

# j.inst.korean.electr.electron.eng.

전기전자학회 논문지

제 28 권 제 4 호 2 0 2 4 년 1 2 월

사단법인 한국전기전자학회

### [Technical Paper]

- Roundabout Negotiation Service: Scenario Definition/Implementation and Real-Road Validation ● Ji-Min Lee, Seong-Hyun Jang, Yoo-Seung Wang, Sang-Hun Yoon, Byoung-Man An (479)
- High-Precision Indoor Positioning System Based on UWB and Vision Sensor Fusion ● Heeyoung Joo, Jongwha Ahn, Sanghoon Yoon, Seonghyun Jang (489)
- Enhancing DGPS positioning accuracy using CNN-LSTM model ● Ji-Eon Lee, Tae-Hee Lee, Woo-Sung Hwang, Myung-Ryul Choi (497)
- A illumination-aware learnable image filtering model for object detection in low-light conditions ● Chang-Hwan Son, Do-Gyun Kim, Hyun-Jun Ko (503)
- Real-Time Fall Detection for the Elderly Based on Low-Resolution Infrared Images ● Min-Su Kim, Jong-won Seok (511)
- Blind Heavy Rain Face Image Restoration Using Multimodal Text-Image Alignment ● Chang-Hwan Son, Yeon-U Choi (520)
- A Study on Power Transformer Health Index Evaluation Using Dimensionality Reduction ● Seung-Yun Lee, Jeong-Sik Oh, Tae-Hun Kim, Jae-Deok Park, Byeong-Hyeon An, Tae-Sik Park (528)
- V2X Communication Technology Classification and LTE-V2X (Rel. 14) Communication Performance Test ● Yong-Tae Kim, Woo-Jai Shin, Jung-Won Lee (540)
- 4-8GHz Injection-Locked Frequency Tripler with Notch Filtering for Quantum Computer Read-Out ● Dohun Kim, Hapsah Aulia Azzahra, Hong Chae, Hyeon-Sik Ahn, Jusung Kim (546)
- Automotive Radar Interference Mitigation via CNN with Logarithmic Preprocessing ● Geonu Kim, Yong-Ho Cho (553)
- Improving 3D Facial Feature Extraction based on Multi-angle 2D Image Generation using StyleGAN ● Hee-Yeol Lee, Seung-Ho Lee (560)
- An Immersive Media Recognition Method Using Feature Information of Multi-view Videos with Depth Information ● Byeongchan Park, Seyoung Jang, Seok-Yoon Kim, Youngmo Kim (566)
- An Illegal Streaming Video Identify Method Using low-frequency component of fast Fourier transform ● Injae Yoo, Seyoung Jang, Byeongchan Park, Seok-Yoon Kim, Youngmo Kim (573)
- Effect of work function variation of the source electrode on electrical properties of N-type organic thin-film transistors ● Jun-Hyuk Park, Woo-Seok Kim, Min-Hoi Kim (580)
- Robust Emotional Speech Recognition using Stochastic Matching Method ● Weon-Goo Kim (585)
- Study on improving 3D object detection performance through point cloud augmentation technique based on Camera+LiDAR ● Seung-Tak Ra, Seung-Ho Lee (593)
- CNN-Based Log-Mel Spectrogram Image Compression Method for Attention Noise Reduction in Speech Emotion Recognition ● Jeong-Yoon Kim, Seung-Ho Lee (600)
- Error Rate-Based Weighted Ensemble Method for Improving the Performance of Deep Voice Detection ● Joong-Chan Lee, Tae-Hee Lee, Woo-Sung Hwang, Myung-Ryul Choi (607)
- FPGA Design of High-speed Template Matching Block for Vision Wafer Alignment System Implementation ● Seung-Jun Lee, Minjoon Kim (613)
- A Study on Communication Planning Algorithm for the Reconnaissance Mission of Unmanned Ground Vehicle ● Seongjun Jo, Jihoon Kim, Youngkyu Cho, Changhyuk Cho (619)
- Investigation of the effect of trap on the electrical properties of organic unipolar device ● Kyungjae Lee, Eunyong Seo, Dong Hyun Kim, Sinhui Min, Ju-Hong Cha, and Donggu Lee (627)
- Study on Charge Balance Improvement in Electro luminescence Quantum Dot Light-Emitting Diodes through Post-Annealing Process ● Jae Yeong Jeong, Seok Hwan Jang, Jaebum Jeong, Seong Woo Jeong, Hae Ju Kwon, Dae Yun Kim, Yeong Uk Kim, Byeong Guk Jeong, Dong Ryeol Whang, Jun Young Kim (635)
- A Cascaded Structure of TOA/DOA Estimator Using DMRS for 5G-NR-V2X ● Seonghyun Jang, Sanghun Yoon (644)
- A study of glare analysis system using a Raspberry Pi camera and glare contrast ● Jun-Hyeok Heo, In-Gu Kang, Min-Sang Kim, Yoonseuk Choi (651)
- Study on Low-Voltage ESD Protection Device with Improved On-Resistance and Holding Voltage ● Jun-Mo Jung, Joo-Young Lee (657)
- A Study on SCR-Based ESD Protection Circuits with Low Trigger Voltage and High Holding Voltage ● Dong-Hyeon Kim, Jae-Yoon Oh, Min-Seo Kim, Yong-Seo Koo (662)
- Asymmetric Residual U-Net for Crack Detection in High-Resolution Contact Lens Images ● Byeong-Ju Park, Jae-Heung Lee (667)
- Changes in the Electrical Characteristics of GaN MIS-HEMT with TMAH Process Under 5 MeV and 25 MeV Proton Irradiation ● Kyeong Min Kim, Kyung Hee Kim, Yeong Hwan Kim, Jong Beom Im, Gyu Ho Choi, Young Jun Yoon (687)
- Optimization of a Single-core/Multi-layer CNN Accelerator using Non-volatile Memory-based IMC ● Gwan-Oh Youn, Shin-Young Kim, Jun-su Heo, Chester Sungchung Park (696)
- CMOS Active Balun-LNA for Quantum Computing Read-out ● Eunseo Chae, Hyeon-Sik Ahn, Hong Chae, Yoonseuk Choi, Jusung Kim (707)
- A Study on a Design of ESPRIT processor with High-Resolution Angle Estimation Capabilities for High Precision Positioning ● JeonHo Kim, Sungjin Lee, Daegi Hong, Jongwha Chong, Kyeongyuk Min (714)

### [Short Papers]

- A Study on Broadening and Rotating Beam Control Based on Phase Calculation Method for Reflectarray Antenna in Low Earth Orbit Satellite ● Sungil Park, Seongmin Pyo, Jinwoo Jung (722)
- A Study on Phase Weight Search Using Vegetative Propagation by Runners Algorithms for Phase-Only Beam Broadening ● Jinwoo Jung, Seongmin Pyo (726)

# 전기전자학회 논문지

j.inst.korean.electr.electron.eng.

