

▲ 펜틀랜드는 컴퓨터가 사람이 즐겁거나 화를 내거나 놀라거나 혐오감을 일으킬 때 얼굴의 움직임을 밝은 색깔모양으로 나타내는 이른바 '동작에너지도'를 만들어 내장된 패턴과 대조하여 회로애락을 인식할 수 있게 만들었다. (사진은 펜틀랜드박사) ① 즐거운 표정과 동작에너지도 ②놀란 표정과 동작에너지도 ③혐오감을 나타내는 표정과 동작에너지도 ④화를 내는 표정과 동작에너지도

첨단과학현장

인공지능을 가진 환경

玄 源 福 <과학저널리스트/본지 편집위원>

2세기에는 어린이들이 집안의 어디서 놀고 있는가, 그리고 혹시 위험한 장난을 하고 있다면 이런 사실을 언제나 주부에게 알려 주는 주택이 등장한다. 또 중요한 모임이 있을 때는 회의에 방해가 되지 않게 걸려 오는 전화나 방문객 등 모든 것을 알아서 처리해 주는 사무실도 등장한다. 2세기에는 운전자가 과로때문에 사고를 낼지 모른다고 감지하면 운전자에게 차를 주행선 밖으로 빼라고 경고할 수 있는 승용차도 선을 보일 것이다.

스마트 룸

1987년의 어느날, 미국 매사추세츠공대(MIT)의 미디어연구소 소장 니콜라스 네그로폰테에게는 시청률조사회사인 아비트론사로부터 이른바 '사람계측기'를 만들어 달라는 요청이 들어왔다. 이는

시청률을 모니터하기 위해 텔레비전 위에 계측기를 올려 놓고 텔레비전을 시청하는 가정의 가족들이 정확하게 언제 텔레비전이 있는 방에 들어와서 어떤 행동을 하는가 체크할 목적이었다. 네그로폰테소장은 그의 제자이며 심리학과 인공지능연구로 박사학위를 받은 알렉산더

펜틀랜드에게 이 일을 맡겼다. 그러나 3년 뒤 연구비를 지원하던 아비트론사가 손을 떼고 말았는데 이유는 그들의 요구대로 얼굴을 인식하는 기계를 만들지 못했기 때문에 그런 것은 아니었다. 오히려 얼굴을 너무 잘 인식하기 때문에 광고주들이 시청자들의 진정한 습성을 너

무 많이 알게 되어 광고비 규모를 줄일 것이라고 걱정했기 때문이다. 예컨대 시청자들은 야구 시합을 구경하면서 모두 잠이 들어 버린다거나 또는 대부분의 저녁뉴스는 보지 않는다는 사실을 알게 되면 필경 광고주들은 이런 시간대의 스폰서가 되겠다는 열의가 식어 버릴 것이다. 그러나 펜틀랜드팀은 연구를 계속하여 마침내 기계가 회로에락을 포함한 사람의 표정을 분석할 수 있는 기술을 개발하는데 성공했으며 벌써부터 여러 분야의 관심을 모으기 시작했다.

사람의 얼굴, 표정 그리고 제스처를 인식하는 컴퓨터시스템을 이용하면 앞서의 주택이나 사무실이나 승용차처럼 행동하는 환경을 조성할 수 있게 된다. 이른바 '스마트룸' (지능실)이라고 불리는 이런 곳에는 카메라와 마이크로폰을 갖추고 있어 기록을 이웃 컴퓨터망에게 중계하면 컴퓨터가 '스마트룸'에 있는 사람들의 말과 행동을 평가한다. '스마트룸'은 그 속에 있는 사람들에 관해 알고 있기 때문에 이들에 대해 지능적으로 반응할 수 있게 된다. 펜틀랜드팀은 1991년 MIT에 최초의 '스마트룸'을 만든데 이어 오늘날 '스마트룸'은 보스턴에 3개, 일본에 1개 그리고 영국에 1개 등 모두 5개로 늘어났고 이들은 모두 전화선으로 연결되어 공동실험을 하고 있다. 곧 파리, 뉴욕 및 델러스에도 '스마트룸'이 설치된다.

