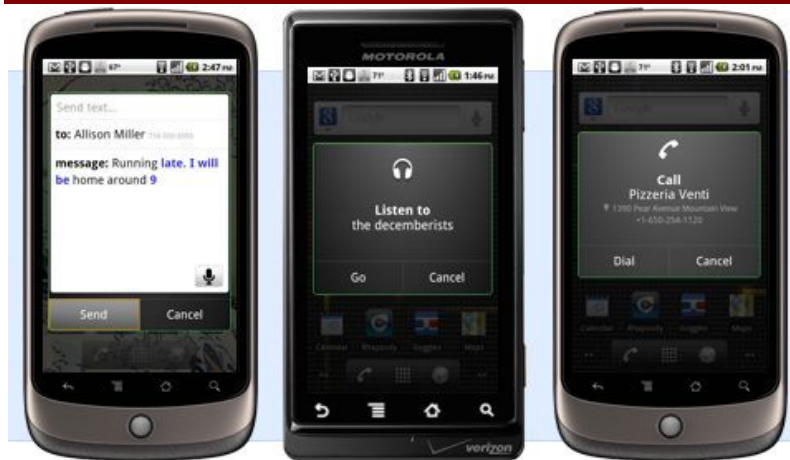
- 그러나 시리가 얼마나 많은 음성 데이터를 웹DB나 윌프람알파로 처리하는지, 아이폰에서도 음성 인식과 데이터 검색이 이루어지는가에 대해서는 알려져 있지 않음
- ‘시리’는 단순한 음성 인식 서비스나 음성 인식 UI를 넘어 많은 산업에 파급효과를 미칠 수 있는 동인으로 작용할 전망
 - 가장 대표적인 것이 바로 TV로, 애플은 2013년 출시가 예상되는 애플 TV에 ‘시리’ 기능을 탑재할 것으로 보임
 - 스티브 잡스 사후에 출간된 공식 전기¹³⁾에도 ‘시리’가 애플 TV에 탑재될 가능성이 언급되어 있어 시장의 관심을 받고 있음
- ‘시리’가 애플 TV에 적용될 경우, 리모콘의 단순한 기능이 음성인식으로 대체되어 고도화될 것으로 전망됨
 - 미국 투자은행 파이퍼 제프리(Piper Jaffray)의 애널리스트는 “애플 TV에 ‘시리’가 탑재되면 프로그램 이름이나 배우의 이름을 TV 리모콘으로 입력해야 하는 귀찮은 일이 단순해질 것”이라고 언급
 - ‘시리’를 통해 콘텐츠 내용에 대해 묻고 답할 수도 있는데, 예를 들어 미식축구 시청 중에 ‘지금 터치 다운을 한 선수의 이름은?’이라는 질문에 ‘시리’가 답을 하는 수준으로 발전할 가능성도 있음

(2) 구글의 음성인식 앱

- 구글의 음성인식 서비스는 2010년 1월 출시된 안드로이드 2.1(Android 2.1)을 통해 본격적으로 가시화됨
 - 구글은 안드로이드 2.1버전의 첫 탑재 단말인 넥서스원(Nexus One)을 자사 브랜드로 출시하며 다이얼, 이메일, 트위터 포스팅 등 다양한 기능의 음성 인식 서비스를 공개했으며, 후속 출시된 최신 안드로이드 단말에도 공통적으로 적용됨
 - 2010년 6월에는 한국어 버전의 음성인식 서비스를 출시, 국내업체 대비 상당한 수준의 음성 인식률을 보여주며 기술력을 입증함
 - 2010년 8월에는 안드로이드 2.2 이상 버전을 탑재한 단말기에서 전화걸기, 문자보내기, 메일작성, 메모, 일정 예약, 알람, 목적지 찾기, 지도 검색 등을 음성으로 실행할 수 있는 ‘보이스액션(Voice Action)’ 서비스를 발표함

13) 책의 내용을 보면, 쓰기 쉬우면서도 다른 전자 단말기와 함께 사용할 수 있는 기기를 만들고 싶어 하던 잡스가 전기의 작가인 아이작슨에게 “당신이 생각할 수 있는 가장 간단한 인터페이스를 갖게 될 것인데, 그 방법을 찾았다”는 구절이 있음

Figure 6. 구글의 '보이스액션' 실행화면



* 사용자가 전화를 거는 명령어 'Call'과 사람이름을 연달아 외치면 번호를 누르지 않고도 전화를 걸 수 있음. 메일과 문자를 보낼 경우에도, 명령어와 받을 사람 이름을 말한 뒤 전달할 내용을 이어서 말하면 음성이 문자로 변환돼 메일이나 텍스트 메시지로 작성됨

자료: 구글

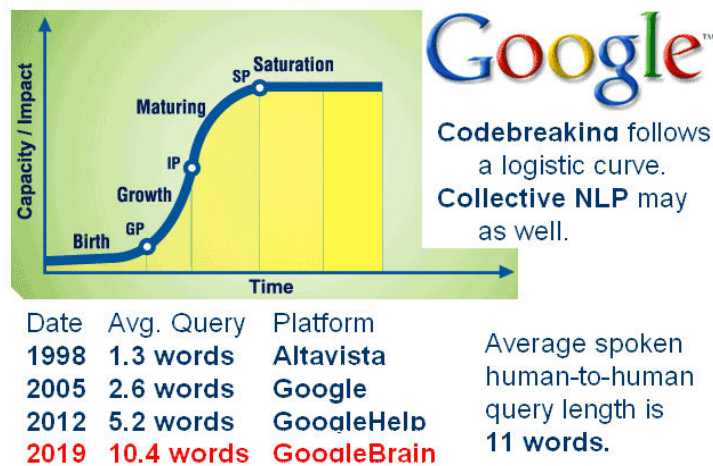
- 구글의 음성인식 기술은 기계에 직접 말을 가르치는 것이 아니라 인간 음성의 파형을 통계적으로 분석해 현재 이용자가 의도하는 바를 유추해냄
 - 따라서 충분한 이용자 기반과 인프라 없이는 선불리 시도할 수 없는 방식임
- 구글은 여러 장소에 분산된 서버를 묶어 한 대의 컴퓨터처럼 사용하는 클라우드 컴퓨팅을 활용해 음성인식 서비스의 속도와 정확도를 높이는 연구를 진행함
 - 구글은 스마트폰에 입력된 음성 데이터를 클라우드 서버로 전송하고, 서버에서 음성인식을 수행한 후 그 결과를 스마트폰으로 재전송하는 방식을 채택하고 있음
 - 음성인식률 95%를 표방하고 있는 구글은 2010년 기준 클라우드 서버에 성별, 연령별, 사투리로 구분한 총 2,300억 개의 영어 단어를 음성 데이터로 저장하고 있는 것으로 집계됨
 - 구글의 음성인식 기술 책임자 마이클 코언(Michael Cohen) 박사의 설명에 따르면 현재 구글의 IDC(인터넷 데이터 센터)로 하루에 들어오는 음성량은 한 사람이 2년 동안 쉬지 않고 얘기하는 양과 맞먹는 것으로 나타남
 - 따라서 세계 각국에 설치된 구글 IDC의 전산자원을 극히 일부분만 활용해도 대용량 음성 데이터를 바탕으로 정확도를 향상시킬 수 있음
 - 클라우드 서버 이용 방식은 이용자의 사용량이 늘어날수록 음성 데이터가 축적되기 때문에 향후 안드로이드 기반의 모바일 기기 보급이 확산될 경우, 구글의 음성 데

이터 확보 및 인식률 향상에 더욱 유리할 것으로 전망됨

- 사람의 경우 평균 8~11단어로 구성된 문장을 사용하는데, 현재 구글의 음성인식 기술은 5단어 정도의 문장을 인식하는 수준임
- 구글은 모바일 환경에서 엄청난 수의 안드로이드폰 사용자들을 통해 다양한 언어의 음성 데이터를 모을 수 있으며, 이를 통해 수집된 음성을 가공하여 성능 개선에 활용하는 선순환 구조를 형성할 것으로 전망됨

Figure 7. 구글 음성기술 향후 개발 계획

The Conversational Interface (CI): Circa 2015-2020 Developmental Attractor



자료: The Conversational Interface: Our Next Great Leap Forward(2007)

