

인간과 로봇의 다양한 상호작용을 위한 휴대 매개인터페이스 ‘핸디봇’

A Portable Mediate Interface ‘Handybot’ for the Rich Human-Robot Interaction

황정훈*, 권동수

(Jung-Hoon Hwang and Dong-Soo Kwon)

Abstract : The importance of the interaction capability of a robot increases as the application of a robot is extended to a human's daily life. In this paper, a portable mediate interface Handybot is developed with various interaction channels to be used with an intelligent home service robot. The Handybot has a task-oriented channel of an icon language as well as a verbal interface. It also has an emotional interaction channel that recognizes a user's emotional state from facial expression and speech, transmits that state to the robot, and expresses the robot's emotional state to the user. It is expected that the Handybot will reduce spatial problems that may exist in human-robot interactions, propose a new interaction method, and help creating rich and continuous interactions between human users and robots.

Keywords : human-robot interaction, service robot, mediate interface, Handybot, icon language

I. 서론

사람을 가장 많이 닮은 기계인 로봇은 과학기술의 발전에 따라 진보하고 있다. 로봇은 지난 수십 년간 3D업종이라 불리는 더럽고, 위험하고, 어려운 작업들에 있어서 점차 사람들의 수고를 대신하여왔고, 이제 일상 생활에서 사람들을 돕고, 심지어 감정적으로 상호작용할 수 있도록 개발되고 있다.

이와 같은 이유들로 인간-로봇 상호작용에 관련된 연구들이 많이 이루어지고 있으며, 주로 로봇이 사람과 같은 언어적, 시각적 능력을 갖도록 하여 사용자와 로봇이 사람간의 상호작용과 유사한 상호작용을 할 수 있도록 하려고 한다. 하지만 로봇은 사람처럼 대화하고, 표정을 이해하고, 제스처를 표현하거나 이해하고, 촉감 상호작용도 할 수 있지만 또 한편으로는 사람과 다르게 그래픽 유저 인터페이스(GUI)나 스위치, 무선 신호등과 같은 디지털의 상호작용 수단도 가지고 있다. 마이크와 GUI등을 갖는 휴대장치와 같은 매개체를 이용한 상호작용은 이러한 로봇의 장점들을 활용하여 로봇의 위치에 관계없이 로봇과 사용자가 상호작용할 수 있도록 한다(그림 1). 이러한 매개체 역할의 인터페이스를 매개인터페이스(mediate interface)라 하겠다.

본 연구에서는 휴대용 매개인터페이스로 개발된 핸디봇(Handybot)을 소개한다. 핸디봇은 KAIST에서 일상생활에서의 사용을 목적으로 하는 서비스 로봇과 사용자의 상호작용을 돕기 위해 로봇과 인간 사이를 매개하는 인터페이스로서 개발된 것이다. 사용자의 입장에서는 로봇을 대신하여 상호작용할 수 있는 또 하나의 로봇이기 때문에 가지고 다닐 수 있는 로봇이란 의미에서 핸디봇(Handybot)이라 명명하였다. 대

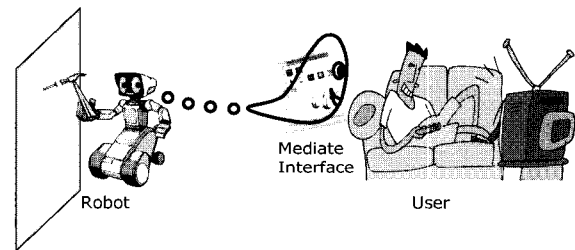


그림 1. 매개인터페이스를 통한 인간-로봇 상호작용의 개념도.
Fig. 1. Schematic diagram of human-robot interaction using a mediate interface.

상이 되는 로봇은 전형적인 가정환경에서 사용되며, 사용자나 환경과의 상호작용을 통해 지속적인 학습이 가능하고, 사용자와 감정적인 상호작용도 가능한 로봇이다. 핸디봇은 이와 같은 기능의 로봇이 사용자와 작업중심의 상호작용과 함께 감정적 상호작용을 할 수 있도록 매개하는 역할을 수행한다.

기존의 연구에도 매개인터페이스라 할 수 있는 장치들이 존재하였다. Fong, Thrope와 Baur는 PDA를 이용하여 로봇과 텍스트로 대화하고, 현재의 지도와 카메라 정보를 사용자에게 보여주고, 수동의 제어를 가능하게 하였다[1]. Keskinpala, Adams와 Kawamura는 PDA를 카메라 및 센서 정보와 몇 개의 버튼으로 이루어진 원격제어기로 이용하였다[2]. Perzanowski 등은 PDA상의 화면에 지도를 보여주며 터치스크린을 사용하여 로봇을 제어하였고[3], Chronis와 Skubic은 PDA를 이용하여 로봇의 네비게이션을 위한 지도를 그리는데 사용하였다[4]. 이러한 연구들은 PDA를 이용하여 로봇을 원격제어하는 시도들이었다. 본 연구에서는 이러한 시도들에 더하여 작업중심의 매개 상호작용에 적합한 아이콘 언어와 감정 상호작용 채널을 추가하여 사용자가 로봇의 위치와 상관없이 다양한 상호작용을 할 수 있도록 하였다.

2장에서는 본 연구와 관련된 이슈 및 관련 연구들을 살펴

* 책임저자(Corresponding Author)

논문접수 : 2007. 5. 15., 채택확정 : 2007. 6. 25.

황정훈, 권동수 : 한국과학기술원 기계공학과

(hwangjh@robot.kaist.ac.kr/kwonds@kaist.ac.kr)

※ 본 연구는 산자부지원의 21C 프론티어 연구개발 프로그램 중 지능형 로봇 개발과제와 BK21과제의 지원에 의해 연구되었음.

보고, 3장에서는 매개인터페이스 핸디박의 설계와 구조에 대해 살펴보고, 4장에서는 작업중심의 상호작용 채널인 아이콘 언어를 설명한다. 5장에서는 핸디박의 감정 상호작용 채널에 대해 살펴보고 6장에서는 논문의 결론을 맺도록 한다.

