

# 센서 융합을 통한 물체 거리 측정 및 인식 시스템

## Object detection and distance measurement system with sensor fusion

이 태 민\*, 김 정 환\*, 임 준 흥\*

Tae-Min Lee\*, Jung-Hwan Kim\*, Joonhong Lim\*

### Abstract

In this paper, we propose an efficient sensor fusion method for autonomous vehicle recognition and distance measurement. Typical sensors used in autonomous vehicles are radar, lidar and camera. Among these, the lidar sensor is used to create a map around the vehicle. This has the disadvantage, however, of poor performance in weather conditions and the high cost of the sensor. In this paper, to compensate for these shortcomings, the distance is measured with a radar sensor that is relatively inexpensive and free of snow, rain and fog. The camera sensor with excellent object recognition rate is fused to measure object distance. The converged video is transmitted to a smartphone in real time through an IP server and can be used for an autonomous driving assistance system that determines the current vehicle situation from inside and outside.

### 요 약

본 논문에서는 자율주행 자동차에 물체를 인식하고 거리를 측정하는데 효율적인 센서 융합을 제안한다. 자율주행 자동차에 사용되는 대표적인 센서는 레이더, 라이다, 카메라이다. 이 중 라이다 센서는 차량 주변의 맵을 만드는 역할을 한다. 하지만 날씨 조건에 성능이 하락하고 센서의 가격이 매우 비싸다는 단점 있다. 본 논문에서는 이러한 단점을 보완하고자 비교적 저렴하고 눈, 비, 안개에 지장 없는 레이더 센서로 거리를 측정하며 차량 주변을 관찰한다. 물체 인식률이 뛰어난 카메라 센서를 융합하여 물체 인식 및 거리를 측정한다. 융합된 영상은 IP서버를 통해 실시간으로 스마트폰에 전송되어 현재 차량의 상황을 내부, 외부에서 판단하는 자율주행 보조 시스템에 사용될 수 있다.

*Key words* : YOLOv3, Vision, Radar, Sensor fusion, FMCW, radar

### 1. 서론

자동차 산업의 끊임없는 혁신과 기술 개선으로 자동차가 일상 생활의 필수 요소가 되었다는 것은

부정할 수 없는 사실이다. 이런 이유로 모든 차량 탑승자의 사망 또는 중상을 예방하기 위해 안전이 최우선 과제이다. 세계 보건기구 WHO는 교통사고로 매년 225만명이 사망한다고 발표했다. 많은 자

\* Dept. of Electronic Engineering, Hanyang University

★ Corresponding author

E-mail : [jhlim@hanyang.ac.kr](mailto:jhlim@hanyang.ac.kr), Tel : +82-31-400-4043

※ Acknowledgment

This research was supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (Grant No.2017R1D1A1B03033806).

Manuscript received Mar. 6, 2020; revised Mar. 15, 2020; accepted Mar. 18, 2020.

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.

동차 회사들은 높은 안전성을 갖춘 자동차를 제작하는 데 필요한 새로운 기능과 프레임 워크를 지속적으로 개발하고 있다.

최근 자율주행 시장동향을 보면 카메라, 레이더, 라이다, 초음파와 같은 여러 가지 센서를 융합하여 사용한다. 센서들을 이용해 물체를 식별하고 주행 환경도 판단한다. 그 중 보행자, 자동차, 동물과 같이 움직이는 것을 판단하는 것이 가장 중요하다. 최근 물체인식 분야 중 딥러닝 알고리즘이 크게 발달되고 있다. 딥러닝은 기계학습 분야 중 하나이며 컨볼루션신경망(Convolution Neural Network) 인 AlexNet[1]을 기점으로 R-CNN[2], Faster R-CNN[3], Mask R-CNN[4] 등의 연구가 활발하게 진행되고 있다.

자동차의 레이더는 자율주행에 핵심적인 센서 중 하나이다. 왜냐하면 안개, 비, 눈 등 모든 환경 조건에 가장 높은 내성을 가지고 있기 때문이다. 일반적으로 사용되는 FMCW 레이더는 주파수 변조 연속파 신호이다. FMCW는 비교적 신호 처리 요구 사항에 유연성이 뛰어나다. FPGA와 같은 저렴하면서 특수 가속 엔진이 있는 마이크로프로세서와 ASIC를 충족할 수 있다.

