

主題

# 서비스 로봇을 위한 인지적 인간-로봇 상호작용

성균관대학교 지능시스템연구센터 이석한, 최종무, 백승민, Hadi Moradi

차 례

1. 서 론
2. 서비스 로봇에서의 HRI 기술 동향
3. 서비스 로봇을 위한 HRI의 주요 과제들
4. 통합적 접근: HRI 구현을 위한 핵심 기술
5. 맺음말

## 1. 서 론

가정 환경에서 인간과 공존하며 다양한 서비스를 제공하는 로봇의 실현으로 인간 삶의 질은 획기적으로 향상될 것이다. 인간과 로봇의 공존은 상호작용의 문제를 제기한다. 이는 조난현장의 구조 로봇을 원격으로 조작하는 것에서, 뇌파로 휠체어를 조작하는 인터페이스에 이르기까지 광범위한 영역에서 고려되는 문제이지만, 이 글에서는 가정이나 사무실과 같은 일상적인 생활공간에서 서비스를 제공하는 로봇과 관련한 인간-로봇 상호작용(Human-Robot Interaction, 이하 HRI) 문제에 초점을 둔다.

그림 1은 인간-로봇 상호작용 기술과 관련된 거시적 로드맵을 보여준다[1]. 서비스 로봇은 숙련된 작업에 대한 보조, 위험한 일이나 일상의

허드렛일에 대한 대체, 오락이나 교육, 인간과 로봇의 협동작업, 비서업무, 그리고 생활의 동반자에 이르기까지 다양한 수준의 응용 분야로 확대될 것이다[1-3]. 단순한 작업에 대한 보조는 마치 가전제품을 조작하듯이, 리모콘, 핸드폰, 웹 등의 도구를 통해 이루어질 수 있지만, 응용의 범위가 확대될수록, 인간과 인간의 상호작용에서 이루어지는 것과 같이, 음성이나 제스처, 표정 등과 같은 의사소통 수단에 대한 의존도와 요구가 점차 높아질 것이다. 향후에 인간 삶의 동반자로서의 로봇은 어린아이에서부터 노인에 이르기까지 인지능력, 교육수준, 연령, 성별, 문화 등 모든 특성을 지닌 사용자들과의 의사소통이 가능해야 한다. 이는 단순한 명령에 대한 이해를 넘어서, 정서나 의도 그리고 사회적 맥락에 대한 이해까지도 요구된다[4,5].

Technology	<b>HRI: Verbal Pictorial Gesture Mental &amp; Professional Emotional &amp; Social</b> Where- Face Speech Body Who- Face (frontal) Speaker Face (3D) Gait Behavior & body What- Command Gesture Dialogue Why- Facial expression Behavior Context understanding Platform: Mobile Robot Platform Humanoid Platform Robot Behavior Behavioral and Deliberate Hybrid Cognitive Engine Dependability/Safety: Soft Robot Self-Repair Self-Adaptation					From Navigation to Manipulation  Higher Level of Autonomy  Natural HRI  Emphasis on Dependability and Robustness				
	Robot Companion Robot Secretary Human-Robot Cooperative Task Execution Personal and Public Entertainment and Tutoring Replacing Human Labor for Daily Chore and Unsafe Work Assisting Human for Skilled Work					Power Supply  Interaction with Intelligent Space and Network				
	PC Power Broadband Batteries Transducers	3G-80G 10Mbps 200 WH/kg	10G-200G 20Mbps 300 WH/kg	30G-600G 100Mbps 400 WH/kg	100G-1200G 300Mbps 500 WH/kg	Smart Actuators	Modernization and Standardization  Cost Price Value			
	1995	1998	2001	2004	2007	2010	2013	2016	2019	

(그림 1) 인간-로봇 상호작용 거시적 로드맵 (Macro Roadmap)

현재의 HRI 관련 연구들은 인간의 명령과 행위를 이해하는 기본적인 능력을 로봇에게 부여해주는 음성인식, 얼굴 표정 해석, 제스처 인식과 같은 요소기술에 집중하고 있다. 그러나 이들 요소기술들은 아직 많은 과제가 남아 있는 불완전한 기술들인 반면에, 가정이나 사무실과 같은 일상환경은 초보적인 인지능력을 가지고 있는 로봇에게는 너무나도 다루기 어려운 복잡한 환경이다. 따라서 이러한 간격을 보완할 수 있는 연구가 시급하다. 그러므로 인간-로봇 상호 작용의 효과적인 구현은 음원추적, 얼굴인식, 얼굴검출, 표정인식, 제스처인식 등과 같은 각 요소기술의 개발만이 아니라, 이러한 요소 기술들을 통합, 보완하여 주어진 과제를 수행하는 기술의 개발을 동시에 추구해야 할 것이다.

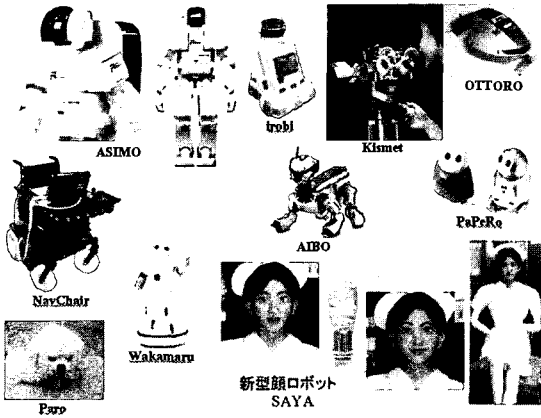
이 논문에서는 서비스 로봇에서의 HRI 연구의 최신 동향을 살펴보고, HRI의 주요 이슈들과 인지적 관점에서의 HRI 실현을 위한 주요한 연구 과제들을 개관한다. 그리고 통합적 접근 방법의 하나로써, 인지적 로봇 엔진의 필요성과 이를 호출자 식별 문제에 적용한 사례를 제시한다.