독특한 표정

'스마트룸'은 사람들의 행동을 분석하려고 시작하기에 앞서 먼저 그들의 위치를 파악해야 한다. 펜틀랜드팀은 방에서 움직이는 사람들을 추적할 수 있는 사람 탐지기를 개발했다. '스마트룸'은 또 방에 있는 사람이 누구이며 무슨 말을 하

고 있는가 알아야 한다. 종전에도 과학자들은 컴퓨터가 말을 이해할 수 있는 알고리즘(일정한 계산의 기준을 정하기 위한 일련의 규칙)을 발명하기는 했으나 이것은 사용자가 마이크를 착용하거나 기계 옆에 가까이 있는 경우에만 통용되었다. 그래서 연구팀은 심지어 소음이 많을 때도 자유롭게 움직이는 사람들의 말을 컴퓨터가 해독하는 방법을 개발했다.

'스마트룸'은 또 말하는 사람이 누구인가를 알아야 한다. 사람을 판별하는 가장 빠른 길은 당사자의 얼굴을 인식하는 것이다. 오늘날 세계에는 60억에 가까운 사람들이 살고 있으나 어떤 사람도 얼굴이 완전히 같은 사람은 없다. 그렇다면 얼굴이 몇군데나 서로 다를까? 얼핏 생각하기에는 다른 점이 엄청나게 많을 것 같지만 펜틀랜드팀은 간추려서 1백가지의 서로 다른 요인 밖에 없다는 사실을 알게 되었다. 예컨대 코가 넓적하다거나 눈이 크다거나 입이 비뚤어졌다거나 사람마다 특징이 있는가 하면 얼굴에서의 코나 눈이나 또는 입의 위치가 저마다 다르다. 이런 1백가지 요소의 조합이 사람마다 독특한 모습을 보여 주게 되는 것이다. 펜틀랜드팀은 컴퓨터에게 이 1백가지의 '얼굴의 특색'을 인식시킨 뒤 수천명의 얼굴사진을 내장한 '웨이스베이스' (얼굴 모습을 저장한 것)와 대조하면서 사람의 얼굴을 가려내게 했다. 물론 '웨이스베이스'에는 같은 인물도 여러 각도와 서로 다른 표정을 지을 때 찍은 사진들을 수록하고 있는데 가장 규모가 큰 것은 3천명의 사진 7천5백장을 내장하고 있다. 이런 방법으로 '스마트룸'은 사람의 얼굴을 99%의 정확성을 갖고 인식할 수 있게 되었다.

펜틀랜드팀은 또 눈과 입을 포함한 얼

굴 근육의 움직임은 통해 노여움, 혐오감, 놀라움, 행복감, 슬픔, 두려움 등 6가지의 기본적인 표정을 기계가 읽을 수 있는 소프트웨어도 개발했다. 그런데 얼굴표정을 읽는다는 것은 얼굴을 판별하는 것 만큼이나 중요하다. 예컨대 교육 프로그램에서는 교육을 받는 학생이 지루한 표정을 하는가를 알아야 하기 때문이다. 그래서 '스마트룸'은 일단 사람의 얼굴을 판별한 뒤 표정을 분석한다. 컴퓨터는 카메라가 기록한 얼굴의 움직임과 여러가지 표정을 지을 때의 얼굴 움직임을 그린 지도와 비교한다. 실상 하나하나의 얼굴표정은 근육동작의 독특한 집합과 관련되어 있다. 예컨대 미소를 지을 때 입 가장자리를 빼죽거리고 이마의 일부를 치켜 올린다. 그러나 가짜 미소를 지을 때는 입만 움직인다. 펜틀랜드팀이 개발한 시스템은 소규모의 집단에서 98%의 정확도로 얼굴표정을 판단했다.