(3) 내비게이션

- 최근 운전 중 화면을 주시하며 글자를 입력해야 하는 내비게이션의 안전성을 높이기 위해 음성인식 기술의 도입이 증가하고 있음
- 음성인식 내비게이터가 설치되어 있는 자동차의 계기판은 차 안에서 가장 잡음이 심한 공간으로, 최근 끝점 검출(Endpoint Detection)기술¹⁴⁾ 등 잡음제거 기술의 발달로 음성인식 성능이 향상되고 있음
- 스트래티지 어널리틱스(Strategy Analytics)에 따르면, 2015년까지 전 세계 출시 자동차의 47%에 음성인식 기술이 탑재될 것으로 전망됨
- 포드(Ford) 자동차는 마이크로소프트와 협력해 음성으로 휴대전화와 엔터테인먼트 시

14) 지금까지 개발된 끝점 검출기술은 잡음 기술을 효과적으로 제거하는 음성개선 모듈을 기반으로, 에너지 레벨을 기준으로 1차 음성 시작부분을 찾고 일정 구간 내에 유성음의 비율을 계산해 최종 음성의 시작점을 결정함. 음성의 끝 부분에서도 비슷한 방법으로 끝점을 결정하는데 이런 방식은 발화 전후의 돌발성 잡음에서도 끝점을 효과적으로 검출하기에 용이함

시스템을 조작할 수 있는 음성인식 시스템 '싱크(Sync)'를 주요 차량에 탑재하고 있음

- 이 기능을 사용하면 운전 중 음성으로 음악 선곡, 목적지 검색 등이 가능하며, 현재 다양한 가격대의 차종에 적용되고 있음(국내 출시된 차종 중에는 '포커스'와 '익스플로러'에 적용)
- 2012년 이후부터는 운전 중 문자메시지 이용에 따른 위험을 줄이기 위해 문자메시지 음성인식 기능을 장착하기로 발표함

Figure 8. 포드사의 음성인식 시스템 '싱크' 소개 페이지



자료: 포드

- 국내 내비게이션 업체 파인디지털은 2011년 8월 세계 최초로 150만 단어 수준의 대어휘 인식이 가능한 음성인식엔진 FineSRTM 7.0을 탑재한 내비게이션 '파인드라이브 iQ3D2000v'를 출시함
- 목적지 검색뿐만 아니라 DMB 채널 전환과 화면 밝기 및 음량 등을 음성으로 조작할 수 있으며, 음성인식율을 90%까지 향상시킴

Figure 9. 음성인식 내비게이션 '파인드라이브' 실행화면



자료: 파인디지털

나. 개인화 서비스

(1) 보안 및 금융 분야

- 보안 및 금융 분야에서는 이용자 음성으로 신원이나 감정, 심리 상태를 파악하여 본인인증, 신용평가 활용 서비스 등을 제공하고 있음
- 러시아 국영은행 스베르뱅크(Sberbank)는 ATM 거래 시 질문에 대한 이용자의 답변을 판단하여 거래 가능 여부를 결정하는 서비스를 시험 도입함
 - 신용카드 발급 등 모든 신용거래를 ATM에서 가능하게 만들기 위해 개발한 시스템으로 스피치테크놀로지센터(SPC)가 러시아 사법당국이 보유한 거짓 진술자의 녹취록을 참조하여 개발함

Figure 10. 스베르뱅크의 음성인식 ATM



자료: 뉴욕타임즈(2011.6)

(2) 교육 분야

- 음성인식 기술은 외국어 학습 시 억양 및 발음 교정을 위해 활용되고 있으며, 최근 의사소통 중심의 실용 영어가 중요시 되면서 교육 분야에서의 적용이 더욱 확대될 것으로 전망됨
- 특히 음성인식을 활용한 발음교정 프로그램이 잇따라 출시되고 있는데, 원어민의 발음을 듣고 이용자가 주어진 단어 혹은 문장을 발화하면, 프로그램이 직접 음절 단위로 발음, 억양, 강세, 속도 등을 분석해 잘못된 부분을 교정해 주는 형태임

Figure 11. 음성인식 발음교정 소프트웨어 용례



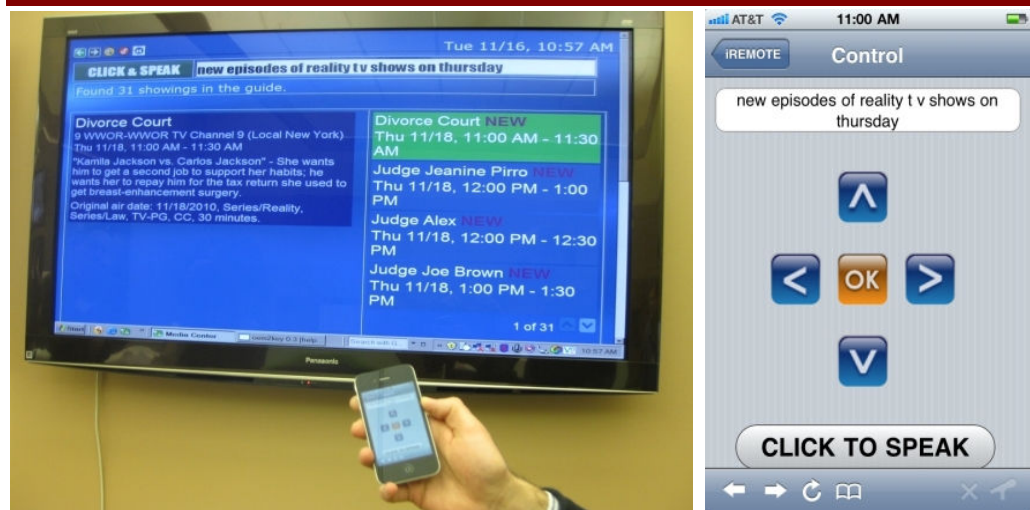
* 원어민과 사용자의 발음을 비교 분석하여 결과를 제시하고, 오류 유형 및 상세한 교정 정보를 제공함
 자료: 닥터스피킹

다. 기타

(1) TV 프로그램 검색

- TV 프로그램 검색은 대표적인 음성인식 활용 사례로, 2010년 미국 이동통신사 AT&T는 스마트폰을 통해 음성으로 TV 프로그램을 검색할 수 있는 서비스를 시연했음
 - AT&T는 왓슨(Watson)의 음성인식 기술을 적용, 자사 IPTV 서비스 'U-Verse' 가입자들을 상대로 음성인식 리모컨 애플리케이션을 제공함

Figure 12. 구글 음성기술 향후 개발 계획



* 좌: 스마트폰 음성인식 애플리케이션을 이용한 IPTV 서비스 실행 화면

우: 스마트폰 음성인식 애플리케이션

자료: AT&T

- 2011년 마이크로소프트(MS)는 게임콘솔 '엑스박스 360(Xbox 360)'에 음성인식 기술을 탑재, 자사 검색엔진 Bing과 키넥트(Kinect)를 접목해 말만 하면 원하는 콘텐츠가 TV에 뜨는 음성 검색 서비스를 소개함
- 한편 업계 전문가들은 향후 애플 TV에도 음성인식 서비스 '시리'가 탑재될 것으로 전망하고 있는데, 이 경우 단순히 음성인식 수준을 넘어 이용자의 음성을 이해하고 콘텐츠에 관한 세부 정보를 제공하는 형태¹⁵⁾까지 발전할 것으로 전망됨

(2) 콜센터 음성처리

- 음성인식 기술을 자동응답시스템(ARS)에 적용, 소비자가 직접 원하는 상담업무를 말로 전달할 수 있는 시스템들이 개발되고 있음

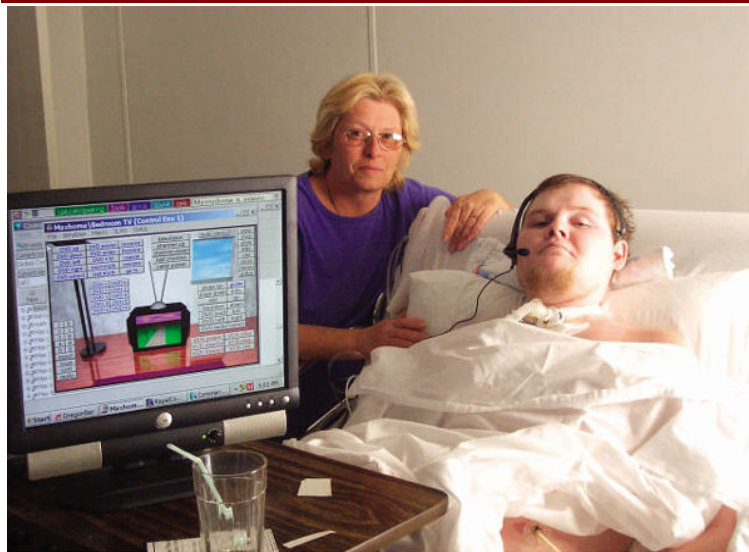
15) 예를 들면, 미식축구 경기 시청 중 “지금 터치다운을 한 선수의 이름은?” 이라는 사용자의 질문에 답하는 수준까지 발전될 것으로 전망됨