## 2. 서비스 로봇에서의 HRI 기술 동향

본 장에서는 이제까지 개발된 로봇들과 현재 진행 중인 서비스 로봇 프로젝트를 기준으로 인간-로봇 상호작용 기술 동향을 소개한다. 그림 2는 다양한 서비스 로봇을 보여준다.

서비스 로봇 로드맵 상에서 가장 먼저 출현할 것으로 예상되거나 이미 상품화된 예로는 청소, 잔디깎이와 같은 단순한 작업보조 로봇이 있다. 이탈리아의 Ambrogio[6]와 미국 IRobot사의 Roombavac[7]은 각각 자동으로 잔디를 손질하거나, 집안을 청소하기 위한 목적으로 설계된 서비스 로봇들이다. 청소로봇의 경우는 국내의 가전 및 로봇 회사들에서 경쟁적으로 제품화하고 있다. 이들 로봇은 버튼을 누르거나 숫자를 디스플레이 하는 정도의 매우 단순한 상호작용만을 요구한다.

경비, 안내와 같은 원격조작이나 간단한 음성 및 화상 처리를 기반으로 한 서비스 로봇도 이미 많은 곳에서 개발을 마쳤다. 전시회장에서 관람객들을 안내하는 로봇으로는 CMU의 Minerva[8]와 KIST에 의해 만들어져 현대 박물관에서 안

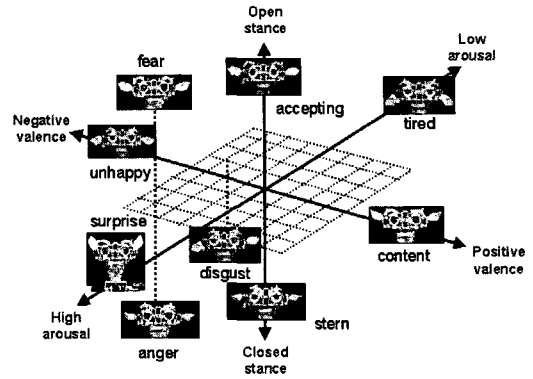


(그림 2) 다양한 서비스 로봇들

내를 수행하는 Jiny가 대표적인 예이다. 이러한 유형의 로봇들은 주로 대화형 디스플레이와 음성 명령어를 이용하여 관객들과의 상호작용을 수행한다. CMU의 Minerva는 입술, 눈썹 그리고 두 개의 카메라가 달린 얼굴을 가지고 있으며, 얼굴검출 기술을 이용하여 자신을 주시하는 사람을 확인할 수 있고, 기본 안내 멘트를 말할 뿐만 아니라 이를 변형해서 표현하는 것이 가능하다.

NEC의 PaPeRo[9,10], 유진로보틱스의 아이로비[11]와 같은 로봇은 가정 내에서 간단한 음성 및 화상인식을 통해 정보제공, 교육 등을 수행하는 상품화된 로봇의 예이다. PaPeRo는 동반자형의 개인용 로봇으로서, 음원 위치 추적, 음성 인식 및 합성, 얼굴 검출, 추적, 인식 기능을 내장하고 있다. PaPeRo는 얼굴을 검출하기 위해 우선 배경 영상과 움직이는 얼굴 영역의 차이를 이용하여, 머리 영역의 후보를 선정하고, 실제 얼굴 여부를 검증하기 위해 스테레오 카메라로부터 얻어진 거리 정보를 이용한 후에 템플릿 매칭을 수행한다. 음원의 위치 추적을 위해서는 복수개의 마이크로폰에 도달한 음성신호의 상대적인 위상 차이가 이용된다.

외형이 사람이나 동물과 유사한 로봇들이 나타나면서 생명체의 감성까지 모사하는 연구들도



(그림 3) KISMET의 감성모델

주목 받기 시작했다. 감성의 인식과 표현은 더 자연스러운 상호작용의 필수 요소이기 때문이다. MIT의 KISMET은 감성엔진을 가지고 있고, 얼굴표정을 통해 감성 표현도 가능한 3세 정도의 아기의 반응을 모델로 만들어진 로봇이다[12]. KISMET 연구의 목적은 로봇에게 사회적 상호작용 능력을 심어주는 데 있다. 사람을 포함한 환경 인식과 감성 반응을 보이는 행동의 구현을 위해 심리학적 생물학적 연구들을 이용했으며, 그림 3과 같은 감성모델을 가지고 있다. 와세다의 WE 시리즈[13]도 KISMET과 유사하게 심리학적 연구결과를 바탕으로 내부적인 감성모델과 감성변화를 구현하였으며 감성에 따른 사람의 얼굴표정을 모사했다. 도쿄 이과대의 SAYA는 인공 피부와 근육을 모사한 구동장치를 이용하여 사람의 표정을 더 현실적으로 표현하는 로봇이다[14,15]. 로봇에 감성을 부여하는 연구들은 로봇이 사람과 공존하면서 자연스러운 상호작용을 할 수 있도록 하고 사람에게 하나의 생명체로 인식되도록 한다.

서비스 로봇이 상호작용에서 실제 생명체나 동반자처럼 느껴지기 위해서는 감성뿐만 아니라 개성도 중요한 요소이다. 로봇의 개성은 미리 주어지기도 하지만, 사람과의 상호작용을 통해 변해가는 것이 자연스럽게 느껴진다. 실제 소년의

애완 강아지robot AIBO[16]는 주인과의 상호작용에 따라 성격이 변하도록 프로그램 되어 있다. 일본 AIST에서 개발된 PARO는 이동성은 없으나 사람의 쓰다듬거나 만지는 행동에 반응하도록 만들어져, 사람의 정서적, 심리적 치료에 쓰인다 [17].

