

# 실외에서 로봇의 인간 탐지 및 행위 학습을 위한 멀티모달센서 시스템 및 데이터베이스 구축

엄태영<sup>\*</sup>, 박정우<sup>\*\*</sup>, 이종득<sup>\*\*</sup>, 배기덕<sup>\*\*</sup>, 최영호<sup>\*\*</sup>

## Multi-modal Sensor System and Database for Human Detection and Activity Learning of Robot in Outdoor

Taeyoung Uhm<sup>\*</sup>, Jeong-Woo Park<sup>\*\*</sup>, Jong-Deuk Lee<sup>\*\*</sup>, Gi-Deok. Bae<sup>\*\*</sup>, Young-Ho Choi<sup>\*\*</sup>

### ABSTRACT

Robots which detect human and recognize action are important factors for human interaction, and many researches have been conducted. Recently, deep learning technology has developed and learning based robot's technology is a major research area. These studies require a database to learn and evaluate for intelligent human perception. In this paper, we propose a multi-modal sensor-based image database condition considering the security task by analyzing the image database to detect the person in the outdoor environment and to recognize the behavior during the running of the robot.

**Key words:** Robot Database, Human Activity Learning, Outdoor Database, Human Detection

### 1. 서 론

로봇이 영상으로부터 객체를 찾고 인식하는 연구는 최근 로봇비전 분야의 주요 연구이다. 특히, 실외 환경에서 로봇의 주요한 인식 대상은 사람이며 행위를 인식하기 위한 연구가 많이 수행되고 있다. 이러한 로봇의 사람을 찾고 행위 인식을 위한 연구는 주로 학습에 의한 방법을 사용하기 때문에 최근에는 이를 위한 다양한 데이터베이스에 관심이 모아지고 있다.

사람을 찾고 행위 인식을 위한 데이터베이스는 대표적인 PETS[1]를 비롯하여 BEHAVED[2], i-LIDS

[3], ViSOR[4]와 같이 객체 및 사람 추적 연구와 관련하여 제공하고 있다. 각각의 데이터베이스는 실내외 환경에서의 사람과 자동차 혹은 바닥에 방치된 가방 등의 특정 상황인식 등의 목적을 가지고 구축되었다. 이러한 데이터베이스들은 해마다 많은 학회와 워크숍을 통해 발표하고 공개되어 꾸준히 질과 양을 늘려가고 있다. 하지만 대부분의 데이터베이스들은 고정된 카메라에서 취득한 영상을 주로 제공하고, 움직이는 카메라에서도 흔들림을 포함한 영상을 제공한다 [5,6]. 또한 컬러 영상 혹은 컬러와 깊이를 동시에 제공하지만 그 외의 다른 종류의 영상들을 포함하지는 않는다[7]. 한편, 데이터를 웹에서 수집하기 위한 연

\* Corresponding Author: Taeyoung Uhm, Address: (37666) Jigok-RO 39, Nam-Gu, Pohang, Gyeongbuk, Korea, TEL: +82-54-279-0451, FAX: +82-54-279-0449, E-mail: uty@kiro.re.kr

Receipt date: Oct. 30, 2018, Revision date: Dec. 5, 2018  
Approval date: Dec. 10, 2018

<sup>†</sup> Field Robotics R&D Division, Korean Institute of Robot and Convergence

<sup>\*\*</sup> Field Robotics R&D Division, Korean Institute of Robot and Convergence

(E-mail: idealcreator@kiro.re.kr)

(E-mail: artofgene@kiro.re.kr)

(E-mail: bgd9047@kiro.re.kr)

(E-mail: rockboy@kiro.re.kr)

\* This research was supported by the ICT R&D program of IITP, Development of multimodal sensor-based intelligent systems for outdoor surveillance robots (2017-0-00306).

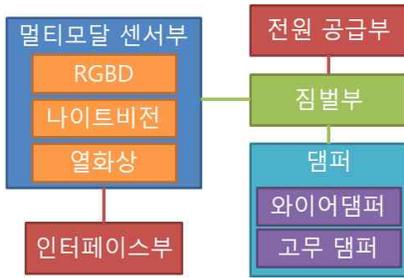


Fig. 1. Multi-modal sensor system configuration.

구[8]도 있지만 실외환경과 로봇을 위한 데이터베이스를 특정하기는 쉽지 않다.