- 현재는 소비자가 ARS 안내 멘트를 모두 확인하고 원하는 상담업무를 받을 수 있지만, 음성인식 기술이 적용될 경우 서비스 메뉴를 알지 못해도 예를 들어 “oo은행으로 100만원 이체”라고 말하면 계좌이체를 완료할 수 있음
- 국내 콜센터 및 지능형 통신장비 전문 업체 브리지텍은 세계적인 음성인식 전문 업체 니앙스(Nuance)사와 제휴를 통해 자연어 음성인식 기술을 적용한 ARS 솔루션을 개발해 보급하고 있음

(3) 환경 제어

- 음성인식 기술은 신체 활동이 어려운 중증 장애우들을 위한 환경 제어 장치에도 탑재돼 활용되고 있음
- 미국에서 사용되고 있는 ‘멀티미디어 맥스(Multimedia Max)’는 컴퓨터 기반의 환경 제어 장치로, 음성인식 프로그램인 드래곤 내추럴리 스피킹(dragon naturally speaking)¹⁶⁾이 탑재되어 있어 컴퓨터의 작동과 프로그램의 실행, 맥스 프로그램의 사용이 모두 음성 명령으로 가능함

Figure 13. 음성인식 환경제어 장치 ‘멀티미디어 맥스’ 사용 모습



자료: 멀티미디어 맥스 홈페이지

16) 니앙스 커뮤니케이션의 음성인식 플랫폼으로, 화자의 연속적인 발화(분당 100단어 속도)를 인식할 수 있음

5. 향후 전망 및 시사점

가. 가장 직관적인 인터페이스로 다양한 IT 기기와 융합

- 음성은 인간이 가진 가장 직관적이고 간편한 입력 도구임
 - 스티브 잡스가 아이폰에 ‘넥스트 터치 스크린’ 인터페이스로 음성인식에 기반한 ‘시리’를 탑재한 것은 음성의 ‘직관성’에 주목했기 때문으로 풀이됨
 - 잡스는 아이폰 이전 터치스크린에 사용되던 스타일러스 펜의 번거로움을 “신은 인간에게 10개의 스타일러스 펜(손가락)을 줬는데, 더 이상의 스타일러스가 왜 필요하느냐¹⁷⁾”고 지적
 - 유사한 관점에서 음성은 터치스크린보다도 더 직관적인 인터페이스 수단이나, 지금까지는 음성을 인식하고 문맥에 맞게 해석하는 기술이 실용화에 이를 정도로 발달하지 못했음
 - 그러나 음성인식 기술의 정확도 향상, 문맥을 해석해 유저의 상황에 맞는 결과를 도출하는 시맨틱 검색(Semantic Search) 기술의 발달은 음성인식의 활용성을 크게 증가시켰음
- 음성인식 기술은 다른 IT 기기나 서비스와 결합됐을 경우에 그 파급력이 극대화될 것으로 전망
 - ‘시리’가 고성능 시맨틱 검색 엔진인 ‘울프람알파’와 결합되어 한 차원 진일보된 형태의 음성인식 서비스를 보여줬듯이, 앞으로 음성인식 기술은 다양한 기기나 서비스와 융합됐을 때 파급력이 극대화될 전망
 - 음성인식 기술은 TV나 내비게이션, 가전기기, 홈네트워킹, 금융서비스 등 다양한 분야에 적용될 수 있을 것으로 예상됨
 - 이 경우, 각종 전자 기기를 음성으로 제어하거나 기존에는 불가능했던 형태의 명령을 내릴 수 있게 됨
 - 예를 들어, 차를 몰다가 ‘가장 가까운 주유소’, ‘가장 저렴한 주유소’ 등의 질문만으로도 목적지를 찾아갈 수 있고, TV를 시청하다가 ‘지금 출연 중인 배우의 프로필은?’과 같은 음성 명령으로 즉각적으로 정보 검색이 가능해짐
- 향후 음성인식 기술은 인공지능, 정황인식 등 지능형 서비스의 결합으로 사용자 경험을 최적화하고, 서비스의 활용성을 증대시킬 수 있을 것으로 전망됨
 - 시장조사기관 가트너(Gartner)에 따르면, 현재 IT 벤더들은 제품에 대한 소비자들의 기대치가 변화하고 있다는 사실에 주목하고, 향후 출시할 제품에 음성인식 기술 탑재

17) “We don’t need a stylus when God gave us ten of them!”

재 의무화를 고려하고 있음

- 그러나 음성만으로는 정확한 정보 전달이 어려울 수 있으므로 터치, 동작인식 등 다른 입력 방식과 음성인식을 결합해서 활용하는 방안도 모색해야 할 필요가 있음
 - 미국 이동통신사 AT&T의 지도검색 애플리케이션 ‘스픽포잇(Speak4It)’의 경우 ‘레스토랑’이라 말하고 지도 위에 동그라미를 그리면 영역 내의 음식점을 표시해주는 서비스를 제공하고 있음

나. IT 소외계층으로 사용자 저변 확대

- 음성인식 기술은 기기 조작에 익숙하지 않은 노인, 장애우 등 IT 소외 계층에게 보다 편리한 서비스와 기능을 제공할 수 있는 매개로 작용할 전망
 - 음성은 인간이 가진 가장 기본적인 도구 중 하나로, 언어의 문맥을 이해하는 음성 인식 기술을 통해 노인이나 아이들이 IT 기기의 복잡한 조작법과 사용법을 익히지 않아도 자유롭게 활용할 수 있음
 - 한편, 손을 자유롭게 움직일 수 없는 장애우나 시각 장애우들 또한 음성 인식 기술을 사용해 IT 기기를 활용할 수 있게 될 전망
- 기업 입장에서조차 음성인식 기술을 활용함으로써 고객 저변을 IT 기기 소외 계층으로까지 확장시킬 수 있을 것으로 예상됨
 - 미쓰비시 전기(Mitsubishi)는 장애우를 위한 음성인식 엘리베이터 시스템 ‘터치리스 콜(Touchless Call)¹⁸⁾’을 개발했으며, 인텔, GE 등은 음성인식 기술을 활용해 노인이나 장애우의 보조기기를 개발하는 ‘CAST(Center for Aging Service Technologies)’ 프로젝트를 수행하고 있음

18) 사용자가 버튼을 누르는 대신 승강장에서 엘리베이터를 호출하고, 엘리베이터에 탑승하면 원하는 층을 음성으로 지정하는 시스템

□ 참고문헌

- 신성장동력산업용 대어휘 음성인식 기술 동향 및 응용, 전자통신동향분석 제 23권 제 1호, 2008.2
- 아이폰4S - 음성인식기술(Siri)의 부각, 메리츠증권증권, 2011.10.17
- ‘음성인식’, IT 업계의 Next Big Thing으로 부상 중...플랫폼 Big 3의 음성인식 UI 경쟁 구도 분석, 스트라베이스, 2011.6.27
- 자연어 음성인식 기술을 이용한 음성 대화 서비스 개발동향, 전자통신동향분석 제 26권 제 5호, 2011.10
- SERI 경영노트-음성인식 기술의 재발견, 삼성경제연구소, 2011.9.1
- Apple's Siri: Game-Changer, Not Gimmick, AllThingsD, 2011.10.10
- A Russian A.T.M With an Ear for the Truth, New York Times, 2011.6.8
- AT&T Ups the Ante in Speech Recognition, CBS News, 2010.11.18
- 50 years of progress in Speech Recognition Technology, Sadaoki Furui, Plenary Speech, 2007

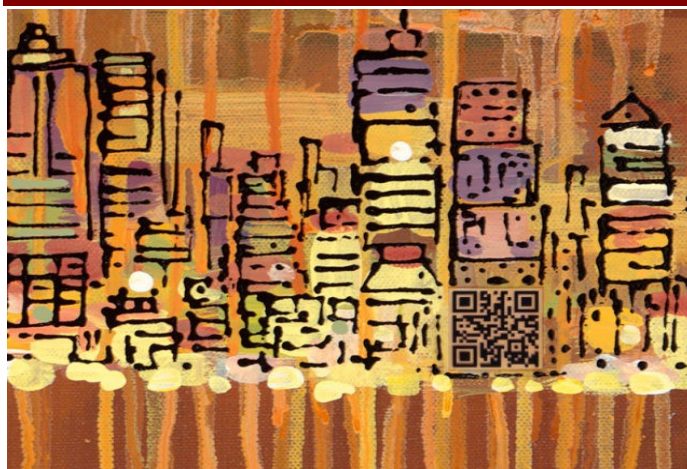
1. 인터랙티브 예술 ‘앰비언트 미디어 포털’, QR 코드와 회화 예술의 만남

QR 코드의 활용도가 여러 분야로 확산되고 있는 가운데, 이를 자신의 회화 작품에 삽입하여 새로운 인터랙티브 예술로 승화시킨 사례가 등장해 눈길을 끌고 있다. 화가이자 모바일 디자이너인 조스 토레스(Jose Torres)는 자신의 도시 풍경화에 QR 코드를 삽입해 이를 스캔한 관객들이 그림 주제와 관련한 애니메이션을 그림과 함께 즐길 수 있도록 했다.

