

# 기계 장치와의 상호작용을 위한 실시간 저비용 손동작 제어 시스템

## A Real Time Low-Cost Hand Gesture Control System for Interaction with Mechanical Device

황 태 훈\*, 김 진 현\*

Tae-Hoon Hwang\*, Jin-Heon Kim\*

### Abstract

Recently, a system that supports efficient interaction, a human machine interface (HMI), has become a hot topic. In this paper, we propose a new real time low-cost hand gesture control system as one of vehicle interaction methods. In order to reduce computation time, depth information was acquired using a time-of-flight (TOF) camera because it requires a large amount of computation when detecting hand regions using an RGB camera. In addition, fourier descriptor were used to reduce the learning model. Since the Fourier descriptor uses only a small number of points in the whole image, it is possible to miniaturize the learning model. In order to evaluate the performance of the proposed technique, we compared the speeds of desktop and raspberry pi2. Experimental results show that performance difference between small embedded and desktop is not significant. In the gesture recognition experiment, the recognition rate of 95.16% is confirmed.

### 요 약

최근에, 효율적인 상호작용을 지원하는 시스템 인 휴먼 머신 인터페이스(HMI)가 인기를 끌고있다. 본 논문에서는 차량 상호작용방법 중 하나로 새로운 실시간 저비용 손동작 제어 시스템을 제안한다. 계산 시간을 줄이기 위해 RGB 카메라를 사용하여 손 영역을 감지할 때 많은 계산이 필요하므로 TOF (Time-of-Flight) 카메라를 사용하여 깊이 정보를 취득한다. 또한, 푸리에 기술자를 사용하여 학습 모델을 줄였다. 푸리에 디스크립터는 전체 이미지에서 적은 수의 포인트만 사용하므로 학습 모델을 소형화 할 수 있다. 제안 된 기법의 성능을 평가하기 위해 데스크탑과 라즈베리 pi 2의 속도를 비교했다. 실험 결과에 따르면 소형 임베디드와 데스크탑의 성능 차이는 크지 않다. 제스처 인식 실험에서 95.16 %의 인식률이 확인되었다.

*Key words : Gesture Control, Fourier Descriptor, TOF Camera, Real-time, Hand Gesture*

---

\* Bachelor, Dept. of Computer Engineering, Seokyeong University

★ Corresponding author

E-mail : jinheon@skuniv.ac.kr, Tel : +82-2-940-7747

※ Acknowledgment

This research is support by Ministry of Culture, Sport and Tourism(MCST) and Korea Creative Content Agenct(KOCCA) in the Culture Technology(CT) Research & Development Program 2019

Manuscript received Dec. 9, 2019; revised Dec. 27, 2019; accepted Dec. 30, 2019.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## I. 서론

최근 음성인식, 동작 인식과 같은 첨단기술이 차량 내 장치에 적용되면서 차량 내 장치와 운전자 간의 효율적이고 직관적인 상호작용 방법에 대한 연구 관심이 증대되고 있다. 종래의 연구는 스위치나 버튼 등을 운전자가 접근하기 쉽도록 배치하고 디자인하는 연구가 진행되어왔다[1]. 하지만 HMI에 점점 더 많은 기능이 추가되고 통합됨에 따라 장치 사용의 복잡성이 증가하고 있다[2]. 장치 사용의 복잡성을 단순화하는 접근법으로 인간의 서로 다른 감각들을 직관적이고 효율적으로 사용하는 것이 중요해지고 있으며, 그 방법으로는 필기인식, 영상인식, 음성인식 등이 있다. 그 중에서도 영상인식 방법은 손동작과 같이 운전자의 움직임 인식을 하여 차량 내 장치를 제어한다.

차량 내 영상인식 방법과 관련되어 손의 형태인식, 움직임 추적과 같은 연구가 활발하게 진행되고 있다[3]. Khurana etc. [4]은 영상의 형태를 인식하기 위한 방안으로 Zernike Moment와 손 모양의 기하학적 특징을 결합하였다. 19개의 손 모양에 대한 공분산으로 구성된 특징 세트의 최소 유클리드 거리가 분류에 사용되었다. 앞서 말한 방법을 사용하여 정적 이미지에서 98.74%의 인식률이 달성되었다. 그러나 이 방법은 유사한 형태의 손모양은 고려하지 않았다. Agarwal etc. [5]은 HOG(Histogram of Oriented Gradients)와 함께 기하학적 특징을 사용하였으며 손 모양을 인식하는데 SIFT(Scale Invariant Feature Transform)기능을 사용하였다. 이는 손 모양 분류 시 다중 클래스 SVM(Support Vector Machine)을 사용하여 92.6%의 인식률을 달성하였다. Gamal etc. [6]는 푸리에 기술자의 저주파 성분에 객체의 대략적인 형태가 있는 특징 이용하여 적은 특징만으로 학습 모델을 소형화하고 연산 속도를 개선하였다[7-9].

