

# HRI 중심의 로봇 콘텐츠와 소프트웨어 통합 기술

박광현 (광운대학교)

## I. 서론

만들면 팔리는 시대에서 기능이나 성능만을 중요시하는 시대를 거쳐 최근에는 사용자의 감성에 맞는 제품의 시대로 변함에 따라 점점 감성 공학을 비롯한 휴먼기술이 중요한 시대가 되었다<sup>[1,2]</sup>. 즉 생산성과 신뢰성, 기능성을 중요시하는 양의 시대에서 편리성, 쾌적성, 다양성을 중요시하는 질의 시대로 변화하고 있다. 이는 다른 말로, 하드웨어 중심의 개발에서, 소프트웨어 중심으로, 다시 콘텐츠 중심으로 개발 중심이 이동하고 있다는 것을 뜻한다. 모바일 폰이 단말기 개발에서 다양한 기능 개발로, 이제는 모바일 콘텐츠 개발에 힘을 쏟고 있듯이 말이다.

우리가 다른 사람들과 많은 시간을 보내는 과정에서 서로의 의견을 나누다거나 자신의 감정을 전달함에 있어서 상호작용 방식은 매우 중요하다. 심리학에서는 얼굴을 마주보고 하는 상면 대화에서 표정이 매우 중요하다고 알려져 있다<sup>[3-5]</sup>. Mehrabian은 메시지의 단지 7%만 순수한 언어(Linguistic Language)에 의해 전달되며 신체 움직임이나 목소리와 같은 준언어(Paralanguage)에 의해서는 38%, 얼굴 표정에 의해서는 55%의

메시지가 전달된다고 보고한 바 있다<sup>[3,5]</sup>. 따라서 서비스 로봇의 개발에 있어서도 사람이 사용하기에 편리한 인간 친화적 인터페이스에 대한 중요성이 높아지면서 다양한 연구가 진행되고 있으며, 특히 사람이 일상생활에서 사용하는 음성, 제스처, 표정 등 시각, 청각을 중심으로 하는 인터페이스가 활발히 연구되고 있다. 이와 같은 인간-로봇 상호작용의 궁극적인 목적은 사람과 사람이 대화하듯이 사람이 로봇을 다루는 것이며, 인간의 의도를 어떻게 측정하고 인식하느냐가 가장 어려운 문제로 부각되고 있다<sup>[1]</sup>.

그러면, 서비스 로봇에서 중요한 요소기술 중의 하나인 인간-로봇 상호작용(HRI) 기술이 개발되면, 우리 주변이 각종 로봇으로 둘러싸여 부르지만 하면 나타나 심부름을 해줄 것인가? 사람들을 안내하고 가사를 돕고 때로는 친구처럼 대화와 놀이의 상대가 되며, 건강 상태를 자동으로 체크하면서 우리를 돌봐줄 수 있을까? 그렇다고 대답하기는 어려울 것이다. 왜냐하면 우선, 로봇은 다른 시스템과는 달리 사용되는 하드웨어 플랫폼이 장난감부터 복잡한 휴머노이드까지 다양성의 범위가 매우 넓기 때문에, 하나의 로봇을 위해 개발된 HRI 모듈을 그대로 다른 로봇에 활용

하기가 쉽지 않다. 또한, 모바일 폰에서 음성 인식이나 얼굴 인식이 된다고 하여도 이를 사람들에게 편리함과 서비스를 제공하기 위해서는 다양한 모바일 콘텐츠가 필요하듯이, 진정한 서비스 로봇을 위해서는 HRI 모듈을 기반으로 사용자에게 필요한 로봇 콘텐츠를 서비스할 수 있어야 한다. 예를 들어, 교육용 로봇에서는 HRI 기술을 활용하여 학생의 영어 음성을 인식할 수 있지만, 정답을 얘기했을 때는 칭찬을 하고 오답을 말했을 때는 격려를 하며 전체 교육 과정을 이끌어 갈 수 있는 교육용 로봇 콘텐츠가 필요하고, 유아를 위한 상호교감형 로봇 연극에서는 유아의 몸짓과 소리를 인식할 수 있지만, 연극을 진행하기 위한 로봇 콘텐츠가 필요하다.

