

HCI 및 감성로봇을 위한 오감 인식기술 동향

성 하 경

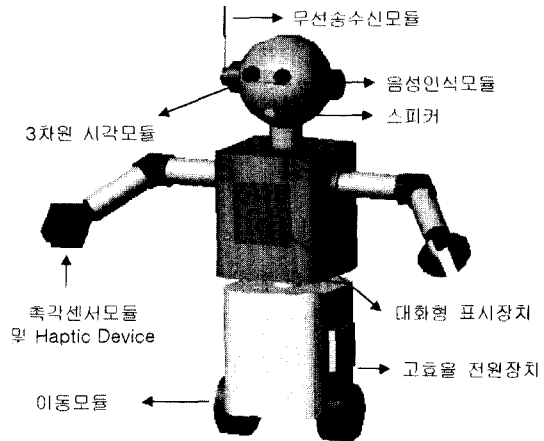
전자부품연구원

I. 서 론

일반적으로 인간과 컴퓨터의 상호작용(HCI: Human Computer Interface)은 사람들이 쉽고 편하게 컴퓨터 시스템과 상호 작용할 수 있는 문제에서 시작하여 보통 GUI(Graphic User Interface)로 대표되어지는데, 본 고에서는 인간이 느끼는 색상과 소리, 향기와 맛, 촉감등을 느낄수 있는 감성로봇과의 상호작용에 대하여 다뤄 보고자 한다.

특히 로봇기술(Robot Technology: RT)은 최근에 국가 전략 산업으로 육성되어야 한다는 소리가 높아지고 있기 때문에 지금까지 많이 활용되어졌던 산업용 로봇의 범위를 벗어나서 일상 생활에 가깝고 인간과의 교감이 큰 감성로봇에 대한 관심이 고조되고 있다. 이에 따라 로봇기술은 산업용에서 비 산업용 로봇이 차지하는 잠재력이 점차 인식되기 시작하여 완구로봇을 중심으로 서비스 로봇, 교육 로봇 등이 양산화 될 준비를 하고 있다. 오감의 인식기술은 단일 기술로도 계속 연구가 진행되어 왔지만 인간과의 교감기능이 필수적인 감성로봇에서는 이 기술들이 모두 필요하므로 이의 융합화가 또 하나의 기술 분야가 되고 있다.

<그림 1>은 감성기능을 가진 휴먼 인터페이스형 로봇의 개념도를 나타낸 것인데 향후 적어도 10년 이내에는 이러한 로봇이 현재의 PC같은 개념으로 어떠한 형태이든 개인형으로 존재하리라고 예상된다. 본 고에서는 오감 기술 중에서 최근에 가장 활발하게 연구되고 있는 촉각, 음성,



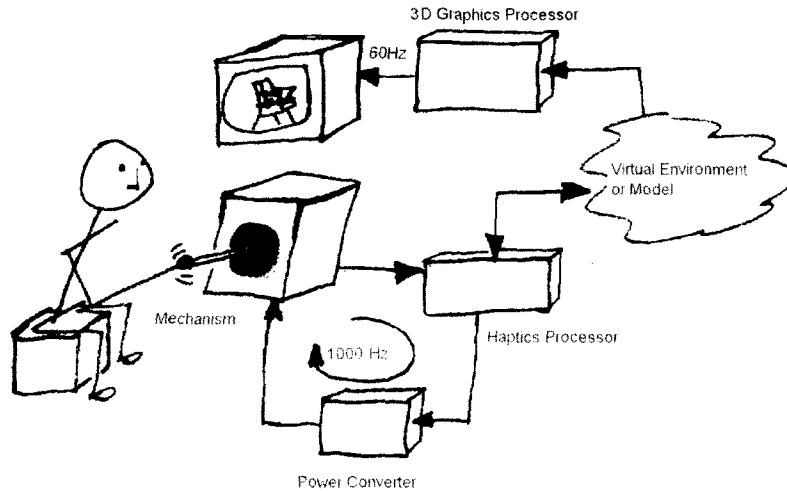
<그림 1> 휴먼 인터페이스형 로봇의 개념도

시각기술에 대하여 논하여 보고자 한다.