'스마트룸'은 어떻게 보면 짐사처럼 행동한다. 주인이 안보이는 곳에서 끊임 없이 주인을 도울 기회를 찾는다. 그러나 사람은 이따금 함께 여행하며 필요한 것은 미리 챙겨 주는 개인보좌역을 더 선호할 때가 있다. 이런 보좌를 받기 위해 '스마트룸'의 카메라, 마이크로폰 그리고 컴퓨터를 옷 속에 내장하는 방법이 있다. 이런 스마트옷은 만나는 사람의 이름이나 다음 모임의 장소와 같은 개인적인 정보를 제공할 수 있다. 또 이런 옷은 컴퓨터와 대부분의 가전제품과 대체할 수 있다. '스마트룸'과 마찬가지로 핵심적인 아이디어는 옷이 주인의 환경을 알고 있기 때문에 주인을 돕기 위해 지적으로 반응할 수 있다는 것이다.

MIT의 스마트옷 사업은 1992년 알렉스 펜틀랜드와 로잘린드 피카드의 도움

으로 태드 스타너와 스티브 만이 착수했다.

스마트 옷들

이들이 컴퓨터, 카메라, 마이크로폰 및 다른 센서를 옷 속에 꾸며넣을 때 컴퓨터는 '스마트룸'에서의 수동적인 3인칭의 입장에서 능동적인 1인칭의 유리한 입장으로 전환된다. 그래서 스마트옷은 사용자의 행동과 보다 친근하게 관여할 수 있어 지능적인 개인보좌역의 역할을 맡는다. 예컨대 카메라를 안경테 속에 내장하면 얼굴인식 소프트웨어가 사용자의 귀에 대고 만나고 있는 사람의 이름을 속삭여 주어 사교에 도움을 줄 수 있다. 구두 속에 지

구축위센서를 내장하면 고

공비행항법 소프트웨어를 이용하여 방향을 쉽게 찾을 수 있다. MIT의 스마트옷의 원형은 사용자에게 프라이버시와 편의성을 제공하기 위해 아직도 기성제품인 HMD(머리에 쓰고 컴퓨터가 만드는 영상을 통해 가상현실감을 체험할 수 있는 장치)를 사용한다. 그 중앙처리장치는 방해가 되지 않게 소형으로 설계되었고 인터넷과의 접촉을 유지하기 위해 무선통신을 사용한다. MIT팀의 목표는 오늘날 안경과 시계처럼 언제나 착용할 수 있게 작고 가벼운 장치를 개발하여 언제든지 컴퓨터와 접근할 수 있게 하자는 것이다.

'스마트룸' 기술은 실용면에 새로운 지평을 열기 시작한다. 먼저 미군과 미 연방사법당국은 얼굴인식기술을 보안분야



▲ 미국 수화법은 복잡한 손의 제스처로 구성되고 있으나 그중에서 많은 부분이 MIT의 컴퓨터과학자 태드 스타너가 개발한 시스템을 이용하여 실시간으로 통역할 수 있게 되었다. 이 사진에서 카메라가 스타너의 '자전거'를 나타내는 손의 동작을 등록하고 있다. 컴퓨터시스템이 그의 동작을 모델로 하고 기존의 사인모델과 비교한다. 만약에 부합되면 기계는 '자전거'라고 말한다.

에 응용할 생각이다. 그중의 하나가 출입용 보안시스템. 출입이 허용된 군 요원들의 얼굴을 시스템에 저장한 뒤 예컨대 누구든지 핵잠수함 내부로 들어가려 할 때 얼굴인식 소프트웨어가 출입이 허용된 사람인가를 가려낸다. 지금까지 실험결과 98%의 정확도를 보여 주고 있으나 소프트웨어의 신뢰도를 100%까지 끌어 올려 암호가 내장된 신분증보다 더 안전한 것으로 만들 생각이다. 그런데 얼굴은 카드와는 달라서 잃어 버리거나 훔칠 수 없다. 특히 눈을 둘러싼 뼈의 구조는 사람의 얼굴에서도 가장 바꾸기 어려운 부분이므로 변장으로도 속일 수 없다. 백악관 당국은 이 기술을 이용하여 테러리스트와 마약범을 감시하는데 관심을 보이고 있다. 오클라호마폭탄사