서비스 로봇에 있어서 HRI의 중요성은 의심의 여지가 없지만, HRI 기술을 보다 잘 활용하기 위해서는 이를 뒷받침할 수 있는 로봇 콘텐츠 기술도 수반되어야 하며, 개발된 HRI 모듈과 로봇 콘텐츠가 다양한 로봇 플랫폼에 서비스될 수 있도록 하는 소프트웨어 통합 기술도 중요해지고 있다. 본 원고에서는 이를 위해 개발된 로보이드 스튜디오에 대해 살펴보고 HRI 모듈의 활용에 대

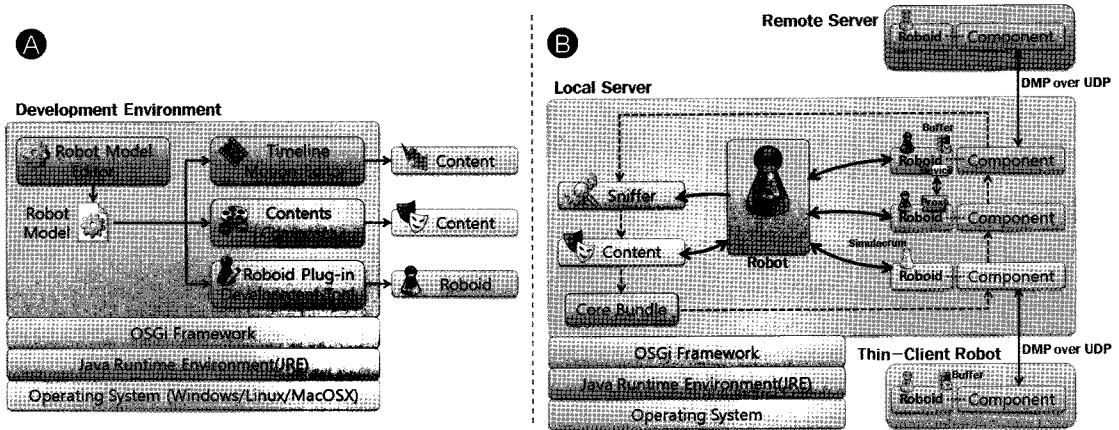
해 알아보려고 한다.

## II. 로보이드 스튜디오

로보이드 스튜디오는 나만의 로봇 콘텐츠를 손쉽게 제작할 수 있는 도구이다. 누구라도 30분 정도면 나를 위해 웃고, 말하고, 춤추는 로봇을 만들어 볼 수 있다. 내 목소리로 말하고 상대방의 반응에 따라 행동하는 나의 분신 같은 로봇을 만들고 메일을 통해 로봇 콘텐츠를 보내고 받을 수 있다<sup>[6]</sup>.

이 원천기술은 표준 로봇제어 프로토콜을 수용한다. 인터넷이 이처럼 활성화된 것도 웹 브라우저와 통신하는 표준적인 프로토콜이 존재하기 때문이다. 다만 로봇제어 프로토콜은 이러한 데이터 전송기술에 비해, 실시간성, 병렬처리성, 신뢰성 등에서 한 차원 높은 기술이다. 이 프로토콜을 지키는 로봇, 즉 로보이드라면 마치 휴대폰을 눌러 친구와 통화하듯이 전 세계 어디에 존재하더라도 원하는 동작을 지시할 수 있다.

