

밀리미터파 레이더 기반 손동작 인식 시스템 및 알고리즘에 관한 연구

이영석*

Study on the Hand Gesture Recognition System and Algorithm based on Millimeter Wave Radar

Youngseok Lee*

요약 본 연구에서는 손동작을 인식하기 위하여 밀리미터파 기반 레이더에서 얻어진 손동작의 주파수 반향 특성을 이용하는 알고리즘 및 시스템을 제안하였다. 제안된 시스템은 밀리미터파 아날로그 송수신부, ADC부 및 신호 처리부가 원 칩으로 구현된 시스템을 이용하여 데이터를 수집하도록 구성하였고, 제안한 알고리즘은 반사된 주파수 영상의 제르니케 모멘트로부터 얻어진 전역 및 지역 디스크립터로 구성된 BoF에서 K-means 클러스터링을 이용하여 코드 워드를 생성하고 SVM을 이용한 손동작 분류를 수행하였다. 수행 결과는 혼동행렬에서 얻어진 정밀도, 민감도 및 정확도를 이용하여 평가였다. 정확도 평가에서는 제안한 방법은 GZM방법 및 LZM방법과 비교하여 성능 평가를 위한 인덱스에서 제안한 방법이 95.6%의 성능을 보였고 비교한 나머지 두 방법은 88.4% 및 84%을 나타내어 제안한 방법이 기존의 두 방법에 비하여 7~8% 성능이 향상되었으며 정밀도 및 민감도에서도 나머지 두 방법에 대하여 향상된 성능을 나타내었다. 제안한 방법은 소형화된 밀리미터파 기반 레이더를 이용하여 동작 인식을 할 수 있는 임베디드 시스템의 응용가능성을 보여준다.

Abstract In this paper we proposed system and algorithm to recognize hand gestures based on the millimeter wave that is in 65GHz bandwidth. The proposed system is composed of millimeter wave radar board, analog to data conversion and data capture board and notebook to perform gesture recognition algorithms. As feature vectors in proposed algorithm, we used global and local zernike moment descriptor which are robust to distort by rotation of scaling of 2D data. As Experimental result, performance of the proposed algorithm is evaluated and compared with those of algorithms using single global or local zernike descriptor as feature vectors. In analysis of confusion matrix of algorithms, the proposed algorithm shows the better performance in comparison of precision, accuracy and sensitivity, subsequently total performance index of our method is 95.6% comparing with another two methods in 88.4% and 84%.

Key Words : Pattern recognition, Millimeter wave radar, Zernike moment, Hand gesture, Bag of features, Global and local descriptor

1. 서론

2차원 평면 또는 3차원 공간에서 인간의 동작을 인식하고 의미를 이해하려는 시도는 패턴 인식 분야에서 가장 오래된 연구 분야이며 현재도 활발한 연구가 진행되고 있다. 최근 들어 인간의 동작을 인식하려는 시도는 인간-컴퓨터 상호 작용 분야의 연구에도 적용되

어 콘텐츠 기반 영상 인식, 영상 감시, 스마트 기술 및 로봇 분야에 이르기까지 다양한 분야로 확장되고 있다 [1]. 그러나 많은 발전을 이룩한 지난 연구들에도 불구하고 현실적으로 인간 동작의 복잡하고 동일한 의미를 전달하는 다양한 변형 동작의 존재는 인간의 동작과 관련된 방대한 데이터 셋을 요구하여 이를 처리하기

본 연구는 2018학년도 청운대학교 학술연구조성비 지원에 의하여 연구되었음

*Corresponding Author :Dept. of Electronics, Incheon campus, Chungwoon University

Received June 05, 2019

Revised June 21, 2019

Accepted June 23, 2019

위한 더 많은 연구가 필요한 실정이다. 인간을 동작을 인식하기 위한 시스템이 강건하기 위해서는 동작의 특성에 영향을 미치는 요소들 즉, 영상과 센서 간의 각도의 영향들, 배경 잡음, 변형된 동작들(gesture variations)에도 불구하고 균질의 동작 특징을 추출하