하지만 이러한 접근법들은 모두 비주얼 영상에서 처리하였는데, 비주얼 영상은 손 영역 검출 시 많은 연산량을 필요로 한다. 연산량이 많은 알고리즘은 차량 내 상호작용 장치와 같이 사용자가 볼 수 없도록 설계되어야 하는 소형 기계장치에서 실시간으로 구현하는데 어려움이 있다. Kyu etc. [10]는 TOF 카메라에서 취득한 저 해상도 깊이 영상을 이용하여 적은 연산만으로 손 영역을 추출하였

다[11]. 이는 비주얼 영상을 사용 시 많은 연산량으로 인한 성능 저하문제를 깊이 영상을 이용함으로써 해결하였다[12-15].

본 논문에서는 손동작 인식 시 필요한 모든 과정을 단순화 하여 소형 임베디드에서도 실시간으로 구현 가능하게 하고자 한다. 손 영역 추출 과정을 TOF 카메라를 이용하여 연산을 단순화 하며, 손 영역의 특징점을 푸리에 기술자를 이용하여 추출하였다. 푸리에 기술자는 전체 영상 중 일부 좌표만으로 손 모양 학습이 가능하므로 학습 모델의 크기를 최소화 할 수 있다. 제안한 알고리즘의 유효성을 평가하기 위하여 알고리즘의 실험을 데스크톱과 라즈베리 파이2 두 환경에서 수행하였다. 이는 하드웨어 성능차이가 큰 두 환경에서 소모되는 자원의 차이가 크지 않음을 보이기 위함이다. 그리고 알고리즘이 소형화됨에 따라 발생하는 인식률 저하문제를 실험을 통하여 확인하였다. 실험 결과 제안한 방법의 인식률이 기존의 방법과 크게 차이가 나지 않음을 확인하고 수행시간 또한 크게 차이가 나지 않음을 확인하였다.

본 논문의 구성은 다음과 같다. 2장에서 본 논문의 기본 이론이 되는 푸리에 기술자에 대하여 정의한 다음, 3장에서 제안하는 동작 인식 알고리즘을 설명한다. 여기에서는 제안한 알고리즘의 전체적인 개요와 손 영역 추출, 푸리에 기술자를 이용한 손 특징 추출, 그리고 손 모양 및 움직임 판별의 순서로 서술한다. 그리고 4장에서는 제안한 알고리즘으로 실험한 결과를 하드웨어 성능이 다른 두 환경에서 동작하여 성능을 평가하고 5장에서 결론을 맺는다.

## II. 본론

### 1. 이론

푸리에 기술자는 객체의 외곽선 정보를 이용하여 모양을 기술하는 모양 기술자 중 하나이다. 푸리에 기술자는 위치, 배율, 회전 및 시작점에 관계없이 폐곡선을 나타내는데 사용할 수 있다. 객체의 외곽선 좌표 중에서 임의의 시작점을 기준으로 연결되어있는 연속된 외곽선의 좌표를  $n$  실수부는 허수부로 취하는 복소 함수는 다음과 같이 정의된다.