로보이드 스튜디오는 로봇을 공부하지 않은



〈그림 1〉 로보이드 프레임워크의 전체 구성. A) 개발환경, B) 런타임 환경

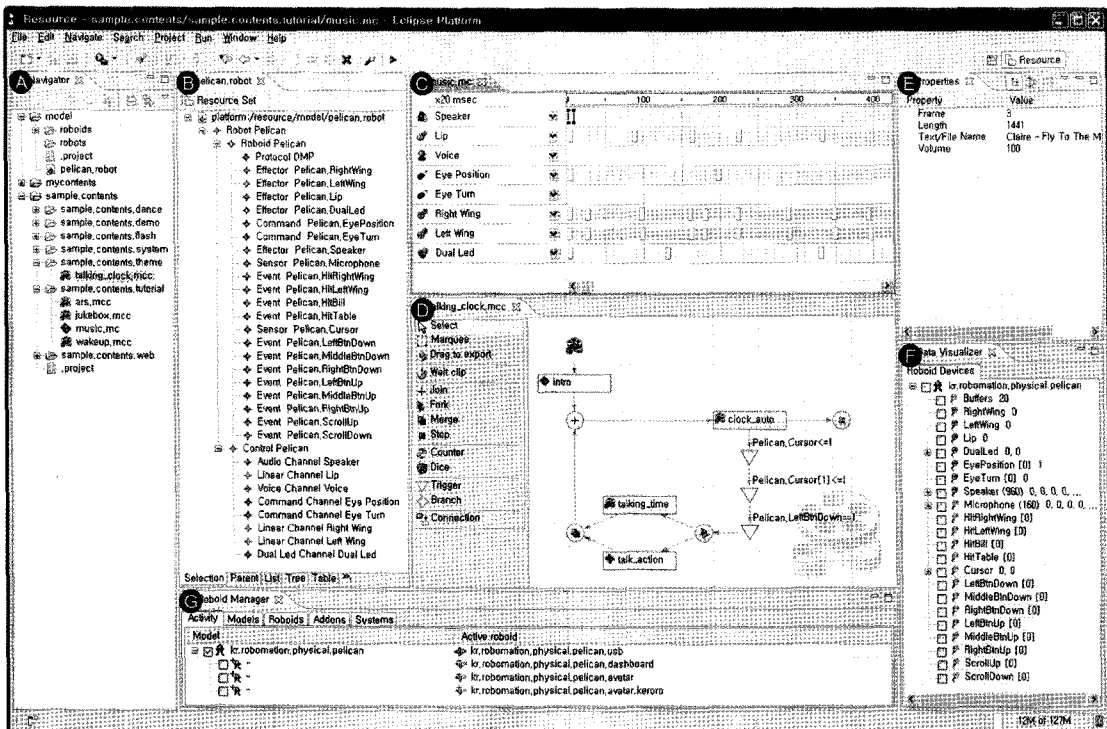
사람은 물론 프로그래밍을 잘 모르는 일반인도 쉽게 사용할 수 있다. 시간 축을 따라 동작을 지시하는 편집 도구(타임라인 모션 에디터)와 이러한 동작을 모아 사용자와 상호작용 하면서 상황에 따라 지능적으로 행동하도록 하는 저작 도구(콘텐츠 컴포저)가 있으며, 로봇 연구자를 위한 로봇 모델링 도구 및 프로그래밍 통합 개발 환경까지 모두 갖추고 있다.

이클립스를 기반으로 만들어진 로보이드 스튜디오는 2008년 5월에 첫 번째 버전이 공개되었으며, 현재 버전 1.3까지 발표되었다(www.roboidstudio.org). 로봇 관련 업체, 교육업체뿐만 아니라 일반인을 포함하여 국내외에서 사용되고 있으며, 새로운 기술을 계속 추가하여 발전

시켜 나가고 있다. <그림 1>은 로보이드 프레임워크의 전체 구성을 나타낸 것이며, 로보이드 스튜디오의 전체 화면은 <그림 2>와 같다.