본 논문에서는 실외환경에서 다양한 날씨 및 밤낮의 환경에 대한 영상을 제공하기 위해 멀티모달센서 모듈을 설계 및 제작하고 이를 이용하여 다중의 사람을 찾고 행위를 인식할 수 있는 데이터베이스를 구축하여 소개한다. 또한 이를 위해 고려된 사항을 기술한다.

본 논문의 구성은 다음과 같다. 2장에서 데이터베이스 구축을 위해 설계 및 제작된 멀티모달센서 모듈을 소개하고 3장에서 로봇에 모듈을 탑재하여 데이터를 취득한 방법 및 고려 사항을 소개한다. 그리고 4장에서는 획득한 데이터베이스를 보이고 5장에서 결론을 맺는다.

## 2. 멀티모달센서 모듈

### 2.1 멀티모달센서 시스템 구성

멀티모달센서 시스템의 구성은 아래의 Fig. 1과 같다. 위에서 설명한 멀티모달센서부를 고주파 진동 감쇄를 위해 짐벌부에 연결하고 이를 저주파 진동 감쇄를 위해 댐퍼에 연결한다. 짐벌을 위한 전원공급부와 영상데이터 처리를 위한 인터페이스부를 포함한다.

### 2.2 멀티모달센서 모듈 설계

멀티모달센서 모듈의 조건

- 낮과 밤에도 사용이 가능한 다중 영상을 출력해야한다.
- 실외환경(눈/비 포함)에서 사용이 가능해야한다.
- 로봇의 주행 중에도 카메라의 흔들림을 최소화해야한다.

위의 조건에 맞게 설계된 1차 멀티모달센서 모듈의 인터페이스는 아래의 Fig. 1과 같다. 모듈은 RGB-D 카메라, 나이트비전 카메라 2개, 열화상 카메라를 포함하고 POE 인터페이스와 USB 3.0을 사용하여 카메라의 전원 및 데이터를 처리한다. 한편, 나이트비전 카메라는 12VDC 전원이 추가로 필요하다. <Fig. 2

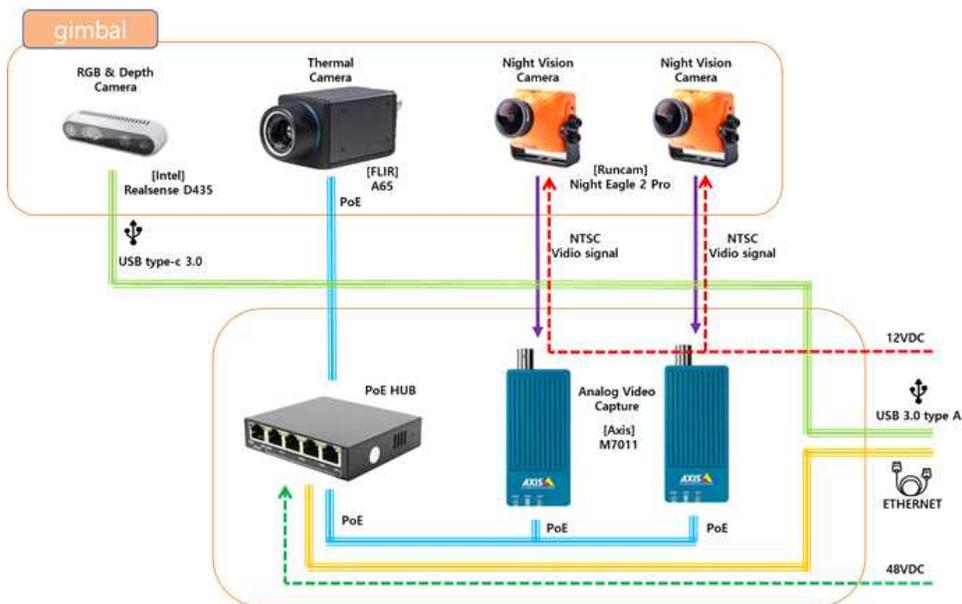


Fig. 2. Multi-modal sensor module interface.