II. 촉각 기술동향

촉각 기술은 감성로봇에 있어서 2가지의 분야가 있는데 첫 번째로는 인간의 촉감을 로봇이 인식하는 기술이 있으며 또한 로봇 또는 기계가 인간에게 촉감을 전달해주는 분야가 있다. 촉감을 인식하는 기술은 보통 터치 센서를 이용하고 이를 인터페이싱 하여 인식하는 기술로써 센서의 설계 및 제작기술과 인터페이싱 기술로 대표된다.

또 한가지는 촉감을 인간에 전달해주는 기술이 있다. 이를 보통 Force Feedback 기술이라고 지칭하는데 이는 인간과 상호 작용하는 컴퓨터



〈그림 2〉 햅틱 디바이스의 구성도

시스템에서 매우 중요한 역할을 담당하며, 시각이나 청각 정보와는 달리 힘 되먹임 장치(Force Feedback Device) 또는 햅틱 디바이스(Haptic Device)를 통하여 물체의 무게나 단단함 등의 촉감을 전달하는 기술이다. 특히, 이 기술은 가상현실(Virtual Reality) 시스템 분야에 해당하는 기술로서, 사람의 기능과 한계가 고려되어야 하며(Ergonomics), 힘 반향(Force Reflecting) 기능을 가지며, 관성이나 마찰 등이 최소화 되어야하고, 시스템의 안정성 향상을 위한 정밀위치제어 및 실시간 제어가 가능해야 한다.

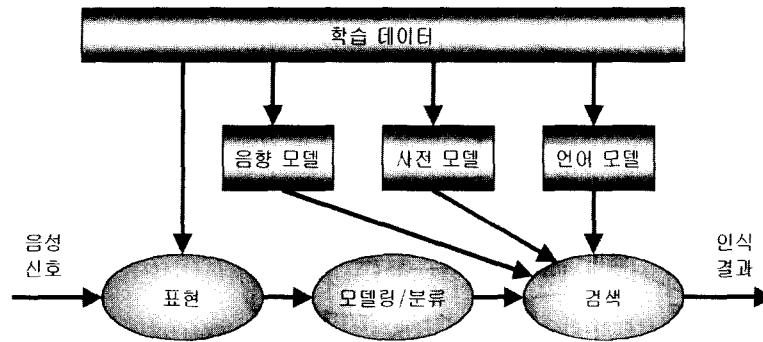
선진국에서는 공학분야 뿐만 아니라, 의학, 미술 등의 여러 분야에서 햅틱 디바이스에 대하여 많은 연구가 수행되고 있고, 시제품 단계의 연구 성과물들이 쏟아져 나오고 있으며, 일부 비디오 게임기를 포함한 Entertainment 산업 분야에서는 Tool-Type Interface 장치인 조이스틱, 마우스, 트랙볼 등이 출시되고 있다. 또한 인터넷 콘텐츠를 기반으로 한 Entertainment 산업, 물건을 구입하기 전에 마우스 커서를 이용하여 실제 물체를 만지는 것과 유사한 촉감을 전달하여 주는 인터넷 전자 상거래 산업, 조직 세포를 만지듯이 수행하는 내시경, 복강경 등의 외과 의료산업 등의 발달이 예상된다.

촉감은 물체의 판별과 조작에 매우 중요한 역할을 하는데 대부분의 경우 작업의 성능을 결정하는 중요한 요소가 된다. 예를 들어, 수술을 할 경우에 내시경을 통한 시각적인 되먹임만으로는 불충분하고 직접 손으로 찾는 것이 더 효율적이다. 촉감 인터페이스 장치는 원격작업이나 장애인을 위한 장치 개발을 시초로 연구되기 시작했는데 현재에도 세계 여러 곳의 연구소에서 활발하게 연구가 진행중이다.

미국의 Harvard 대학의 응용과학부에서는 원격작업에서의 정교한 조작을 위한 촉감감지 및 표시 장치를 개발하고 있으며 원격작업을 수행할 때 작업대상에 대한 형상 정보의 역할에 대한 연구와 촉감 인터페이스를 이용한 형상정보 취득방법 개발을 수행하고 있다.