제28권 제4호 2024년 12월

### [논문]

- 회전교차로 주행 협상 서비스 시나리오 정의/구현 및 실도로 실증 ● 이지민, 장성현, 왕유승, 윤상훈, 안병만 (479)
- UWB 및 비전 센서 융합 기반 고정밀 실내 위치 추정 시스템 ● 주희영, 안종화, 윤상훈, 장성현 (489)
- CNN-LSTM 모델을 이용한 DGPS 위치 정확도 향상 기법 ● 이지언, 이태희, 황우성, 최명렬 (497)
- 저조도 환경에서의 객체 검출을 위한 조영 인식 학습 가능한 이미지 필터링 모델 ● 손창환, 김도균, 고현준 (503)
- 저해상도 적외선 이미지 기반 실시간 노인 낙상 검출 ● 김민수, 석종원 (511)
- 멀티모달 텍스트-이미지 정렬을 활용한 블라인드 폭우 얼굴 영상 복원 ● 손창환, 최연우 (520)
- 차원축소 기법을 이용한 전력용 변압기 건전도 지수 평가에 관한 연구 ● 이승윤, 오정식, 김태훈, 박재덕, 안병현, 박태식 (528)
- V2X 통신 기술 분류와 LTE-V2X 통신 성능 테스트 ● 김용태, 신우재, 이정원 (540)
- 양자컴퓨터 Read-out을 위한 4-8GHz 노치필터링을 적용한 위상 주입 잠금 주파수 삼배기 ● 김도현, Hapsah Aulia Azzahra, 채홍, 안현식, 김주성 (546)
- 로그 전처리를 적용한 컨볼루션 신경망 기반 차량 레이더 간섭 경감 ● 김건우, 조용호 (553)
- StyleGAN을 활용한 다각도 2D 이미지 생성 기반의 3D 얼굴 특징점 추출 개선 ● 이희열, 이승호 (560)
- 깊이 정보가 포함된 다시점 영상의 특징정보를 이용한 몰입형 미디어 인식 방법 ● 박병찬, 장세영, 김석윤, 김영모 (566)
- 고속 푸리에 변환의 저주파 성분을 이용한 불법 스트리밍 영상 식별 방법 ● 유인재, 장세영, 박병찬, 김석윤, 김영모 (573)
- 소스 전극의 일함수 변화가 N형 유기 박막 트랜지스터의 전기적 특성에 미치는 영향 ● 박준혁, 김우석, 김민희 (580)
- 확률적 매칭 방법을 이용한 강인한 감정 음성 인식 ● 김원구 (585)
- Camera+LiDAR 기반의 포인트 클라우드 증강 기법을 통한 3D 객체 감지 성능 향상 연구 ● 라승탁, 이승호 (593)
- 음성 감정 인식에서의 어텐션 노이즈 감소를 위한 CNN 기반의 Log-Mel 스펙트로그램 이미지 압축 기법 ● 김정윤, 이승호 (600)
- 딥 보이스 탐지 성능향상을 위한 오인률 기반 가중치 앙상블 기법 ● 이충찬, 이태희, 황우성, 최명렬 (607)
- FPGA 기반 고속 템플릿 매칭 블록 설계를 통한 비전 웨이퍼 얼라인먼트 시스템 구현 ● 이승준, 김민준 (613)
- 무인자동차의 정찰 임무 수행을 위한 통신 계획 알고리즘에 관한 연구 ● 조성준, 김지훈, 조용규, 조창혁 (619)
- 유기 단극 소자의 트랩에 의한 전기적 특성 연구 ● 이경재, 서은용, 김동현, 이주완, 민신희, 차주홍, 이동구 (627)
- 후열처리 공정을 통한 양자점 전계 발광다이오드의 전하 균형 향상 연구 ● 정재영, 장석환, 정재범, 정성우, 권해주, 김대윤, 김영욱, 정병국, 황동렬, 김준영 (635)
- 5G-NR-V2X 통신 기반 DMRS를 이용한 TOA/DOA 추정기 구조 ● 장성현, 윤상훈 (644)
- 라즈베리파이 카메라와 휘도 대비를 활용한 눈부심 분석 시스템 ● 허준혁, 강인구, 김민상, 최윤석 (651)
- 개선된 온-저항 및 홀딩 전압을 갖는 저전압급 ESD 보호소자에 관한 연구 ● 정준모, 이주영 (657)
- 낮은 트리거 전압과 높은 홀딩 전압을 가진 SCR 기반의 ESD 보호회로에 관한 연구 ● 김동현, 오재윤, 김민서, 구용서 (662)
- 고해상도 콘택트렌즈 이미지의 크랙 검출을 위한 Asymmetric Residual U-Net ● 박병주, 이재홍 (667)
- 고장차량 감지를 위한 V2V 통신 기반 응용 서비스 구현 ● 정태완, 민해식, 김태원 (674)
- 5 MeV 및 25 MeV 양성자 조사에서 TMAH 공정을 적용한 GaN MIS-HEMT의 전기적 특성 변화 ● 김경민, 김경희, 김영환, 임종범, 최규호, 윤영준 (687)
- 비휘발성 메모리 기반 IMC를 활용한 단일 코어/다중 레이어 CNN 가속기 최적화 ● 윤관오, 김신영, 허준수, 박성정 (696)
- 양자 컴퓨터 Read-out을 위한 CMOS Active Balun-LNA ● 채은서, 안현식, 채홍, 최윤석, 김주성 (707)
- 고정밀 측위를 위한 고해상도 각도 추정 기능을 갖는 ESPRIT processor 설계에 관한 연구 ● 김전호, 이성진, 홍대기, 정정화, 민경욱 (714)