□ QR 코드를 통해 인터랙티브 예술인 ‘앰비언트 미디어 포털(AMP)’로 새롭게 탄생한 미술 작품

- QR 코드를 자신의 그림에 삽입해 그림 감상뿐만 아니라 작품 주제와 관련된 애니메이션이나 영화, 사진 등의 다양한 미디어를 함께 경험할 수 있는 인터랙티브 예술이 등장함
 - 토니 타지(Tony Taj)라는 예명으로 활동 중인 화가이자 모바일 디자이너 조스 토레스는 자신의 미술 작품과 관련된 애니메이션이나 사진 등을 QR 코드와 연결된 페이지에 올리고 관객들이 즐길 수 있도록 함
 - 스마트 디지털 기술이 예술과 관객 사이의 일종의 통로(Portal) 역할을 한다는 의미에서 토레스는 이 예술 기법을 ‘앰비언트 미디어 포털(AMP)’이라고 명명함
 - 토레스의 도시 그림들에 삽입된 QR 코드를 스마트폰으로 스캔한 관객들은 그림 속 도시의 이야기가 담긴 애니메이션을 감상할 수 있으며, 이를 통해 관객들은 그림의 주제에 대한 보다 깊은 이해가 가능함

Figure. QR 코드가 삽입된 조스 토레스의 AMP 예술 작품



자료: www.kickstarter.com

- 토레스는 디지털 기술의 적용이 예술적인 성취를 저해하는 것이 아니라 오히려 상승 작용을 불러올 수 있다고 밝힘
 - 토레스는, “나는 눈속임을 하려 한 것이 아니라 그림을 보면서 얻을 수 있는 시각적 경험을 보다 확장하길 바랐다”며, “사람들이 그림을 스캔하고 애니메이션을 보며 즐거워하는 것을 보면 무척 만족스럽다”고 말함
- 토레스는 자신의 AMP 예술이 기술 발전에 따라 QR 코드 사용에 머무르지 않고 NFC(근거리 무선통신) 기술 등 다양한 기술을 적용하여 변화해 나갈 수 있을 것으로 기대함

□ 참고문헌

- ‘QR Codes Help Paint the Deeper Pictures Behind Art’, PCWorld, 2011.11.3

2. 옷과 종이에도 적용할 수 있는 만능 터치 센서

정전식 터치 스크린 방식이 사실상 적용되기 어려운 종이나 옷에 적용할 수 있는 새로운 방식의 터치 센서 기술이 개발되었다. 독일의 뮌헨 대학 및 하수 플래트너 연구소(Hasso Plattner Institute)의 연구진들은 해저 심해 케이블의 결함 여부를 탐지하는 데 사용되어온 TDR(Time Domain Reflectometry) 방법을 응용, 원하는 어떤 소재에서든지 손가락의 접촉 여부를 판단해낼 수 있는 센서 시스템을 구현해 냈다.

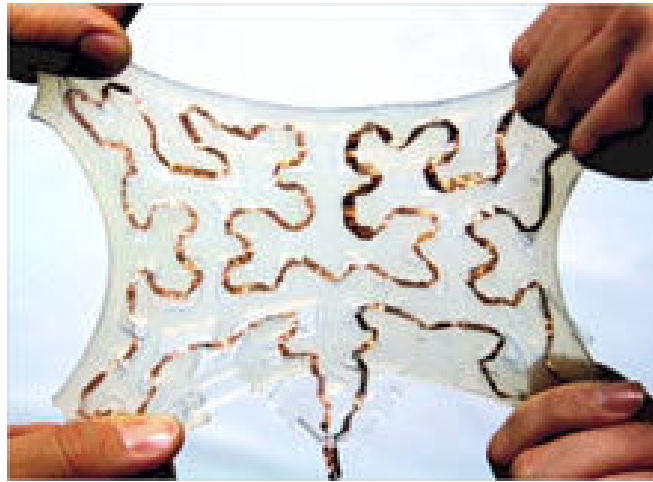
□ 새로운 방식의 터치 센서 기술 개발

- 옷과 종이 등 일상생활에서 사용하는 거의 모든 물체에서 손가락 등에 의한 접촉 여부를 판단하고 감지해 낼 수 있는 터치 센서 기술이 독일 연구진들에 의해 개발됨
 - 뮌헨 대학 및 독일의 하수 플래트너 연구소(Hasso Plattner Institute)의 연구진들이 공동으로 개발에 참여
- 독일 연구진들은 이번 터치 센서 기술 구현에 TDR(Time Domain Reflectometry, 이하 TDR) 기술을 이용한 것으로 알려짐
 - TDR은 오랫동안 수중 케이블의 결함 여부를 탐지하는 데 사용되어왔던 방식임
 - 케이블 등의 도체에 짧은 전기 펄스를 흘려보내고, 특정 지점이나 예상 지점으로부터 반사되어 돌아오는 시간을 측정함
 - 해당 지점에 손상이 있거나 혹은 전기 전도 상의 변화가 있을 경우 예상했던 시간과 실제 전기 펄스가 진행한 시간에 차이가 생기게 되고, 이 시간차를 통해 문제가 발생한 지점의 위치를 확인 할 수 있는 원리임
 - 하수 플래트너 연구소의 컴퓨터 과학 교수인 패트릭 보디쉬(Patrick Baudisch)는 “1960년대에 이미 TDR 기술이 전선의 손가락 접촉 여부를 판단하는 데 사용될 수 있음이 알려져 있었다”며, “매우 짧은 거리에서의 초단기 시간 지연을 측정하는 기술이 최근 정밀해지면서 TDR 기술을 인터랙티브 애플리케이션에 활용할 수 있는 가능성이 생기게 되었다”고 밝힘
- TDR 방식을 응용한 터치 센서 기술은 직관적인 구현을 가능케 함
 - 두 가닥의 구리 띠를 연결 한 종이의 어떤 지점에 손가락을 가까이 가져갔거나 접촉이 발생했는지를 확인할 수 있음
 - 연구진들은 시연에서 종이 표면에 구리 띠를 붙이고 이를 전기 신호 발생기와 검출기에 연결함
 - 구리 띠가 연결된 종이에 손가락을 가까이 가져가거나 접촉했을 때 발생하는 전기

신호를 통해 손가락의 움직임을 측정할 수 있게 됨

- 측정된 전기 신호의 파형(waveform)을 컴퓨터를 통해 분석하여 접촉이 발생한 지점을 결정할 수 있음
- 특히 이번 기술은 적용된 소재의 표면이 구겨지거나 늘어나도 된다는 장점이 있음