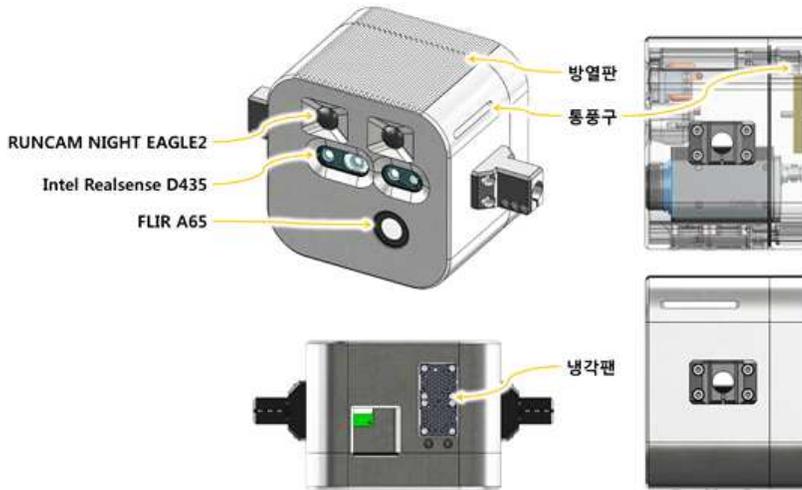


Fig. 3. Multi-modal sensor module Head design.

참조>

Fig. 3에서와 같이 센서모듈의 헤드는 4개의 카메라를 포함한 발열을 위한 공기 순환과 보호유리를 사용한 방수를 고려하여 총 1.3kg의 무게로 설계되었다. 각각의 카메라들은 방수용 실을 사용하였고, 하단의 공기 유입과 각도가 있는 공기 순환 설계로 발열에 대비하였다. 즉, 멀티모달센서 헤드 내부에는 여러 종의 카메라와 기타 전장 모듈이 배치되기 때문에 발열이 심하게 발생하고, 발열이 심할 때 제대로 열 배출을 하지 못할 경우 카메라가 오동작 하거나 영구적인 손상을 일으켜 센서의 기능을 제대로 수행하지 못할 수 있다. 더욱이 야외에서 장시간 구동해야 하는 로봇에 장착하여 사용할 경우 직사광선을 직접적으로 받는 상황을 가정하면 오동작 및 작동 불능의 가능성이 매우 높아지므로 카메라에서 발생하는 열을 제대로 방출하기 위해 내부 공기순환을 통한 통풍 구조를 헤드 설계에 적용하였다. 우선 헤드의 뒷 커버 및 앞 커버의 양 측면에 통풍을 위한 홀을 적용하였다. 이때 우천 시 수직방향으로 떨어지는 빗물이 쉽게 유입되지 못하도록 모든 홀은 대각선 방향으로 아래를 향하도록 설계하고 헤드의 하단에 냉각용 팬을 배치하여 상대적으로 차가운 외부의 공기를 내부로 끌어들이어 헤드 내부의 공기가 강제적으로 순환할 수 있도록 하였다.

또한 마지막 조건인 카메라 흔들림 보정을 위해서 상용 짐벌(MOZA LTE 2)을 활용하여 저주파 진동을 감쇄하고 댐퍼를 활용하여 고주파 진동을 감쇄하는 장치를 설계하였다. <Fig. 4 참조> 짐벌의 수용하

중은 4.8 kg이며 댐퍼는 개당 최대 3.6 kg의 수용하중을 갖고 있다. 이렇게 제작된 최종 시스템을 멀티모달센서를 포함한 모듈화하여 다양한 로봇 플랫폼에 이식 가능하도록 하였다.

위와 같은 장치들을 통합하여 최종 구성된 멀티모달센서 모듈은 Fig. 5와 같이 설계되었다. 이 모듈은

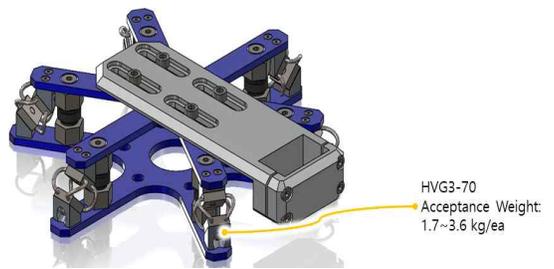


Fig. 4. Wire based vibration damper.

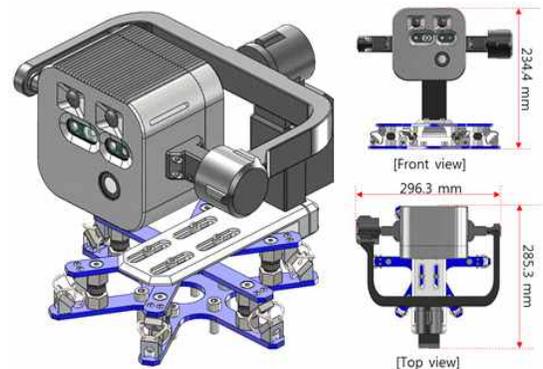


Fig. 5. Wire based vibration damper.