$$z(k) = x(k) + jy(k), k = 0,1,2,\dots,K-1 \quad (1)$$

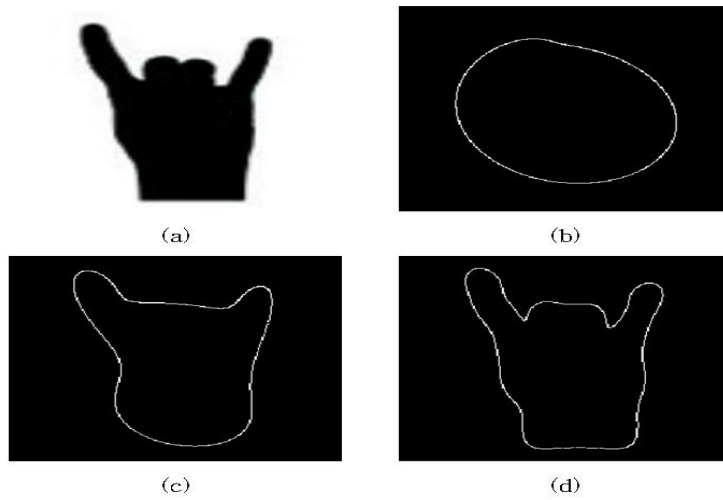


Fig. 1. Hand shape according to the low frequency components when fourier descriptor(FD) reconstruction (a) Original Image, (b) FD = 5, (c) FD = 10, (d) FD = 30.

그림 1. 푸리에 기술자 (FD) 복원시 저주파 성분에 따른 손 모양 (a) Original Image, (b) FD = 5, (c) FD = 10, (d) FD = 30

위의 식 (1)은 2차원 공간상의 픽셀 좌표를 1차원 복소수 함수의 형태로 표현한 것이다. 여기서 K는 외곽선 점의 개수를 나타낸다. 이를 이산 푸리에 변환하면 식 2와 같은 주파수 공간의 함수로 변환된다.

$$Z(u) = \sum_{k=0}^{K-1} z(k) \exp(-\frac{j2\pi uk}{K}) \quad (2)$$

식 (2)와 같이 만들어진 주파수 공간의 1차원 복소 함수  $Z(u)$ 를 푸리에 기술자라 한다. 일반적으로 푸리에 기술자는 저주파 성분에 원본 함수의 대략적인 모양 정보를 포함하고 있다. 그림 1은 전체 푸리에 기술자 중 각각 0.5%, 1% 3%를 사용하여 재구성한 그림이다.

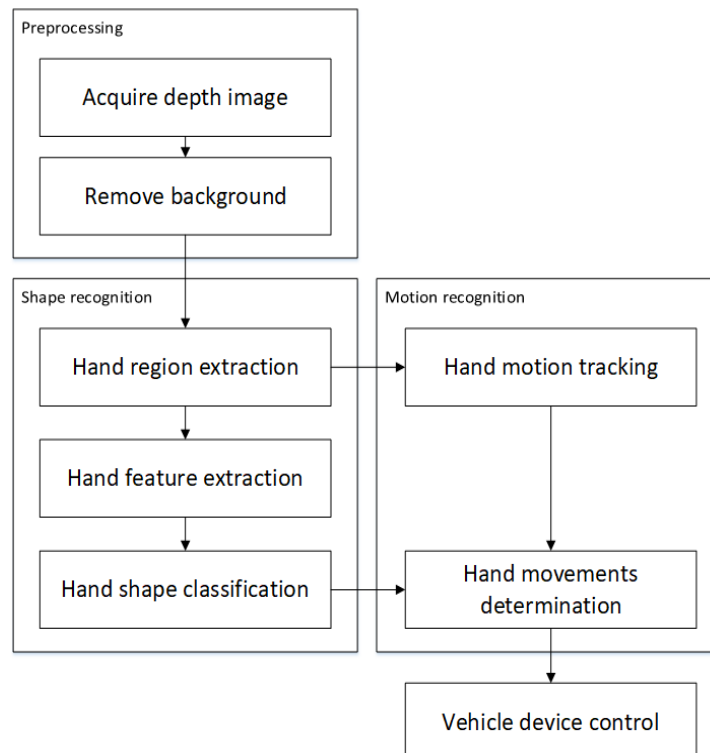


Fig. 2. Flowchart of proposed hand gesture control system.

그림 2. 제안된 손 제스처 제어 시스템의 흐름도

2. 제안한 방법

(1) 제안한 시스템의 개요

본 논문에서는 차량 내 장치를 손동작으로 제어하기 위한 깊이 정보 기반의 저비용 손동작 인식 시스템을 제안한다. 그림 2에 제안하는 손동작 인식 시스템의 전체적인 흐름도를 나타내었다.