II. 매개인터페이스 핸디박 관련연구

1. 모달리티 관련연구

사람의 상호작용에서 멀티모달 상호작용은 매우 자연스럽게 흔히 관찰되는 것이다. 멀티모달 상호작용이 필요한 이유 중의 하나는 음성 상호작용만으로는 의도를 명확하게 전달하지 못하는 경우가 많기 때문이다. 시각정보 같은 것들이 음성정보와 함께 사용되기도 하고, 의도나 감정을 표현할 때 음성표현과 함께 제스처를 사용하기도 한다. 로봇에게는 멀티모달 상호작용이 필요한 중요한 이유들이 더 존재한다. 표정, 음성이나 제스처 인식에 관한 현재의 기술들은 아직까지 큰 폭으로 변화하는 다양한 입력들에 대해 안정성이 부족하다. 예를 들어 음성 인식률은 화자로부터의 거리에 민감하게 변화한다. 따라서 기본 인식 시스템의 실패에 대비한 방법이 필요하다. 이러한 이유로 로봇은 다양한 모달리티의 정보를 인식하고 표현할 수 있는 능력이 요구된다 하겠다.

멀티모달 상호작용은 Bolt의 연구로부터 시작되었다[5]. Bolt는 "Put that there"라는 음성정보와 특정 물건과 장소를 가리키는 지시동작을 이용하여 명령을 생성하였다. Bischoff와 Graefe는 그들의 로봇 'HERMES'에 음성과 키보드, 그리고 이메일을 사용하여 명령을 내렸고 전달된 명령은 자연어 처리 기술을 이용하여 로봇이 이해할 수 있는 언어로 번역되었다[6]. Perzanowski 등에 의해 개발된 'Coyote'는 음성명령, 제스처, 그리고 터치스크린 입력 등의 멀티모달 정보를 이용하여 제어되었다[3]. 이러한 연구들은 모두 사용자와 로봇간의 상호작용을 쉽고 효과적으로 만들기 위하여 다양한 모달리티들을 사용하고자 노력하였다.

매개인터페이스 핸디박은 로봇의 멀티모달 상호작용 능력을 증가시키기 위해서 사용된다. 음성 인터페이스와 함께 GUI를 사용하고, 로봇이 멀리 있을 때에도 사용자의 표정과 음성감정을 인식할 수 있는 인식기들을 가지고 있도록 하였다. 또한 핸디박은 휴대가 가능하고 사용자의 손위에서 작동할 것이기 때문에 핸디박상에 존재하는 마이크나 카메라와 같은 센서들과 사용자와의 거리가 항상 일정한 간격을 유지할 것이므로 음성인식이나 표정인식에서 유리하다. 따라서 핸디박은 로봇이 사용자와의 상호작용을 위해 사용자와 가까이 올 필요를 없애기 때문에 로봇의 효율성을 높이는 데도 기여를 할 것으로 본다. 매개인터페이스 핸디박의 설계 및 구조는 3장에서 자세히 설명하였다.

2. 작업중심 언어 관련연구

앞서 얘기한 바와 같이 멀티모달 상호작용의 실제적인 이유는 다른 모달리티들-주로 시각정보-과 관련된 상황의 경우 음성 상호작용만으로는 충분하지 못한 경우가 많기 때문이다. 즉 음성과 시각정보의 결합을 이용하면 사람들이 상호작용을 보다 잘 이해할 수 있도록 할 수 있다. 이러한 연구들의 일환이 비주얼 언어(visual language)들이다. Pfeiffer는 LEGO로봇과 같은 모바일 로봇을 제어하는 프로그래밍 언어

의 용도로 Altaira라는 비주얼 언어를 제시하였었다[7]. Cox, Risley와 Smedley도 같은 종류의 로봇에 대해 유사한 비주얼 언어를 사용한바 있다[8]. 그보다 앞서 Leifer, Loos와 Lees가 제안한 비주얼 언어는 보다 높은 수준의 언어를 지양하였다[9]. 그들이 사용한 비주얼 언어는 언어장애자들을 위해 개발된 것으로 자연어와 유사한 수준의 표현이 가능하였다. 하지만, 이러한 연구들은 모두 로봇의 프로그래밍을 위해 개발되었으며 직접적인 상호작용을 위한 것이 아니었다.

본 연구에서는 로봇을 프로그래밍하는 것이 아닌 로봇과 상호작용하는 것에 초점을 맞추었다. 인간-로봇 상호작용에서 효율을 낮추는 가장 큰 요인 중의 하나는 그들이 얘기하는 내용의 기저를 이루고 있는 배경지식(grounding [10])에 관련된 것이다. 만약 인간과 로봇이 배경지식을 공유하고 있지 못하다면 상호작용은 지루하고, 어렵고, 오해하기 쉽게 될 것이다. 음성언어는 작은 정보량으로도 매우 효과적으로 정보를 전달할 수 있지만 그만큼 의미가 모호한 경우가 많다. 따라서 대화에 포함된 대상의 특징을 가장 효과적으로 표현할 수 있는 정보인 대상의 실제 이미지를 이용하여 상호작용에 사용함으로써 사용자와 로봇이 작업중심의 상호작용을 하는데 필요한 배경지식(작업의 대상이 되는 물체나 장소와 같은 것들)을 쉽게 공유할 수 있도록 하기 위해 아이콘 언어가 제안되었다. 아이콘 언어는 로봇의 지식에 기반하고 있다. 아이콘들은 로봇이 외부에 대해 갖는 표상을 형상화 한 것으로 이루어져 있다. 사용자는 로봇과 상호작용 하기 위해 그 아이콘들을 탐색하고 이용할 수 있도록 하였다. 작업중심의 아이콘 언어는 4장에서 보다 상세히 기술하였다.

3. 감정 상호작용 관련연구

Nomura, Kanda와 Suzuki의 연구를 보면 로봇에 대해 부정적인 견해를 갖고 있는 잠재적 사용자들은 로봇과의 상호작용을 회피하려 하는 경향을 보인다고 한다[11]. 이러한 결과는 감정적으로 로봇을 받아들이지 못하면 로봇과의 상호작용도 성공적이지 못할 것이라는 것을 보여준다. 한편 Mitsui, Shibata, Wada와 Tanie의 실험은 로봇이 정신 치료에 사용될 수 있음을 보여주었다[12]. 애완 로봇과의 상호작용은 환자들의 감정상태와 생리적 상태를 긍정적으로 변화시켰다. 이러한 결과들은 로봇과의 감정 상호작용이 다른 종류의 상호작용에도 큰 영향을 미칠 수 있음을 보여준다.