### [단편]

- 저궤도 위성 반사배열 안테나를 위한 위상 제어 기반 빔 확장 및 회전에 관한 연구 ● 박성일, 표성민, 정진우 (722)
- 위상 전용 빔 확장을 위한 VPR 알고리즘을 이용한 위상 가중치 탐색에 관한 연구 ● 정진우, 표성민 (726)



j.inst.Korean.electr.electron.eng.(2024년 12월/제28권 제4호)

## 목 차

### 논 문

화전교차로 주행 협상 서비스 시나리오 정의/구현 및 실도로 실증 .....	이 지 민, 장 성 현, 왕 유 승, 윤 상 훈, 안 병 만 (479)
UWB 및 비전 센서 융합 기반 고정밀 실내 위치 추정 시스템 .....	주 희 영, 안 종 화, 윤 상 훈, 장 성 현 (489)
CNN-LSTM 모델을 이용한 DGPS 위치 정확도 향상 기법 .....	이 지 언, 이 태 희, 황 우 성, 최 명 렬 (497)
저조도 환경에서의 객체 검출을 위한 조명 인식 학습 가능한 이미지 필터링 모델 .....	손 창 환, 김 도 균, 고 현 준 (503)
저해상도 적외선 이미지 기반 실시간 노인 낙상 검출 .....	김 민 수, 석 종 원 (511)
멀티모달 텍스트-이미지 정렬을 활용한 블라인드 폭우 얼굴 영상 복원 .....	손 창 환, 최 연 우 (520)
차원축소 기법을 이용한 전력용 변압기 건전도 지수 평가에 관한 연구 .....	이 승 윤, 오 정 식, 김 태 훈, 박 재 덕, 안 병 현, 박 태 식 (528)
V2X 통신 기술 분류와 LTE-V2X 통신 성능 테스트 .....	김 용 태, 신 우 재, 이 정 원 (540)
양자컴퓨터 Read-out을 위한 4-8GHz 노차필터링을 적용한 위상 주입 잠금 주파수 삼배기 .....	김 도 현, Hapsah Aulia Azzahra, 채 흥, 안 현 식, 김 주 성 (546)
로그 전처리를 적용한 컨볼루션 신경망 기반 차량 레이더 간섭 경감 .....	김 건 우, 조 용 호 (553)
StyleGAN을 활용한 다각도 2D 이미지 생성 기반의 3D 얼굴 특징점 추출 개선 .....	이 희 열, 이 승 호 (560)
깊이 정보가 포함된 다시점 영상의 특징정보를 이용한 몰입형 미디어 인식 방법 .....	박 병 찬, 장 세 영, 김 석 윤, 김 영 모 (566)
고속 푸리에 변환의 저주파 성분을 이용한 불법 스트리밍 영상 식별 방법 .....	유 인 재, 장 세 영, 박 병 찬, 김 석 윤, 김 영 모 (573)
소스 전극의 일함수 변화가 N형 유기 박막 트랜지스터의 전기적 특성에 미치는 영향 .....	박 준 혁, 김 우 석, 김 민 회 (580)
확률적 매칭 방법을 이용한 강인한 감정 음성 인식 .....	김 원 구 (585)
Camera+LiDAR 기반의 포인트 클라우드 증강 기법을 통한 3D 객체 감지 성능 향상 연구 .....	라 승 탁, 이 승 호 (593)
음성 감정 인식에서의 어텐션 노이즈 감소를 위한 CNN 기반의 Log-Mel 스펙트로그램 이미지 압축 기법 .....	김 정 윤, 이 승 호 (600)
딥 보이스 탐지 성능향상을 위한 오인률 기반 가중치 양상별 기법 .....	이 중 찬, 이 태 희, 황 우 성, 최 명 렬 (607)
FPGA 기반 고속 템플릿 매칭 블록 설계를 통한 비전 웨이퍼 얼라인먼트 시스템 구현 .....	이 승 준, 김 민 준 (613)
무인자살차량의 정찰 임무 수행을 위한 통신 계획 알고리즘에 관한 연구 .....	조 성 준, 김 지 훈, 조 용 규, 조 창 혁 (619)
유기 단극 소자의 트랩에 의한 전기적 특성 연구 .....	이 경 재, 서 은 용, 김 동 현, 이 주 완, 민 신 희, 차 주 흥, 이 동 구 (627)
후열처리 공정을 통한 양자점 전계 발광다이오드의 전하 균형 향상 연구 .....	정 재 영, 장 석 환, 정 재 범, 정 성 우, 권 해 주, 김 대 윤, 김 영 욱, 정 병 국, 황 동 렬, 김 준 영 (635)
5G-NR-V2X 통신 기반 DMRS를 이용한 TOA/DOA 추정기 구조 .....	장 성 현, 윤 상 훈 (644)
라즈베리파이 카메라와 휘도 대비를 활용한 눈부심 분석 시스템 .....	허 준 혁, 강 인 구, 김 민 상, 최 윤 석 (651)
개선된 온-저항 및 홀딩 전압을 갖는 저전압급 ESD 보호소자에 관한 연구 .....	정 준 모, 이 주 영 (657)
낮은 트리거 전압과 높은 홀딩 전압을 가진 SCR 기반의 ESD 보호회로에 관한 연구 .....	김 동 현, 오 재 윤, 김 민 서, 구 용 서 (662)
고해상도 콘택트렌즈 이미지의 크랙 검출을 위한 Asymmetric Residual U-Net .....	박 병 주, 이 재 흥 (667)
고장차량 감지를 위한 V2V 통신 기반 응용 서비스 구현 .....	정 태 완, 민 해 식, 김 태 원 (674)
5 MeV 및 25 MeV 양성자 조사에서 TMAH 공정을 적용한 GaN MIS-HEMT의 전기적 특성 변화 .....	김 경 민, 김 경 희, 김 영 환, 임 종 범, 최 규 호, 윤 영 준 (687)
비휘발성 메모리 기반 IMC를 활용한 단일 코어/다중 레이어 CNN 가속기 최적화 .....	윤 관 오, 김 신 영, 허 준 수, 박 성 정 (696)
양자 컴퓨터 Read-out을 위한 CMOS Active Balun-LNA .....	채 은 서, 안 현 식, 채 흥, 최 윤 석, 김 주 성 (707)
고정밀 측위를 위한 고해상도 각도 추정 기능을 갖는 ESPRIT processor 설계에 관한 연구 .....	김 전 호, 이 성 진, 홍 대 기, 정 정 화, 민 경 욱 (714)