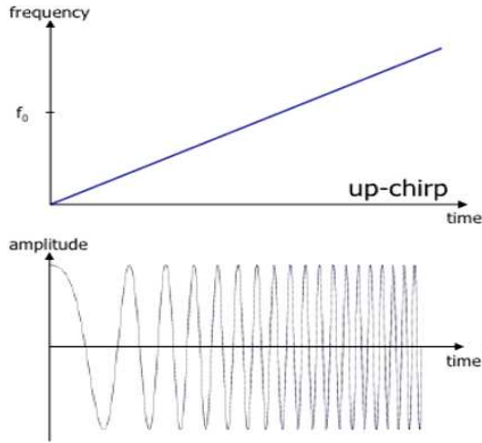


그림 1. 주파수변조 연속파형의 주파수 및 시간 표현

Fig. 1. Frequency and time representation of Frequency modulation continuous wave form

는 것이 필요하다. 따라서 인간 동작을 인식하는 문제는 특정한 동작들로 구성된 클래스를 기준으로 동작을 레이블링(labeling)하는 과정으로 정의된다[2].

이와 같은 과정은 전체적인 동작을 인식하여 특징을 추출하는 전역 특성(global feature) 또는 지역 특성(local feature)을 추출하여 특성 벡터를 구성하는 방법이 일반적이다. 그리고 동작 특성을 손동작과 같이 신체 일부의 동작으로 제한하고 수 개의 동작으로 의미를 축소하면 적절한 특성 디스크립터(descriptor)를 사용하여 위에 기술한 특성 가운데 한 가지로도 동작의 특성을 표현할 수 있다[3].

본 연구에서는 인간의 동작 특성을 인식하기 위하여 대략 60GHz 대역(55~65GHz)을 의미하는 밀리미터파(millimeter wave) 기반 레이더에서 얻어진 데이터의 주파수 반향 특성을 이용하는 알고리즘 및 시스템을 제안하였다. 제안된 시스템은 밀리미터파 아날로그 송수신부, ADC(analog to digital conversion)부 및 신호처리가 원 칩으로 구현된 시스템을 이용하여 테

이터를 수집하도록 구성하였다. 구현된 시스템에서 알고리즘의 구성은 지역 및 전역 특성을 추출하고 동작과 관련된 특성들을 제르니케 모멘트(zernike moment) 디스크립터를 이용하여 추출한 후, 동작 인식을 위하여 미리 만들어진 동작과 관련된 특성 집합 BoF(Bag of Features)를 비교하여 동작을 인식하도록 알고리즘을 구성하였다.

구현된 시스템 및 알고리즘은 4개의 손동작에 대하여 특성에 따른 알고리즘의 성능을 분석하기 위하여 전역 및 지역 특성을 제르니케 모멘트를 이용하여 추출하고 동작별 주파수 특성을 수집하여 BoF를 구성하였다. 특성에 따른 성능 평가는 두 가지 특성을 각각의 BoF간의 유클리디안 거리를 측정하여 인식을 구분하는 방법을 사용하였고, 거리에 따른 손동작의 인식 정도를 실험을 통하여 비교하였다. 또한 제르니케의 모멘트 차수의 변화에 따른 인식률의 변화를 실험적으로 측정하였다.

2. 밀리미터파 레이더 시스템의 구성

밀리미터파 대역은 최근 반도체 기술의 발달로 부품의 소형화 및 경량화가 이루어져 센서 및 무선통신 분야에서 새로운 주파수 자원으로 각광을 받고 있다. 60GHz 대역의 밀리미터파는 직진성 및 빠른 감쇄특성을 이용하여 레이더 방식으로 주변의 상황을 인지하려는 연구가 시도되고 있다[4].

기본적으로 레이더는 송신 단에서 특정 주파수 대역의 신호를 전송하고 물체로부터 반사되는 파형을 수신하여 주파수 분석을 통하여 반사된 물체의 각도, 거리 및 속도와 관련된 정보를 추출하는 것이다. 최근의 밀리미터파 레이더 시스템은 간단한 시스템과 응용성을 특징으로 그림 1의 FMCW (frequency modulated continuous waveform) 방식의 상용 레이더가 널리 사용되고 있다. FMCW 방식은 시간에 따라 중심 주파수가 증가하는 chirp 주파수 신호를 연속적으로 발생시켜 물체에서 반사되는 주파수 성분과 시간 정보사이의 주파수 차이인 비트 주파수(beat frequency)를 분석하여 거리를 측정하는 방식이다. 물체가 움직이는 경우에 비트 주파수는 도플러 주파수의 함수로서 표현할