먼저, 취득한 깊이 영상에서 손의 특징을 추출하기 위한 전 처리과정을 거친다. 취득한 깊이 영상에서 일정 범위를 넘어서는 값은 배경이라 판단하고 제거한다. 배경이 제거된 손 영역의 모양을 인식하기 위한 방안으로 손 모양의 외곽선을 이용한 푸리에 기술자를 사용한다. 푸리에 기술자로 계산된 손 모양에 SVM알고리즘을 적용시켜 손의 형태를 학습시킨다. SVM의 결과를 바탕으로 손의 모양을 결정하며, 일정 구간 동일한 손 모양이 출력되면 동작의 시작점이라고 판단한다. 동작이 시작되면 앞서 추출한 손 영역을 이용하여 손의 움직임을 계산한다. 손의 움직임은 동작이 시작된 시점을

기준으로 일정 구간의 정보를 수집하여 어떠한 동작인지를 판단한다. 손의 형태와 손의 움직임을 판별하는 두 가지 학습모델의 결과를 바탕으로 손의 동작을 판별 한다.

(2) 배경 제거




손 영역을 추출하기 위한 전 처리 과정으로 입력 받은 깊이 영상의 배경을 제거한다. 계산의 단순화를 위하여 TOF 카메라와 가장 가까운 거리에 위치한 물체를 손으로 가정한다. 입력받은 깊이 영상의 잡음 제거를 위하여 메디안 필터(median filter)를 적용한 후 깊이 값이 작은 점들 중 중심에 가까운 점을 기준으로 하여 플러드 필(floodfill)알고리즘을 수행한다. 대략적인 손의 영역이 검출되면 잡음으로 인한 구멍을 채우기 위하여 모폴로지(morphology) 열기, 닫기 연산을 수행한다. 깊이 영상으로부터 추출된 손 영역의 영상은 그림 3과 같다.



Fig 3. Hand region extracted through preprocessing. (a) Input depth image, (b) Extracted hand region  
 그림 3. 전처리를 통해 추출된 손 영역 (a) 입력 깊이 이미지, (b) 추출된 손 영역

Table 1. Shape recognition rate by hand gesture.

표 1. 손동작에 의한 형상 인식율

|                  | Rock  | Scissor  | Paper   |
|------------------|---|--|---|
| shape            |  |  |  |
| Recognition rate | 83.6%   | 97.2%  | 95.3%   |

(3) 손 모양 인식

배경이 제거된 깊이 영상의 손 영역에서 손의 외곽선을 이용한 푸리에 기술자로 손의 특징을 추출한다. 학습 모델의 소형화를 위하여 계산된 전체 푸리에 기술자 중 일부 저주파 성분만을 사용하여 학습하였으며, 학습 모델은 패턴 인식, 자료 분석 시 보편적으로 사용되는 SVM을 사용한다. SVM에서 나오는 결과를 바탕으로 현재 손 모양을 결정하며, 이 손 모양이 일정시간 동안 지속적으로 출력되면 동작의 시작점이라고 판단한다.

(4) 손동작 판별

손동작은 입력 영상에서 배경이 제거된 손 영역을 바탕으로 손 영역의 전체 평균이 다음 프레임(frame)에 어떠한 방향으로 이동하였는지를 계산한다. 손 영역의 이동은 크기가 1로 미리 정규화한 8방향 단위 벡터를 이용하여 손의 이동을 재정의 하였다. 손의 전체적인 움직임은 동작의 시작점을 기준으로 일정 구간 동안의 움직임을 SVM알고리즘으로 학습하며, 이 결과와 손 모양을 종합하여 어떠한 손동작을 하는지 판별한다.

는 30fps로 처리하였다. 실험의 검증을 위한 손 모양은 주먹, 가위, 보 세 가지를 사용하였으며, 손의 움직임은 상, 하, 좌, 우, 회전 등 총 6가지를 사용하여 인식률을 확인하였다.

표 1은 본 논문에서 제안한 시스템의 손 모양 인식률을 나타낸 표이다. 각 손 모양을 학습시키기 위하여 20개의 푸리에 디스크립터를 사용하였다. 실험 방법은 각 손 모양 영상을 1000회 반복 실험을 통하여 인식률을 산출하였다. 실제 손 촬영 시 일어날 수 있는 손 모양의 변형을 반영하기 위하여 손을 대략 0~30°비틀어 손 모양의 변형을 주고 실험을 하였다. 실험결과 가위와 보는 95%이상의 인식률이 나왔으나 주먹은 85%의 인식률에도 미치지 못하였다. 가위와 보는 손의 특징이 명확하여 인식률이 높게 나온 것으로 보이며, 주먹은 다른 손 모양과 명확히 구분되는 특징이 없어 조금이라도 변형이 일어나면 다른 손 모양으로 인식되는 것을 확인하였다.

