

ISSN 2287-5026 (Print)
ISSN 2288-159X (Online)

Journal of the Institute of Electronics and Information Engineers

2026 **8** 제 63 권 8호

Vol.63, No.8 August 2026

AI Signal Processing

- 61 Wavelet-Based All-in-One Dual-Branch Adverse Weather Image Restoration with Bidirectional Subband Interaction / Park Su Yeon and Eom Il Kyu
73 UWB Radar–Based Classification of Marine Debris / Jihyeon Yoo and Sungbin Cho, and Jinhwan Koh
84 Improving Detection Performance via R-VAE-Based Synthetic Data Augmentation for Indoor Fire Scenarios / Gyeongmin Kim and Jinhwan Koh
98 Geometric Feature-Based Adaptive Voxel Sampling Technique for LiDAR-Based 3D Object Detection / Seung-Tak Ra and Seung-Ho Lee
106 Water Level Measurement in a Multi-Bent Pipeline Environment Using Ultrasonic Based CNN / Wanyeol Gu, Suyeon Lee, and Jinhwan Koh
116 Local Small LLM-based AI Agent System with Finite State Controller for SCADA Operation Assistance / Jeong-Yoon Kim and Seung-Ho Lee

System and Control

- 127 Development of an AI Model for Quality Prediction and Root Cause Tracing of Cathode Materials Based on FeatureMAP Image Analysis / Jeong-Yun Song, Soo-Rim Kim, Hyun-Woo Oh, and Yong-Kwi Lee

Industry Electronics

- 139 Integrating Non-Local Neural Networks into Deep CNN Image Denoising via Residual Learning / Young-Ro Kim
47 Comparative Study on 2D and 3D CVA Measurement Differences According to Distance Conditions Using an RGB-D Camera / Eun Seong Park and Yongjun Chang

Communications

- 3 Weighted Binary Search Tree for NDN Lookup / Gyubin Kim and Hyesook Lim

Semiconductor and Devices

- 15 Low-Power Tri-Layer Synaptic Device with Filament Control by AI2O3 Insertion / Yu-Bin Kim, Sung-Ho Kim, Dong-Min Kim, and Hi-Deok Lee
23 A Complete Quadruple-Node-Upset Recoverable Latch with Enhanced Quintuple-Node-Upset Self-Recoverability / Suhyeon Kim, Jaeyong Jang, and Taehui Na
32 Analysis of Performance and Resource Utilization for FPGA-Based QR Decomposition Architectures / Suk-Jae Yoon and Hoyoung Yoo
40 Signal Integrity Analysis of Multi-branch Transmission Lines in the Multi-Layer Ceramic of a Probe Card Using Factorial Design of Experiments / Youngjae Kim, Moonjung Kim, and A-Rang Jang

Computer and Information

- 53 Forgery Detection and Characteristic Analysis of Scanned Document Utilizing Paper Fingerprint / YongSoo CHOI

전자공학회 논문지

2026
8

제 63 권
8호

IEIE
사단
법인
대한전자공학회

www.theieie.org

Vol.63, No.8 August 2026

The Institute of Electronics and Information Engineers (IEIE)
Room #907, The Korea Science Technology Center The first building, 22,
Teheran-ro 7 Gil, Gangnam-gu, Seoul, Republic of Korea



전자공학회 논문지

•이 논문집은 한국연구재단 우수등재학술지임.



차 례

2026년 8월

제63권 제8호

TIC / 통신

[미래네트워크]

- 3 NDN 이름 탐색을 위한 가중치 기반 이진 탐색 트리
..... 김규빈, 임혜숙

SD / 반도체

[반도체 소자 및 재료]

- 15 Al_2O_3 삽입을 통한 필라멘트 제어 기반 저전력 삼중층 시냅스 소자 구현
..... 김유빈, 김성호, 김동민, 이희덕

[SoC 설계]

- 23 향상된 오중 노드 업셋 자가회복성을 갖는 사중 노드 업셋 완전회복 래치
..... 김수현, 장재웅, 나태희
- 32 FPGA 기반 QR 분해 아키텍처의 성능 및 리소스 사용량 분석
..... 윤석재, 유호영

[PCB & Packaging]

- 40 요인 실험계획법을 적용한 프로브 카드 Multi-Layer Ceramic 다중 분기 전송선의 신호 무결성 분석
..... 김영재, 김문정, 장아람

CI / 컴퓨터

[멀티미디어]

- 53 전자화문서의 종이지문 특성치 분석과 위·변조 검출
..... 최용수

AISP / 인공지능 신호처리

[영상 신호처리]