수 있기 때문에 속도를 측정할 수 있다. 레이더에서 방출되는 밀리미터파와 물체 사이의 각도는 위상 비교 모노 펄스(phase-comparison mono-pulse) 방식을 이용하여 구현할 수 있다. 위상 비교 모노 펄스 방식은

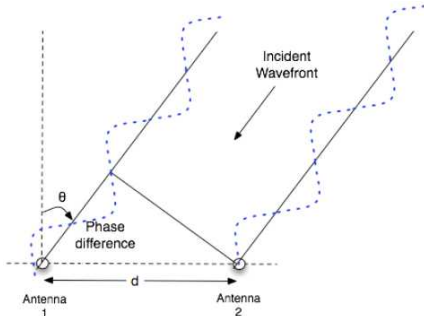


그림 2. 위상 비교 모노펄스 방식
Fig. 2. Phase-comparison mono-pulse method

그림 2에 나타낸 바와 같이 거리가 d 만큼 떨어진 2개의 수신 안테나로 수신된 두 신호에 대하여 위상차를 구하여 입사각 θ 를 식 (1)과 같이 구할 수 있다.

$$\theta = \text{asin}(\rho \frac{\text{phase difference}}{2\pi d}) \quad (1)$$

위 식에서 ρ 는 캐리어 주파수의 역수이다. 밀리미터파 레이더에서 사용하여 시간-주파수 평면 톱니파(time-frequency plane sawtooth wave)는 도플러 주파수 측정을 고려하여 시간에 따른 파수의 기울기가 충분히 크도록 하여 도플러 주파수가 거리에 의한 주파수의 영향을 받지 않도록 하는 것이 관건으로서 검출된 비트 주파수가 움직이는 않는 물체의 대하여 레이더와 물체 사이의 거리에 해당하므로 여러 개의 물체를 동시에 검출할 수 있는 장점이 있다.

$$S = f_b \frac{cT}{2B} \quad (2)$$

$$v = f_d \frac{c}{2f_0} \quad (3)$$

위식에서 f_b 는 비트 주파수, T 는 시간-주파수 평면 톱니파의 시간 폭, B 는 주파수 폭, f_d 는 도플러 주파수, f_0 는 캐리어 주파수 그리고 c 는 광속을 의미한다. 또한 식 (2)와 식 (3)에 의해 얻어지는 물체로부터의 거리 S 와 및 상대속도 v 는 그림 3과 같이 다수의 시간

-주파수 톱니파의 반향 신호에 대한 FFT(fast Fourier transform)를 수행하고 도플러 주파수를 검출하기 위하여 도플러 FFT를 수행한다.

3. 동작인식 알고리즘의 구성

제안한 동작 알고리즘은 미리 정의된 손동작에서 반사된 밀리미터파로 구성된 시간-주파수 평면의 영상에서 특징을 추출하고, 이로부터 동작을 인식하는 일련의 특징 추출 및 인식 처리과정을 수행한다. 특징 추출은 제르니케 모멘트를 이용하여 시간-주파수 평면 영상의

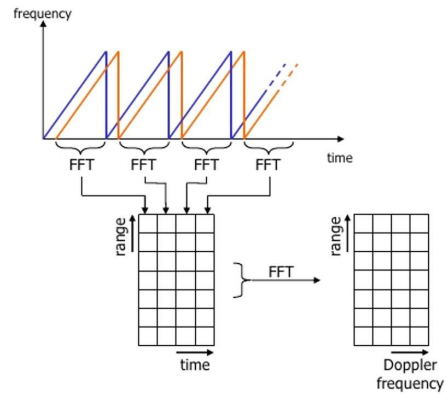


그림 3. 주파수 변조 연속 레이더 신호처리 구조
Fig. 3. Frequency modulation continuous wave radar signal processing structure

전역(global) 및 지역(local) 특징을 추출하는 과정으로 구성되어 있다.

제르니케 모멘트는 영상의 회전(rotation)과 천이(translation) 및 스케일링(scaling)의 변화에 강건한 특성 정보를 나타낼 수 있는 대표적인 영상 모멘트 기술자(descriptor)로서 단위원에 전개되는 2차원 직교 다항식인 식 (4)와 같은 제르니케 다항식으로부터 표현을 할 수 있다[5].

$$V_{nm}(\rho, \theta) = R_{nm}(\rho) e^{jm\theta} \quad (4)$$

식 (4)의 제르니케 다항식으로부터 2차원 함수 $f(x, y)$ 에 대하여 얻어지는 차수 n 의 제르니케 모멘트는 식 (5)와 같이 나타낼 수 있다.