이러한 사실은 KTH에서 수행한 지능로봇 연구에서도 확인되고 있다. 오피스 환경에서의 'Fetch-and-carry' 형태의 작업에서 잠재적 사용자들을 대상으로 수행한 실험의 결과는 감정 상호작용이 바람직함을 보이고 있다[13]. 또한 CMU의 접대로봇 'Valerie'는 다양한 개성을 가지고 방문객들과 호감가는 상호작용을 수행하기도 하였다 [14]. 따라서 로봇의 감정 상호작용 능력은 상호작용의 성공을 위해서 매우 중요하다 할 수 있다. 또한 감정은 로봇 자신이 판단, 계획, 그리고 학습과 같은 인지적 작업을 수행함에 있어서 중요한 영향을 미치는 부분이기도 하다.

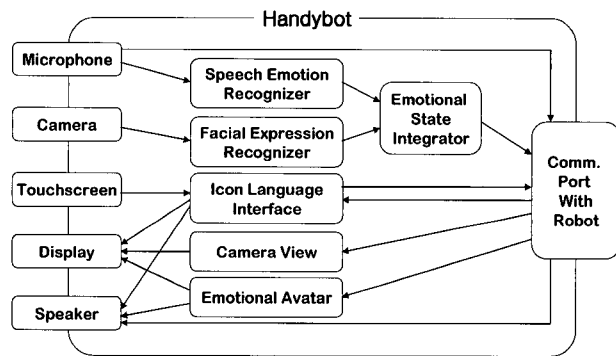
따라서 지능형 서비스 로봇을 위한 매개인터페이스는 로봇과 사용자간의 감정 상호작용 채널을 갖는 것이 필요하겠 다. 핸디박은 다양한 감정 상호작용 채널을 갖는다. 표정인식과 음성감정 인식을 수행하고, 로봇의 다양한 감정을 표현할

수 있도록 아바타를 이용하여 감정을 표현할 수 있다. 핸디봇을 통한 인간-로봇 감정 상호작용에 대해 5장에서 보다 상세히 다루었다.

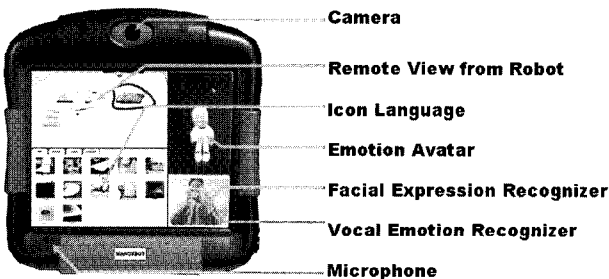
III. 휴대용 매개인터페이스 핸디봇의 설계와 구조

핸디봇이라는 이름은 핸드캐리 로봇(hand-carry robot)에서 비롯되었다. 핸디봇의 설계는 지능형 서비스 로봇의 상호작용 범위를 확장하는 데에서 시작하였다. 대상이 되는 로봇은 가정환경에서 청소나 물건 가져오기, TV프로그램 정보 제공 등과 같이 간단한 작업을 수행하는 로봇이다. 다양한 모달리티로 상호작용할 수 있으며 사용자의 감정상태를 이해하고 자체적인 감정을 생성할 수 있는 로봇을 목표로 한다. 핸디봇은 이러한 로봇의 기능을 공간적으로 확장시킨다. 사용자의 음성명령을 로봇으로 전송하고, 음성감정과 표정을 인식하고, 추가적으로 아이콘 언어, 로봇으로부터의 원격지 장면을 사용자에게 제공하며, 로봇의 감정상태를 아바타의 형태로 표현할 수 있다.

핸디봇의 하드웨어는 충분한 컴퓨팅 파워와 이동성을 확보하기 위하여 노트북 컴퓨터를 개조하여 만들어졌다. 무게는 2.89kg이며 사이즈는 32.0cm, 31.5cm, 4.5cm이다. 핸디봇은 사용자의 얼굴과 표정을 인식하기 위하여 카메라를 가지고 있다. 카메라는 30frames/sec의 리프레쉬율을 갖는 1.3M픽셀의 CCD를 사용하였다. 핸디봇의 화면은 12.1인치에 1024, 768 픽셀들을 갖는 터치스크린을 사용하였다. 또한 무선 네트워크를 사용하며, 리눅스를 OS로 사용하고 있다. 그림 2(a)는 핸디봇의 구조를 보여주고 있고, (b)는 핸디봇상의 다양한 인터페이스들을 보여주고 있다.



(a) The structure of Handybot



(b) Various interfaces on Handybot

그림 2. 휴대 매개인터페이스 핸디봇.

Fig. 2. Portable mediate interface Handybot.

핸디봇 상에서의 표정인식과 음성감정 인식는 몇 가지 장점들을 갖고 있다. 사용자가 핸디봇을 통해 로봇과 상호작용하고자 할 때 로봇은 원격지에서 다른 일을 하고 있을 수 있는 상황을 예상할 수 있다. 이때 핸디봇에서 사용자의 표정 및 음성 감정을 인식한다면 로봇의 작동에 필요한 컴퓨팅 파워의 부담을 줄여줄 수 있다. 또한 각각의 핸디봇은 제한된 소수의 사용자들과만 로봇과 상호작용할 수 있도록 하면 되기 때문에 로봇에서 사용되는 사용자 독립형 인식방법들 보다 상대적으로 간단하고 정확한 사용자 종속형 인식방법들을 사용할 수 있다. 또한 부수적으로 핸디봇 자체가 로봇과는 다른 감정적 특징들을 갖도록 할 수도 있겠다.

핸디봇은 위와 같은 다양한 모달리티들을 이용해 로봇과 사용자가 두 종류의 상호작용을 할 수 있도록 한다. 첫번째는 작업중심의 상호작용 채널이며, 두번째는 감정 상호작용 채널이다. 다음 장들에서 각각에 대해 자세히 설명하였다.

IV. 핸디봇의 작업중심 상호작용 채널

1. 아이콘 언어의 설계

매개인터페이스를 사용하는 가장 주된 이유는 로봇이 사용자의 시야에서 벗어난 경우에도 상호작용을 할 수 있도록 하기 위함이다. 예를 들어 사용자와 로봇이 각기 다른 방에 있어도, 사용자는 매개인터페이스를 통해 로봇이 보는 것을 볼 수 있다. 그리고 로봇과의 거리에 상관없이 매개인터페이스를 통해 음성 상호작용이나 GUI를 이용한 상호작용을 할 수 있다. 그러나 음성 상호작용 방법이 사용자에게 가장 자연스러운 방법임에는 틀림없지만, 현재의 기술의 한계로 인해 사용자와 로봇간의 음성 상호작용의 효율과 효과가 떨어지는 것이 사실이다. 본 연구에서는 이러한 문제점을 보완하고, 로봇이 갖고 있는 배경지식을 사용자와 공유하기 위한 방법으로 아이콘 언어를 개발하였다. 이 아이콘 언어는 음성 상호작용과 같이 쓰일 수 있다(그림 3).