### 단 편

저궤도 위성 반사배열 안테나를 위한 위상 제어 기반 빔 확장 및 회전에 관한 연구 .....	박 성 일, 표 성 민, 정 진 우 (722)
위상 전용 빔 확장을 위한 VPR 알고리즘을 이용한 위상 가중치 탐색에 관한 연구 .....	정 진 우, 표 성 민 (726)

\* 이 학술지는 정부재원(과학기술진흥기금)으로 한국과학기술단체총연합회의 지원을 받아 출판되었음

# Camera+LiDAR 기반의 포인트 클라우드 증강 기법을 통한 3D 객체 감지 성능 향상 연구

## Study on improving 3D object detection performance through point cloud augmentation technique based on Camera+LiDAR

라 승 탁\*, 이 승 호\*\*

Seung-Tak Ra\*, Seung-Ho Lee\*\*

### Abstract

In this paper, we propose a method to improve the performance of 3D object detection using a camera+LiDAR-based point cloud augmentation technique. First, we perform pseudo-LiDAR generation and log-scale intensity reconstruction using camera images. Log-scale intensity reconstruction includes distance-related patterns, so that the learning model can learn spatial structures and distance information between objects more effectively. Second, we augment the point cloud by fusion of pseudo-LiDAR and LiDAR to fill in the low-resolution or undetected areas of LiDAR with pseudo-LiDAR points, thereby increasing the overall number of points and clearly expressing the outline of the object. The augmented point cloud provides more points and various features, and the PV-RCNN model can effectively utilize them. In order to evaluate the objective reliability of the proposed technique, we compared the performance with other deep learning models by calculating mAP using the KITTI dataset. The experimental results show that the performance of 3D object detection using the camera+LiDAR-based point cloud augmentation technique proposed in this paper is superior to that of existing methods. In particular, mAP increased by about 0.37 compared to PV-RCNN, which was used as a learning model. The efficiency of the technique proposed in this paper was proven as it showed excellent results in the experiment.

### 요 약

본 논문에서는 카메라+LiDAR 기반의 포인트 클라우드 증강 기법을 통해 3D 객체 감지 성능을 향상하는 방법을 제안한다. 첫 번째로, 카메라 이미지로 Pseudo-LiDAR 생성 및 로그 스케일 방식의 intensity 재구성을 수행하며 로그 스케일 방식의 intensity 재구성은 거리와 관련된 패턴을 포함하므로 학습 모델이 공간적 구조와 객체 간의 거리 정보를 더 효과적으로 학습할 수 있도록 한다. 두 번째로, Pseudo-LiDAR와 LiDAR 융합으로 포인트 클라우드를 증강하여 LiDAR의 낮은 해상도나 감지되지 않은 영역을 Pseudo-LiDAR 포인트로 채워 전체적인 포인트 수를 늘리고, 객체의 윤곽선을 명확하게 표현할 수 있도록 한다. 증강된 포인트 클라우드는 더 많은 포인트와 다양한 특성을 제공하며, PV-RCNN 모델은 이를 효과적으로 활용할 수 있게 된다. 제안하는 기법의 객관적인 신뢰성을 평가하기 위해 KITTI dataset으로 다른 딥러닝 모델들과 mAP를 산출하여 성능을 비교하였다. 실험 결과, 본 논문에서 제안한 Camera+LiDAR 기반의 포인트 클라우드 증강 기법을 통한 3D 객체 감지의 성능이 기존의 방법들보다 우수한 성능을 나타내었다. 특히 학습 모델로 쓰였던 PV-RCNN과 비교하여 mAP가 약 0.37 증가하였다. 실험에서 우수한 결과를 나타내므로 본 논문에서 제안한 기법의 효율성이 입증되었다.

*Key words : Autonomous Driving, Lidar, 3D Object Detection, Sensor Fusion, Point Cloud Augmentation*

\* Dept. Electronic Engineering, Hanbat National University

★ Corresponding author

E-mail : shlee@cad.hanbat.ac.kr, Tel : +82-42-821-1137

※ Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT)(No. NRF-2022R1F1A1066371)

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) support program(IITP-2024-RS-2022-00156212) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation)

Manuscript received Nov. 25, 2024, revised Dec. 9, 2024, accepted Dec. 22, 2024.

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

자율주행 시스템에서의 객체 감지 성능은 차량의 주행 안전성에 핵심적인 요소로 작용하며, 이를 위해 고성능의 센서 융합 기법[1, 2]은 필수적이다. 현재 자율주행 기술에서는 주변 환경을 정확하게 파악하기 위해 LiDAR (Light Detection and Ranging)와 카메라를 함께 사용하는 것이 일반적이다. LiDAR는 차량 주변의 거리 정보를 고해상도로 제공하지만, 데이터 밀도가 낮고 작은 객체나 먼 거리의 물체를 감지하는 데 한계[3]가 있다. 반면 카메라는 텍스처와 색상 같은 시각적 정보를 풍부하게 제공하며, 특히 원거리 물체나 LiDAR가 놓치는 세부 정보까지 포착할 수 있다. 이러한 두 센서의 융합은 자율주행 시스템의 3D 객체 감지 성능을 향상하는 데 중요한 역할을 할 수 있다. 최신 센서 융합 방법들은 카메라 이미지에서 2D instance segmentation을 수행한 결과를 LiDAR 측에 융합하는 방법[4], LiDAR의 포인트 클라우드에서 3D instance segmentation을 수행 후 두 센서의 segmentation 결과를 융합하는 방법[5] 등과 같이 두 센서의 장단점을 극대화하려는 방법이 제안되고 있다.