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) V^*_{nm}(x,y) \quad (5)$$

위 식에서 $|m| \leq n, n > 0$ 이고 *는 공액 복소수를 의미한다.

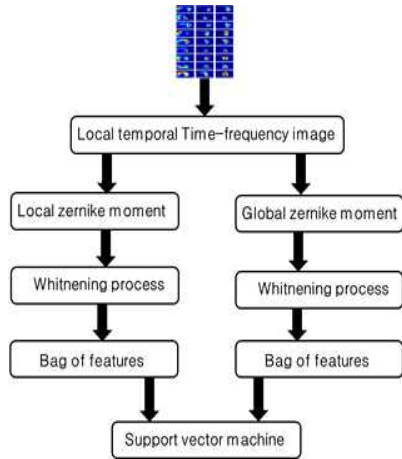


그림 4. 동작 인식 알고리즘 흐름도
Fig. 4. Gesture recognition algorithm flow chart

특성을 추출하기 위한 제르니케 모멘트의 차수는 따라 0에서 5까지 변화시켜 12개의 모멘트 값을 성분으로 하는 특징 벡터를 추출하도록 구성하였다. 제르니케 모멘트의 적용은 시간-주파수 평면 영상에 대하여 전역적 특성을 추출하기 위한 전역 제르니케 모멘트 특징 벡터와 부분적 특성을 추출하기 위한 부분 제르니케 모멘트 특징 벡터로 구분하여 추출하였다. 각각의 손동작에 대하여 얻어진 특징벡터들은 불필요한 표현 성분들(redundancy components)을 제거하기 위하여 상호 공분산 행렬을 단위행렬로 만들어 주는 화이트닝 과정을 수행한다. 본 연구에서는 화이트닝 과정을 주성분 분석(principal component analysis)을 통하여 선형 직교 공간에 주어진 특징 벡터들을 투사하는 방식으로 구현하였다. 손동작을 학습하고 인식하기 위하여 BoF(Bag of Features) 방법을 사용하였다[6]. BoF 방법은 같은 의미를 갖는 손동작이 사람의 특성에 따라 서로 다른 특징 벡터를 생성하기 때문에 개인별로 생성된 특징 벡터를 사용하는 경우에 발생하는 인식오율을 줄일 수 있다. BoF 과정은 다른 대상과 다른 조건에서 얻은 손동작과 관련된 시간-주파수 평면의 영

상을 클러스터링하여 코드 워드(code word)를 얻는 과정, 각 손동작의 코드 워드를 모은 코드 북을 생성하는 과정을 포함하며 본 연구에서는 k-means 알고리즘을 이용하여 클러스터링하고 클러스터링의 중심을 코드 워드로 사용하였으며 4개의 손동작에 대한 클러스터링 중심을 코드워드의 집합인 코드 북으로 생성하였다. 코드 북을 생성한 후 코드 북과 임의의 영상에서 얻어진 특성 벡터간의 히스토그램은 손 동작 학습 및 인식을 위한 특성 벡터로 SVM(support vector machine) 분류기에 사용된다.

4. 실험 및 결과 고찰

실험을 위한 밀리미터파 송수신 장치로 TI(Texas Instrument) 사의 IWR1443 BOOST 밀리미터파 송수신 평가보드를 사용하였다[7]. 수신된 데이터들은 TI 사의 밀리미터파 개발 보드에서 수집되고 데이터 캡처 보드와 연결된 노트북으로 이동하여 Matlab 환경에서 손동작 인식 실험을 실시하였다.