본 연구에서의 아이콘은 단순한 시각표상이 아니고 로봇이 외부세계에 대해 내부적으로 갖고 있는 표상이다. 따라서 일반적인 그래픽 사용자 환경에서 사용되는 아이콘이라기

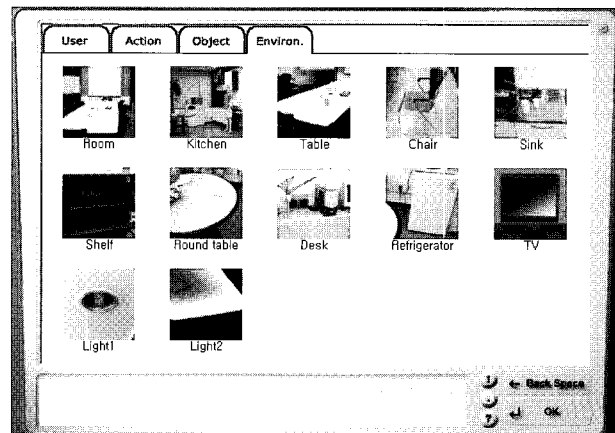


그림 3. 핸디봇을 위한 아이콘 언어 인터페이스. 아이콘들이 로봇이 외부세계에 대해 갖는 표상으로 이루어져 있음.

Fig. 3. Icon-language for the Handybot. Icons are robot's representation of external world.

보다는 일종의 심볼이라고 할 수 있으며 로봇 내부에서는 하나의 심볼로서 사용된다. 각각의 아이콘들은 시각적 요소와 음성적 요소로 이루어져 있다. 아이콘의 시각적 부분은 로봇이 관찰하고 기록한 외부세계에 대한 화상으로 이루어지며, 음성적 부분은 사용자가 외부세계에 대한 사용자의 교육으로부터 얻어진다.

아이콘들은 네가지 그룹으로 나눌 수 있다. 액션(action), 사용자(user), 장소(place) 그리고 사물(object)이 그것이다. 액션 그룹은 로봇이 할 수 있는 것들로서 미리 정의되어 있다. 본 연구에서 로봇의 액션은 다음과 같이 정의하였다.

- A에서 B로의 이동
- 물건을 잡거나 놓기
- 어떤 장소나 물건을 보기
- 스위치를 누르기

아이콘은 심볼로서 인식되기 때문에 동사나 명사, 형용사 같은 언어적 요소로 해석될 수 있다. 아이콘 들의 조합은 한국어나 영어의 문법적 순서 등에 기초하여 문장을 형성할 수도 있다. 예를 들어 '탁자 위의 컵을 가져와'라는 표현은 다음과 같은 아이콘들의 조합으로 표현될 수 있다.

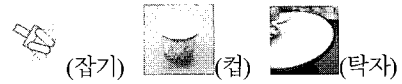


그림 4(a)는 이와 같이 로봇의 표상을 아이콘으로 이용한 아이콘 언어만을 사용하여 상호작용하는 예를 보여주고 있다. 먼저 사용자가 “나에게 음료 한컵 다오”라는 명령을 위해 아이콘들을 차례로 눌러서 입력하여 대화를 시작한다. 다음에는 로봇이 음료를 가져와 “컵을 내려놓겠습니다.”라고 말한다. 아이콘 언어는 또한 음성 대화와 함께도 사용이 가능하다. 그림 4(b)는 음성 언어와 아이콘 언어를 함께 사용하여 위와 같은 작업을 로봇에게 부여하는 것을 보여준다. 사용자는 특정 음료에 해당하는 아이콘을 누르며 “나에게 음료 한컵 다오”라는 음성 명령을 내린다. 이때 아이콘 언어 인터페이스에서는 로봇이 인식한 결과를 아이콘 문장으로 함께 표현한다. 이후의 과정은 아이콘 언어만 사용한 상호작용에서와 같게 진행된다.

이와 같은 과정으로 사용자의 의도를 아이콘 문장의 형태로 입력 받으면 그것으로부터 사용자의 의도를 인식하는 과정이 필요하겠다.

2. 아이콘 언어의 해석

일단 사용자가 아이콘 언어를 이용하여 입력을 생성하면 아이콘 문장이 만들어지고 각각의 아이콘 문장은 로봇이 수행할 수 있는 특정한 범주의 작업으로 할당될 수 있다. 앞에서 정의한 로봇의 액션으로 수행 가능한 범주의 로봇의 작업들은 다음과 같다.

- Task 1: 특정한 장소나 물건이 있는 곳으로 이동하기
- Task 2: 특정 물건을 가져오기
- Task 3: 특정 물건을 A에서 B로 옮기기
- Task 4: 특정 장소나 물건을 보여주기
- Task 5: 정해진 스위치를 켜기(끄기)
- Task 6: 사용자에게 와서 물건을 특정 장소로 옮기기

하나의 아이콘 문장은 하나의 작업으로 할당된다. 하지만 아이콘 언어의 표현에 있어서 사용상의 편의를 위해 문법적 제한을 두지 않는다면 같은 작업을 위해 다양한 형태의 표현이 가능하겠다. 또한 실제의 경우 사용자들은 몇몇 정보를 빼먹기도 한다. 제안된 아이콘 언어 인터페이스는 사용자의 편의를 위해 불완전 문장과 다양한 순서로 표현할 수 있도록 하였다. 따라서 다양한 형태의 아이콘 문장들을 해석하기 위해 문법적 모델을 사용하지 않고 확률적 모델을 사용하여 아이콘 문장의 해석을 구현하였다.

아이콘 문장이 로봇에게 제시되면 로봇은 입력된 문장을 작업으로 할당하는 과정에서 다음과 같은 문제를 갖게 된다.

- 1) 같은 의미를 갖는 다른 표현들: 사람들은 그들의 의도를 같은 방식으로 표현하기 보다는 사람마다 다르게, 그리고 때마다 다르게 표현한다.
- 2) 불충분한 입력: 사람들은 종종 축약된 문장을 사용한다. (예: '나에게 커피 한컵 가져다오'라는 표현 대신 '커피좀'이라고 표현하곤 한다.)

예를 들어 로봇이 부엌에 있는 컵을 가져오길 원한다면 그것은 다음과 같이 다양하게 표현될 수 있다.