FMCW 레이더의 주요 목적은 비트 주파수 또는 중간 주파수로 알려진 주파수 변조 연속파 신호를 사용하여 목표 범위, 속도 및 각도를 측정하는 것이다 [5]. 이 신호는 주로 선형으로 변조되며 전파 지연과 도플러 주파수를 사용하여 모델링할 수 있다[6].

미국 및 유럽에서 운전자의 편의나 능동안전에 대한 평가 프로그램이 있다. 법규가 강화되고 있어 전방 충돌을 예방하는데 관심이 많아지고 있다. 전방 충돌 감지를 고급화시켜 앞으로 자율주행 3단계 이상의 성능을 구현하기 위해서 레이더와 카메라 센서의 거리 및 물체 인식 기술이 잘 융합되어야 한다. 카메라 센서는 물체 인식은 잘 되지만 거리 측정에 어려움이 있고 밤이나 기상상황이 안 좋을 때 인식도 어렵다. 그러나 레이더 센서는 물체 검출은 감지는 쉽지만 거리 측정에 특화 되어 있고 기상에 영향이 없다. 그러므로 레이더 데이터와 카메라 데이터를 융합하면 물체의 거리 및 인식에 정확성을 높일 수 있다. 본 논문에서는 센서 융합을 통해 물체 인식 및 거리 측정을 하는 알고리즘을 제안한다. 2장에서는 레이더 센서의 거리측정 기술을 소개하고, 3장에서는 딥러닝을 이용한 물체

인식을 소개한다. 4장에서는 센서융합 및 데이터 전송 알고리즘을 소개하며, 5장에서는 실험 결과를 설명한다. 마지막으로 6장에서 결론을 맺는다.

## II. 레이더

### 1. FMCW 레이더 신호처리

FMCW 레이더의 기본 개념은 처프라는 신호이다. 이 신호는 진폭-시간 도표 및 순간 주파수-시간 도표인 그림 1와 같이 시간에 따라 선형적으로 증가하는 주파수 값을 갖는다. 처프 신호의 특성은 톱니 주파수를 생성하는 것이다.

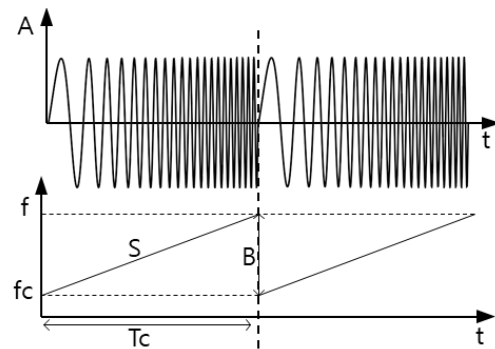


Fig. 1. amplitude-time and frequency-time graph of the chirp signal.

그림 1. 처프 신호의 진폭-시간, 주파수-시간 그래프

초기 주파수 값은 경사의 시작인 주파수  $f_c$ 로 표기했다. 논문의 개발은 76-81GHz의 주파수 범위를 가진 밀리미터파 레이더 모듈을 사용하여 수행하였다. 경사의 시작과 끝 사이의 주파수 간격을 나타내는 주파수 스위프 대역폭은 B이다. 처프 지속 시간  $T_c$ 는 할당된 처프 사이클 시간을 나타낸다. 마지막으로 다음과 같이 정의할 수 있는 처프 경사 S는

$$S = \frac{B}{T_c} \quad (1)$$

로 구할 수 있다.

식 (1)은 단위 시간당 처프 신호 주파수의 변화율을 나타낸다. 처프가 증폭되고 전송되면 수신기 안테나는 동일한 처프 신호의 반사 지연 버전을 검출한다. 저잡음증폭기(Low-Noise Amplifier)를 지난 후 이 수신 신호는 원래의 신호와 혼합되어 비트 주파수 또는 IF 신호를 생성한다. 마지막으로 이

IF 신호는 저역통과필터에 적용되어 고주파수성분을 제거하고 디지털화된다. 생성된 디지털 신호는 후처리 DSP 알고리즘을 위해 다른 장치로 전송된다. DSP 알고리즘은 수학적 레이아웃을 기반으로 주파수 영역의 신호를 분석하여 측정할 대상들의 범위, 속도, 각도와 같은 여러 물리적 특성을 도출한다.