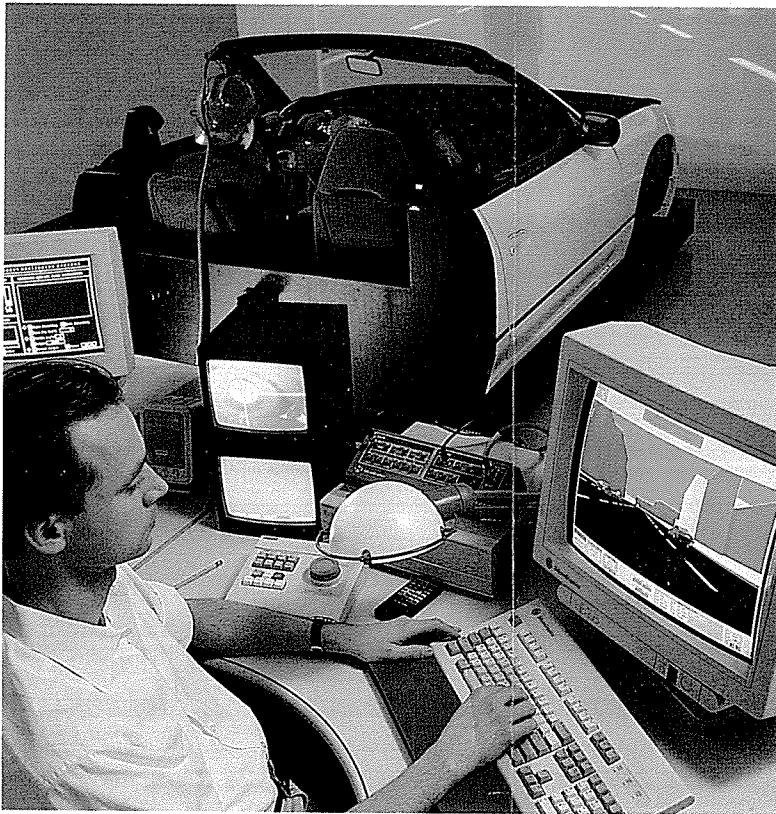
건의 경우처럼 용의자는 일을 저지르기 전에 여러 번 목표물을 답사한다. 얼굴인식카메라를 주요한 곳에 설치하여 자주 나타나는 용의자얼굴을 가려내게 할 수 있다. 또 카메라를 세관검사대에 설치하면 변장과 가짜여권 그리고 위조비자사용을 일삼는 마약밀매자의 얼굴을 정확히 판별할 수도 있다.

다양한 응용분야

21세기 초에는 현금자동출납기에도 얼굴인식기술이 진출한다. 카메라와 연결된 얼굴인식프로그램은 고객의 얼굴을 가려낼 수 있어 종래의 비밀번호와 대체할 수 있다. 미국 매

서추세츠주 정부의 자동차국은 이 기술을 이용하여 운전면허증을 분실했다고 주장하면서 재발급을 요구하는 사람이 본인인가 아닌가를 확인할 생각이다. 한편 멀티미디어분야로 진출하여 교육용 CD롬을 사용하는 어린이들이 지루하거나 어리둥절한 표정을 보일 때 소프트웨어가 지체없이 적절한 대응을 하게 만들 수도 있다.

한편 펜들랜드팀은 이 기술을 이용하여 수화(手話)를 동시통역하는 시스템을 개발했다. 그런데 미국 수화법(ASL:미국 및 캐나다 영어권지역에서 사용)은 농아자들이 사용하는 정교한 손의 제스처인데 이 제스처는 매우 복잡해서 '스마트룸'의 능력을 시험하는데는 안성맞춤이었다. 이들은 ASL을 해석하는 시



▲ 닛산사의 어원 보어가 모의코스에서 주행하는 운전자를 모니터하고 있다. 센서는 그의 손 발 그리고 눈을 기록한다. 운전자가 무엇을 하려고 하는가 헤아려 보기 위해 이 기록은 전형적인 운전동작 모델과 비교한다. 컴퓨터는 운전자의 동작이 시작된 불과 몇초 뒤에 그 행동을 판별할 수 있다. 예컨대 운전자가 차선을 비껴려고 할 때 눈에 보이지 않는 곳에 트랙이 있다면 차는 운전자에게 경고할 수 있다.