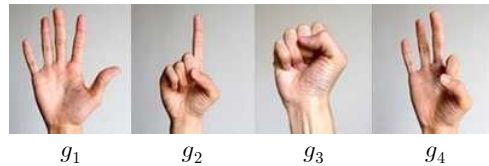


그림 5. 인식을 위하여 사용된 손동작들
Fig. 5. Hand gestures for recognition

그림 5는 실험에서 메시지를 표시하기 위하여 미리 약속된 4 가지의 서로 다른 손동작을 나타내고 있으며 손동작들은 30cm의 거리에서 밀리미터파 레이더의 축을 중심으로 회전을 달리하는 4가지 손동작에 대하여 200회 반복된 손동작에서 수집한 데이터를 60%는 특징 벡터의 클러스터링 과정과 히스토그램 생성 과정을 거쳐 훈련 데이터로서 사용하고 나머지 40%의 데이터에 대하여 동작 인식 실험을 수행하였다. 그림 6은 실험을 위하여 사용한 밀리미터파 레이더 시스템인 IWR1443 보드와 나머지 데이터 수집 및 알고리즘 동작을 위한 보드 및 노트북을 나타내고 있다.

제안한 방법을 이용한 손동작 인식의 결과는 라벨을 가진 기계 학습 분류 모델의 성능을 측정하기 위하여 사용하는 혼동 행렬(confusion matrix)을 이용하였다

[8-10]. 표 1은 제안한 방법을 이용하여 얻은 분류 결과에 대한 혼동 행렬이고 표 2 및 표 3은 비교를 위하여 사용한 전역 제르니케 방법(GZM) 및 지역 제르니케 방법(LZM)에 대한 혼동 행렬 결과이다.

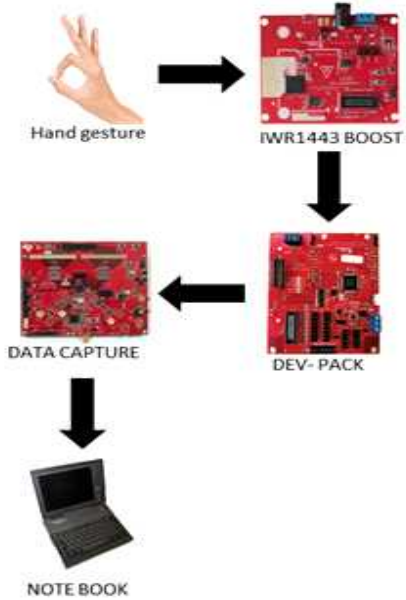


그림 6. 손동작 인식을 위한 실험 장치 설치
 Fig. 6. Experimental set up for gesture recognition

표 1. 제안한 방법에 대한 혼동 행렬
 Table 1. Confusion matrix for proposed method

Proposed method		Prediction			
		g_1	g_2	g_3	g_4
Input	g_1	18	0	0	2
	g_2	1	18	1	0
	g_3	0	0	20	0
	g_4	2	1	0	17

표 2. GZM 방법에 대한 혼동 행렬
 Table 2. Confusion matrix for GZM method

GZM		Prediction			
		g_1	g_2	g_3	g_4
Input	g_1	15	0	0	5
	g_2	1	15	1	3
	g_3	0	2	18	0
	g_4	6	0	2	10

표 4는 표 1, 표 2 및 표 3의 혼동 행렬로부터 사용한 방법들의 성능을 평가하기 위하여 계산된 분류 성능 특성으로서 각 클래스에 대하여 TP(true positive), TN(true negative), FP(false positive) 및 FN(false negative)을 나타내고 있다.

표 3. LZM 방법에 대한 혼동 행렬
 Table 3. Confusion matrix for LZM method

LZM		Prediction			
		g_1	g_2	g_3	g_4
Input	g_1	12	0	0	8
	g_2	2	15	2	1
	g_3	1	1	16	2
	g_4	7	1	1	11

표 4. 혼동 행렬에서 계산된 분류 특성
 Table 4. Computed Classification characteristics

Method	Class	TP	TN	FP	FN
Proposed method	g_1	18	57	3	2
	g_2	18	59	1	2
	g_3	20	59	1	0
	g_4	17	58	2	3
GZM	g_1	15	53	7	5
	g_2	15	58	2	5
	g_3	18	57	3	2
	g_4	10	54	8	8
LZM	g_1	12	50	10	8
	g_2	15	58	2	5
	g_3	16	57	3	4
	g_4	11	49	11	9

표 5. 제안한 방법과 GZM 및 LZM의 성능 비교
 Table 5. Performance comparison of methods

	Performance index (%)		
	Precision	Sensitivity	Accuracy
Proposed	91.3	91.2	95.6
GZM	72.1	73.8	88.4
LZM	69.2	67.5	83.0