## 1. 타임라인 모션 에디터 (Timeline Motion Editor)

타임라인 모션 에디터는 로봇에 포함된 모든 디바이스의 동작과 멀티미디어 데이터를 동기화하여 로봇의 기본 동작을 만들기 위한 그래픽 도구이다. 단순히 마우스 클릭과 드래깅, 프로퍼티 뷰의 값을 변경함으로써 원하는 동작을 추가할 수 있으며, 로봇의 동작과 멀티미디어 데이터가 정확히 언제 발생해야 하는지 쉽게 설정할 수 있



<그림 2> 로보이드 스튜디오의 전체 화면. A) 로봇 콘텐츠 파일 리스트, B) 로봇 모델 에디터, C) 타임라인 모션 에디터, D) 콘텐츠 컴포저, E) 프로퍼티 뷰, F) 시뮬레이션 정보, G) 로보이드 매니저

다. <그림 2> c에서 수평축은 1/50초 간격의 프레임 수를 나타내며, 수직축은 로봇 모델에서 설정한 콘트롤 채널이다. 각 콘트롤 채널 내에 있는 사각형은 콘트롤 채널의 종류에 따라 키 프레임을 나타내거나 멀티미디어 데이터 혹은 동작 데이터의 시작 프레임을 나타낸다. 매 1/50초마다 로봇 동작과 오디오 데이터를 같이 보내기 위해 오디오 데이터는 960 바이트로 나누어지고 동작 데이터는 키 프레임 간에 보간법으로 만들어진다.

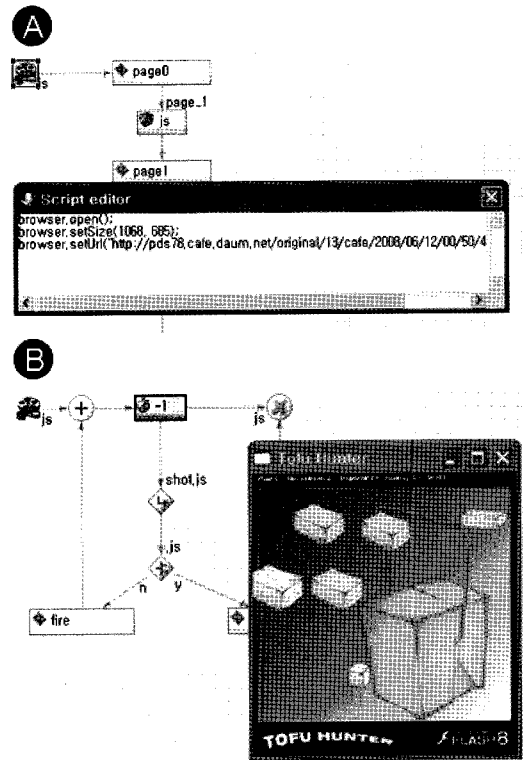
## 2. 콘텐츠 컴포저 (Contents Composer)

콘텐츠 컴포저는 타임라인 모션 에디터와 달리 사용자의 행동이나 조건에 따라 반응하는 동적인 로봇 콘텐츠를 제작하기 위한 그래픽 도구이다. 콘텐츠 컴포저는 타임라인 모션 에디터에서 만든 모션 클립이나 콘텐츠 컴포저에 의해 만들어진 다른 모션 콘텐츠를 포함할 수 있으며, 브랜치, 트리거 등의 다양한 구성요소들과 연결하여 로봇 콘텐츠를 구성할 수 있다.

일반적으로 로봇 콘텐츠나 태스크는 유한상태 기계(Finite State Machine) 혹은 플로우 차트로 표현이 된다. 콘텐츠 컴포저는 이 두 가지를 동시에 표현할 수 있어 거의 모든 콘텐츠 혹은 태스크를 기술할 수 있으며, 최소한의 필요한 구성요소만 채택함으로써 사용법을 배우는 시간과 작업 시간을 최소화할 수 있도록 하였다.