III. 음성인식기술 동향

인간이 의사전달을 하는 가장 자연스러운 형태가 음성이다. 따라서 음성을 매체로 한 휴먼머신 인터페이스의 필요성이 크게 대두되었으며, 1950년대 이후 음성인식에 대한 연구가 활발히 진행되어왔다. 여기서, 음성인식이란 마이크 등



〈그림 3〉 음성인식 시스템의 구성요소

을 통하여 얻어진 음향학적 신호를 단어나 단어의 집합 또는 문장으로 변환하는 과정을 말하며, 〈그림 3〉에서 음성인식 시스템의 구성요소를 보여준다. 오늘날에는 음성인식기술과 자연어처리기술과의 결합으로 단순히 인식의 차원을 넘어 주어진 음성의 뜻을 파악하고 그에 맞는 대응을 수행하는 음성이해의 단계에 와 있다.

음성기반 인터페이스는 인간과 컴퓨터 간의 상호작용의 편의를 증가시켜준다. 그리고, 전화 또는 휴대형 컴퓨터와 같이 원격에서 데이터입출력을 해야하는 상황에서 효율성을 높여주며, 중간작업자가 필요없게 되어 비용절감을 가져올 수 있다. 음성인식기술은 현재 다양한 업무에 활용되기 시작했는데, 단독 또는 화자인증 기술과 조합된 음성인식기술이 원격 데이터 등록, 중앙데이터베이스 원격 접속, 이동전화 단말기나 PDA 같은 휴대 장비의 데이터처리등의 애플리케이션에 응용되고 있다.

음성인식기술은 미국등 선진국들을 중심으로 활발하게 연구되고 있는데, 미국의 경우 음성인식과 관련한 연구개발은 대부분 국방부에서 비롯되었고, 미 공군과 방위고등연구계획국(DARPA)의 주도아래 수억달러가 음성인식프로젝트에 할당되었다. 미국의 국방부문의 경우 음성인식기술은 R&D단계에서 시제품이 선보이는 단계로 옮겨가고 있고 IBM, AT&T, Lucent Technologies, Microsoft 등의 컴퓨터 및 통신관련 기업체들은 독자적인 대용량 음성인식 시스템을 개발하고, 자사의 관련제품에 응용하고 있다.

유럽의 경우 ESPRIT(European Strategic Program for Research and development in Information Technology)라는 IT분야의 공동연구개발프로그램을 진행시켰다. ESPRIT에서는 비디오게임용 음성인식 및 음성처리 하드웨어와 소프트웨어의 개발과 시험을 목표로 하는 IVORY, 전화나 직접금융거래 등 금융부문에서 음성처리 소프트웨어의 유용성을 높이는 것을 목표로 하는 OVID 등 학계와 산업계의 경험을 통합하여 많은 음성관련연구가 진행되었다. 영국에서는 제5세대 컴퓨터 프로젝트인 앨비(Alvey) 프로그램을 통하여 국가연구소와 산업체연구소의 협력으로 음성인식관련 연구를 수행하였으며, 현재는 ITI프로젝트를 통하여 음성 인식 및 데이터베이스 구축에 대한 연구를 진행하고 있다. 일본의 경우는 1986년 이래 15년간의 장기 계획으로 자동통역전화개발을 추진해오고 있으며, 1987년에는 인간과 기계와의 구어체 대화를 목표로 하는 Advanced man-machine interface through spoken language라는 국가단위프로젝트를 통하여 대화체 언어이해 및 소음환경에서의 음성인식에 관한 연구가 진행되어 많은 결과가 도출되었다.

국내는 1980년대 초부터 본격적으로 음성인식에 대한 연구가 이루어졌으며, 아직 선진국에 비하여 다소 뒤처지는 경향을 보이고 있다.