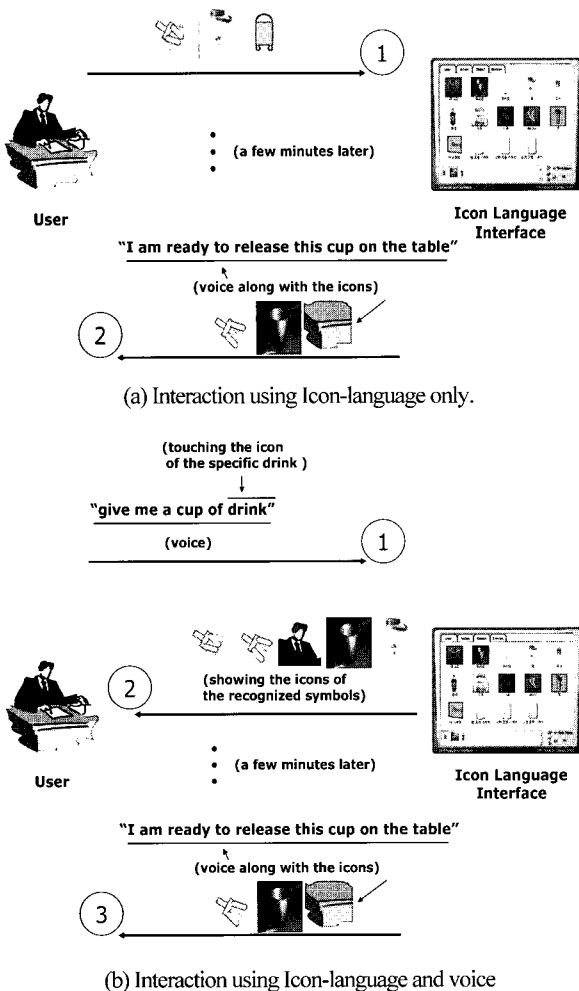


그림 4. 아이콘 언어 인터페이스를 이용한 상호작용예.
Fig. 4. Examples of the Icon-language interface interaction.

아이콘 문장	사용자의 의도
이동 부엌 잡기 컵 이동 사용자	← “부엌으로 가서, 컵을 잡아, 나에게로 돌아오라”
부엌 잡기 컵 이동 사용자	← “부엌에서, 컵을 잡아 나에게 돌아오라”
부엌 컵	← “부엌에 있는 컵쯤”
컵	← “컵쯤”

위와 같은 문제들을 갖는 문장들을 해석하기 위해 마르코프 프로세스(Markov process)가 사용되었다[15]. 확률론적 방법으로서 마르코프 프로세스는 패턴분류와 같은 문제에 많이 사용되고 있다. 이 모델은 또한 입력의 개수가 변화하는 것에도 유연하게 사용될 수 있다(그림 5).

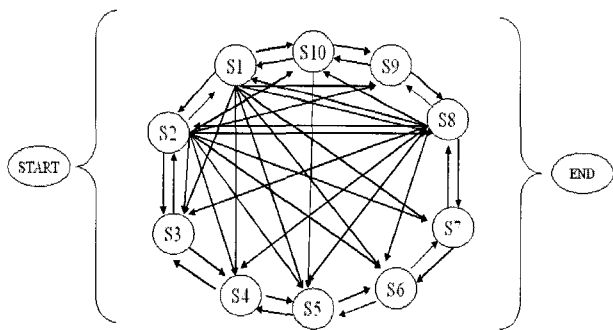
1차 마르코프 프로세스에서 상태변이는 직전 상태에만 관련되어 있다. 이것이 잘 알려진 마르코프 가정이다. 이 관계는 $\{S_1, S_2, \dots, S_n\}$ 와 같은 순서의 상태들에 대해 (1)과 같이 표현된다.

$$P(S_n | S_{n-1}, S_{n-2}, \dots, S_1) = P(S_n | S_{n-1}) \quad (1)$$

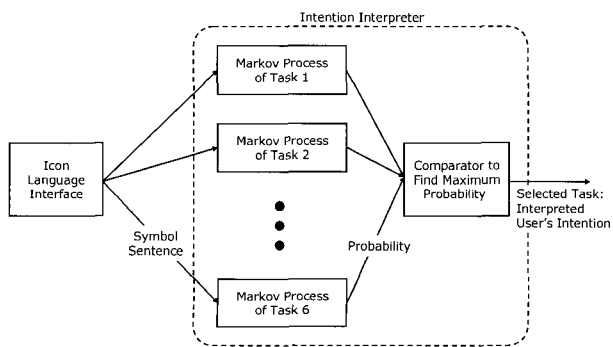
따라서 전체 문장의 확률은 (2)와 같이 계산된다.

$$P(S_1, \dots, S_n) = \prod_{i=2}^n P(S_i | S_{i-1}) \quad (2)$$

제안된 모델은 시작 상태와 종료 상태만 서로 연결되지 않



(a) 1st order Markov process model used in Interpreter. Every state has connection with each other except start and end state



(b) Structure of the interpreter

그림 5. 아이콘 언어 해석기.

Fig. 5. Interpreter of Icon-language.

고 다른 상태들과 모두 연결을 갖고 있다. 그림 5는 제안된 모델을 보여준다. 앞 절에서 언급한 바와 같이 모든 아이콘들은 네가지 그룹으로 분류될 수 있다. 하지만 각 작업간의 구별성을 확보하기 위하여 분류를 좀더 세분화 하였다. 네가지 분류 중 사물 그룹은 다시 잡을 수 있는 것과 누를 수 있는 사물로 나뉘지고, 액션 그룹은 이동, 돌아오기, 잡기, 놓기, 보기, 누르기로 세분화하였다.

입력 문장이 “부엌 이동” 과 같다면 문장의 특정 작업 할당 확률은 (2)를 이용하여 (3)과 같이 계산된다.

$$P(\text{“부엌 이동”}) = P(\text{장소시작})P(\text{이동장소})P(\text{종료이동}) \quad (3)$$

앞 절에서 기술된 여섯가지의 작업에 대응하여 여섯개의 모델이 만들어 졌고, (2)를 이용하여 여섯가지 작업에 대한 할당 확률을 계산할 수 있겠다. 이중 가장 높은 할당 확률을 갖는 작업을 사용자의 의도로 인식하게 된다.