최근에는 카메라 기반의 Pseudo-LiDAR[6] 기술이 주목받고 있다. Pseudo-LiDAR는 카메라 이미지를 통해 depth 정보를 예측하여 생성되는 가상 포인트 클라우드로, 특히 원거리 또는 LiDAR로는 인식이 어려운 객체에 대한 정보를 제공하여 LiDAR의 약점을 보완하는 데 유리하다. 하지만 depth 추정 오차로 인해 생성된 데이터의 정확도가 낮을 수 있다. 이러한 카메라와 LiDAR 각 센서의 장단점은 센서 융합을 통해 서로의 단점을 보완하여 객체 감지 성능을 한층 더 개선할 수 있다. 카메라+LiDAR 융합 기법은 자율주행 연구의 중요한 주제로

자리 잡고 있으며, 다양한 연구들이 이를 개선하고 최적화하는 방향으로 발전하고 있다.

본 논문은 카메라+LiDAR 기반의 포인트 클라우드 증강 기법을 통해 3D 객체 감지 성능을 향상시키는 방법을 제안한다. Pseudo-LiDAR를 이용하여 기존 LiDAR 포인트 클라우드의 부족한 부분을 보완하고, 이를 활용한 증강된 포인트 클라우드를 통해 더욱 정밀한 3D 객체 감지를 구현한다. 이에 따른 결과는 향후 다양한 센서 융합 기술 개발과 자율주행 시스템의 안전성을 높이는 데 기여할 것으로 예상된다.

## II. 본론

### 1. 전체 개요

본 논문에서 제안하는 Camera+LiDAR 기반의 포인트 클라우드 증강 기법을 통한 3D 객체 감지 성능 향상 연구의 전체 개요도는 그림 1과 같다. 첫 번째로, 카메라 이미지로 획득한 depth 정보를 기반으로 intensity

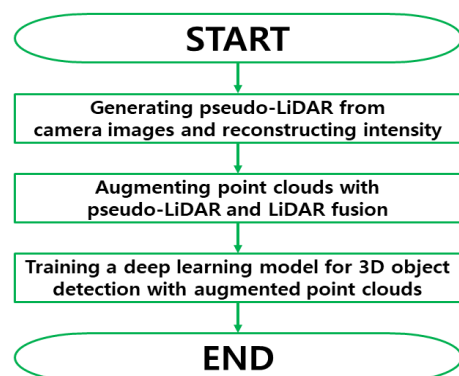


Fig. 1. A complete overview of the proposed method in this paper.

그림 1. 본 논문에서 제안한 방법의 전체 개요도

를 재구성한 Pseudo-LiDAR를 생성한다. 두 번째로, Pseudo-LiDAR와 기존의 LiDAR를 융합하여 부족한 영역을 보완할 수 있는 증강된 포인트 클라우드를 획득한다. 마지막으로, 카메라+LiDAR 센서로 융합된 포인트 클라우드를 이용하여 3D 객체 감지 딥러닝 모델을 학습한다.

## 2. 카메라 이미지로 Pseudo-LiDAR 생성 및 intensity 재구성

Pseudo-LiDAR는 그림 2와 같이 카메라 이미지를 사용해 depth 정보[7]를 추정하여 기존의 LiDAR 센서와 유사한 가상의 포인트 클라우드를 생성하는 방법이다. LiDAR는 포인트 클라우드를 제공하는 센서이기에 Pseudo-LiDAR로 생성된 가상의 포인트 클라우드는 실제 LiDAR와 결합할 때 각 센서의 장점을 극대화할 수 있다. 이 방법은 저비용의 RGB 카메라 이미지를 활용할 수 있으며, 원거리의 물체 감지 같은 특정 상황에서 LiDAR의 한계를 보완할 수 있다.

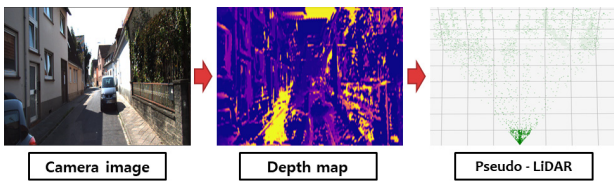


Fig. 2. Generating Pseudo-LiDAR from Camera Images.  
그림 2. 카메라 이미지로 Pseudo-LiDAR 생성

Pseudo-LiDAR의 저자는 depth map을 생성하기 위해 단안 카메라만을 사용한 DORN(Deep Ordinal Regression Network)[8]과 스테레오 카메라를 사용한 PSMnet(Pyramid stereo matching network)[9]의 2가지 딥러닝 기반 네트워크를 참조하였다. DORN은 다중 스케일 특징과 순서형 회귀를 활용하여 단일 이미지에서 depth를 예측하는 딥러닝 네트워크이고, PSMnet은 disparity 추정을 위해 삼 네트워크를 활용하여 스테레오 이미지에서 depth를 예측하는 스테레오 비전 접근 방식이다. Pseudo-LiDAR 논문의 KITTI dataset[10]을 통한 실험에서는 두 방법 모두 효과적이지만 PSMnet을 사용한 스테레오 비전 방식이 정확도 측면에서 단안 카메라 접근법을 능가하였다. KITTI dataset은 스테레오 이미지와 정밀한 차량 위치 데이터를 제공하여 스테레오 기반 depth 추정 연구에 적합한 환경을 제공한다. 만약 활용할 dataset이나 자율주행 환경이 스테레오 이미지를 제공하지 않는다면 DORN과 같은 단안 카메라 기

반의 기법을 활용하여 depth map을 생성하여야 한다.

본 논문에서는 스테레오 매칭 기반의 depth map을 생성하여 각 픽셀의 depth를 추정하고, 이를 통해 가상의 포인트 클라우드인 Pseudo-LiDAR를 생성한다. 스테레오 매칭 기반은 식 1과 같이 좌우 이미지 쌍의 disparity를 계산해 픽셀의 depth를 추정하며  $f$ 는 초점 거리,  $B$ 는 두 카메라 간의 베이스라인 거리를 의미한다. disparity는 PSMnet을 통해 계산되며, PSMnet은 두 개의 입력 이미지(왼쪽, 오른쪽)를 동시에 처리하는 삼 네트워크 구조를 사용하여 두 이미지 간의 차이를 효과적으로 학습할 수 있다.

$$depth = \frac{f \cdot B}{disparity} \quad (1)$$

이렇게 생성된 depth map과 이미지의 픽셀 좌표를 사용하여 각 픽셀을 3D 공간의 점으로 투영할 수 있으며, 이에 대한 식은 식 2, 식 3과 같다. 식에 따라 픽셀 좌표( $u, v$ )는 3D 좌표( $x, y, z$ )로 변환될 수 있으며  $c_x, c_y$ 는 카메라 중심 좌표이며,  $f_x, f_y$ 는 초점거리를 나타낸다.