로봇의 외형이 사람과 유사할수록 친근감을 줄 수 있기 때문에, KAIST의 AMI[18], 독일의 HERMES[19], Mitsubishi의 Wakamaru[20] 등과 같이 상반신이 휴머노이드 형태인 로봇도 많이 개발되고 있다. 기본적인 음성 인식, 합성뿐만 아니라 각종 센서와 스테레오 카메라를 이용한 환경인식과 사람처럼 두 팔을 이용한 제스처 표현이 가능하다. Wakamaru는 약 10,000개의 단어인식과 음성합성과 제스처를 이용한 표현이 가능하고, 전방향 카메라로 자신의 위치와 움직이는 사람의 위치를 확인하여 템플릿 매칭으로 얼굴을 인식하는 기능을 가지고 있다.

혼다의 Asimo[21], 소니의 Qrio[22]와 같이 이동방식도 사람과 같은 이족 보행이 가능하며, 계단, 문턱도 이동할 수 있는 사람과의 공존에 한 발 더 다가선 로봇도 개발되고 있다. Asimo는 50여개의 호출, 인사, 질문과 관련된 문장을 이해하고 반응할 수 있으며, 30여개의 신체 동작과 관련한 명령을 수행할 수 있다. 또한 얼굴인식, 움직이는 사람을 따라가는 기능, 손으로 지시하는 장소를 파악하는 등의 간단한 제스처 인식 기능을 갖고 있다.

이러한 휴머노이드 로봇들에게는 현재까지 개발된 음성, 비전, 촉각인식 등의 요소기술들을 응용한 멀티-모달 상호작용 혹은 통합적 인지 기능이 필수적이다.

눈동자의 움직임을 감지하여 휠체어 로봇이 이동하거나, 뇌파와 신경신호를 감지하고 전기적으로 변환하여 로봇팔 등을 움직이는 등의 특수한 인간-로봇 상호작용 기술이 장애인 혹은 노

인의 일상 생활을 돕기 위한 로봇들에게 적용된다. Michigan 대학의 Navbelt 와 NavChair 는 보행시에 장애물을 탐지하고, 경보음을 울릴 수 있도록 시각장애인을 보조하는 기능을 제공한다 [23]. 최근에는 신체장애자에게 조작 능력을 제공하기 위하여 로봇 머니플레이터를 제공하는 시도가 이루어지고 있으며, 더 나아가서 인체의 신경을 팔이나 다리의 손상된 부분을 대체하기 위한 로봇 제어 시스템과 연결하는 연구가 진행되고 있다. 그 외에 군사용, 의료용 로봇의 원격조작 및 힘 피드백이 되는 햅틱 인터페이스 개발도 여러 연구기관에서 이루어지고 있다.

### 3. 서비스 로봇을 위한 HRI의 주요 과제들

서비스 로봇을 위한 인간-로봇 상호작용의 주요한 개별 요소 기술로는 크게 시각 정보처리 기술, 음성 정보처리 기술 그리고 감성 및 인지 관련 기술로 구분 할 수 있다. 본 장에서는 개별 요소 기술의 주요 과제들을 살펴본다.

인간이 정보 전달을 위해 언어 이외에 동작, 표정과 같은 비언어적 수단을 이용하는 것처럼, 시각 정보처리 기술은 중요한 의미를 가진다. 시각 정보처리 기술에는 얼굴검출 및 인식, 신체 영역 검출 및 추적, 표정인식, 독순(lip reading), 시선 추정, 제스처 인식 등이 포함된다. 이 중에서 얼굴과 신체 영역의 검출은 주로 사용자의 존재나 위치를 파악하기 위해 사용될 수 있다. 얼굴검출 기술은 가장 활용도가 높은 기술 중의 하나로서, 복잡한 배경에서도 사용자의 위치를 비교적 손쉽게 추정할 수 있는 단서로 사용되는데, 원거리의 측면 얼굴까지도 검출하는 기술의 개발이 요구되고 있다. 신체 영역의 검출은 복잡한 배경에서의 영역 분할이라는 매우 어려운 문제와

관련되기 때문에 운동정보나 색상등의 부가적인 정보에 의존하는 접근 방법이 많이 쓰이고 있다. 복잡한 배경에서 실시간으로 다양한 자세의 신체를 확인하는 기술 개발이 요구된다. 얼굴인식 기술은 미 Identix, Eyematic Interfaces, 독일의 Cognitec, 삼성중기원, 블루닉스, 한신텔레리전트 등 다수의 국 내외 기업들이 사용자 인증이나, 출입 통제 등을 목적으로 상용화 하고 있는 기술이다. 그러나 많은 연구들이 제한된 조명 조건에서 자세나 표정 등이 제한된 상태의 얼굴 인식을 대상으로 하기 때문에, 인간과 로봇의 자연스러운 상호작용을 위해서는 아직 해결해야 할 과제들이 남아 있다. 특히, 조명 등의 환경 변화에 강인하게, 로봇을 직접 바라보지 않고 있는 다양한 자세의 사용자를 인식할 수 있는 기술이 요구된다. 이와 관련하여 독일의 Max Plank 연구소 등은 2차원 입력 영상만이 주어지는 경우에도 미리 저장된 3차원 얼굴 모형과의 정합을 통하여 자세와 조명 문제를 동시에 해결하는 접근 방법을 시도하고 있으나, 아직 실시간 처리는 곤란한 상황이다. 한편, 3차원 카메라의 발전은 3차원 얼굴인식 문제를 보다 쉽게 해결해 줄 것이다. 신원을 확인하기 위하여 걸음걸이 등의 행동을 이해하는 기술이 연구되고 있지만 실용적으로 적용되고 있지는 못하다.