3. 아이콘 언어 구현 및 실험결과

각각의 작업에 대응되는 마르코프 프로세스는 사용자들에게 여러 가지 작업을 아이콘 언어를 이용하여 표현하게 하는 실험을 통해 구할 수 있었다. 각각의 사용자들이 같은 작업을 모두 같은 방식으로 표현하지 않기 때문에 각각의 작업들에 대해 자연스럽게 다양한 아이콘 문장 표현을 얻을 수 있었다.

실험은 다음과 같이 수행되었다. 로봇에 대한 특별한 훈련이나 경험이 없는 보통의 사람들에게 로봇에게 여러 가지의 작업을 시키도록 하였다. 제시된 작업내용의 단순한 복제를 방지하기 위하여 피실험자에게는 작업의 내용을 주지 않고 특정한 상황이 주어졌다. 또한 여러 가지 물건들이나 장소 등을 피실험자나 로봇이 모두 알고 있는 상황을 가정하였다. 다음의 문장들이 실험에서 작업 2, 즉 물건 가져오기 범주로 사용된 예제 문장들이다.

- 당신은 지금 매우 목이 마릅니다. 당신은 로봇에게 마실 것을 가져오도록 만들고 싶습니다.
- 테이블 위에 꽃이 있습니다. 당신은 그것을 가져오게 하고 싶습니다.
- 싱크 위에 사과가 있습니다. 당신은 그것을 먹고 싶습니다.

실험에는 56명의 피실험자가 참여하였다. 71%가 남성이었고, 평균 나이는 30.9세였다. 연령대는 20세에서 59세까지이며 25~29세의 그룹이 가장 많은 41.1%의 비중을 차지했다. 33.9%의 피실험자는 직장인이고 나머지는 학생이었다.

실험으로부터 여섯가지의 작업을 위해 모두 1051개의 명령 샘플들을 얻을 수 있었다. 즉 각각의 작업당 175개 정도의 명령 샘플들을 얻었고 이를 이용하여 여섯개의 전이확률표를 얻었다. 표 1은 작업 3을 위한 전이확률을 보여준다.

위와 같은 실험을 통해 구현된 아이콘 언어 해석기를 검증하기 위하여 두 가지의 테스트를 수행하였다. 하나는 학습한 자료의 부분을 이용한 작업 할당 테스트, 또 하나는 학습하지 않은 아이콘 문장들을 이용한 작업 할당 테스트 이다. 첫 번째 테스트의 수행 결과는 표 2에 표시된 컨퓨전 행렬(confusion matrix)을 통해 확인할 수 있다.

표 1. 작업 3과 관련된 천이확률표.

Table 1. Transition Probabilities of task 3 model.

		i^{th} state					
		S1	S2	S3	S4	...	End
$i-1^{th}$ state	Start	0.536	0.050	0.028	0.000	...	0.000
	S1*	0.000	0.016	0.000	0.000	...	0.056
	S2	0.253	0.000	0.000	0.000	...	0.019
	S3	0.167	0.000	0.000	0.000	...	0.000
	S4	0.000	0.000	0.000	0.000	...	1.000
	S5	0.007	0.000	0.000	0.035	...	0.852
	:	:	:	:	:	:	:
	S10	1.000	0.000	0.000	0.000	...	0.000

표 2. 컴퓨터전 행렬: 학습에 참여한 그룹의 아이콘 문장 입력시의 작업 할당 성공률.

Table 2. Confusion matrix: Success rate of interpretation of the subject in learned group.

		인식결과						계	성공율
		작업 1	작업 2	작업 3	작업 4	작업 5	작업 6		
사용자의도	Task1	74	5	-	1	-	-	80	92.5
	Task2	-	74	5	-	-	1	80	92.5
	Task3	-	1	74	-	-	5	80	92.5
	Task4	5	-	-	75	-	-	80	93.8
	Task5	-	-	-	-	80	-	80	100.0
	Task6	-	2	-	-	-	78	80	97.5

표 3. 컴퓨터전 행렬: 학습에 참여치 않은 그룹의 아이콘 문장 입력시의 작업 할당 성공률.

Table 3. Confusion matrix: Success rate of interpretation of the subject in un-learned group.

		인식결과						계	성공율
		작업 1	작업 2	작업 3	작업 4	작업 5	작업 6		
사용자의도	Task1	54	1	-	1	-	1	56	96.4
	Task2	-	51	-	-	-	5	56	91.1
	Task3	-	2	52	-	-	2	56	92.9
	Task4	2	-	-	50	2	1	55	90.9
	Task5	-	-	-	-	37	1	38	97.4
	Task6	-	1	3	-	-	34	38	89.5

두번째 테스트를 수행하기 위해서는 새로운 데이터를 필요로 한다. 새로운 데이터는 앞서 수행한 것과 유사한 실험을 통해 획득하였다. 8명의 피실험자 중 2명이 여성이고 평균 나이는 23.8세였다. 5명은 로봇과 관련된 경험이 전혀 없었다. 표 3을 보면 확인 할 수 있지만 학습된 그룹의 입력과 학습의 그룹의 입력에 모두 높은 비율로 정확하게 작업을 할당할 수 있음을 확인하였다.

이 같은 결과들은 제안된 해석방법이 사용자의 의도를 정확하게 추정할 수 있음을 보이고 있다. 예를 들면 부엌에 있는 바나나를 가져오는 작업에 대해 피실험자들은 다음과 같

은 아이콘 문장들을 사용하였다. “부엌 이동 식탁 바나나 잡기 돌아오기 놓기”, “바나나 잡기 돌아오기”, 또는 “식탁 이동 바나나 잡기 돌아오기”. 이처럼 다양한 표현들은 제안된 인식기를 통해 모두 올바르게 인식되었다. 하지만 너무 불충분한 입력에 대해서는 잘못 해석하는 경우도 있었다. 콜라를 가져오도록 하는 작업에 대해 한 피실험자가 “잡기 콜라”라는 아이콘 문장을 사용하였고 다른 작업으로 인식되었다. 하지만 이 같은 문장은 사람의 경우도 원래의 의도를 바르게 인식할 수 없을 것이다.

이와 같은 결과들은 아이콘 언어가 음성언어의 사용 없이도 상호작용을 통해 로봇에게 특정한 작업을 전달하기에 쉽고 충분한 방법임을 잘 보여주고 있다.