Figure. 구리띠를 종이에 연결해 TDR 방식 터치 센서를 구현한 모습



자료: MIT Technology Review

□ 이번 기술은 일상생활 분야에서 그 응용분야가 다양할 것으로 기대되고 있음

- 옷은 물론 종이와 테이블에 이르기까지 다양한 소재와 재질에 터치 센서를 부여할 수 있다는 점이 매력적이거나, 일상생활에 활용되기 위해서는 관련 장비의 소형화가 먼저 이루어져야함
 - 연구진은 “현재의 기술 단계가 걸음마에 불과한 것은 사실”이라면서도, “전기 신호 발생 및 검출 기능과 접촉 위치 판별 기능이 칩 하나로 구현될 수 있다”고 자신함
 - 실제로 연구진들은 향후 몇 달간 TDR 터치 센서 시스템을 단일 칩에 구현하는 것을 목표로 실험을 진행할 예정임
- 아이폰 등의 기존 정전식 터치 스크린 방식에 비해 훨씬 간단한 것으로 평가받고 있는 이번 터치 센서 기술은 헤드폰 케이블 등에 실제로 활용될 수 있을 것으로 보임
 - 대형 멀티 터치 디스플레이 개발 업체인 퍼셉티브 픽셀의 CEO인 제프 한(Jeff Han)은 “TDR 방식을 응용한 터치 센서 기술은 매우 훌륭하다”며, “현재의 헤드폰 케이블을 변형하지 않고도 터치 입력 기능을 부여하는 데 사용될 수 있을 것”이라고 평가

□ 참고문헌

- ‘A Versatile Touch Sensor’, MIT Technology Review, 2011.11.01

3. 디지털 음악 시장의 미래, 자동차·수익모델·소셜 등 3대 변수로 지각 변동 예고

음악 시장이 콘텐츠 디지털화의 침범으로 주목받고 있지만, 디지털 음악 시장의 성장세가 전반적인 음악 시장의 쇠퇴를 상쇄하지 못하고 있다는 분석이 나오고 있다. 이에 대해 IT 전문 미디어 기가옴(GigaOm)은 △자동차를 통한 스트리밍 음악 서비스의 활성화 △월정액 기반 가입형 서비스와 광고 기반 무료 서비스로 양분된 스트리밍 음악 서비스 수익 모델 △페이스북(Facebook) 등과 연계된 소셜 음악 서비스의 확산 등을 향후 디지털 음악 시장을 좌우할 지표로 언급했다.

□ 2010년 전 세계 디지털 음악 시장 규모 59억 달러...음악 시장 침체 상쇄하기엔 역부족

- 전체 음악 시장의 40%를 점유하고 있는 디지털 음악 시장의 매출 규모가 2010년 기준 59억 달러로 집계돼, 1999년 전성기 매출 146억 달러에 한참 못 미치는 것으로 나타남
- 시장조사업체 가트너(Gartner)는 2011년 전 세계 디지털 음악 시장 규모가 63억 달러 수준으로 증가할 것이며, 2012년에는 68억 달러, 2015년에는 77억 달러까지 증가할 것으로 전망함
 - 디지털 음악 시장의 꾸준한 성장세에도 불구하고 전체 음악 시장의 침체를 상쇄할 만큼의 성장 속도를 보이고 있지 않아 시장 전망은 부정적임

□ 디지털 음악 시장 성장의 3대 변수

- 한편, IT 전문 미디어 기가옴(GigaOm)은 향후 디지털 음악 시장 전망에 있어 △자동차 기반 음악 서비스 △수익 모델 △소셜 음악 서비스 등을 변수로 지목하고 있음
 - 스마트폰이 모바일 스트리밍 음악 서비스의 성장을 주도하는 가운데, 미국인이 가장 많이 음악을 청취하는 장소인 자동차 역시 스트리밍 음악 서비스의 성장을 주도하는 서비스 플랫폼으로 자리매김할 전망이다
 - 월정액 기반의 가입형 서비스와 광고 기반 무료 서비스 등으로 양분된 스트리밍 음악 서비스 업계의 수익모델 역시 향후 디지털 음악 시장의 성장 양상에 변수로 작용할 것으로 예상됨
 - 다양한 스트리밍 음악 서비스와의 연계를 통한 소셜 음악 서비스를 제공하려는 페이스북(Facebook) 등을 중심으로 한 음악 콘텐츠 이용이 증가하면서, 기존 라디오 방송을 대체하는 현상도 확산될 것으로 전망됨

□ 참고문헌

- ‘Online music revenue to reach USD 6.8 bln in 2012 - study’, Telecompaper, 2011.11.8
- ‘The 6 ways digital music is being disrupted’, GigaOM, 2011.11.10

4. 사진의 미(美)적 수준을 판단하는 이미지 인식/분류 기술

사진의 미(美)적 수준에 따라 좋은 사진을 자동으로 찾아내 분류하고 정리해주는 기술이 제록스(Xerox)의 연구진들에 의해 개발되었다. 사람들이 좋은 사진으로 꼽는 사진들에는 일반화할 수 있는 원칙이 있다는 것을 전제로 하는 기술로, 인터넷 상에 공개된 사진 앨범들을 통해 좋은 사진의 공통 요소를 뽑아내어 이를 알고리즘화하는 원리로 구현되며 2012년 베타서비스를 목표로 하고 있다.

□ 미(美)적 수준에 따라 사진을 자동으로 분류 정리해주는 기술 개발

- 사람이 직접 사진을 분류 정리하지 않아도 사진의 내용은 물론 그 미적 수준에 따라 자동적으로 분류 정리해주는 기술이 개발되었음
 - 복사기로 유명한 미국 제록스(Xerox)社 산하 제록스 유럽 연구소(Xerox Research Center Europe)에서 개발한 기술임
 - 이 기술을 이용하면 가족 여행에서 찍은 수백 장의 디지털 사진 중 앨범에 남겨 오랫동안 간직해야 할 사진을 찾기 위해 잘 나온 사진들을 직접 고르는 대신, '잘 나온' 사진을 컴퓨터 기술을 활용하여 자동적으로 분류 정리할 수 있게 됨
- 현재는 프로토 타입 기술이 개발된 상황이며, 다양한 파트너 업체들과 함께 2012년 베타 테스트를 목표로 하고 있음
 - 프랑스의 그레노블(Grenoble)에 위치한 제록스 유럽 연구소의 컴퓨터 시각 연구 그룹 책임자인 크레이그 손더스(Craig Saunders)는 "제록스 파트너들과의 베타 테스트가 내년에 예정되어 있다"며, "이들 파트너 업체들 중에는 그래픽 디자인 업체 및 온라인 사진 앨범 업체 등 다양한 사진 관련 업체들이 포함되어 있다"고 밝힘

□ 좋은 사진의 기본 요소를 정의하고 이를 분석하여 인지/분류할 수 있는 시스템 구현

- 제록스의 연구진들은 미(美)적 분류 시스템을 개발하기 위해 페이스북이나 플릭커(Flicker) 등에 공개된 온라인 사진 앨범들 중에서 높은 평가를 받은 사진들의 공통 특징을 연구하였음
 - 이들 특징은 사진 전문가들이 좋은 사진의 요소로 꼽는 것들과 상당부분 일치했는데, 예를 들어 인물 사진의 경우 얼굴은 간접광으로 처리되어야 하며 인물에 초점이 맞춰지도록 배경은 흐릿하거나 단색이어야 함
 - 마찬가지로 좋은 해변사진은 느린 셔터 속도를 통해 얻을 수 있는 물방울처럼 부서지는 파도 이미지를 포함하고 있었음
 - 또한 사진을 3등분했을 때 수직과 수평으로 3분의 1 지점에 피사체를 위치시키는 황금분할비를 따른 사진들이 좋은 사진으로 꼽힌 경우가 많았음

- 손더스는 “우리는 어떤 요소들이 사진을 ‘좋은 사진’으로 만드는 가를 알기 위해 노력했다”고 말함
- 또한 연구진들은 사진의 이미지를 인지하기 위한 ‘시각 언어’도 개발하였음
 - 등그런 형태는 바퀴일 수 있으며, 녹색이 가득한 지역은 풍경일 수 있다는 식으로 수많은 이미지 요소를 시각 언어로서 정의하였음
 - 시각 언어 정의에 필요한 정보는 인터넷 상의 사진들에 붙어있는 텍스트 내용을 분석해 얻어졌음
- 원하는 ‘내용’이나 이미지를 포함하고 있는 사진을 시각 언어 등의 이미지 인식 기술을 통해 분류한 후, 이들을 다시 좋은 사진의 요소를 가지고 있는 사진들로 분류하는 방식임
 - 이미지 인식 기술을 활용하기 때문에 사진파일에 태그가 포함되지 않아도 됨
 - 예를 들어, 이 기술이 ‘새’들을 찍은 사진들에 적용될 경우, 새의 이미지가 사진의 많은 부분을 차지하고 있으며 배경이 단색이거나 간단해서 새 이미지를 잘 인식할 수 있는 사진들을 좋은 사진으로 분류하여 제시하게 됨
- 제록스는 이 기술을 바탕으로 한 클라우드 기반 사진 분류/검색 서비스를 내년에 론칭할 계획
 - 제록스는 최근 5백만 장에 달하는 사진을 1초도 안 되는 시간 내에 비슷한 이미지끼리 분류 정리할 수 있는 기술을 선보이기도 했음
 - 또한 일종의 사진 분석 페이스북 애플리케이션인 ‘케이트픽스(Catepix)’를 시장에 선보이기도 했음
 - 케이트픽스는 사용자의 페이스북 사진을 분석하여 이를 분류하고 이를 근거로 하여 사용자의 ‘성격’에 대해 말해주는 애플리케이션임
 - 제록스는 이번 미(美)적 분석 알고리즘 기술을 활용해 클라우드에 저장된 사진을 분류하고 검색하는 서비스를 내년에 시작할 계획임
- 참고문헌
 - ‘Beauty Now in the Eye of the Algorithm’, MIT Technology Review, 2011.11.17.