사용자의 정서 상태나 의도에 대한 이해를 위한 표정인식 및 시선 추정 등은 주요한 요소기술의 하나이다. 얼굴의 개별적인 특징 요소의 추출 혹은 저장된 모형과의 정합은 비교적 높은 해상도의 영상입력을 요구한다. 그러나 통상적인 상호작용은 거리를 두고 이루어지기 때문에 요소 기술에서 요구하는 만큼의 정보를 얻지 못하는 경우가 많게 된다. 이와 관련하여 축적된 일련의 영상으로부터 높은 해상도의 영상을 얻는 기법 등이 유용하게 적용될 수 있을 것이다. 제스처 인식은 보다 직접적으로 호출, 인사, 지시 등의

행위를 이해하기 위해 사용된다. 제스처 인식은 Sony의 Eyetoy 등과 같이 매우 초보적인 범위로만 상용화 되어 있으며, 아직 많은 연구가 필요한 상황이다. 제스처 인식 기술은 여전히 피부 색상이나 움직임 정보 등의 초보적인 단서를 사용하는 방법에 많이 의존하고 있으며, 윤곽선의 추출 및 추적, 신체 모형과의 정합 등 다양한 접근 방법들이 시도되고 있다. 그러나 가정환경의 지능 로봇에 적합한 제스처 인식 방법은 아직 연구가 많이 필요한 상황이다. 특히, 실시간에 복잡한 배경으로부터 신체 영역을 분할 해 내고, 자세를 추정하는 기술이 하나의 핵심적인 과제이다.

음성 정보처리에는 음원 추적, 음성인식 및 합성, 화자 인식 등을 포함한다. 음성 정보처리에는 IBM, Microsoft, AT&T, NTT, ATR, MIT 등 많은 기관에서 연구되고 있는 오랜 역사를 가지고 있는 기술로서, 전화를 이용한 정보검색 서비스, 휴대폰 서비스, 번역기 등 다양한 분야에 상용화되고 있다. HRI 분야에서의 음성 정보는 사용자의 위치 파악 및 명시적인 명령어 입력 수단으로 주로 사용되고 있다. 음성인식 기술은 소규모, 고립단어, 화자중속에서 대규모, 연속문장, 화자독립 기술로 확장되고 있다. 제한된 전문분야를 영역으로 하여 마이크로폰을 사용하는 경우에는 받아쓰기가 가능한 정도로 인식성능이 향상되었음에도 불구하고, 잡음 환경에서, 원거리 음성인식은 아직도 과제로 남아있는 상황이다. 소수의 가족만을 대상으로 하는 서비스 로봇의 경우에는 화자적응기술을 사용하는 것이 가능한 반면에, 공공 서비스를 수행하는 안내 로봇과 같은 경우에는 화자 독립 음성인식 기술이 주요한 문제가 된다. 음원의 위치 추정 기술 역시도 실제의 환경에서 요구되는 잡음 환경에서의 복수 음원 위치의 동시 추정은 아직 해결되지 않고 있다. 음성 정보처리와 관련하여 배경 잡음과 음성신호를

실시간으로 분리하는 기술은 환경 잡음의 문제를 해결하는 하나의 중요한 기술이 될 것이다.

감성은 사회적 맥락과 관련된 매우 복잡한 현상으로 인간의 행동, 의사소통 및 상호작용에 핵심적인 요소이다. 예를 들어 휴머노이드 로봇이 음성 합성을 통해 사용자에게 정보를 전달할 때, 음성의 기복과 얼굴 표정의 변화를 적절히 사용하면 좀더 자연스럽게 느껴질 것이다. 또한 로봇이 사람의 얼굴 표정과 음성에서 감성 정보를 얻을 수 있다면 그 상황에 맞는 서비스를 제공하기에 유리할 것이다. 이렇듯 감성 상호작용 기술은 로봇을 단순한 기능만을 수행하는 기계가 아닌 사람과 공존하는 독립된 사회 구성원으로 느껴지게 하는 데에 없어서는 안 될 요소기술이다.

인간들 사이의 의사소통은 시각 정보와 청각 정보가 동시에 사용되는 경우가 많다. 예를 들어, 사용자가 손으로 특정한 사물을 지시하면서 음성으로 '저 컵을 가져와'하고 명령을 내리는 경우라면, 로봇은 음성인식과 제스처 인식을 통합하여 상황을 이해해야만 적절한 서비스를 수행할 수 있을 것이다. 이 때 시각이나 청각 단서 하나만을 사용해서는 제한적인 정보만을 얻게 된다. 따라서 효율적인 시청각 정보의 융합은 자연스러운 상호작용에 중요한 기술이 될 것이다. 한편, 영상을 이용한 독순(lip reading) 기술은 잡음이 매우 심한 환경에서 급격히 저하되는 음성인식 성능을 보완하는데 사용될 수 있다. 더 나아가 복합적인 입력정보를 상황 이해(context awareness)하고 적응 및 학습을 통해 로봇의 행동을 개선해 나간다면 생활의 동반자에 가까운 로봇이 구현될 것이다. 로봇의 상황 이해 능력을 극대화 하기 위한 유비쿼터스 네트워크 환경 연구나 로봇 기술을 적용한 지능형 빌딩, 지능형 자동차와 같은 임베디드 로봇 연구도 근래 많이 이루어지고 있다.