시스템을 구축하기 위해 먼저 관련된 손 동작의 사례를 관찰하면서 각 신호의 모델을 만들었다. 이 시스템을 이용하여 실시간으로 99.2%의 정확성을 가지고 ASL의 40단어 세트를 해석할 수 있게 되었다. 만약에 이해하는 어휘의 규모를 늘릴 수 있다면 현재 들을 수 있는 사람들을 위해 도입되고 있는 음성인식시스템처럼 귀가 먼 사람을 위해서도 신뢰성 있는 인터페이스를 만들 수 있을 것 같다.

미리막는 운전사고

그런데 자동차 운전자들도 '스마트룸' 기술의 혜택을 볼 수 있다. 미국의 경우 보통 근로자들은 매주 10시간을 차에서 보낸다. 해마다 4만명 이상의 자동차 운

전자들이 교통사고로 사망하는데 사고의 대부분은 운전자의 잘못에 기인한다. MIT연구팀은 닛산자동차 케임브리지연구소의 과학자와 함께 자동차 내부를 모방한 '스마트룸'을 만들었다. 이들의 목적은 운전자가 하고 있는 동작을 모니터하고 이들에게 도로방향, 운전수칙과 같은 유용한 정보를 제공할 수 있는 차량을 개발한다는 것이다. 운전자들이 차를 주행하고 회전하고 정지하거나 또는 가속하거나 차선을 바꿀 때 취하는 행동을 포함하여 운전모형을 모으기 위해 이들은 운전자들이 모의코스를 주행할 때의 손과 다리의 움직임에 관찰했다. 이렇게 완성된 모델을 이용하여 운전자의 행동을 되도록 빨리 분류하는 시스템을 만들

었다. 그런데 놀라운 사실은 이 시스템은 운전자의 행동이 개시되자마자 0.5초 이내에 86%의 정확도를 가지고 그 행동을 분류할 수 있다는 것이다. 따라서 이런 시스템이 실용화되면 운전자가 잘못된 행동을 할 때 거의 동시에 경고를 주거나 주의를 환기시켜 사고를 미리 막을 수 있게 된다.

우리는 얼굴이 익은 사람을 만나도 갑자기 이름이 기억나지 않아 난감할 때가 흔히 있다. 현재 MIT팀은 아는 사람의 얼굴을 인식한 뒤 귀에다 대고 그의 이름을 속삭여 줄 수 있는 안경을 만들고 있다. 이들은 또 시청자가 쳐다보고 있는가 또는 보고 있지 않는가를 알 수 있는 텔레비전 스크린을 개발하고 있다. 이밖에도 주인을 인식할 수 있는 신용카드를 개발하여 도용을 막을 수 있는 길을 열어 줄 생각이다.

한편 MIT미디어연구소의 아론 보빅은 스포츠에서 사용되는 인간의 동작을 해석하는 소프트웨어를 만들고 있다. 이 소프트웨어를 이용한 텔레비전 카메라는 예컨대 미식축구에서 쿼터백스니크(쿼터백이 센터로부터 공을 패스받아 그대로 직접 라인 한가운데로 공격하는 플레이)와 엔드런(공을 가진 선수가 방어진의 엔드를 우회하여 달려서 빠져 나가는 플레이)과 같은 2개의 플레이간을 차별화하면서 액션을 추적할 수 있게 된다.

이리하여 21세기에는 '스마트룸' 기술이 모든 분야로 번져나가 개인용 컴퓨터에서 승용차와 주방기구에 이르는 모든 기계들이 사람을 인식하고 선호하는 것이 무엇인가 알게 된다. '스마트룸' 기술이 더욱 더 발전하면 생물과 무생물간의 차이는 실제로 모호해져 분명하게 구별하기 어렵게 된다. ⑤7