방수가 고려되었으며, 4개의 카메라에서 4종의 영상 취득이 실시간으로 가능하다. 멀티모달센서 모듈은 최소 285mm\*296mm\*319mm 크기의 공간과 짐벌 구동을 위한 배터리 탈부착 공간을 차지한다. 또한 고정을 위해 밑면에 6개의 볼트 홈을 사용하여 로봇 플랫폼에 장착한다. <Fig. 6 참조>

이렇게 로봇 플랫폼에 장착된 모듈은 진동을 극복한 영상 데이터를 제공한다. 즉, 과속방지턱과 같이 언덕을 오르내리는 저주파 진동과 일반 도로 주행에서의 고주파 진동을 보정하는 효과를 보인다. 이러한 영상들은 로봇의 학습을 위한 목적으로 획득하고, 동기화되어 데이터베이스화 된다. Fig. 7에서는 스테레오 깊이 카메라(Realsense D435, Intel), 나이트 비전 카메라(Night Eagle 2 pro, Runcam)와 열화상 카메라(A65, FLIR)에 의해 동기화된 영상 데이터 예시를 보여준다. 깊이 영상은 32bit의 영상을 표현하였다.

2.3 멀티모달센서 모듈 진동 보정 효과

멀티모달센서 모듈 각각의 센서는 캘리브레이션 방법[9]을 이용하여 보정되며 보정된 영상은 제안된



Fig. 6. The multi-modal sensor module mounted on robot platform.

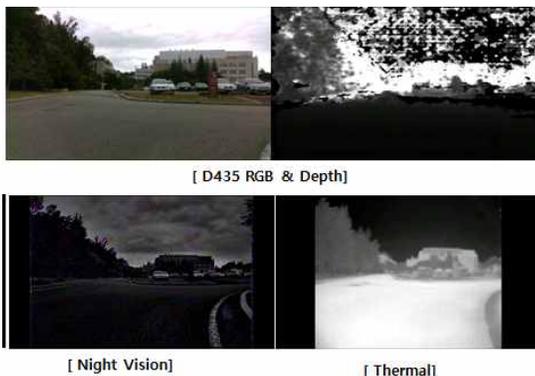


Fig. 7. An example of multi-modal sensor data.

짐벌과 댐퍼 장치를 이용하여 진동에 강한 결과를 보인다. Fig. 8과 같이 일반포장도로 및 자갈밭 주행에서 저주파 및 고주파 진동을 감쇄한다.

3. 사람 추출 및 추적을 위한 데이터베이스

3.1 데이터베이스 구축 고려사항

로봇이 실외환경에서 사람을 찾고 추적하는 알고리즘 학습을 위해서는 실외 환경의 영향을 최소화하기 위해서 다양한 배경을 갖는 반복된 동작을 기본으로 영상 데이터베이스를 구축하였다. 이를 위해 다음과 같은 요소를 고려하였다.

- 전문 배우의 고정된 동작별 반복 시연
- 다수 촬영에 의한 장소 변화
- 배우의 움직임과 로봇의 움직임을 동시 발생에 의한 배경 변화
- 맑음/비에 의한 날씨변화
- 주간/야간에 의한 일조량 변화
- 로봇과 감시 카메라 시점에 따른 변화
- 한 장의 이미지에는 한 명의 사람만 등장 한다.



(a)



(b)



(c)

Fig. 8. Result of vibration damping. (a) Original, (b) Proposed method, (c) Experimental Platform

Table 1. The behavior definition in normal and abnormal situations

	Algorithm	
	Normal situation	Abnormal situation
Behaviors	Lying down	Sit down and fall
	Standing	Standing and falling
	Walking	Walk down
	Running	Run down
	Sitting on a chair / floor	Lie seizure
	Holding a cup / bag / umbrella	-

3.2 데이터베이스 구성

위와 같은 조건으로 현재까지 획득한 학습용 데이터베이스는 다음의 Table 1에서와 같이 정상상황과 이상상황을 구분하여 실외환경에서 발생 가능한 행위를 정의하고 이를 기반으로 데이터베이스를 구축하였다. 각 데이터는 총 4회의 촬영으로 구축되었으며, 각각의 목적에 따라 장소와 시간을 정하여 촬영하였다. 그 구성은 Table 2와 같다. 이후에는 RGB영상을 우선하여 태깅작업을 수행하고 있다. (태깅 완료 후 공개 예정)

3.3 데이터베이스 사례

아래는 각기 다른 환경에서 획득한 데이터베이스

사례를 보인다.