2. 사거리 추정

FMCW 레이더 사용하는 용도로 첫 번째이자 가장 많이 사용하는 것은 사거리 추정이며, 이는 레이더 모듈과 측정 대상 사이의 거리에 해당한다. 그러나 사거리가 처프 신호와 어떤 관련이 있는지 완전히 이해하려면 믹서 블록에서 나오는 IF 신호를 분석해야한다. 두 정현파 입력을 통한 믹서 블록 출력은

$$X_{out}(t) = \sin[(\omega_1 - \omega_2)t + (\theta_1 - \theta_2)] \tag{2}$$

와 같이 표현 될 수 있다. 여기서  $\omega_1 - \omega_2$  및  $\theta_1 - \theta_2$  는 두 입력 신호의 순간 주파수와 위상의 차이이다. 이 다이어그램에서 레이더 앞의 단일 대상은  $S \cdot \tau$ 의 일정한 주파수 값을 갖는 IF신호를 생성한다. 여기서 S는 처프의 기울기이고  $\tau$ 는 RTT이다. 최고 주파수 값인 IF신호는 레이더가 감지 할 수 있는 최대 거리를 결정한다. 그러나 작업 흐름을 고려할 때 이 IF신호는 추가 DSP를 위해 디지털화되어야한다. 이러한 이유로 IF대역폭은 ADC(Analog-to-Digital Converter)의 샘플링 주파수  $f_s$ 로 제한 된다.

따라서 최대 범위 거리  $R_{max}$ 는

$$R_{max} = \frac{f_s c}{2S} \tag{3}$$

로 샘플링 주파수  $f_s$ 에 정비례한다.

III. 딥러닝 YOLO v3

1. YOLO v3 개념

convolutional neural networks를 사용하는 알고리즘 중 YOLO는 가장 빠른 물체 감지 알고리즘 중 하나이다. 가장 정확한 물체 감지 알고리즘은 아니지만 높은 정확도가 필요하지 않고 실시간 감지가 필요할 때 매우 적합하다.

다른 물체 인식 알고리즘과 비교하여, 이 알고리즘은 클래스 레이블을 예측할 뿐만 아니라 객체의 위치도 감지한다. 따라서 이미지를 카테고리 분류 할뿐만 아니라 이미지 내에서 여러 객체를 감지 할 수도 있다. 이 알고리즘은 단일 신경망을 전체 이미지에 적용한다. 즉, 이 네트워크는 이미지를 영역으로 나누고 각 영역의 경계 상자 및 확률을 예측한다. 이러한 경계 상자는 예측 된 확률로 가중치가 적용된다.

2. YOLO 아키텍처 설계

YOLO는 컨볼루션 레이어만 사용하므로 완전 컨볼루션 네트워크 FCN에 해당된다. YOLO v3 논문에서 저자는 Darknet-53이라는 특징 추출기의 새롭고 더 깊은 아키텍처를 제시한다[7]. 이로부터 알 수 있듯이 53 개의 컨볼루션 레이어로 구성되며 각 레이어에는 배치 정규화 레이어와 Leaky ReLU 활성화가 있다. 풀링 형태는 사용되지 않으며 stride 2의 컨볼루션 레이어가 특징 맵을 샘플링하는 데 사용됩니다. 이것은 풀링에 의한 특징이 낮은 수준의 손실을 방지하는 데 도움이 된다. 그림 2은 YOLOv3의 Darknet-53 아키텍처이다.