서비스 로봇의 기능 및 목표가 변화함에 따라서 HRI의 요구사항도 발전되어야 한다. 첫째, 명

령수행을 목표로 하는 경우 HRI 기능은 추상적이고 불완전한 인간의 명령에 대하여 상호작용으로 인간 의도를 분명하게 파악하는 기능이 요구된다. 둘째, 인간과 협동하여 작업을 수행하는 것을 목표로 하는 경우에는 작업, 임무 및 인간의 능력에 대한 이해가 필요하며, 인간의 현재 상태를 감지 및 이해하여 임무를 공동으로 성공 시킬 수 있는 상호작용이 요구된다. 셋째, 인간을 감성적, 정서적으로 지원하는 목표를 갖는 경우, 로봇은 인간의 감성을 이해하는 기능이 요구되는데, 인간 인지에 대한 모델링을 통해 인간이 자세하게 불평하지 않고 불완전한 표현을 해도 대처할 수 있는 능력과 개인의 특성에 따른 학습 기능이 요구된다. 넷째, 동반자로서의 서비스 목표를 갖는 경우, 앞의 세 가지 기능에 추가하여, 지적, 정서적 조력이 가능한 기능이 요구되며, 한 예로 개인의 특성에 맞는 깊이 있는 대화가 가능해야 한다. 이와 같은 HRI의 요구사항의 실현을 위해서는 지각적 수준의 상호작용을 넘어서는 상당한 수준의 로봇 지능이 필요하다.

#### 4. 통합적 접근: HRI 구현을 위한 핵심 기술

서비스 로봇의 자연스러운 인간-로봇 상호작용 실현을 위해서는 우선 시각, 청각 등의 개별 단서에 기반한 요소기술들을 조명 변화, 배경 잡음, 가려짐, 변형, 신호의 혼합 등과 같은 다양한 환경 요인들로부터 강인하게 동작하도록 하는 과제가 요구된다. 그러나, 인간-로봇 상호작용은 다양한 감각 단서에 의존하면서 동시에 운동정보의 해석과 같은 하위 수준의 단서에서부터 상황의 이해나, 대화 상대에 대한 인지 모형에 이르는 상위 수준의 정보가 통합된다는 점에서, 분명 개별 요소기술과는 다른 수준의 문제가 제기된

다. 즉, 자연스러운 인간-로봇 상호작용을 실현하는 하나의 접근 방법으로서, 특정한 단위 요소 기술에 의존하는 것이 아니라, 각 개별 요소기술을 최대한 효과적으로 통합하여 주어진 과제를 해결하는 것에 초점을 둔 방법과 기술의 개발이 요구된다. 본 절에서는 HRI의 개별 요소기술을 통합하는 하나의 접근 방법으로서 CRE를 제안하고, 이를 서비스 로봇의 호출자 식별에 적용한 사례를 제시한다.

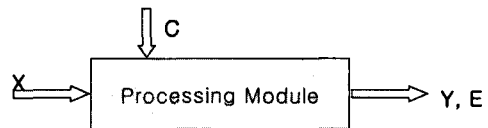
#### 4.1 CRE (Cognitive Robotic Engine)

기존의 로봇 제어구조(control architecture)는 크게 상위 수준의 인지과정을 강조하는 숙고형(deliberative)과 하위 수준의 인지과정을 강조하는 반사형(reactive)으로 구분된다[24]. 숙고형 시스템은 센싱-플래닝-행위의 순차적 과정을 통하여 주어진 목적을 달성하는 문제 해결과정을 갖는다. 이러한 접근 방법은 복잡한 행동을 계획하고 수행할 수 있는 가능성을 제시하고 있으나, 정확한 환경모델에 근거해야만 예측 가능한 성능을 얻을 수 있으므로 동적인 요소와 불확실성이 많은 환경에서의 동작에 문제가 있을 수 있다. 이와는 달리, 행위기반의 접근 방법은 분산된 개별 모듈들이 각자 환경에서 얻어진 정보로부터, 즉각적인 반응을 산출하는 방식에 기초한다. 복잡한 플래닝 없이도 실제 곤충의 움직임과 같은 자연스러운 행위를 보여 줄 수 있다는 점에서, 많은 가능성을 제시하고 있지만, 추론, 판단과 결정, 계획과 같은 좀 더 복잡한 상위 인지과정을 실현하기 곤란하다는 단점을 지닌다. 지능형 서비스 로봇에게 요구되는 기능의 실현에 적용하기에는 기존의 두 가지 접근 방식 모두 한계를 가진다. 따라서 본 연구에서는 인간과의 상호작용에 강인하고 상황 적응적이며 효율적인 서비스 로봇의 구현을 위하여 기존의 접근 방식의 한계를 극복할 수 있는 새로운 로봇 제어방식인 인지

로봇엔진(Cognitive Robotic Engine)을 제안하고자 한다. CRE는 계획모듈(planning module)과 반사모듈(reactive module)을 근본적으로 나누어 생각하지 않고, 모든 모듈들에게 입출력정보와 각 모듈에서 소비될 것으로 예상되는 시간 및 연산자원(resource)을 속성으로 부여한다.

각 모듈의 입출력 정보에 따라 로봇에 자국이 입력 되어 추론 및 행동이 이루어지는 연결형태가 결정된다. 로봇에게 미션이 주어지면 이 연결형태에 따라 스트레스 호르몬의 전파개념으로 미션을 수행하는 데에 필요한 모듈들이 선택된다. 만약 리소스의 제약이 무시된다면 들어오는 자극에 따라 데이터 흐름 개념으로 모든 모듈들이 활성화 되어 처리과정이 이루어진다. 이 경우 빠르게 나오는 결과를 선택하여 행동할수록 계획모듈들보다 반사모듈들의 결과를 따라 행동하게 되어 동적인 환경에 대한 반응이 빨라지고, 늦더라도 잘 계획된 결과를 선택하여 행동할수록 반응은 느리지만 주어진 미션을 효율적으로 수행할 수 있을 것이다. 더 나아가 로봇이 가지고 있는 자원의 제약을 고려하여 주어진 자원에 맞는 최적의 데이터 흐름을 선택하는 모듈이 필요하고, 이 모듈의 선택 방식에 따라 로봇의 행동양식이 변한다. 이러한 행동양식이 로봇의 개성으로 표현될 수 있고 사람과의 상호작용을 통해 변해갈 때 적응과 학습이 이루어지는 과정이라 볼 수 있다.