Fig. 9와 같이 첫 데이터는 동일한 장소에서 로봇의 시점과 2.5m의 CCTV 시점에 해당하는 DB를 오전부터 야간까지 시간별 5회, 액터 4명의 고정된 동작 시퀀스를 촬영하였다.

두 번째 데이터는 산업단지에서 로봇의 주행 중 10명의 액터들이 다양한 자세로 누워있는 DB를 3회의 시퀀스로 촬영하였다. <Fig. 10 참조>

세 번째 데이터는 산업단지에서 로봇의 시점과 2.5 m cctv의 시점으로 4명의 액터들이 고정된 동작 DB를 9회의 시퀀스로 촬영하였다. 같은 환경에서 맑은 날, 천둥 및 비오는 날, 흐린 날의 날씨 환경을 포함하였다. <Fig. 11 참조> 마지막으로 공원의 다양한 배경에서 촬영하여 데이터를 획득하였다. 로봇의 시점과 2.5m cctv의 시점으로 4명의 액터들의 고정된 동작을 촬영하였다. <Fig. 12 참조> 촬영된 영상 중 선별 후 20만장 태깅을 수행 중이다. 현재까지 수집된 데이터의 태깅이 수행된 후 휴먼 디택션 학습을 위해 사용될 예정이다.

4. 결 론

최근 로봇에 대한 관심이 많아지고 있으며, 로봇을 학습시키기 위한 데이터베이스에 대한 구축 사례도 폭발적으로 증가하고 있다. 본 논문에서는 이러한 관심에 적합한 영상 데이터베이스 구축용 멀티모달

Table 2. The database configuration

Take	Purpose		Conditions	
	Situation	Learning	Time / Weather	Place
1	Normal & Abnormal	Detection / Recognition	Day&Night/Sun	Park
2	Abnormal	Detection	Day/Sun	Driveway
3	Normal & Abnormal	Detection / Recognition	Day/Rain	Driveway
4	Normal & Abnormal	Detection / Recognition	Day/Sun	Park



Fig. 9. Take 1: Day and Night time database in the lawn playground.



Fig. 10. Take 2: Lying database in the driveway.



Fig. 11. Take 3: Heavy rain and thunder in the driveway.

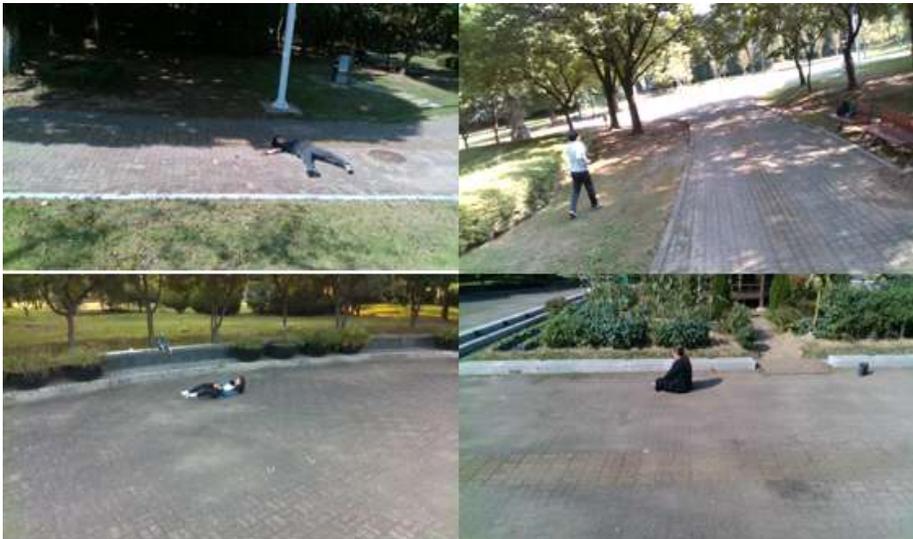


Fig. 12. Take 4: Daytime in the park.