- 61 양방향 부밴드 상호작용을 이용한 웨이블릿 기반 올인원 이중 분기 악천후 영상 복원
..... 박수연, 엄일규
- 73 UWB 레이더 기반 해양 쓰레기 분류
..... 유지현, 조성빈, 고진환
- 84 R-VAE 기반 실내 화재 합성 데이터 증강을 통한 판별 성능 향상
..... 김경민, 고진환
- 98 LiDAR 기반 3차원 객체 검출을 위한 기하학 특성 기반의 적응형 복셀 샘플링 기법
..... 라승탁, 이승호

[음향 및 신호처리]

- 106 초음파를 이용한 CNN 기반 다중 굴곡 관 환경 수위 측정
..... 구완열, 이수연, 고진환
- 116 SCADA 운영 보조를 위한 유한상태제어기 적용 로컬 소형 LLM 기반 AI agent 시스템
..... 김정윤, 이승호

SC / 시스템 및 제어

[스마트팩토리]

- 127 FeatureMAP 이미지 분석 기반의 양극재 품질 예측 및 불량 원인 추적 AI 모델 개발
..... 송정윤, 김수림, 오현우, 이용귀

IE / 산업전자

[신호처리 및 시스템]

- 141 잔차 학습 기반 심층 합성곱 신경망의 Non-Local 연산 통합을 통한 영상 잡음 제거 개선
..... 김영로

[컴퓨터 응용]

- 149 RGB-D 카메라 기반 실시간 자세 측정에서 2D와 3D 방식의 정확도 비교
..... 박은성, 장용준

논문 2026-63-8-10

LiDAR 기반 3차원 객체 검출을 위한 기하학 특성 기반의 적응형 복셀 샘플링 기법

(Geometric Feature-Based Adaptive Voxel Sampling Technique
for LiDAR-Based 3D Object Detection)

라 승 탁*, 이 승 호**

(Seung-Tak Ra and Seung-Ho Lee[©])

요 약

본 논문은 LiDAR 기반 3차원 객체 검출에서 제한된 복셀 수 제약 하에 검출에 유효한 복셀을 선별하기 위한 기하학 특성 기반 적응형 복셀 샘플링 기법을 제안한다. 제안 방법은 복셀 내부 포인트 분포와 인접 복셀 간 국소 구조 관계를 바탕으로 여섯 가지 기하학 특성을 정량화한다. 이후 각 특성을 정규화하고 가중 결합하여 후보 복셀의 중요도 점수를 산출하며, 동일한 최종 보존 복셀 수 조건에서 점수가 높은 상위 복셀만을 선택하여 검출기의 입력으로 사용한다. 이를 통해 객체 경계와 구조 정보를 더 잘 보존하는 입력 복셀 집합을 구성하고자 하였다. 제안한 방법의 성능을 검증하기 위하여 PV-RCNN과 KITTI 데이터셋을 활용하여 실험을 수행하였다. 데이터는 학습 세트 3712개, 검증 세트 3769개 장면으로 구성되며 주요 평가 지표로는 KITTI의 Overall 3D AP11 moderate를 사용하고 실용성 분석을 위해 프레임당 평균 추론 시간을 함께 측정하였다. Overall 3D AP11 moderate는 Car, Pedestrian, Cyclist의 세 클래스를 대상으로 moderate 난이도에서의 3차원 검출 성능을 균형 있게 비교할 수 있으며, KITTI 환경에서 대표 성능 지표로 널리 사용된다는 장점이 있다. 반복 실험의 평균 결과에 따르면, 후보 복셀 수 상한을 (48K, 100K)으로 설정한 경우 Overall 3D AP11 moderate는 72.15로 나타나 기존 모델의 70.56 대비 1.59 향상되었다. 또한 평균 추론 시간은 기존 모델 79.48ms, 제안 방법 81.74ms로 측정되어 약 2.26ms의 제한적인 시간 증가만으로 성능 향상을 달성하였다. 추가로 SECOND, PointPillars, Part-A2에 대한 확장 실험에서도 각각 3.18, 0.71, 0.34의 성능 향상이 확인되어, 제안 방법이 특정 검출기에 한정되지 않고 다른 복셀 기반 검출기로도 확장될 수 있음을 보였다.