$$z = depth(u, v) \quad (2)$$

$$x = \frac{(u - c_x) \cdot z}{f_x}, y = \frac{(v - c_y) \cdot z}{f_y} \quad (3)$$

마지막으로  $x, y, z$  형태의 3D 좌표로 생성된 포인트 클라우드를 LiDAR 포맷과 유사한  $x, y, z, intensity$ 의 4차원 구조로 변환하게 된다. intensity는 레이저 빛이 물체에 반사된 강도를 나타내며, 주로 금속과 같은 고반사율 표면이나 근거리에서 1에 가까운 값을 나타내고 아스팔트 같은 저반사율 표면이나 원거리에서는 0에 가까운 값을 나타낸다. 기존 Pseudo-LiDAR는 보편적으로 intensity를 1 혹은 0으로 초기화하여 사용하게 된다. 하지만 이러한 동일 값의 초기화는 학습 모델이 intensity에 대한 추가적인 정보를 학습할 수 없고, 거리와의 상관관계가 없어진다.

본 논문에서는 intensity를 거리 값  $z$ 에 따라 로그 스케일로 변환하여 재구성하였으며 그 식은 식 4와 같다. 식 4에서의 70은 최대 약 70m까지의 거리를 활용하려는 뜻이며, 약 70m 이상의 음수로 출력되는 값들은 0으로 처리하였다. 로그 스케일 방식은 LiDAR에서 가까운 객체가 더 강한 반사를 생성하는 특성을 모방하여, 가상 포인트 클라우드의 물리적 현실감을 높인다. 또한 원거리의 거리 값도 적절히 반영할 수 있도록 intensity를 부

드럽게 조정하기 때문에 거리가 먼 객체가 감지되지 못하는 문제를 완화할 수 있다. 따라서 본 논문의 로그 스케일 방식의 intensity 재구성은 거리와 관련된 패턴을 포함하므로, 학습 모델이 공간적 구조와 객체 간의 거리 정보를 더 효과적으로 학습할 수 있게 된다. 이는 특히 객체가 겹쳐 있거나 복잡한 환경에서 성능 향상에 도움을 줄 수 있다.

$$intensity = \log_{70}(70 - z) \tag{4}$$

### 3. Pseudo-LiDAR와 LiDAR 융합으로 포인트 클라우드 증강

본 논문에서는 그림 3과 같이 Pseudo-LiDAR와 LiDAR를 효과적으로 결합하기 위해 데이터 레벨 융합 (data-level fusion) 방식을 채택한다. 먼저 두 센서의 포인트 클라우드는 기본적으로 다른 방식으로 생성된 데이터이기 때문에 좌표계 및 범위 정규화를 수행하여 동일한 데이터 형식으로 변환한다. 그 후, 각 포인트 클라우드의 밀도를 복셀화하여 비교하고 LiDAR 포인트 클라우드에서 밀도가 적은 복셀을 보완하는 방식으로 융합을 진행한다.

Pseudo-LiDAR는 카메라 기반으로 생성되므로 시야 범위가 더 넓어 원거리의 객체까지 포인트 클라우드가 생성되어 LiDAR의 거리적 스캔 한계를 보완할 수 있으며, 이를 통해 LiDAR의 낮은 해상도나 감지되지 않은 영역을 Pseudo-LiDAR 포인트로 채워 전체적인 포인트 수를 늘릴 수 있다. 또한 객체의 윤곽선 등의 세부적인 3D 형상을 캡처하여, 모델이 객체를 정확히 인식하도록 할 수 있다. 따라서, 융합된 포인트 클라우드는 더욱 균일한 밀도를 가지며, 3D 객체 감지 성능을 향상시킬 수 있다.

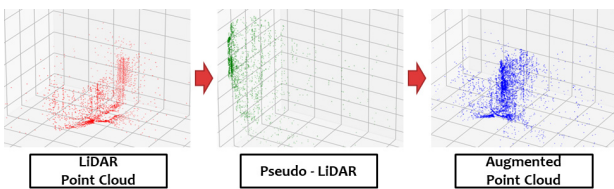


Fig. 3. Augmenting Point Clouds with Pseudo-LiDAR and LiDAR Fusion.

그림 3. Pseudo-LiDAR와 LiDAR 융합으로 포인트 클라우드 증강

### 4. 증강된 포인트 클라우드로 3D 객체 감지 딥러닝 모델 학습

Pseudo-LiDAR와 LiDAR 융합으로 증강된 포인트 클

라우드는 딥러닝 기반의 3D 객체 감지 모델의 입력으로 활용되며 학습된다. 본 논문에서는 Voxel-based 및 Point-based 네트워크 둘의 장점을 합친 PV(Point-Voxel)-RCNN 딥러닝 모델[3]을 적용하여 3D 객체 감지의 학습을 진행하였다. Voxel-based 접근법[11]은 융합된 포인트 클라우드를 격자 형태로 변환하여 모델 입력으로 사용하며, 객체의 구조적 특징을 포착하는 데 유리하다. 반면, Point-based 네트워크는 융합된 포인트 클라우드를 그대로 사용하여 객체의 세부 정보를 더욱 정밀하게 파악할 수 있다.

PV-RCNN의 핵심은 Voxelization과 Point-wise Feature 학습의 결합이다. Voxel 방식은 희소 포인트 클라우드를 효율적으로 처리할 수 있지만, 세부 정보 손실 문제가 있다. Point-wise 방식은 개별 포인트의 세밀한 정보를 학습할 수 있으나, 계산량이 많아 대규모 데이터 처리에 비효율적이다. PV-RCNN은 이 둘을 결합하여 Voxelization을 통해 계산량을 줄여 효율성을 극대화 하고, Point-wise feature를 활용해 미세 정보를 복구하여 정확성을 높였다. PV-RCNN 모델의 구조는 그림 4와 같다.