센서 모듈을 설계하고 로봇에 인식 가능하도록 제작하였으며, 이를 이용하여 사람의 추출 및 추적에 적합한 영상 데이터를 획득하였다. 특히, 실외환경에서 로봇의 학습을 위해 효율적으로 데이터베이스를 구축하기 위한 요소들을 분석하여 고려하였다. 현재까지 구축된 데이터베이스는 개발된 알고리즘의 학습

및 테스트용으로 사용되며, 향후에는 초상권 등의 협의 후 공개될 예정이다. 이는 국내 로봇 및 감시 시스템의 학습에 활용될 뿐만 아니라 관련 연구 영역을 활성화 시켜줄 것으로 기대된다. 더욱이, 국내 학습용 데이터베이스가 체계적으로 다양한 목적에 맞도록 구축이 이루어져야 할 것이다.

## REFERENCE

- [ 1 ] PETS, <https://motchallenge.net/workshops/bmtt-pets2017/> (accessed June, 12, 2018).
- [ 2 ] BEHAVE, <http://groups.inf.ed.ac.uk/vision/BEHAVEDATA/INTERACTIONS/> (accessed June, 22, 2018).
- [ 3 ] i-Lids, [www.ilids.co.uk](http://www.ilids.co.uk). (accessed July, 2, 2018).
- [ 4 ] ViSOR, <http://imagelab.ing.unimore.it/visor/> (accessed July, 9, 2018).
- [ 5 ] G. Moon and J. Rue, "Remote Person Recognition Test Database for Intelligent Video Surveillance," *Journal of Information Security*, Vol. 22, No. 4, pp. 38-45, 2012.
- [ 6 ] R. Fenrich and J.J. Hull, "Concern in Creation of Image Database," *Proceedings of Third International Workshop on Frontiers in Handwriting Recognition*, pp. 112-121, 1993.
- [ 7 ] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, "Places: A 10 Million Image Database for Scene Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 40, No. 6, pp. 1452-1464, 2018.
- [ 8 ] S. Bae, H. Lee, and D. Cho, "Design and Implementation of a Web Crawler System for Collection of Structured and Unstructured Data," *Journal of Korea Multimedia Society*, Vol. 21, No. 2, pp. 199-209, 2018.
- [ 9 ] T. Uhm, G. Bae, J. Lee, and Y. Choi, "Multi-modal Sensor Calibration Method for Intelligent Unmanned Outdoor Security Robot," *Proceedings of the Sixth International Conference on Green and Human Information Technology*, pp. 215-220, 2018.



엄 태 영

1997년 3월~2004년 2월 경기대학교 전자공학과 학사  
2004년 3월~2006년 2월 한양대학교 전자통신컴퓨터공학과 석사  
2006년 9월~2014년 8월 한양대학교 전자통신컴퓨터공학과 박사

2014년 9월~2016년 4월 한국과학기술연구원 박사 후 과정  
2016년 5월~2017년 4월 한국과학기술연구원 위촉연구원  
2017년 5월~현재 한국로봇융합연구원 필드로봇연구본부 선임연구원

관심분야: 영상처리, 휴먼로봇인터랙션, 인공지능, 자율주행, 증강현실



박 정 우

2001년 3월~2008년 2월 부경대학교 지능기계공학과 학사  
2011년 3월~2013년 2월 부산대학교 제어시스템공학과 석사  
2008년 11월~현재 한국로봇융합연구원 필드로봇연구본부 선임연구원

관심분야: 기계 설계, 로봇 제어, 유압 시스템



이 종 득

2002년 3월~2009년 2월 대구대학교 전자시스템공학과 학사  
2015년 3월~현재 경북대학교 대학원 전자공학부 석사과정  
2010년 6월~현재 한국로봇융합연구원 필드로봇연구본부 주임연구원

관심분야: 모바일 로봇, 임베디드 제어 시스템



배 기 덕

2009년 3월~2015년 2월 서울과학기술대 기계시스템 디자인공학과 학사  
2015년 3월~2017년 2월 서울과학기술대 기계설계로봇공학과 석사

2017년 4월~현재 한국로봇융합연구원 필드로봇연구본부 주임연구원

관심분야: 영상처리, 소프트웨어 개발, 지능제어, 자율주행



최 영 호

1996년 3월~2003년 2월 한양대학교 전기전자공학과 학사  
2003년 3월~2008년 2월 포항공과대학교 지능로봇공학 박사  
2008년 2월~현재 한국로봇융합연구원 필드로봇연구본부 책임연구원

관심분야: 지능제어, 인공지능, 자율주행, 무인시스템