표 4의 분류 성능 특성을 이용하여 각 방법의 분류 정밀도(precision)는 $TP/(TP+FP)$ 로, 분류 민감도(Sensitivity)는 $TP/(TP+FN)$, 그리고 정확도(accuracy)는 $(TP+TN)/((TP+TN+FP+FN))$ 을 이용하

여 계산할 수 있고 모든 성능 지표는 100%에 가까울 수록 우수한 성능을 의미한다. 표 5는 제안한 방법과 나머지 두 방법의 성능을 정밀도, 민감도 및 정확도를 이용하여 평가한 결과이다. 분류기의 성능을 결정하는 정확도 성능 평가에서 제안한 방법이 나머지 두 방법보다 약 7~8%의 성능 향상이 이루어진 것을 확인할 수 있다.

5. 결론

본 연구에서는 손동작을 인식하기 위하여 밀리미터 파 기반 레이더에서 얻어진 손동작의 주파수 반향 특성을 이용하는 알고리즘 및 시스템을 제안하였다. 제안된 시스템은 밀리미터파 아날로그 송수신부, ADC부 및 신호처리부가 원 칩으로 구현된 시스템을 이용하여 데이터를 수집하도록 구성하고 제안한 알고리즘을 수행하였다. 수행 결과는 제안한 방법과 지역 및 전역 제르니케 모멘트를 하나만 사용한 경우에 대하여 혼동행렬을 구성하고 정밀도, 민감도 및 정확도에 대하여 평가하였다. 정확도 평가에서는 제안한 방법 95.6%의 성능을 보였고 GZM 방법과 LZN 방법을 이용하는 경우 88.4% 및 83%의 성능을 보여 제안한 방법이 나머지 두 가지 방법보다 7~8% 성능이 향상되었으며 정밀도 및 민감도에서도 향상된 성능을 나타내었다. 제안한 방법은 소형화된 밀리미터파 기반 레이더를 이용하여 동작 인식을 할 수 있는 임베디드 시스템의 응용가능성을 보여준다.

REFERENCES

[1] J. Preece, Y. Rogers, H. Sharp, D. Benyon, S. Holland and T. Carry, *Human-computer interaction. Addison-Wesley Ltd.*, 1994.
 [2] S.S. Rautaray, A. Siddharth and A.Agrawal. "Vision based hand gesture recognition for human computer interaction: a survey," *Artificial intelligence review*, vol.43, no. 1, 43.1, pp. 1-54, 2015.
 [3] Nishikawa, Atsushi, A. Ohnisi, and Fumio Miyazaki. "Description and recognition of human gestures based on the transition of

curvature from motion images," *Proceedings Third IEEE Int'l Conference on Automatic Face and Gesture Recognition*, 1998.
 [4] Nagatsuma, Tadao, et al. "Millimeter-wave and terahertz-wave applications enabled by photonics," *IEEE Journal of Quantum Electronics*, vol. 52.no. 1, pp. 1-12, 2015.
 [5] Chen Weiguo, and Jinchuan Lu. "Human Shape Recognition Algorithm Design Based on Hu Moments and Zernike Moments," *2016 4th International Conference on Machinery, Materials and Information Technology Applications*. 2017.
 [6] Ramesh Bharath, Cheng Xiang, and Tong Heng Lee. "Shape classification using invariant features and contextual information in the bag-of-words model," *Pattern Recognition* vol. 48. no. 3, pp. 894-906, 2015.
 [7] www.ti.com/lit/ds/symlink/iwr1443.pdf
 [8] <http://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>
 [9] T. Townsend, "Theoretical analysis of an alphabetic confusion matrix," *Perception & Psychophysics* vol. 9, pp. 40-50, 1971.
 [10] A. Hay, "The derivation of global estimates from a confusion matrix," *International Journal of Remote Sensing* vol. 9, pp.1395-1398, 1988.

저자약력

이 영 석 (Young-Seok Lee)

[정회원]



- 1995년 2월 : 서울시립대학교 대학원 전자공학과 (공학석사)
- 1998년 2월 : 서울시립대학교 대학원 전자공학과 (공학박사)
- 1998년 3월 ~ 현재 : 청운대학교 인천캠퍼스 전자공학과 교수

〈관심분야〉 디지털신호처리, 계산신경과학, 임베디드 시스템