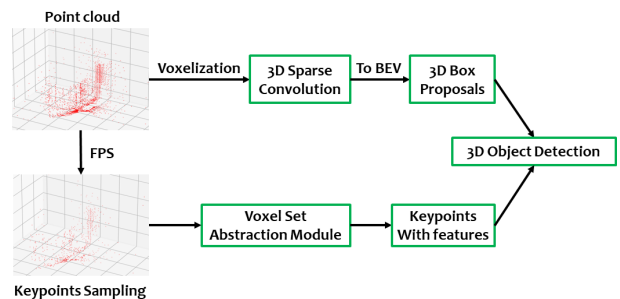


Fig. 4. PV-RCNN model structure.

그림 4. PV-RCNN 모델 구조

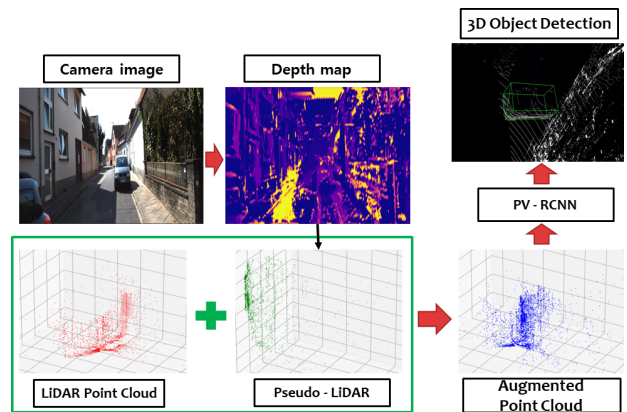


Fig. 5. The entire 3D object detection process.

그림 5. 전체 3D 객체 감지 과정

Pseudo-LiDAR와 LiDAR 융합으로 증강된 포인트 클라우드는 더 많은 포인트와 다양한 특성을 제공하며, PV-RCNN 모델은 이를 효과적으로 활용할 수 있다. Voxelization 단계에서 데이터 밀도가 높아져 Voxel Feature Encoding의 성능이 향상되며, Point-wise feature 학습에서 Pseudo-LiDAR의 intensity 정보나 LiDAR의 정확한 거리 정보를 활용하여, 각 포인트의 특성을 세밀히 보완할 수 있다. 그림 5는 Pseudo-LiDAR와 LiDAR 융합으로 증강된 포인트 클라우드의 전체 3D 객체 감지 과정을 나타낸다.

## 5. 성능 실험

### 가. 실험 환경

본 논문에서 제안한 Camera+LiDAR 기반의 포인트 클라우드 증강 기법을 통한 3D 객체 감지 성능 향상 연구 실험에 사용된 하드웨어 사양은 Intel Xeon(R) Silver 4215R와 Nvidia RTX 3090 4개, 128GB RAM으로 구성되어 있다. 또한 개발 환경은 Ubuntu 20.04 OS, CUDA 11.3, cudnn 8.2.0, pytorch 1.12.1, MMDetection3D 1.4.0 등의 라이브러리를 사용하였다. 3D 객체 감지의 학습 및 테스트에는 KITTI dataset[10]을 사용하였다. 실험의 경우 train set을 3,712개의 train samples과 3,769개의 validation samples로 나눈 후 학습은 Epoch 40으로 동일하게 설정하여 학습하였고, 테스트는 validation samples로 진행하였다. 3D 객체 감지 딥러닝 모델은 PV-RCNN 모델을 기반으로 하였다. Car 클래스의 경우 IoU는 0.7, Pedestrian과 Cyclist 클래스는 IoU 0.5로 고정하고 AP 11 측정 항목과 Moderate 기준으로 테스트를 진행하였다. 측정된 mAP는 5번 반복된 실험을 통한 평균값이다.

### 나. 실험 결과

본 논문에서 제안한 Camera+LiDAR 기반의 포인트 클라우드 증강 기법을 통한 3D 객체 감지 성능 향상 연구에 대한 실험 결과는 표 1과 같다. 표 1은 기존 기법들과의 3D 객체 감지 성능 비교 결과를 나타낸다. MVX-Net[1]은 LiDAR와 카메라 데이터 융합을 활용한 3D 객체 감지의 기초 모델로 LiDAR의 정확한 거리 정보와 카메라의 풍부한 RGB 정보를 Early Fusion 방식으로 결합하여, 객체 감지 성능을 향상시켰다. Part-A2[2]는 포인트 클라우드의 구조적 특징을 활용해 객체 감지 성능을 향상시켰다. 이 모델은 객체의 Parts 정보를 추출하고 이를 효율적으로 집계하여, 전체 객체 수준의 특

성을 강화하여 기존 방식보다 복잡한 구조를 가진 객체도 정확히 감지가 가능하다. 또한 intensity를 로그 스케일로 변환하여 재구성하는 방식의 정량적 평가를 위해 intensity를 1로 초기화하여 Pseudo-LiDAR를 생성한 경우의 실험도 진행하였다.

실험 결과 본 논문에서 제안한 Camera+LiDAR 기반의 포인트 클라우드 증강 기법을 통한 3D 객체 감지의 성능이 기존의 방법들보다 우수한 성능을 나타내었다. 특히 학습 모델로 쓰였던 PV-RCNN[3]과 비교하여 mAP가 약 0.37 증가하였다. Pseudo-LiDAR와 LiDAR 융합으로 증강된 포인트 클라우드는 Voxel Feature Encoding 단계에서 더 다양한 공간 정보를 제공하기에 모델이 더 나은 물체 후보를 생성한 것으로 예상된다. 또한 멀리 있는 객체에서 희소성이 문제가 되는 LiDAR의 한계를 Pseudo-LiDAR가 보완한 것으로 예상된다. intensity를 1로 초기화하여 Pseudo-LiDAR를 생성한 경우 고정된 intensity 값으로 인해 객체 간의 세밀한 차이를 학습하기 어렵고 거리와의 상관관계가 없어진다. 이에 따라 상대적으로 단순한 환경에서는 성능 차이가 크지 않을 수 있으나, 복잡한 환경에 대해 성능이 저하된 것으로 예상된다. 그림 6은 제안된 방법의 KITTI dataset 3D 객체 감지 결과를 나타낸다.

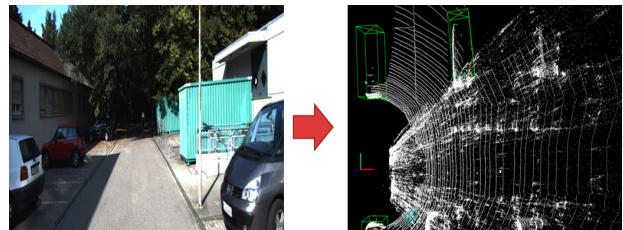


Fig. 6. KITTI dataset 3D object detection results of the proposed method.