Type	Filters	Size	Output
Convolutional	32	3 x 3	256 x 256
Convolutional	64	3 x 3 / 2	128 x 128
Convolutional	32	1 x 1	
Convolutional	64	3 x 3	
Residual			128 x 128
Convolutional	128	3 x 3 / 2	64 x 64
Convolutional	64	1 x 1	
Convolutional	128	3 x 3	
Residual			64 x 64
Convolutional	256	3 x 3 / 2	32 x 32
Convolutional	128	1 x 1	
Convolutional	256	3 x 3	
Residual			32 x 32
Convolutional	512	3 x 3 / 2	16 x 16
Convolutional	256	1 x 1	
Convolutional	512	3 x 3	
Residual			16 x 16
Convolutional	1024	3 x 3 / 2	8 x 8
Convolutional	512	1 x 1	
Convolutional	1024	3 x 3	
Residual			8 x 8
Avgpool		Global	
Connected		1000	
Softmax			

Fig. 2. Multiple-object velocity estimation workflow. 그림 2. 다중 물체 속도 추정 워크플로우

YOLO는 입력 이미지의 크기와 다르게 여러 종류로 검색한다. 그 이유는 물체의 크기가 항상 같지 않기 때문이다. 이미지의 크기를 일괄적으로 나누어 처리 할 경우 GPU가 병렬로 처리 할 수 있기 때문에 속도를 향상 시킬 수 있다. 그리고 고정된 높이와 폭의 모든 이미지를 가져야하며 여러 이미지를 큰 배치로 연결하는 작업이 필요하다.

네트워크는 일정한 크기로 이미지를 다운 샘플링한다. 예를 들어, 네트워크의 보폭이 32 인 경우 크기 416×416의 입력 이미지는 크기 13×13의 출력을 생성한다. 일반적으로 네트워크 모든 레이어의 보폭은 출력의 계수와 같으며 모든 레이어는 입력 이미지보다 작다.

### 3. 출력 내용 처리방법

입력 모양의 이미지 배치는 (m, 416, 416, 3)이다. 출력은 인식 된 클래스와 함께 바운딩 박스 목록입니다. 각 바운딩 박스는 6 개의 숫자  $((p_c, b_x, b_y, b_h, b_w, c))$ 로 표시된다.  $c$ 를 80 차원 벡터로 확장하면 각 경계 상자는 85 개의 숫자로 표시됩니다. 모든 물체 탐지기와 동일한 방식으로 컨볼루션 레이어가 학습한 특징은 분류기/회귀기로 전달되어 탐지 예측(바운딩 박스의 좌표, 클래스 레이블 등)을 만든다.

YOLO에서는 1×1 컨볼루션 레이어를 사용하여 예측을 수행한다. 가장 먼저 주목할 것은 출력 값인 특징 맵이다. 1×1 컨볼루션을 사용 했으므로 예측 맵의 크기는 정확히 특징 맵의 크기이다. YOLO v3에서 이 예측 맵을 해석하는 방법은 각 셀이 고정 된 수의 바운딩 박스를 예측할 수 있다. 예를 들어 피쳐 맵에  $(B \times (5+C))$  항목이 있는 경우,  $B$ 는 각 셀이 예측할 수 있는 바운딩 박스 수를 나타낸다. YOLO논문에 따르면, 이들  $B$ 바운딩 박스 각각은 특정 종류의 물체를 검출하는 것을 전문으로 할 수 있다. 각 경계 상자에는  $5+C$  속성이 있으며, 각 경계 상자의 중심 좌표, 치수, 객관성 점수 및  $C$  클래스 신뢰도를 설명한다.

YOLO v3는 모든 셀에 대해 3개의 바운딩 박스를 예측한다. 객체의 중심이 해당 셀의 수용 필드에 속하는 경우 특징 맵의 각 셀 경계 상자 중 하나를 통해 객체를 예측할 것으로 기대한다. 이것은 YOLO 훈련 방법과 관련 있다. 단 하나의 바운딩 박스만이 주어진 객체를 감지하는 역할을 한다. 먼저 이 바운딩 박스가 속하는 셀을 확인해야한다. 이를 위해 입력 이미지를 최종 특징 맵과 동일한 크기의 격자로 나눈다.

## IV. 센서 융합 및 웹 프레임 워크

### 1. 센서 융합

레이더에서 촬영한 데이터를 시리얼 포트를 연결

하여 PC에 전송한다. 레이더 데이터는 MATLAB을 통해 GUI로 영상화한다. 하지만 레이더 데이터는 탐류 형태로 들어오기 때문에 2D형태의 영상이 생성된다. YOLO v3로 학습한 영상은 물체 인식된 좌표에 맞게 상자를 만들어 주고 물체 이름을 기재한다. 마지막으로 영상 융합을 위해 YOLO 영상을 레이더 영상의 거리에 맞춰 탐류로 만드는 작업을 한다. 탐류는

$$Affine = \begin{bmatrix} \cos\theta & -\sin'\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

와 같은 어파인 변환을 이용하여 구성된다.