표 2는 손 모양에 따른 움직임 인식률을 나타낸 표이다. 각 움직임들은 동작의 시작점을 기준으로 10 프레임을 뽑아내어 움직임을 판별하였다. 손의 움직임은 각 동작 영상을 100회의 반복 실험을 통하여 인식률을 산출하였다. 인식결과 평균 95.16%로 비교적 우수한 인식률을 보였다.

표 3은 하드웨어 성능 차이가 큰 두 환경에서 동일한 영상을 수행하였을 때의 수행 속도를 비교한 표이다. 실험 방법은 한 동작의 영상을 데스크톱과 라즈베리 파이 2에서 수행한 후 한 프레임 당 소모

III. 실험결과

본 논문에서 실험환경은 Pentium PC(Core™ i3-4170, 3.7GHz)와 라즈베리 파이2를 사용하였으며 Linux 운영체제에서 Python2.7을 사용하였다. 테스트한 깊이영상의 크기는 320x240이며 프레임 속도는

Table 2. Motion recognition rate by hand gesture.

표 2. 손동작에 의한 움직임 인식률







| motion shape |  |  |  |  |  |  |
|--------------|---|---|---|--|---|---|
| Rock         | 98%   | 95%   | 98%   | 90%  | 95%   | 97%   |
| Scissor      | 96%   | 90%   | 80%   | 87%  | 97%   | 98%   |
| Paper        | 100%  | 100%  | 96%   | 98%  | 100%  | 98%   |

Table 3. Compare program execution time

표 3. 프로그램 실행 시간 비교

|                       | Desktop     | Raspberry Pi2 |
|-----------------------|-------------|---------------|
| Execution time        | 0.033       | 0.105         |
| Frame Per Second(FPS) | 30(maximum) | 10            |

된 시간을 계산하여 실행시간을 산출하였다. 실험 결과 수행시간은 데스크톱에 비해 떨어졌지만 소형 임베디드에서도 10fps의 준수한 성능으로 충분히 실시간으로 동작가능한 수준임을 확인하였다.

#### IV. 결론

본 논문에서는 손동작 인식을 위한 고성능 하드웨어를 사용할 수 없는 환경에서 사용가능한 실시간 저비용 손동작 인식 시스템을 제안하였다. 손영역 추출 연산 단순화를 위하여 TOF센서를 사용한 깊이 영상을 취득하였다. 취득한 영상의 배경을 제거한 후, 푸리에 기술자를 이용하여 손의 특징을 추출하였다. 학습 모델의 소형화를 위하여 전체 푸리에 기술자중 저주파 성분의 소수 점만을 사용하였다. 손의 움직임 판별하기 위하여 손 영역의 평균이 다음 프레임에 어떠한 방향으로 움직였는지 8방향 단위벡터로 표현하였다. 손의 움직임과 손의 모양인식을 위하여 SVM알고리즘을 이용하였으며, 인식 결과는 약 95.16%로 비교적 우수한 인식률을 보여주었다.

현재의 시스템은 사람에 따라 한 동작을 수행하는 시간이 다름을 반영하지 않았다. 이는 사람에 따라 한 동작 행할 시 발생하는 속도 차이가 인식률을 크게 저하시킬 수 있다. 향후 서로 길이가 다른 동작 신호를 정규화 하는 방법에 대한 연구가 더 필요할 것으로 기대되며, 위의 문제점들에 대하여 제안한 방법을 적용하고 비교 실험 및 고찰이 필요할 것이라고 판단된다.

#### References

[1] T Dukic, L Hanson, K Holmqvist and C Wartenberg "Effect of button location on driver's visual behaviour and safety perception", *Journal Ergonomics*, vol.48, no.4, pp.399-410, 2005. DOI: 10.1080/00140130400029092

[2] M. Zobl, M. Geiger, B. Schuller, M. Lang and G. Rigoll, "A real-time system for hand gesture controlled operation of in-car devices," *Proceedings of IEEE Multimedia and Expo2003*, Vol.3, 2003. DOI: 10.1109/ICME.2003.1221368

[3] B. yu, S. Y. Park, Y. S. Kim, I. G. Jeong, S.