5. 美 도서관 e-북 대여 확산 열기, 아마존의 킨들 이용자 대여 도서관 서비스가 기폭제

PC, e-리더 등 기존의 e-북을 이용하기 위한 단말에 이어 스마트폰, 태블릿 PC의 보급이 확대되면서 공공 도서관의 e-북 대출이 활성화되고 있다. 특히 아마존(Amazon)이 지난 9월 공공 도서관과 협력 하에 킨들(Kindle) 이용자 대상 도서관 e-북 대출 서비스 '킨들렌딩라이브러리(Kindle Lending Library)'를 출시한 이후 도서관 e-북 대출건수가 증가한 것으로 파악되고 있다.

- 美 공공 도서관, e-북 대출건수 급증세...시카고 공공 도서관 2011년에만 대출 건수 380% 증가
 - 최근 美 공공 도서관을 통해 e-북을 대출해 이용하는 사례가 크게 증가하고 있음
 - 시애틀 공공 도서관은 2010년에만 e-북 대출건수가 92% 증가했으며, 시카고 공공 도서관 역시 2011년 동안 e-북 대출건수가 380% 폭증한 것으로 나타남
 - 뉴욕 공공 도서관은 e-북 대출 수요가 꾸준히 증가하면서 2009년 이래 e-북 관련 예산을 4배 증액하는 등 대응에 나섬
- 아마존, 도서관 e-북 대출 서비스 통해 e-북 대여 확산에 기여
 - 아마존은 지난 9월 공공 도서관과 연계한 킨들 이용자 대상 도서관 e-북 대출 서비스 '킨들렌딩라이브러리(Kindle Lending Library)'를 출시해 도서관 e-북 대출 활성화의 기폭제 역할을 하고 있음
 - '킨들렌딩라이브러리'는 e-북 유통업체 오버드라이브(OverDrive)의 디지털 서비스를 제공하는 도서관의 웹 사이트에서 도서관 ID로 e-북을 대출한 후 해당 e-북을 'Send to Kindle' 옵션을 통해 아마존 계정에 접속하여 모든 종류의 킨들 단말 및 애플리케이션에서 이용 가능한 서비스
 - '킨들렌딩라이브러리' 출시를 기점으로 뉴욕 공공 도서관과 시애틀 공공 도서관의 e-북 대출 건수가 30% 이상 증가한 것으로 드러남
 - 아마존은 최근 연간 79달러에 무료 배송 서비스를 이용할 수 있는 '아마존 프라임(Amazon Prime)'에 가입한 킨들 단말 보유자를 대상으로 아마존이 직접 e-북을 대여해주는 '킨들 오너스 렌딩 라이브러리(Kindle Owner's Lending Library)'도 출시함
- 일부 메이저 출판사, 도서관 e-북 대출 지원 거부...인기 도서 e-북 대출 제한
 - 도서관을 통한 e-북 대출 서비스 이용이 확산되면서, e-북 판매로 인한 디지털 매출 증대를 기대해왔던 출판사 진영은 당혹감을 드러내고 있음

- 사이먼 & 슈스터(Simon & Schuster), 맥밀란(Macmillan), 햏체트(Hatchette) 등 일부 대형 출판사들은 인기도서의 e-북 대출을 제한하는 등 급히 대응에 나서고 있음

□ 참고문헌

- ‘Kindle lending library mere "exercise of brute economic power" says Authors Guild’, Ars Technica, 2011.11.15
- ‘Libraries ramp up e-book lending’, USA Today, 2011.11.14

6. 신생 소셜 TV 서비스 ‘커넥 TV’, 美 메이저 방송사 진영 동참으로 관심 집중

TV 시청 중 SNS를 통해 지인들과 의견을 교류하는 소셜 TV 시청 행태가 확산되고 있는 가운데, 최근 미국 메이저 방송사들이 신생 소셜 TV 서비스 ‘커넥 TV(ConneCTV)’에 대거 동참을 선언하면서 소셜 TV 서비스에 대한 관심이 증폭되고 있다.

□ ABS, CBS, Fox, NBC 등 美 메이저 지상파 방송사의 ‘커넥 TV’ 참여...253개 채널 지원

- TV 시청자를 위한 실시간 SNS를 지향하는 신규 소셜 TV 서비스 ‘커넥 TV’에는 ABC, CBS, Fox, NBC 등 美 메이저 방송사가 투자했으며, ‘커넥 TV’의 지원 TV 채널도 253개에 달하는 것으로 나타남
 - 미국 내 7,600만 가구에 달하는 커버리지를 자랑하는 ‘커넥 TV’의 10개 제휴사는 자사 방송 채널 및 홈페이지 등을 통해 ‘커넥 TV’를 홍보할 방침임
 - ‘커넥 TV’에 동참을 선언한 방송사 중 다수는 ‘커넥 TV’에 상당한 금액을 투자한 것으로 파악되고 있으나, 구체적인 투자액 규모는 공개되지 않음
- 현재 초대받은 이용자에 한해 서비스의 일부 내용만 공개하고 있는 ‘커넥 TV’는 오는 2012년 공식 출시를 계획하고 있음
 - ‘커넥 TV’ 이용자는 같은 TV 프로그램을 시청하는 다른 이용자와 간편하게 상호 작용할 수 있으며, TV 프로그램과 관련된 다양한 콘텐츠도 제공받게 됨
- ‘커넥 TV’의 공동 설립자인 랜 애런(Ian Aaron)은 “전 세계적인 스마트폰, 태블릿 PC의 확산에 따라 50억 명에 달하는 시청자들의 TV 시청 방식이 변모하고 있다”며, “TV 프로그램 시청 중 최고의 순간을 시청자들이 서로 공유할 수 있는 SNS가 되는 게 ‘커넥 TV’의 목표”라고 밝힘

□ 최초의 메이저 방송사 연합 소셜 TV 서비스 ‘커넥 TV’

- 그동안 지상파 방송사, 케이블 TV 채널 등 특정 방송사가 시청자 견인을 위해 자체적인 소셜 TV 애플리케이션을 선보이거나 소셜 애플리케이션 제공업체 측에서 특정 방송사와 제휴를 체결한 사례는 있었지만, ‘커넥 TV’와 같이 소셜 TV 서비스 제공을 위해 여러 메이저 방송사가 손을 잡은 것은 이례적인 사례로 평가됨
 - 음악 전문 채널 MTV와 VH1은 각각 자사 TV 프로그램과 관련된 페이스북, 트위터 대화 내용 중 가장 인기 있는 코멘트를 실시간으로 정리해서 보여주는 ‘와치위드(WatchWith)’, ‘코스타(CoStar)’ 등의 소셜 TV 애플리케이션을 선보인 바 있음

- 위성 TV 사업자 디렉TV(DirecTV)는 소셜 TV 플랫폼 사업자 미소(Miso) 및 갯글루(GetGlue)와 IPTV 사업자 에이티엔티(AT&T)는 4개의 소셜 TV 플랫폼 사업자와 제휴를 체결하기도 했음

□ 참고문헌

- ‘TV Station Groups Plan Social TV Push with ConnecTV’, Hollywood Reporter, 2011.11.15
- ‘Broadcasters Connect With ConnecTV’, Broadcasting & Cable, 2011.11.15

7. 선명한 컬러 동영상을 저전력으로 구현하는 디스플레이 개발

태블릿 PC에 주로 사용되고 있는 LCD 디스플레이와 e-리더 시장을 차지하고 있는 e-ink 디스플레이 방식의 장점만을 결합한 새로운 디스플레이 기술이 퀄컴(Qualcomm)社를 통해 상용화를 앞두고 있다. 미라솔(Mirasol)이라고 명명된 이 기술은 2004년 최초 개발된 이래 드디어 2012년 상용화를 앞두고 있다. 기존 LCD 디스플레이 방식에 비해 10분의 1 이하의 전력 소비는 물론 밝은 빛 아래에서도 LCD 디스플레이에서나 가능했던 풍부한 색감을 구현할 수 있으며 초당 30 프레임에 달하는 영상 재생도 감당할 수 있다.