CRE는 로봇의 제어구조(control architecture)와 제어과정(control processing)으로 구성된다. 우선 구조는 다음과 같이 정의된다. 구조를 이루는 개별 구성요소는 처리모듈(Processing Module, 이하 PM)이다.



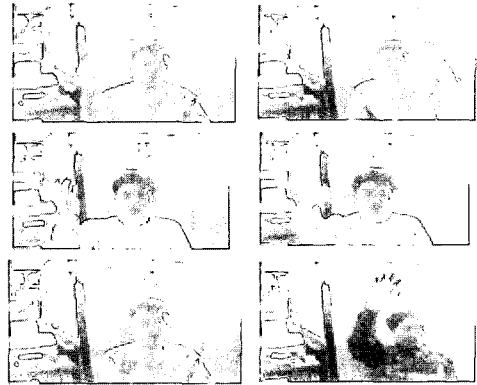
(그림 4) 처리 모듈

PM은 그림 4와 같이 외부로부터의 정보에 기반하여 행위를 산출하는 하나의 기능적 단위이다. 입력  $X$ 는 센서로부터 얻어진 외부 환경으로부터의 입력이거나, 다른 PM이 출력으로부터 얻어진 정보가 될 수 있다. 입력  $C$ 는 실행시간 조건, 신뢰도 등을 반영하는 제약조건이다. 출력  $Y$ 는 다른 PM의 입력이 되거나, 아니면 행동제어 모듈의 입력이 될 수 있다. 출력  $E$ 는 모듈의 자기 평가 결과로서 모듈 수행의 비용과 신뢰도 등을 포함한다. 각 PM은 주어진 입력으로부터 출력을 만드는 각기 다른, 이질적인 시스템으로 간주할 수 있다. 따라서 서로 다른 PM은 각자의 내부적인 정보의 표현(representation)이 가능하다. 예를 들자면, 어떤 PM은 HMM으로, 또 다른 PM은 규칙기반시스템으로 구현될 수 있다.

전체의 CRE는 PM 네트워크와 이 네트워크를 제어하는 시스템으로 구성된다. PM 네트워크는 PM들의 연결로 정의된다. PM 네트워크의 각 노드는 하나의 PM을 의미한다. 모든 PM들은, 개념적으로는, 서로 연결될 수 있다. 과제의 수행은, PM 네트워크의 활성화로 이루어진다. 네트워크의 활성화는 센서 입력과 직접 연결된 PM 노드의 활성화에서 목표 노드로의 전방향 과정(forward processing)과, 역으로 목표 노드의 활성화로부터 센서 입력 노드까지 활성화되는 후방향 과정(backward processing)으로 구분된다.

#### 4.2 사례 연구: 호출자 식별 시스템

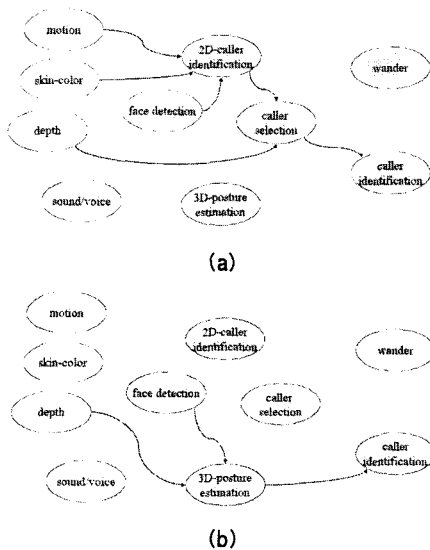
본 연구에서는 HRI를 위한 통합적 접근을 호출자 식별 문제에 적용한 사례를 살펴본다. 호출자 식별이란, 다수의 사람이 존재하는 복잡한 환경에서 서비스를 요구하기 위하여 호출하는 사용자를 로봇이 어떻게 판단하는가 하는 문제이다. 로봇을 호출하는 사람의 행동은 음성 명령, 응시, 손짓 등의 다양한 시·청각 단서를 통하여 이루어질 수 있다. 그림 5는 호출 제스처의 사례를 보



(그림 5) 호출 제스처의 사례들

여준다. 만일 잘 통제된 실험실 환경에서 한 명의 사용자와 로봇이 상호 작용하는 경우라면, 음원의 방향 과 같은 단일한 단서만으로도 호출자의 위치를 찾고 서비스를 수행하는 것이 가능할 것이다. 그러나 일상 가정환경은 실내외의 소음, 다수 사람의 공존 등의 복잡한 요인이 존재하기 때문에 음원 정보만으로 호출자를 판단하기는 매우 어렵다. 시각 정보의 경우에도, 단순한 상황이라면 얼굴검출과 같은 요소 기술로써 사용자의 위치를 파악하고, 제스처 인식 기술을 통하여 호출 의사를 파악하는 것이 가능할 것이다. 그러나 실제의 경우에는 우선 시야에 항상 호출자가 보여지는 것을 보장하기도 어려울 뿐만 아니라, 얼굴의 위치나 제스처와 같은 단서의 경우에도 다수의 사람이 공존하거나, 상호작용하는 복잡한 상황에서는 단독으로 사용되어 호출자를 정확하게 결정할 수 없게 된다. 이처럼 호출자 식별과 같은 서비스 제공의 초보적인 기능조차도, 음성이나, 시각 어느 하나의 단서만으로는 해결될 수 없으며, 또한 하나의 지각 단서 내에서도 다양한 수준의 단서들 - 시각 단서의 경우에는 움직임, 색상, 텍스처 등의 하위 수준 단서에서부터, 얼굴 위치 등의 상위 수준 단서에 이르기까지 - 이 통합되어야만 한다. 즉, 자연스러운 호출자 식별의 서비스 기능을 제공하기 위해서는, 얼굴 검출,





(그림 6) 호출자 식별의 두 가지 시나리오

제스처 인식, 음원 추정과 같은 개별 요소 기술들을 조명 변화, 복잡한 배경, 배경 잡음, 가려짐과 변형과 같은 내. 외부의 잡음에 강인하게 동작하도록 하는 요소 기술 개발의 노력과는 별도로, 자연스러운 상호작용을 실현하는 하나의 접근 방법으로서, 특정한 단위 요소 기술에 의존하는 것이 아니라, 각 개별 요소기술을 최대한 효과적으로 통합하여 주어진 과제를 해결하는 것에 초점을 둔 방법과 기술의 개발이 요구된다.