V. 핸디밋의 감정 상호작용 채널

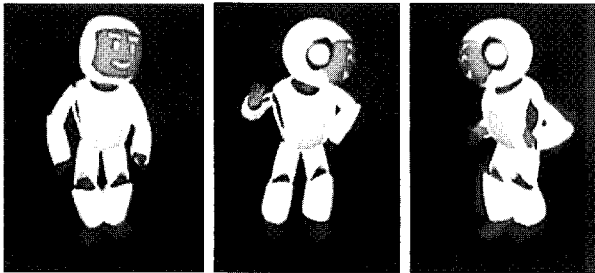
핸디밋은 작업중심의 상호작용 채널과 함께 사용자가 로봇과 감정적으로 상호작용 할 수 있는 채널을 갖고 있다. 감정 상호작용은 앞서 언급한 바와 같이 사용자와 로봇의 친밀도를 향상시키고 작업의 효율을 높일 수 있는 중요한 요소라 하겠다. 핸디밋은 사용자의 감정을 인식하는 인터페이스로 표정 인식기와 음성 감정 인식기를 갖고 있고, 로봇의 감정을 표현하는 인터페이스로 아바타를 가지고 있다. 본 논문에서는 핸디밋의 감정 상호작용 채널을 위해 개발된 표정 인식기와 음성 감정 인식기, 감정 표현 아바타에 대해 간단하게 소개하도록 하겠다.

핸디밋에서 사용된 표정 인식기는 Bien, Lee 등에 의해 개발되었다[16,17]. 특정 사용자가 들고 다니며 사용하는 핸디밋의 특징으로 인해 핸디밋에서의 표정인식기는 소수의 사용자들에 대해서만 동작하면 되며, 핸디밋상의 카메라와 사용자의 얼굴 사이의 거리의 변화가 작은 편이다. 인식 과정은 세단계로 이루어진다. 첫번째 단계에서는 아다부스트 기반의 하르(Adaboost-based Haar) 방법에 의해 얼굴을 찾고, 다음으로는 소벨엣지(Sobel edge) 방법을 이용하여 몇가지의 특징점을 찾는다. 마지막으로 가버 웨이브렛 신경망(Gabor wavelet neural networks)을 이용하여 얼굴 표정을 분류한다. 핸디밋에서 사용된 얼굴 인식기는 초당 5회이상의 표정인식 성능을 갖고 있어 얼굴 표정을 연속적으로 인식할 수 있다.

핸디밋의 음성감정 인식기는 Kim, Hyun, Kwak 등에 의해 개발되었다[18]. 표정인식기의 경우와 마찬가지로 핸디밋상에서의 음성감정인식기도 특정 사용자에게 한정되어 사용되는 이점을 가지며, 추가적으로 사용자가 핸디밋을 들고 다니며 사용하기 때문에 화자와 마이크 사이가 가까운 이점도 있다. 사용자의 음성에서 드러나는 감정은 두가지로 나눌 수 있는데 문맥상에 존재하는 감정과 음성자체에 내포된 감정상태이다. 이 중에서 핸디밋의 음성감정 인식기는 대화의 문맥과 상관없이 소리의 특징만을 이용한다. 음성의 특징 중의 하나인 주파수 특성을 살펴보면 사람의 음성의 주 대역은 100~5000Hz 사이이지만 정상시의 목소리와 화가 났을 때의 목소리는 그 주 대역이 상이하여 화가 났을 경우 정상시보다 고음을 내게 된다. 이와 같이 이 범위는 사람의 감정 상태에 따라 변화하게 된다. 핸디밋의 음성감정 인식기는 감정에 따른 주파수 대역의 차이를 이용하여 감정을 인식한다. 인식기



(a) Facial expressions of emotional avatar



(b) Gesture expressions of emotional avatar

그림 6. 핸디봇상의 아바타를 이용한 다양한 감정표현의 예, Lee, Kim 등의 연구에서 인용[19].

Fig. 6. Various emotional expressions of avatar on Handybot, referred from the Lee, Kim et al's paper[19].



그림 7. 핸디봇을 통한 로봇과의 감정 상호작용 데모; 데모 중에는 핸디봇상에 아이콘 언어 인터페이스 대신 OX퀴즈용 GUI가 보여졌음.

Fig. 7. Emotional interaction demo using Handybot; OX quiz GUI was displayed during the demonstration instead of Icon-language interface.

의 성능은 사람이 다른 사람의 음성감정을 78%의 정확도로 인식할 때 인식기는 76%의 정확도로 인식할 수 있었다.

핸디봇 상의 표정인식기와 음성감정인식기에서 인식된 결

과는 각각 평상, 기쁨, 슬픔, 화남의 4가지 차원을 갖는 감정 상태 벡터로 출력된다. 핸디봇에서는 다시 이 두 인식기의 인식 결과를 벡터합한 결과에서 가장 높은 확률 값을 갖는 감정 상태를 사용자의 감정상태로 결정하여 로봇에게 보내게 된다.

로봇이 갖는 감정상태 역시 핸디봇에서 표현된다. 이를 위해 감정표현 아바타가 Lee, Kim 등에 의해 개발되었다[19]. 핸디봇에서의 아바타는 표정 및 제스처를 이용하여 다섯가지의 감정상태를 세가지 단계로 표현할 수 있다. 그림 6은 아바타의 감정표현 예를 보여주고 있다.

감정 인식기들과 감정 아바타로 이루어진 감정 상호작용 채널은 사용자 데모를 통해 검증되었다. 사용자들과 로봇이 OX퀴즈를 푸는 과정에서 핸디봇을 통해 인식된 사용자의 감정은 로봇으로 전달되었고 로봇의 감정은 핸디봇상의 아바타를 통해 표현되었다. 그림 7은 데모 장면들을 보여주고 있다. 데모를 통해 핸디봇의 감정 상호작용 채널이 사용자와 로봇간의 감정상호작용에 효과적으로 사용될 수 있음을 확인하였다.

VI. 결론

로봇과 사용자가 지속적인 상호작용을 하도록 돕기 위해 핸디봇이 개발되었다. 핸디봇은 다양한 인터페이스들을 갖고 있으며 이러한 인터페이스들은 작업중심의 상호작용 채널과 감정 상호작용 채널로 분류할 수 있다.

먼저 작업중심의 상호작용 채널을 보면 음성언어를 보조하거나 때에 따라 대체할 수 있는 인터페이스로서 아이콘 언어가 개발되었다. 아이콘 언어는 로봇이 로봇의 배경지식을 사용자와 공유하고 공유된 정보들을 이용하여 상호작용할 수 있도록 설계되었다. 또한 아이콘언어는 마크로프 프로세스 기반의 해석기를 사용하여 사용자가 큰 자유도를 가지고 로봇에게 입력을 수행할 수 있도록 한다. 또한 아이콘 언어의 유용성은 실험을 통해 확인되었다.