□ 풀 컬러 동영상 구현과 저전력을 동시에 실현한 디스플레이, ‘미라솔’

- 현재 태블릿 PC에 사용되는 LCD 방식과 전자책 디스플레이에 이용되는 e-ink 방식의 장점만을 결합한 새로운 디스플레이 기술이 퀄컴을 통해 상용화를 앞두고 있음
 - 미라솔 디스플레이 기술은 2004년 퀄컴에 인수된 이리다임(Iridigm)社에 의해 개발되었음
 - 퀄컴의 신기술 개발 부문을 이끌고 있는 클레어런스 추이(Clarence Chui)는 “현재 디스플레이 시장은 LCD 방식과 e-ink 방식이 양분하고 있다”면서, “미라솔 디스플레이는 저전력과 풀 컬러를 통해 이들 두 시장을 연결하게 될 것”이라고 말함
- 안드로이드 태블릿에서 5.7인치 풀컬러 화면으로 구현된 퀄컴의 디스플레이 기술은 밝은 빛 아래에서도 LCD 디스플레이나 실제 잡지처럼 풍부하고 선명한 색감을 구현해 냈음
 - 또한 초당 30 프레임에 달하는 영상 재생이 가능해 Angry Birds 등의 게임 플레이에도 문제가 없음
- 인쇄지나 e-ink 디스플레이 방식처럼 환경광(Ambient Light)¹⁹⁾을 사용하기 때문에 소비 전력 또한 LCD의 10분의 1 이하에 불과함
 - 또한 어두운 곳에서는 패널 모서리에 내장된 OLED 광원을 활용하게 됨
 - 美 시장조사 업체 파이크 리서치(Pike Research)의 2010년 연구 결과에 따르면 5.7인치의 미라솔 디스플레이는 같은 크기의 LCD 디스플레이에 비해 약 두 배 이상의 시간 동안 웹 브라우징이 가능한 것으로 나타났음
- 미라솔 디스플레이 기술은 반사광을 원하는 대로 조절할 수 있는 일종의 불완전 거울 방식을 통해 색을 구현하게 됨
 - 미라솔은 디스플레이 표면 아래 배치되어 있는 수십 마이크로미터(100만분의 1미터)의 폭

19) 물체에 투사되는 광 중에서 광원으로부터 직접 오지 않고, 다른 물체에 반사되어 도달하는 간접광

과 수백 나노미터(10억분의 1미터)의 두께를 가진 간섭 변조 소자(Interferometric Modulator)들을 통해 색을 구현하게 됨

- 추이는 “디스플레이 표면과 변조 소자 아래 반사막 사이의 공기층이 색을 결정한다”고 말함
- 변조 소자들은 빨간색, 녹색, 파란색의 세 가지 색깔로 구분되며 미라솔 디스플레이 표면 픽셀은 서로 다른 명암비를 가진 이들 변조 소자 여러 개가 함께 모여 색을 구현하게 됨

□ 미라솔 디스플레이의 상용화 계획도 지속적으로 진행 중

- HTC 및 LG 삼성 등의 모바일 단말기 제조 업체들이 미라솔 디스플레이의 주 구매 고객이 될 것으로 예상됨
 - 추이는 “미라솔 디스플레이가 태블릿 PC는 물론 스마트폰에서도 활용될 수 있을 것”이라고 말함
- 단 2012년 상용화계획은 작년 예상되었던 시장 출하 계획에 비하면 상당히 늦춰진 수준임
- 현재 미라솔 디스플레이를 시험적으로 생산하고 있는 대만의 공장 외에 대규모 생산을 진행할 수 있는 생산 공장이 새롭게 대만에 지어지고 있는 상황임
 - 퀄컴이 9억 7,500만 달러를 투자할 계획인 공장은 2012년 중반에야 완공될 수 있을 것으로 보이기 때문에 미라솔의 실제적인 상용화는 2012년 2분기 이후에나 가능할 것으로 보임

□ 참고문헌

- ‘Vibrant Color Video’, MIT Technology Review, 2011.11.15

8. 美 학교 교육 현장에서 태블릿 PC의 교과서 대체 분위기 확산...e-텍스트북 대중화 임박

최근 美 교육기술전문가들을 대상으로 한 설문조사에 따르면, 앞으로 학교 등 교육 현장에서의 태블릿 PC 활용 비중이 더욱 커질 것으로 전망된다. 교육계 전문가들은 향후 5년 내에 교육 현장에서의 태블릿 PC 사용이 데스크톱이나 노트북의 수요를 앞지를 것으로 전망했으나, 태블릿 PC가 데스크톱의 대체재라기보다는 보완재로서 기능할 것으로 예상했다.

□ 美 학교 교육에서 태블릿 PC 활용 비중은 앞으로 더욱 커질 것

- 교육 현장의 디지털 교육 기술 담당관들은 태블릿 PC가 교육 현장에서 보다 많이 활용될 것으로 전망함
 - 투자은행 파이프 제프레이(Piper Jaffray)의 진 먼스터(Gene Munster) 애널리스트가 교육 기술 컨퍼런스에 참여한 25명의 美 교육구(school district) 기술 담당관을 대상으로 설문을 실시한 결과
- 한편, 먼스터의 설문에 응답한 25명의 美 교육구 기술 담당관들 모두 안드로이드(Android) 기반 태블릿 PC를 사용하지 않는 것으로 나타나, 美 교육 현장의 태블릿 PC 수요가 사실상 아이패드(iPad)에 집중되어 있다는 주장도 함께 제기됨
 - 먼스터는 “물론 애플(Apple)의 시장 선점에 의한 효과도 있겠지만 무엇보다 이미 개인적으로 애플 제품을 사용하고 있는 학생들과 교직원들이 직장과 학교에서도 애플 제품을 사용하길 바라기 때문”이라며, “이는 안드로이드 태블릿 PC가 아이패드에 비해 활용성이 제한적이라는 것을 의미한다”고 분석함

□ 태블릿 PC, 美 교육현장서 노트북·데스크톱 수요 압지를 듯

- 美 교육 관계자들은 향후 5년 내에 교육 현장에서 태블릿 PC의 활용 수요가 노트북이나 데스크톱을 앞지르게 될 것으로 예상했으며, 다만 태블릿 PC가 일반 컴퓨터를 대체하기 보다는 보완재 역할을 하게 될 것으로 전망함
- 애플의 前 판매 운영 담당이었으며 현재 유통업체 J.C의 CEO인 페니 론 존슨(Penny Ron Johnson)은 “지금 학교에 다니는 학생들이 종이책을 사용하는 마지막 세대가 될 것”이라며, 미래에는 태블릿 PC가 교과서를 대체하게 될 것이라고 주장함
 - 시장조사업체 커런트 애널리시스(Current Analysis)의 애널리스트 애비 그린가트(Avi Greengart)는 “많은 기업과 공공기관이 아직 태블릿 PC를 어떻게 활용할지 가늠해 보고 있는 단계”지만, “다른 단말에서 구현하기 어려운 서비스나 애플리케이션도 아이패드용으로 개발되면서 빠르게 확산되고 있다”고 언급함
- 아마존(Amazon)이 최근 출시한 안드로이드 기반 태블릿 PC ‘킨들 파이어(Kindle

Fire)’가 200달러의 저가로 판매되는 등 태블릿 PC 가격의 지속적인 하락세 역시 앞으로 교육 현장에서 태블릿 PC의 인기를 견인하는 데 도움이 될 것으로 전망됨

- 다만, 태블릿 PC의 터치스크린이 타자 입력 등 이용 면에서 다소 제한이 있으며, 외부 입력 장치나 보조 장치를 연결하는 것은 태블릿 PC의 강점인 휴대성을 제한할 우려가 제기되는 등 본격 확산을 위한 해결과제가 남아있음

□ 참고문헌

- ‘Schools expect to have more iPads than computers in next 5 years’, AppleInsider, 2011.10.31

9. 예술, 게임 산업의 핵심 요소로 부상 中...게임 아티스트에 대한 관심도 증가세

최근 게임 업체가 신작 출시를 앞두고 게임의 예술적 표현에 주력하고, 게임 아티스트에 대한 관심도 고조되는 등 게임 업계 내에서 예술 분야의 중요성이 대두되고 있다. 게임의 그래픽 수준과 게이머의 기대 수준이 증가함에 따라 게임 업계의 예술적 요소에 대한 관심은 지속적으로 증가할 전망이다.