그림 6은 호출자 식별 시스템의 동작을 지각-행위들의 네트워크로서 표현한 것이다. 각 모듈은 하나의 지각-행위를 의미하며, 화살표는 모듈간의 의존 관계를 나타낸다. 음원과 시각정보를 사용하는 호출자 식별 과정은 사용되는 단서, 호출 동작이 이루어지는 상황, 각 모듈이 요구하는 자원의 정도와 신뢰도 등 많은 요인들이 관련되는 일련의 과정으로 볼 수 있다. 따라서 하나의 처리 방식이 다양한 상황에서 잘 동작하기는 쉽지 않다. 그렇다고 해서, 모든 변화 가능한 상황을 미리 고려한 알고리즘을 작성하기도 또한 쉽지 않다.

그림 6은 환경의 변화에 따라서, 개별 모듈이나, 모듈의 조합으로 연결된 특정한 정보 처리 방식이 적합하지 않게 되었을 때, 어떻게 시스템이 다른 대안적인 처리 경로를 생성할 수 있는가에 대한 하나의 답을 보여준다. 이 두 그림에서 구체적인 알고리즘은 중요하지 않으며, 처리 과정이 상황의 변화에 의해서 변화하는 것에 초점을 두어 설명한다. 상단의 그림 (a)는 2차원 영상처리에 주로 의존하여 호출자를 식별하는 과정을 나타낸다. 즉, 얼굴검출 알고리즘을 통하여 영상 내에서 사용자의 위치를 추정하고, 운동정보, 피부색 등의 하위 수준 지각 단서를 사용한 제스처 인식으로 호출 의사를 파악할 수 있다. 그런데 만일, 두 명 이상의 호출자가 발생하는 복잡한 상황이라면, 거리 정보에 추가적으로 의존하여 호출자를 선택할 수 있다. 하단의 그림 (b)는 3차원 정보에 주로 의존하여 호출자를 식별하는 과정을 나타낸다. 3차원 거리 정보로부터 신체 자세를 추정하는 알고리즘은 일반적으로 연산량이 많은 알고리즘에 의존하게 된다. 경우에 따라서는 2차원에서의 얼굴검출 정보를 사용하여 연산 시간을 줄일 수 있다.

그림 (a)와 (b)는 호출자 식별이라는 목표를 가지고 있는 지각-행위가 활성화 되었을 때, 이러한 목표를 달성할 수 있는 여러 경로들이 서로 경쟁하는 것으로 해석할 수 있다. 실제의 경로는 그림에서 표시한 것 보다 더 많을 수 있지만, 두 그림은 대표적인 두 개의 경로를 나타낸 것이다. 즉, 단순화 시켜서 보자면, 2차원 정보를 주로 이용하는 (a) 경로와 3차원 정보를 주로 이용하는 (b) 경로로 대표된다. 만일 연산시간의 제한이 없는 이상적인 상황이거나, 개별 모듈들의 처리에 비용이 거의 들지 않는다면, 모든 가능한 모듈들이 다 활성화 될 것이며, 모든 가능한 경로 혹은 과정으로부터 얻어진 결과들의 융합으로 최종적인 호출자 식별이 이루어질 수 있을 것이다.

그러나 실제의 로봇은 동시에 여러 목표를 추구하거나, 제한된 메모리와 처리시간을 갖는 등 항상 자원의 제약을 고려해야만 한다. 따라서 모든 경쟁적인 경로가 다 활성화 되는 것이 아니라, 선택적인 활성화가 이루어지게 된다. 물론 활성화 되는 경로는 자원의 허용 범위에서 복수로 선택될 수 있다. 각 경로에 대한 초기의 우선 순위는 미리 얻어진 실험 결과와, 전형적인 사용 환경을 고려함으로써 설정 될 수 있다. 만일 현재의 가용한 자원을 기준으로 볼 때, 2차원 정보에 주로 의존하는 (a) 경로가 평균적으로 더 좋은 결과를 제공해 준다면, (a) 경로가 더 높은 우선 순위를 가져야 한다. 그리고 이러한 우선순위의 배정으로 인하여 실제의 동작 과정에서 자원의 제약으로 인하여 (a), (b) 두 경로가 배타적으로 활성화 되는 경우라면 (a) 경로만이 선택되게 된다. 그런데, 만일 피부색과 유사한 배경이 많이 존재하고 또 배경에 움직이는 영역이 넓게 분포해 있어서, 피부색 정보와 운동정보를 거의 사용할 수 없게 되는 환경에 실제의 로봇이 놓여지는 경우를 생각하자. 우리가 대안적인 경로를 제공하지 않는다면, 이러한 예외적인, 그러나 실제로는 매우 빈번하게 발생하는 환경의 변화로부터 로봇은 자신의 초보적인 기능을 제대로 발휘하지 못하게 될 것이다. 반면에 제안하는 구조와 같은 대안적인 경로를 가진 경우라면, 전형적인 경로의 실패, 혹은 개별 모듈 출력 결과의 불확실성으로부터 초기 경로의 신뢰도를 예측할 수 있게 될 것이며, 다른 경로의 선택이 가능해질 것이다. 뿐만 아니라, 우리가 예측하는 것보다 더욱 양호한 환경이 주어지는 경우에는 더욱 효과적인 호출자 식별도 가능해질 수 있다. 예를 들자면, 조용한 환경에서 독신자와 주로 상호작용하는 로봇의 경우에는, 청각 정보 만으로도 호출자의 위치와 의사를 동시에 파악하는 것이 가능하다. 이러한 경우에는 미리 설정된 (a) 경로 이외에도, 단

위 모듈의 통계적인 상관에 기초하여, 더욱 단축된 새로운 경로를 쉽게 생성해 낼 수 있게 된다.