여기서  $\alpha$ 와  $\beta$ 는 크기 조정 행렬이고  $\sin\theta$ 와  $\cos\theta$ 는 회전 행렬이다.  $t_x$ 와  $t_y$ 는 평행이동 행렬이다. 어파인된 YOLO 탐류 영상은 레이더 탐류 영상과 융합할 수 있다[9].

영상 융합을 위해서는 두 영상의 정보를 채널별로 나누고, 대상 이미지에서 삽입할 영상의 영역을 추출한다. Mask를 만들기 위해 삽입할 이미지를 회색으로 변경한 후 이진 영상으로 전환한다. 이진 영상으로 만드는 방식은 주변 영역의 평균밝기를 임계값으로 이용하는 지역 가변 이진화를 사용한다.  $T$ 는

$$T(x, y) = \frac{1}{n^2} \sum_{x_i} \sum_{y_i} I(x+x_i, y+y_i) - C \quad (5)$$

로 표현할 수 있다.

$n \times n$ 주변 영역의 평균 밝기에서 상수  $C$ 만큼을 빼서 결정한다. 결과 값은 상수  $C$ 와  $n \times n$ 크기의 설정에 따라 크게 달라진다. Mask와 반대되는 Mask\_inv를 만들기 위해 Mask의 값을 반전시킨다. 마지막으로 and 연산과 Mask, Mask\_inv를 이용하여 대상 영상은 삽입할 부분을, 삽입할 영상은 배경을 제거하여 겹치는 부분이 사라진 상태로 or 연산을 이용해 두 영상을 완전히 융합한다. 최종적으로 융합한 탐류 영상을 원본 사진에 넣기 위해 다시 한 번 어파인 변환하고 어파인한 영상을 비트 연산으로 재 융합한다[10].

### 2. 웹 프레임 워크 개발

Python의 웹 프레임 워크로 널리 사용되는 Django와 Flask 중 Flask는 매우 가벼운 기본 웹 응용 프

로그래밍이다[8]. html을 기반으로 서버에서 IP와 Port를 열어 클라이언트가 접속 가능하도록 한다. 접속 가능한 클라이언트는 컴퓨터, 스마트폰, 태블릿과 같이 인터넷이 접속 가능하면 클라이언트가 될 수 있다. 아무나 접속할 수 없도록 보안도 설정 가능하다.

### V. 실험환경 및 결과

본 논문에서는 Python 환경에서 OpenCV라이브러리를 사용한다. 실험에 사용된 시스템의 사양은 표1 과 같다. 카메라 BRIO 4K PRO와 레이더 AWR1843BOOST를 설치해 실험 데이터를 얻는다. 실험 데이터는 라우터 WR246를 통해 스마트폰으로 전송되며 전송 FPS는 30프레임이다.

Table 1. System specification.

표 1. 시스템 사양

Device	Component	Specification
Radar	Manufacturing Company	Texas Instruments
	Model Name	AWR1843BOOST
	Available Bandwidth	76- to 81-GHz coverage with 4 GHz
	Memory	2MB
	Receive Channels	Four
	Transmit Channels	Three
Camera	Manufacturing Company	Logitech
	Model Name	BRIO 4K PRO
	Resolution	640×480 (30FPS)
	Size	102×27×27mm, 63g
	Angle of view	90°
Router	Manufacturing Company	Cioswi
	Model Name	WR246
	Hardware	MT7621A/ dual-core/ 880Mhz
	Memory	DDR3 512MB
	Size	164×137×35mm
	Protocol standards	IEEE 802.11ac, IEEE 802.11a
	Wireless rate	2.4G: 300Mbps, 5G: 900Mbps

그림 3의 (a)는 주간 촬영한 융합 영상으로 사람과 차를 인식하였으며 레이더 데이터 정보 또한

포함되어 있음을 사각 박스로 확인할 수 있다. (b)는 야간에 주차장에서 촬영한 영상으로 전방의 차가 검출된 결과를 볼 수 있다.