□ 예술, 이제는 게임의 필수 요소...메이저 게임사도 신작 개발에 예술적 표현기법 활용

- 게임이 예술의 일부인가에 대한 의견은 여전히 분분하지만, 업계 내에서는 게임에서 예술이 차지하는 비중이 매우 높다는 의견이 지배적
 - 서던 메소디스트(Southern Methodist) 대학의 닉 하이츠만(Nick Heitzman) 교수는 “사회 각층에서 게임 산업에 대한 관심이 증가함에 따라 게이머의 시각적 요소에 대한 기대 수준을 충족시키는 전문 아티스트 수요도 높아지고 있다”고 언급
- 600억 달러 규모의 전 세계 게임 시장에서 유명 게임 업체는 경쟁력을 높이기 위해 블록버스터 타이틀 제작에 전문 아티스트로 구성된 팀을 투입하는 등 예술적 요소 도입에 주력 중이며, 게임의 예술적 표현력을 높이기 위한 투자도 아끼지 않고 있음
 - 유비소프트(Ubisoft)는 ‘어쌔신 크리드: 래버레이션(Assassin’s Creed Revelations)’의 시대적 배경인 16세기 건축 양식을 표현하기 위해 이스탄불 답사를 진행함
 - ‘젤다의 전설: 스카이워드 소드(The Legend of Zelda: Skyward Sword)’는 인상주의적 표현양식을 채용하는 등 예술적 표현기법을 적극 활용함
 - ‘엘더 스크롤5: 스카이림(The Elder Scrolls 5: Skyrim)’은 방대한 게임 배경과 던전을 디자인하기 위해 40인 규모의 아티스트 팀을 구성

□ 전체 게임 산업의 활성화가 게임 아티스트에 대한 관심으로 이어져

- ‘어쌔신 크리드: 래버레이션’의 라파엘 라코스테(Raphael Lacoste) 아트 디렉터는 “저해상도 시절에도 게임 분야의 예술적 표현 활용은 존재했다”며, “당시 게임 아티스트들은 적은 예산에도 불구하고 흥미로운 것들을 만들어냈다”고 언급함
 - 게임 아티스트들은 게임 배경과 캐릭터의 개념을 구성하고 시각 효과를 디자인하며, 게임에 조명과 텍스처를 입히는 작업을 담당함
 - 초기 비디오 게임들은 투박한 픽셀 그래픽 기반임에도 불구하고 2012년 3월 16일 워싱턴에서 개최되는 ‘비디오 게임 예술 전시회(Art of Video Game Exhibition)’에 전시되는 등 그 예술성을 인정받음

- 게임 아티스트에 대한 업계 수요가 증가함에 따라 게임 아트 관련 교육과정도 증가하는 추세임
 - 교육전문지 프린스턴 리뷰(Princeton Review)의 조사에 따르면, 지난 2년간 대학과 아트 스쿨의 게임 아트 관련 교육과정이 두 배 가량 증가한 것으로 조사됨
- 게임업계의 예술분야 투자도 지속적으로 증가해, 최근 블록버스터 게임은 예술분야에서 4,000만에서 6,000만 달러의 예산을 투자하고 있음

□ 참고문헌

- 'Artist in high demand for video game industry', USA Today, 2011.11.13

10. TV 프로그램을 인식하는 아이패드 애플리케이션 ‘인투나우’

아이패드(iPad)를 이용해 현재 방송 중인 TV 프로그램을 자동적으로 인식하고 이와 관련된 웹 콘텐츠 및 부가 정보, 타게팅 광고를 사용자에게 실시간으로 서비스할 수 있는 아이패드용 애플리케이션이 야후(Yahoo)社에 의해 개발, 출시되었다. ‘인투나우(IntoNow)’라는 이름의 이 애플리케이션은 야후의 음성 인식 기술을 활용해 실시간으로 TV 프로그램을 인식하고 관련 정보를 사용자에게 전달할 수 있다. 야후는 인투나우를 실시간 타게팅 광고 플랫폼으로 육성해 이를 통해 새로운 광고 수익을 창출할 수 있기를 바라고 있다.

□ 시청중인 TV 프로그램을 인식하고 이와 관련된 웹 콘텐츠나 광고를 실시간으로 서비스할 수 있는 아이패드 애플리케이션 ‘인투나우’

- 현재 시청중인 TV 프로그램과 관련된 웹 콘텐츠, 뉴스 정보, 타게팅 광고를 실시간으로 서비스할 수 있는 아이패드 애플리케이션 ‘인투나우’가 야후를 통해 공개됨
 - 인투나우는 2011년 5월 야후에 인수된 동명의 업체가 개발한 TV 프로그램 의견 교환 SNS 애플리케이션을 기반으로 한 것임
 - 기존의 애플리케이션은 단지 페이스북이나 트위터를 통해 시청중인 TV 프로그램에 관한 의견 교환만이 가능했지만, 야후의 텍스트 의미 인식 기술을 통해 실시간 웹 콘텐츠와 광고 전달 플랫폼으로 새롭게 거듭날 수 있었음
 - 인투나우의 창립자였으며 현재 야후의 임원인 아담 케이한(Adam Cahan)은 “이 애플리케이션에 사용된 기술은 시청중인 TV 프로그램을 자동적으로 인식할 수 있다”고 밝힘
 - 실제로 인투나우는 폴란드에 불시착한 비행기에 관한 야후의 CNN 방송 뉴스 리스트나 트위터 멘션들을 자동적으로 제시할 수 있음
 - 인투나우 애플리케이션은 클라우드 서버에 저장되어 있는 방대한 분량의 TV 음성 지문(Audio Fingerprints)을 사용자가 시청중인 프로그램의 음성 지문과 비교해 해당 프로그램을 인식하게 됨
 - TV 프로그램을 인식한 후에는 다중방송 시스템에서 송출되는 문자 자막(Caption)²⁰⁾을 텍스트 형태로 받아 현재 어떤 내용이 방영되고 있는지를 알아내는 방식임
- 인투나우는 아이패드 사용 습관에 최적화된 애플리케이션
- 인투나우는 사용자들이 TV 시청 중에 수동적으로 해야만 했던 일들을 애플리케이션이 자동적으로 대신 처리해주는 것을 목표로 하고 있음

20) TV 프로그램 송출 정보에는 장애우를 위한 캡션정보가 텍스트 형태로 포함되어 있음

- 즉, TV를 시청하면서 프로그램 내용과 관련된 정보를 태블릿PC나 스마트폰들로 찾아보곤 하는 사용자들의 습관을 겨냥한 것임
- 실제로 야후는 최근 연구조사를 통해 아이패드 사용의 30%가 TV 시청 중에 이루어진다는 사실을 알아내기도 함

□ 타깃팅 광고 플랫폼으로서 진화를 최종 목표

- 야후는 인투나우를 사용자의 시청 습관과 관련된 타깃팅 광고 플랫폼으로 진화시켜 새로운 광고 수익을 창출하는 것을 최종 목표로 하고 있음
 - 실제로 인투나우는 TV에서 펩시 광고가 나올 때, 펩시의 무료 음료 쿠폰을 제공하는 시험적인 타깃팅 광고 프로모션을 진행하기도 했음
 - 케이한은 이에 대해 “인투나우는 소비자들이 지금까지의 수동적인 광고 시청에서 벗어나 보다 직접적으로 반응할 수 있게 해줄 것”이라고 말함
- 인투나우는 광고 수입 감소를 우려한 기존 TV 업계의 강력한 반발에 부딪히고 있는 구글(Google) TV와는 달리 보다 유연하게 광고 시장 진입이 가능할 전망
 - 태블릿 PC 시장은 기존 TV 산업과는 별개로 성장하고 있는 데다 그 성장 또한 매우 빠르기 때문에 야후의 광고 시장 진출은 구글 TV에 비해 보다 용이할 전망이다

□ 참고문헌

- ‘Yahoo Gives the iPad the Power to Understand TV’, MIT Technology Review, 2011.11.02.