그림 6. 제안된 방법의 KITTI dataset 3D 감지 결과

Table 1. Results of comparison of 3D object detection performance with existing techniques.

표 1. 기존 기법들과의 3D 객체 감지 성능 비교 결과

No	Method	mAP ↑
1	MVX-Net[1]	63.5
2	Part-A2[2]	68.33
3	PV-RCNN[3]	72.28
4	Pseudo-LiDAR(intensity:1) + LiDAR + PV-RCNN	72.5
5	The proposed method (intensity : log scale)	72.65

### III. 결론

본 논문에서는 카메라+LiDAR 기반의 포인트 클라우드 증강 기법을 통해 3D 객체 감지 성능을 향상하는 방법을 제안하였다. 첫 번째로, 카메라 이미지로 Pseudo-LiDAR 생성 및 intensity 재구성을 수행한다. Pseudo-LiDAR는 LiDAR가 감지하기 어려운 원거리의 객체나, 작은 물체의 세부 정보를 보완할 수 있다. 또한 로그 스케일 방식의 intensity 재구성은 거리와 관련된 패턴을 포함하므로, 학습 모델이 공간적 구조와 객체 간의 거리 정보를 더 효과적으로 학습할 수 있도록 한다. 두 번째로, Pseudo-LiDAR와 LiDAR 융합으로 포인트 클라우드를 증강한다. 이를 통해 LiDAR의 낮은 해상도나 감지되지 않은 영역을 Pseudo-LiDAR 포인트로 채워 전체적인 포인트 수를 늘리고, 객체의 윤곽선을 명확하게 표현할 수 있다. 마지막으로, 증강된 포인트 클라우드로 3D 객체 감지 딥러닝 모델을 학습한다. 증강된 포인트 클라우드는 더 많은 포인트와 다양한 특성을 제공하며, PV-RCNN 모델은 이를 효과적으로 활용할 수 있다. 제안하는 기법의 객관적인 신뢰성을 평가하기 위해 KITTI dataset으로 다른 딥러닝 모델들과 mAP를 산출하여 성능을 비교하였다. 실험 결과, 본 논문에서 제안한 Camera+LiDAR 기반의 포인트 클라우드 증강 기법을 통한 3D 객체 감지의 성능이 기존의 방법들보다 우수한 성능을 나타내었다. 특히 학습 모델로 쓰였던 PV-RCNN과 비교하여 mAP가 약 0.37 증가하였다. 실험에서 우수한 결과를 나타내므로 본 논문에서 제안한 기법의 효율성이 입증되었다.

한편, 본 논문에서는 기존의 3D 객체 감지 과정에서 Pseudo-LiDAR를 생성하고 LiDAR와 융합하는 과정이 추가되었기 때문에 기존의 추론 시간에서 약 0.3초가 추가로 소모되는 성능 향상에 따른 비용이 발생한다. 이러한 추론 시간에 대한 문제는 자율주행에 있어 치명적일 수 있기에 향후 연구과제로 실시간으로 Pseudo-LiDAR를 생성하는 연구가 필요하다고 사료된다. 또한 데이터 레벨 수준의 센서 융합이 아닌 특징 레벨의 센서 융합을 통해 더욱 3D 객체 감지 성능을 향상시킬 수 있는 연구가 필요하다고 사료된다.

### References

[1] Sindagi, Vishwanath A., Yin Zhou, and Oncel Tuzel. "Mvx-net: Multimodal voxelnet for 3d

object detection," *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019.

DOI: 10.48550/arXiv.1904.01649

[2] Shi, Shaoshuai, et al., "From points to parts: 3d object detection from point cloud with part-aware and part-aggregation network," *IEEE transactions on pattern analysis and machine intelligence* vol.43, pp.2647-2664, 2020.

DOI: 10.1109/TPAMI.2020.2977026

[3] Shi, Shaoshuai, et al., "Pv-rcnn: Point-voxel feature set abstraction for 3d object detection," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020.

DOI 10.1109/CVPR42600.2020.01054

[4] Li, Yingyan, et al., "Fully sparse fusion for 3d object detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2024).

DOI: 10.1109/TPAMI.2024.3392303

[5] Xu, Shaoqing, et al., "Multi-sem fusion: multi-modal semantic fusion for 3D object detection," *IEEE Transactions on Geoscience and Remote Sensing* (2024).

DOI: 10.1109/TGRS.2024.3387732

[6] Wang, Yan, et al., "Pseudo-lidar from visual depth estimation: Bridging the gap in 3d object detection for autonomous driving," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019.

DOI: 10.48550/arXiv.1812.07179

[7] Qian, Rui, et al., "End-to-end pseudo-lidar for image-based 3d object detection," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020.

DOI: 10.48550/arXiv.2004.03080

[8] Fu, Huan, et al., "Deep ordinal regression network for monocular depth estimation," *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.

DOI: 10.1109/CVPR.2018.00214

[9] Chang, Jia-Ren, and Yong-Sheng Chen. "Pyramid stereo matching network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.

DOI: 10.1109/CVPR.2018.00567

[10] Geiger, A., Philip Lenz, and Raquel Urtasun. "Are we ready for autonomous driving. The KITTI Vision Benchmark Suite," *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.1109/CVPR.2012.6248074

[11] Zhou, Yin, and Oncel Tuzel. "Voxelnet: End-to-end learning for point cloud based 3d object detection," *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018. DOI: Bookmark: 10.1109/CVPR.2018.00472

---

## BIOGRAPHY

---

### Seung-Tak Ra (Member)



2017 : BS degree in Electronic Engineering, Hanbat National University

2019 : MS degree in Electronic Engineering, Hanbat National University

2022~current : Ph. D degree course of Electronic Engineering Hanbat National University

### Seung-Ho Lee (Member)



1986 : BS degree in Electronic Engineering, Hanyang University

1989 : MS degree in Electronic Engineering, Hanyang University

1994 : Ph. D degree in Electronic Engineering, Hanyang University

1994~current : Professor, Department of Electronic Engineering, Hanbat National University