## 3. 로보이드 스크립트

콘텐츠 컴포저의 각 구성요소 중 조건을 설정할 수 있는 브랜치와 트리거는 로봇 모델에 포함된 디바이스에 제한되어 있다. 따라서 “오전 9 시일 때” 혹은 “마우스 클릭이 홀수 번일 때”,

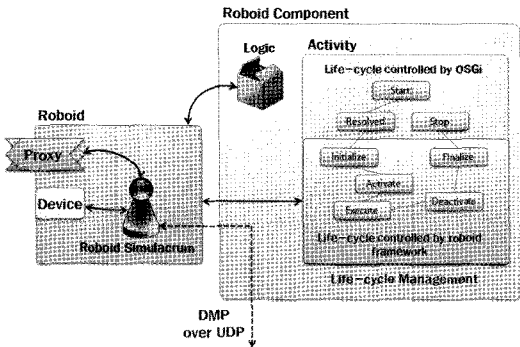


<그림 3> 로보이드 스크립트의 사용 예  
A) 웹 브라우저 연동, B) 플래시 연동

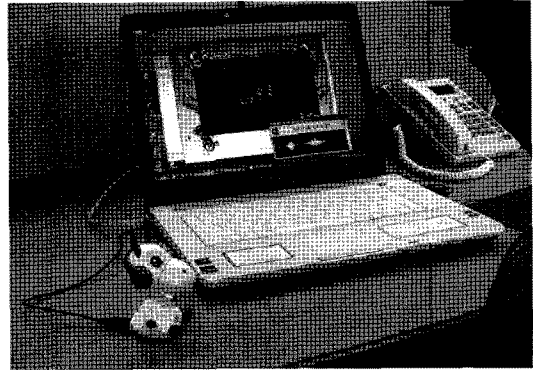
“사람이 카메라 화면에 나타났을 때”와 같이 일반화된 조건을 주기 위해 자바 스크립트를 확장하여 로보이드 스크립트를 사용할 수 있도록 하였다. 또한, 스크립트를 통해 그래픽 인터페이스를 구성하거나 웹 브라우저 및 플래시와 연동할 수 있도록 함으로써 보다 다양한 로봇 콘텐츠를 제작할 수 있는 환경을 제공한다. <그림 3>

## 4. 로보이드 컴포넌트

로보이드는 하드웨어에서 시작한 개념이지만, 로보이드 스튜디오에서는 이 개념을 소프트웨어 디바이스와 알고리즘, 통신 채널을 포함하는 소프트웨어 컴포넌트로 확장하여 사용한다. 로보이



〈그림 4〉 로보이드 컴포넌트의 구조



〈그림 5〉 음성 인식을 이용한 영어 단어 학습

드 컴포넌트는 마치 하드웨어 컴포넌트와 같이 하나의 독립된 기능을 수행하는 단위이며, 다른 컴포넌트와 프록시 디바이스(Proxy Device)를 통해 서로 연결되어 데이터를 교환한다. 〈그림 4〉는 로보이드 컴포넌트의 구성을 나타내며, 로보이드 스튜디오의 코어 번들(Core Bundle)이 각 로보이드 컴포넌트의 라이프 사이클(Life-Cycle)을 관리하여 초기화, 활성화, 실행, 비활성화, 종료를 수행한다.

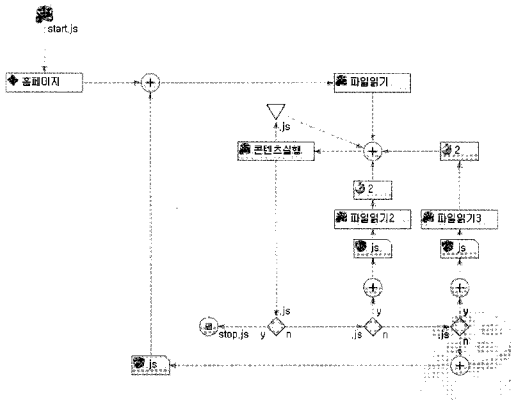
지능형 서비스 로봇에서는 제공하는 서비스의 종류와 성능이 중요한데, 이를 위해 로보이드 스튜디오에서는 버전 1.3에서 33개의 로보이드 컴포넌트를 제공하며, 얼굴 인식, 음성 인식, 웹 서비스, 정보처리 등 보다 다양한 로보이드 컴포넌트가 제작되어 계속 추가되고 있다.