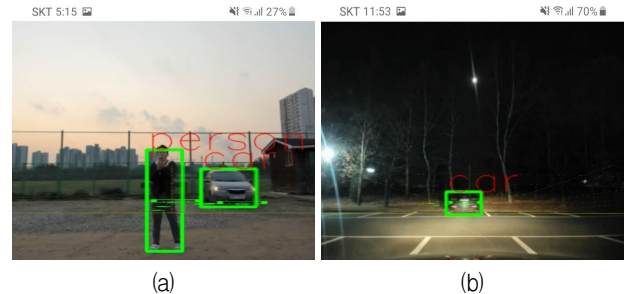


Fig. 3. Object Detection in daytime and nighttime.

그림 3. 주간 야간 물체 감지

### VI. 결론

본 논문은 물체 감지, 거리 측정, 센서 융합, 영상 스트리밍 이러한 모든 요소를 결합한 차량 감시 시스템을 제안한다. 실험 결과 제안한 알고리즘은 발생할 수 있는 사고와 지나간 사고를 정확하게 파악하고 효율적으로 증거자료를 수집할 수 있어 원인 규명할 수 있게 된다. 또한 무인 자동차 개발에 적용할 수 있을 것으로 판단된다.

### References

[1] Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, 2012. DOI: 10.1145/3065386

[2] Ross Girshick, Jeff Donahue, Trevor Darrell and Jitendra Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014. DOI: 10.1109/cvpr.2014.81

[3] Shaoqing Ren, Kaiming He, Ross Girshick and Jian Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.39, No.6, pp.1137-1149, 2017. DOI: 10.1109/TPAMI.2016.2577031

[4] Kaiming He, Georgia Gkioxari, Piotr Dollar and Ross Girshick, "Mask r-cnn," *Proceedings of the IEEE international conference on computer vision*, 2017. DOI: 10.1109/iccv.2017.322

[5] A. G. Stove, "Linear FMCW radar techniques," *IEE Proceedings F-Radar and Signal Processing*, Vol.139, pp.343-350, 1992.

DOI: 10.1049/ip-f-2.1993.0019

[6] Jau-Jr Lin, Yuan-Ping Li, Wei-Chiang Hsu and Ta-Sung Lee, "Design of an FMCW radar baseband signal processing system for automotive application," *SpringerPlus*, Vol.5, pp.42, 2016.

DOI: 10.1186/s40064-015-1583-5

[7] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement," arXiv preprint arXiv:1804.02767 2018.

[8] Gareth Dwyer, Jack Stouffer and Shalabh Aggarwal, *Flask : building Python web services*, Packt Publishing, 2017.

[9] Hyunggi Cho, Young-Woo Seo, B.V.K. Vijaya Kumar and Rangunathan Raj Rajkumar, "A multi-sensor fusion system for moving object detection and tracking in urban driving environments," *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp.1836-1843, 2014.

DOI: 10.1109/ICRA.2014.6907100

[10] Ricardo Omar Chavez-Garcia and Olivier Aycard, "Multiple Sensor Fusion and Classification for Moving Object Detection and Tracking," *IEEE Transactions on Intelligent Transportation Systems*, Vol.17, No.2, pp.525-534, 2016.

DOI: 10.1109/TITS.2015.2479925

## BIOGRAPHY

### Tae-Min Lee (Member)



2016 : BS degree in Computer Engineering, Mokwon University.  
2016~current : MS degree course in Electronic Engineering, Hanyang University.

### Jung-Hwan Kim (Member)



2013 : BS degree in System Control Engineering, Hoseo University.  
2015 : MS degree in Electronic Systems Engineering, Hanyang University.  
2016~current : PhD degree course in Electronic Engineering, Hanyang University.

### Joonhong Lim (Member)



1979 : BS degree in Electronic Engineering, Seoul National University.  
1981 : MS degree in Electrical Engineering, KAIST.  
1986 : PhD degree in Electrical and Computer Engineering, University of Iowa.

1986~1989 : Assistant Professor, Dept. of Electronic Engineering, Korea Aerospace Univ.

1989~current : Professor, Dept. of Electronic Engineering, Hanyang Univ.