Y. Ok, E. J. Lee, "Hand Tracking and Hand Recognition for Human Computer Interaction," *Journal of Korea Multimedia Society*, Vol.14 no.2, pp.182-193, 2011. DOI: 10.5565/rev/elcvia.109

[4] G. Khurana, G. Joshi and J. Kaur, "Static hand gestures recognition system using shape based features," *Recent Advances in Engineering and Computational Sciences*, pp.1-4, 2014. DOI: 10.1109/RAECS.2014.6799633

[5] S. C. Agarwal, A. S. Jalal and C. Bhatnagar, "Recognition of Indian Sign Language using feature fusion," *4th International Conference on Intelligent Human Computer Interaction*, pp.1-5, 2012. DOI: 10.1109/IHCI.2012.6481841

[6] H. M. Gamal, H. M. Abdul-Kader, and E. Sallam, "Hand gesture recognition using fourier descriptors," *8th International Conference on Computer Engineering & Systems*, pp.274-279, 2013. DOI: 10.1109/ICCES.2013.6707218

[7] P. Shukla, A. Garg, K. Sharma and A. Mittal, "A DTW and Fourier Descriptor based approach for Indian Sign Language recognition," *3th International Conference on Image Information Processing*, pp.113-118, 2015. DOI: 10.1109/ICIIP.2015.7414750

[8] T. McElroy, E. Wilson. and G. Anspach, "Fourier descriproe and neural networks far shape classification," *International Conference on Acoustics, Speech, and Signal Processing*, vol.5, pp.3435-3438, 1995.

[9] K. S. Park, D. H. Lee, Y. T. Park, "Hand Gesture Recognition Using Depth Information and Visual Image," *Journal of Korea Institute of Information Technology*, Vol.11, No.7, pp.57-65, 2013.

[10] M. K. Kyu, K. H. Bae, S. H. Cho and T. C. Kim, "Gesture-dependent depth data extraction for low resolution Time-of-Flight camera," *International Conference on Consumer Electronics*, pp.183-184, 2012. DOI: 10.1109/ICCE-Berlin.2012.6336464

[11] E. Kollorz, J. Penne, J. Hornegger and A. Barke, "Gesture recognition with a Time-Of-Flight camera," *International Journal of Intelligent Systems Technologies and Applications*, Vol.5, pp.334-343, 2008.

[12] C. H. Wu and C. H. Lin, "Depth-based hand gesture recognition for home appliance control," *IEEE 17th International Symposium on Consumer Electronics*, pp.279-280, 2013.

DOI: 10.1109/ISCE.2013.6570227

[13] Y. Li, "Hand gesture recognition using kinect," *IEEE 3rd International Conference on Software Engineering and Service Science*, pp.196-199, 2012. DOI: 10.1109/ICSESS.2012.6269439

[14] F. Dominio, M. Donadeo, G. Marin, P. Zanuttigh and G. M. Cortelazzo, "Hand gesture recognition with depth data," *Proceedings of the 4th ACM/IEEE international workshop on Analysis and retrieval of tracked events and motion in imagery stream*, pp.9-16, 2013. DOI: 10.1145/2510650.2510651

[15] N. Kawarazaki, and A. I. B. Diaz, "Gesture recognition system for wheelchair control using a depth sensor," *Proceedings of IEEE Symposium on Computational Intelligence in Rehabilitation and Assistive Technologies*, pp.48-53, 2013.

DOI: 10.1109/CIRAT.2013.6613822

---

## BIOGRAPHY

---

### Tae-Hoon Hwang (Member)



2018 : BS degree in Computer Engineering, Seokyeong University.  
Current research interests : Digital Image Processing, Computer Vision

### Jin-Heon Kim (Member)



1982 : BS degree in Electrical Engineering, Korea University.  
1983 : Research Engineer, OPC  
1984 : MS degree in Electrical Engineering, Korea University.  
1984 : Senior Researcher, Samsung Advanced Institute of Technology

1989 : FAE, ZyMOS Korea Branch

1990 : PhD degree in Electrical Engineering, Korea University.

Current research interests : Computer Vision, Image Signal Processing, Image/Video System