### III. 로보이드 스튜디오의 HRI 컴포넌트 활용 예

#### 1. 음성 인식을 이용한 영어 단어 학습

〈그림 5〉는 유아 및 초등학교 저학년을 대상

으로 영어 단어를 교육하기 위해 개발된 시스템이다. 화면에 한글로 된 단어가 표시되었을 때, 그림 좌하단의 동키 로보이드를 통해 해당하는 영어 단어를 말하면, 음성을 인식하여 정답인 경우에는 칭찬하는 말과 함께 기뻐하는 동작을, 오답인 경우에는 격려하는 음성과 함께 아쉬워하는 동작을 로보이드가 표현하도록 구성되어 있다. 이와 같은 시스템을 기존의 방식대로 개발하고자 한다면, 음성 인식 모듈을 새로 개발하거나 이 시스템에 맞도록 수정하여야 하며, 화면 GUI를 만드는 것부터 음성 인식 모듈과 연동하는 부분, 음성 합성 모듈, 로봇의 동작을 디자인하는 작업, 동작을 표현하기 위해 제어하는 부분을 모두 개발하여야 하기 때문에 개발 기간도 오래 걸릴 뿐만 아니라, 음성 인식 전문가, 음성 합성 전문가, 화면 GUI 개발자, 시나리오 개발자, 콘텐츠 개발자, 로봇 전문가, 소프트웨어 전문가 등 많은 사람들이 모여 협력을 하여야 한다. 하지만, 로보이드 스튜디오를 사용하면 컴포넌트로 제공되는 음성 인식 모듈과 타임라인 모션 에디터, 콘텐츠 킴포우저를 사용하여 공학을 전공한 사람이 아니더라도 순수한 콘텐츠 개발자 혼자 〈그림 6〉과 같이 콘텐츠를 구성하여 전체 시스템을 개

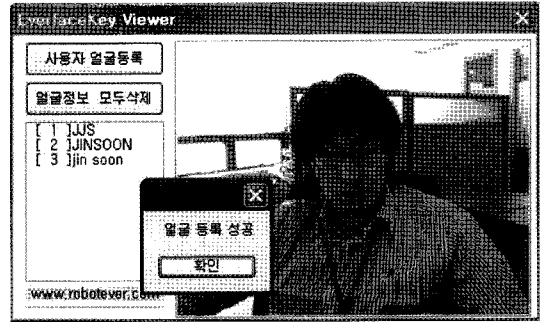


〈그림 6〉 영어 단어 학습용 콘텐츠

말할 수 있다. 또한, 이미 만들어진 콘텐츠에서 로봇의 음성이나 동작을 변경하거나 더 좋은 성능의 음성 인식 컴포넌트로 교체하는 일이 굉장히 빠르고 쉽게 이루어질 수 있다. 따라서 진정으로 콘텐츠 중심의 개발을 할 수 있는 시대가 열린 것이며, 고부가가치 산업인 콘텐츠 산업으로 넘어갈 수 있는 발판이 마련된 것이다.