최근 붐이 일고 있는 서비스(엔터테인먼트)로봇의 경우 대표적인 제품인 소니 AIBO의 경우 고감도의 Stereo Phone을 탑재하여 사용자(주

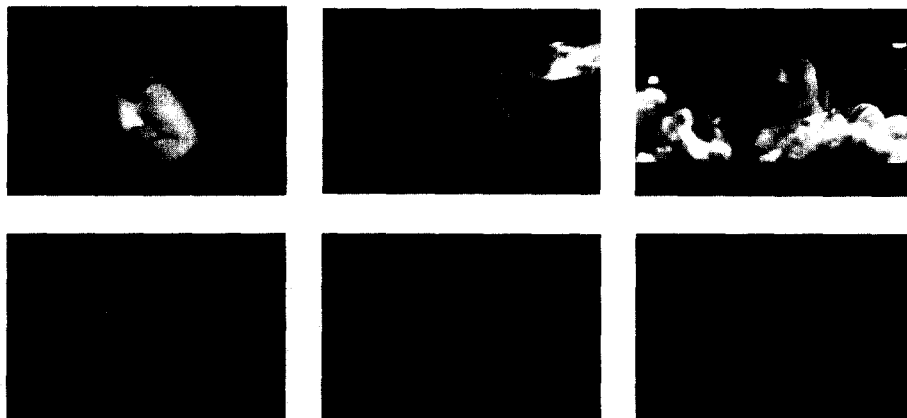
인)의 음성을 타인과 구별하여 반응을 보이는 기능 및 음성인식에 의한 소규모 인식모듈 탑재로 인하여 음성명령에 해당하는 명령을 수행한다. 또한 음성인식을 이용한 Application으로서 워드프로세서에 엔진을 탑재하여 약 95% 정도의 정확성을 가지는 프로그램이 개발되어 장애인 및 컴퓨터 초보자의 음성으로 워딩을 할 수 있을 정도의 엔진이 구축되어 있다.

IV. 시각 기술동향

인간이 외부에 대한 정보를 입력받아서 인지하는 과정에서 시각이 차지하는 비율은 약 70% 이상 대부분을 차지하고 있다. 인간의 시각과 뇌의 연산기능에 해당되는 영상처리 기술은 주로 패턴이 일정한 Machine Vision Application 분야의 검사 장비 제품, Tomography의 원리를 이용한 3차원 단층촬영 분야, DVR에 탑재하여 감시기능을 수행하며 소형 모듈로 제작되어 휴먼로봇, 탐사용 로봇, 혹은로봇에도 탑재되어 사용되어 왔다. 그러나 영상데이터의 방대함과, 연산 알고리즘의 복잡성으로 인하여 실시간이란 측면과 효과적인 리소스 사용이 힘들었던 것이 사실이다. 최근 시각센서(CMOS, CCD)가격이 하락하고 USB, IEEE1394의 등장으로 인하여 기존보

다 10배 이상의 capture bandwidth를 가질 수 있는 기기들의 등장으로 인하여 시각센서를 이용한 기술의 저 가격 및 실시간이란 측면에서 영상처리기술을 응용한 다양한 제품이 선보이고 있다. 시각모듈의 가장 중요한 기능은 인간의 눈, 뇌에 해당하는 물체의 추적, 인지에 있다. 1980년대 급격한 디지털 기술의 발전으로 인하여 Object Tracking 및 Recognition에 관한 많은 연구가 디지털 강국을 중심으로 학계, 연구계, 관련업체를 중심으로 진행되어 왔다. 본 고에서는 Tacking 및 Recognition에 관한 기술동향 및 발전방향에 국한하여 소개하고자 한다.

물체나 인체의 특징점 또는 컬러 모델 분석에 의한 Visual Tracking 기술은 영상처리 기술 분야 중 가장 Active한 분야 중 하나로서 물체의 윤곽(contour)을 추적하는 Snake, Eigen-space Matching 방법, 대규모의 통계학적 가설을 이용한 방법, feature detector를 영상주위로 돌돌 감싸는 방법 등 많은 알고리즘이 소개되어 있다. 그러나 이러한 방법들은 연산량이 상당히 하여 빠른 프로세서를 탑재하지 않으면 실시간으로 구현되기 어려운 단점이 있다. 가끔씩 빠르고 연산량이 적은 알고리즘에 관한 연구도 많이 진행되어 왔다. 또한 컬러영상 기반의 Tracking에 관한 여러 가지 알고리즘도 연구되어 왔으나, 컬러 correlation, blob, region growing, prediction, contour consideration 등의 사용



<그림 4> mean-sift 에 의한 Visual Tracking



〈그림 5〉 HMM 기반 얼굴인식

으로 인하여 연산량이 증가 실시간 측면에서 문제가 되고 있다. 최근에는 robust statistics, 확률 분포에 의한 간결하면서도 연산량이 적은 알고리즘이 개발되고 있는 추세이다.