## 5. 맺음말

본 논문에서는 서비스 로봇의 자연스러운 인간-로봇 상호작용을 실현하는 핵심 사항으로서, 개별 요소 기술의 문제점을 상호 보완하는 통합적 접근 방법을 제안하였다. 얼굴인식, 제스처 인식, 음성인식 등의 개별적인 요소 기술들은 인간-로봇 상호작용에서 차지하는 중요성에도 불구하고, 아직 해결해야 할 과제들이 많이 남아 있다. 따라서 기술의 통합에 초점을 두어 문제를 해결하는 접근 방법은, 잠재적인 사용자들의 요구 수준과 기술의 격차를 해소하는 하나의 열쇠가 될 것이다.

## 참고문헌

- [1] S. H. Lee, "The Emergence of Intelligent Service Robotics," UN-ECE, Oct. 20, 2004.
- [2] Technology readmap: Robot, 산업자원부 한국산업기술 평가원, 2001.
- [3] 'IT 신성장동력' 발전전략, 정보통신부, 2003.
- [4] E. Rogers and R. R. Murphy, Human-Robot Interaction: Final Report for DARPA/NSF Study on Human-Robot Interaction, <http://www.csc.calpoly.edu/~erogers/HRI/HRI-report-final.html>
- [5] T. Fong, I. Nourbakhsh, K. Dautenhahn, "A survey of socially interactive robots," Robotics and Autonomous Systems, Vol. 42, pp.143-166, 2003.
- [6] <http://www.ambrogiorobot.com/>
- [7] <http://www.irobot.com/>

[8] <http://www-2.cs.cmu.edu/~minerva/>  
 [9] <http://www.incx.nec.co.jp/robot/>  
 [10] A. Kobayashi, I. Kume, A. Ueno, Y. Kono, M. Kidode, "A Robot Programming Model for Meditating Between Familiarity-Oriented Behaviors and Environment-Oriented Behaviors," The 7th World Multi-Conference on Systemics, Cybernetics and Informatics, pp.295-302, July 2003.  
 [11] <http://www.yujinrobot.com/>  
 [12] <http://www.ai.mit.edu/projects/humanoid-robotics-group/kismet/kismet.html>  
 [13] <http://www.takanishi.mech.waseda.ac.jp/eyes/>  
 [14] <http://koba0005.me.kagu.sut.ac.jp/newsinfo.html>  
 [15] [http://koba0005.me.kagu.sut.ac.jp/SAYA\\_2004\\_March.pdf](http://koba0005.me.kagu.sut.ac.jp/SAYA_2004_March.pdf)  
 [16] <http://www.sony.net/Products/aibo/>  
 [17] <http://www.mel.go.jp/soshiki/robot/biorobo/shibata/shibata.html>  
 [18] [http://mind.kaist.ac.kr/3\\_re/HumanRobot/HumanRobot.htm](http://mind.kaist.ac.kr/3_re/HumanRobot/HumanRobot.htm)  
 [19] <http://www.unibw-muenchen.de/hermes/>  
 [20] <http://www.sdia.or.jp/mhikobe-e/products/etc/robot.html>  
 [21] <http://asimo.honda.com>  
 [22] <http://www.sony.net/SonyInfo/QRIO/>  
 [23] [http://www.engin.umich.edu/research/mrl/00MoRob\\_19.html](http://www.engin.umich.edu/research/mrl/00MoRob_19.html)  
 [24] R. C. Arkin, Behavior-Based Robotics, The MIT Press, 1998.



**이 석 한**

1972년 : 서울대학교 (학사)  
 1974년 : 서울대학교 (석사)  
 1982년 : Purdue University, West Lafayette, Electrical Engineering (박사)  
 1974년~1978년 : 육군사관학교,

인하대학교 전임강사

1982년~1983년 : Rensselaer Polytechnic Institute(NY) Post Doctoral Fellow  
 1983년~1989년 : Univ. of Southern California 조교수  
 1990년~1997년 : Jet Propulsion Laboratory/ California Institute of Technology/NASA Senior Member of Technical Staff  
 1994년~ : IEEE Robotics & Automation Society 운영 위원 (97~ ), 기술위원회 의장 (94~ )  
 1994년~1995년 : KAIST 초빙교수  
 1994년~1999년 : IEEE Transactions on Robotics & Automation (94~96), Transaction on Neural Network(98~99) Associate Editors.  
 1995년~1997년 : IEEE ISATP, CIRA Conference 조직위원장(97), Program위원장(95,97)  
 1996년~ : KSEA(Korean American Scientists and Engineers Association)Southern California Chapter 회장(96,97) : KSEA Council Member (98~ )  
 1997년~ : IEEE Fellow  
 1998년~ : 제어자동화시스템공학회 이사(98~ ) ; 한국 센서학회 부회장 (2000~ )  
 1999년~ : 기초기술연구회 기획평가위원 ; G7 MEMS 총괄운영위원회 위원  
 1998년~ : 삼성종합기술원 System & Control Sector 연구소장 (전우)  
 2003년~ : 성균관대학교 정보통신공학부 교수  
 2004년 5월~ : 부회장, IEEE Robotics & Automation Society

<관심분야> Robotics, 자동화, 지능시스템, 정보저장, Mechatronics, 처리 및 Mobile System을 위한 Micro 및 Nano Systems(MEMS/NEMS)연구