## 2. 얼굴 인식

〈그림 7〉은 외부에서 개발한 얼굴 인식 컴포넌트를 로보이드 스튜디오에 추가하여 활용하는 예이다. PC에 카메라를 연결하고, 로보이드 스튜디오에서 제공하는 카메라 컴포넌트와 얼굴 인식 컴포넌트, 콘텐츠 컴포우저를 사용하여 콘텐츠를 구성하기만 하면 간단하게 로봇이 나를 알아보고 인사하도록 하거나 뉴스 혹은 날씨 정보를 알려주도록 할 수 있다. 물론 사용자가 GUI 화면을 구성할 필요도 없다. 사용자는 단순히 자신이 구성하고자 하는 시나리오를 그래픽 도구를 이용하여 표현하기만 하면 된다. 기존의 개발 방식에서는 이러한 서비스를 제공하기 위해 각



〈그림 7〉 얼굴인식 컴포넌트

분야의 전문가들이 모여 정해진 시나리오대로 동작하는 로봇을 만들어 사용자에게 제공하였기 때문에 콘텐츠 측면에서는 금방 싫증이 나고 쉽게 관심을 잃어버려서 로봇의 활용도가 낮았다. 하지만 로보이드 스튜디오를 사용하면 사용자가 원하는 시나리오를 쉽게 구성할 수 있기 때문에 지속적으로 새로운 서비스가 가능하다. 내가 집에 들어갈 때 사랑하는 애인의 목소리로 나를 맞이하고, 잠을 잘 때 옆에서 자장가를 불러주는 나만의 로봇을 상상해 보라. 어찌 사랑스럽지 않겠는가?

## IV. 결론

로보이드 기술은 현대 공학의 집합체이다. 로봇 공학을 위시하여 정보통신 기술, 소프트웨어 엔지니어링 기술 등이 집약되었고, 인터넷과 웹 기술이 접목되었으며, 나아가서 문화, 예술, 교육 콘텐츠 개발과 로봇 교육에 적합하도록 설계되었다.

로보이드 스튜디오는 로봇마다 다르게 만들어진 기존의 로봇 개발환경과 달리 동일한 환경에서 간단히 로봇 모델을 교체함으로써 어떤 종류

의 로봇도 지원할 수 있다. 사실 모든 로봇을 지원하는 운영체제를 만든다는 것은 수많은 종류의 로봇을 고려해 볼 때 매우 어려운 일이다. 하지만 웹 서비스가 그러했듯이 공통점을 찾아 표준화하고자 한다면 로봇 발전에 획기적인 전기가 될 수도 있을 것이다. 이 실마리를 로보이드 스튜디오에서 찾아볼 수 있다.

### 참고문헌

- [1] 이순요, 양선모, 감성공학, 청문각, 1996
- [2] 변중남, 박광현, 방원철, “사용자 의도 파악 III - 얼굴 표정을 인식하는 시스템,” 자동화기술, 2003년 2월호.
- [3] Albert Mehrabian, “Communication without words,” Communication: Concepts and Process(Editor: Joseph A. DeVito), pp. 106-114, Prentice-Hall, 1971.
- [4] Albert Mehrabian and James A. Russell, An Approach to Environmental Psychology, MIT Press, 1974.
- [5] F. Hara and H. Kobayashi, “Use of Face Robot for Human-Computer Communication,” IEEE Conference on SMC, pp.1515-1520, 1995.
- [6] 김경진, 박광현, “로봇 소프트웨어 개발 도구 - (5) 로보이드 스튜디오,” 제어·로봇·시스템학회지, 2009년 3월호.

### 저자소개



박 광 현

1994년 2월 KAIST 전자전산학과 학사  
 1997년 2월 KAIST 전자전산학과 석사  
 2001년 2월 KAIST 전자전산학과 박사  
 2001년 3월~2003년 3월 KAIST 인간친화 복지 로봇시스템 연구센터 박사후 연구원  
 2003년 4월~2004년 2월 KAIST 전자전산학과 연구교수  
 2004년 3월~2005년 3월 미국 조지아공과대학교 방문 연구원  
 2005년 4월~2008년 2월 KAIST 전자전산학과 초빙교수  
 2008년 3월~현재 광운대학교 정보제어공학과 조교수  
 주관심 분야 : 지능형 서비스 로봇, 인간-로봇 상호작용, 로봇 소프트웨어, 로봇 콘텐츠