Recognition 기술 중에서 현재는 얼굴인식에 많은 연구가 이루어지고 있다. 얼굴 인식은 은행 ATM 머신이나, 컴퓨터, 제한구역에서의 출입통제, 특정구역서의 사람 인식, 경찰 범죄기록의 사용 등 그 활용분야는 매우 넓다. 얼굴인식 기술에서 가장 중요한 점은 가변환경에서의 적응과 얼굴의 일부분만으로 추론을 할 수 있는 기술, 그리고 연령층, 피부색에 강인한 알고리즘 구현이라고 할 수 있다. 현재까지 Geometric feature-base Model, Template feature-base Model, Model-base 방법이 사용되어 왔다. 주요 feature로는 머리털, 이마, 눈, 코, 입 등을 사용한다. 최근에는 1D-HMM(Hidden Markov Model)개념에서 확장된 pseudo 2D-HMM 방법이 사용되고 있으며, 2D-DCT(Discrete Cosine Transform)계수로 구성된 observation vector를 이용한 방법이 사용되고 있다.

V. 결 론

인간의 오감 중에서 기술개발이 활발하게 진행되고 있는 촉각, 음성, 시각에 대해서 간략하게 알아 보았다. 그러나 진정한 오감을 인간과 로봇이 상호작용을 하기 위해서는 후각과 미각까지

기술개발이 진행되어야 한다. 후각은 여러 가지 센서들이 연구가 되고 있고, 미각에 있어서도 미국 텍사스대에서 최근 마이크로 머신 기술을 이용하여 액체의 맛과 화학성분을 가릴 수 있는 전자혀를 시제품의 형태로 개발되는 정도이다.

인간의 오감을 지닌 감성로봇에 대한 기술은 전자, 기계, 생체, 나노, 정보기술의 융합 기술이고, 미래 사회에는 반드시 이러한 형태의 로봇이 1가정 1대 또는 1인 1대의 보급이 될 것으로 예상되므로 발전가능성이 아주 크며 또한 세계적으로도 미성숙 기술로서 연구개발을 서두른다면 국가적으로도 우리 나라가 세계적으로 선점이 가능한 기술분야가 될 것이다.

참 고 문 헌

- [1] Grigore C. Burdea "Force and Touch Feedback for Virtual Reality"
- [2] Ishii, M. and M. Sato, 1994 "Force Sensations in Pick-And-Place Tasks," Proceedings of ASME WAM, DSC-vol. 55-1, ASME, New York, pp 339-344.
- [3] R. Cole, et al., "The Challenge of Spoken Language Systems : Research Directions for the Nineties," IEEE trans. on Speech and Audio Processing, Vol.3, No.1, pp 1-21, Jan 1995.
- [4] C. Boietet, K. Loken-Kim, "Human-

Machine-Human Interaction in Interpreting Telecommunications,” Proceedings ISSD-93, pp. 247-250, 1993

- [5] K. Sobottka and I. Pitas, “Segmentation and tracking of faces in color images,” Proc. of the Second Intl. Conf. On Auto. Face and Gesture Recognition pp. 236-241, 1996
- [6] P. Fieguth and D. Terzopoulos, “Color-based tracking of heads and other mobile objects at video frame rates,” In Proc. of IEEE CVPR, pp. 21-27, 1997
- [7] Gary R. Bradski “Computer Vision Face Tracking for use in a Perceptual User Interface” Intel Technology Journal Q2' 1998.
- [8] R. Chellappa, C. Wilson, and S. Sirohey, “Human and machine recognition of faces : A survey,” Proceeding of IEEE, vol.83, May 1995
- [9] F. Samaria and S. Young, “HMM based architecture for face identification,” Image and Computer Vision, vol.12, pp. 537-543, October 1994.
- [10] A. V. Nefian and M.H.Hayers, “A Hidden Markov Model for face recognition,” in ICASSP 98, vol.5, oo. 2721-2724, 1998

저자 소개



成夏慶

1960년 9월 13일생, 1986년 2월 한양대학교 기계공학과 학사, 1995년 8월 한양대학교 기계공학과 석사, 2001년 8월 아주대학교 기계공학과 박사수료, 1986년 2월~1987년 6월 : 한국전력, 1987년 6월~1992년 3월 : (주)삼성전기, 1992년 3월~현재 : 전자부품연구원 정밀기기연구 센터장, <주 관심 분야 : 모터설계 및 자동화 설계>