핸디봇의 감정 상호작용 채널은 사용자와 로봇이 멀리 떨어져 있을 때에도 로봇이 사용자의 감정을 이해하고 자신의 감정을 표현할 수 있도록 도와주도록 설계되었으며 사용자 데모를 통해 감정 상호작용을 효과적으로 매개할 수 있음이 확인 되었다.

매개인터페이스 핸디봇은 작업중심의 상호작용 중에 로봇과 사용자간의 거리 문제를 해결하기 위한 기본 목적에 더하여 아이콘 언어나 감정 상호작용 채널과 같은 추가적인 인터페이스들을 통하여 거리에 상관없이 사용자와 로봇이 편하고, 지속적이고, 풍부한 상호작용을 할 수 있도록 한다. 추후 과제로는 다양한 연령층과 다양한 부류의 사용자들에 대해 핸디봇의 사용성을 조사하고 개선하는 과정을 통해 핸디봇의 유용성을 높이고자 한다.

참고문헌

[1] T. Fong, "Collaborative control: A robot-centric model for vehicle teleoperation," Robotics Institute, Doctor of Philosophy, 2001.
 [2] H. K. Keskinpala, J. A. Adams, and K. Kawamura, "PDA-based human-robotic interface," in the proceedings of IEEE

- International Conference on Systems, Man and Cybernetics*, pp. 3931-3936, 2003.
- [3] D. Perzanowski, A. C. Schultz, W. Adams, E. Marsh, and M. Bugajska, "Building a multimodal human-robot interface," *IEEE Journal on Intelligent Systems*, vol. 16, pp. 16-21, 2001.
- [4] G. Chronis and M. Skubic, "Sketch-based navigation for mobile robots," in the proceedings of *The 12th IEEE International Conference on Fuzzy Systems*, St. Louis, pp. 284-289, 2003.
- [5] R. A. Bolt, "Put-that-there: voice and gesture at the graphics interface," in the proceedings of *The 7th annual conference on Computer graphics and interactive techniques*, Seattle, pp. 262-270, 1980.
- [6] R. Bischoff and V. Graefe, "Dependable multimodal communication and Interaction with robotic assistants," in the proceedings of *IEEE International Workshop on Robot and Human Interactive Communication*, Berlin, pp. 300-305, 2002.
- [7] J. J. Pfeiffer, "Altaira: A rule-based visual language for small mobile robots," *Journal of Visual Languages and Computing*, vol. 9, pp. 127-150, 1998.
- [8] P. T. Cox, C. C. Risley, and T. J. Smedley, "Toward concrete representation in visual languages for robot control," *Journal of Visual Languages and Computing*, vol. 9, pp. 211-239, 1998.
- [9] L. Leifer, M. Van der Loos, and D. Lees, "Visual language programming for robot command/control in unstructured environments," in the proceedings of *Fifth International Conference on Advanced Robotics*, Pisa, Italy, pp. 31-36, 1991.
- [10] S. Hamad, "The symbol grounding problem," *Physica D: Nonlinear Phenomena*, vol. 42, pp. 335-346, 1990.
- [11] T. Nomura, T. Kanda, and T. Suzuki, "Experimental investigation into influence of negative attitudes toward robots on human-robot interaction," in the proceedings of *The 3rd Workshop on Social Intelligence Design*, pp. 125-135, 2004.
- [12] T. Mitsui, T. Shibata, K. Wada, A. Touda, and K. Tanie, "Psychophysiological effects by interaction with mental commit robot," in the proceedings of *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Maui, pp. 1189-1194, 2001.
- [13] H. Hüttenrauch, A. Green, M. Norman, L. Oestreicher, and K. S. Eklundh, "Involving users in the design of a mobile office robot," *IEEE Transactions on Systems, Man and Cybernetics, Part C, Applications and Reviews*, vol. 30, pp. 113-124, 2004.
- [14] R. Gockley, J. Forlizzi, and R. Simmons, "Interactions with a moody robot," in the proceedings of *Human-Robot Interaction*, Salt Lake City, Utah, pp. 186-193, 2006.
- [15] J. D. Williams and S. Young, "Partially observable markov decision processes for spoken dialog systems," *Computer Speech and Language*, vol. 21, pp. 393-422, 2007.
- [16] S. W. Lee, Kim, D.-J., Kim, Y. S., & Bien, Z., "Training of feature extractor via new cluster validity-Application to adaptive facial expression recognition," *Lecture Notes in Computer Science*, vol. 3684, pp. 542-548, 2005.
- [17] D.-S. Kwon, W.-C. Yoon, Z. Bien, Y. K. Kwak, K.-H. Hong, K.-S. Park, J.-Y. Yang, J.-H. Hwang, K.-B. Lee, J.-C. Park, Y.-C. Kim, Y.-J. Kwon, S.-W. Lee, and K.-H. Hyun, "Multimodal cognitive and emotional interaction module for human robot interaction," *Presented at The 5th Technical Workshop of Center for Intelligent Robotics*, 2006.
- [18] E. H. Kim, K. H. Hyun, and Y. K. Kwak, "Robust emotion recognition feature, frequency range of meaningful signal," in the proceedings of *IEEE International Workshop on Robot and Human Interactive Communication*, Nashville, pp. 667-671, 2005.
- [19] D.-S. Kwon, J.-H. Hwang, S.-H. Park, Z. Bien, D.-J. Kim, S.-W. Lee, Y. J. Kwon, K.-H. Hyun, Y. K. Kwak, G. H. Lee, and J.-H. Kim, "Mediated interface technologies using multi-modality," *Presented at The Technical Workshop of Center for Intelligent Robotics*, 2004.



황정훈

1997년 연세대학교 기계공학과(공학사).
1999년 카이스트 기계공학과(공학석사).
2007년 카이스트 기계공학과(공학박사).
2007년~현재 카이스트 인간-로봇 상호
작용 핵심연구센터 박사후연구원. 관심
분야는 인간-로봇 상호작용, 인공지능.



권동수

1980년 서울대학교 기계공학과(공학사).
1982년 카이스트 기계공학과(공학석사).
1991년 Georgia Tech 기계공학과(공학박
사). 1991년~1995년 미국 Oak Ridge국립
연구소 선임연구원. 1995년~현재 카이
스트 기계공학과 조교수, 부교수, 교수.
2003년~현재 카이스트 인간-로봇 핵심연구센터 소장. 관심분
야는 인간-로봇 상호작용, 햅틱스, 의료용 로봇.