Abstract

This paper proposes a geometry-aware adaptive voxel sampling method for LiDAR-based 3D object detection under a limited voxel budget. The method quantifies six geometric features from within-voxel point distributions and local neighborhood structure, aggregates them into a voxel importance score, and selects only the highest-ranked candidate voxels under the same final preserved voxel budget, thereby constructing an input voxel set that better preserves object boundaries and structural information. Experiments were conducted with PV-RCNN on the KITTI dataset using 3712 training scenes and 3769 validation scenes. KITTI Overall 3D AP11 moderate was adopted as the primary metric because it enables balanced comparison across the Car, Pedestrian, and Cyclist classes under the moderate difficulty setting, and the average inference time per frame was additionally measured for practicality analysis. With a candidate voxel limit of (48K, 100K), the proposed method achieved an Overall 3D AP11 moderate of 72.15, improving the baseline score of 70.56 by 1.59, while increasing the average inference time only from 79.48ms to 81.74ms. Additional gains of 3.18, 0.71, and 0.34 were also observed on SECOND, PointPillars, and Part-A2, respectively, indicating that the proposed method can be extended to other voxel-based detectors.

Keywords : LiDAR, 3D object detection, Voxel-based detectors, Voxel sampling, Geometric features

*학생회원, **평생회원, 한밭대학교 전자공학과(Dept. Electronic Engineering, Hanbat National University)

[©] Corresponding Author(E-mail : shyolee@hanbat.ac.kr)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 학·석사연계ICT핵심인재양성사업의 연구결과로 수행되었음(IITP-2026-RS-2022-00156212).

Received ; April 24, 2026

Revised ; June 4, 2026

Accepted ; June 11, 2026

I. 서 론

LiDAR 기반 3D 객체 검출은 자율주행, 이동 로봇, 지능형 감시 시스템 등 다양한 응용 분야에서 핵심 기술로 활용되고 있다. 특히 복셀 기반 표현은 불규칙한 포인트 클라우드를 규칙적인 구조로 변환할 수 있으므로 sparse convolution과의 결합이 용이하며, VoxelNet^[1]과 PV-RCNN^[2]은 이러한 복셀 기반 검출기의 대표적인 발전 흐름을 보여준다. 또한 PointNet^[3]과 PointNet++^[4]는 포인트 클라우드의 국소 구조와 계층적 표현 방식이 특징 추출 성능에 직접적인 영향을 미친다는 점을 보여주었으며, 이는 입력 단계에서 어떤 공간 정보를 보존할 것인가가 중요하다는 사실을 시사한다. 여기서 복셀은 2차원 영상의 픽셀을 3차원 공간으로 확장한 개념으로, 일정한 크기의 3차원 격자 셀 하나를 의미한다. 즉, 연속적인 포인트 클라우드를 작은 3차원 셀들로 분할하고 각 셀 내부의 포인트 분포를 하나의 단위로 표현함으로써, 불규칙한 포인트 집합을 규칙적인 격자 구조로 다룰 수 있게 한다. 이러한 특성 때문에 복셀 표현은 3차원 검출 문제에서 연산 효율성과 구조적 일관성을 동시에 확보할 수 있는 장점이 있다.

최근에도 복셀 기반 3D 객체 검출은 빠르게 발전하고 있다. 예를 들어 Voxel Mamba^[5]는 state space model을 이용하여 전체 복셀 공간의 장거리 의존성을 효율적으로 모델링하고, PVTransformer^[6]는 포인트-복셀 변환 과정에서 기존 PointNet pooling의 정보 병목을 완화하기 위해 attention 기반 aggregation을 도입하였다. 이러한 최신 연구들은 복셀 표현 자체의 품질과 공간 정보 보존 방식이 여전히 중요한 연구 주제를 보여준다.

그러나 복셀 기반 검출기의 입력 단계에서는 여전히 제한된 복셀 수 안에서 어떤 복셀을 유지할 것인가가 중요한 문제로 남아 있다. 기존 방식은 입력 포인트를 사전에 무작위로 뒤섞은 뒤, 들어오는 순서대로 복셀을 생성하다가 상한수에 도달하면 동작을 멈추고 나머지를 버리는 방식이기에 복셀의 기하학적 중요도를 직접 반영하는 입력 선별 전략은 상대적으로 부족하였다. 예를 들어 도로 평면이나 단조로운 배경 구조는 많은 영역을 차지하지만 검출에 기여하는 정보량은 제한적일 수 있다. 반대로 보행자 팔, 차량 모서리, 자전거 프레임과 같이 국소적인 구조 변화가 큰 영역은 전체 포인트 수가 많지 않더라도 검출기 입장에서 중요한 단서를 제공한다. 따라서 입력 복셀을 동일하게 취급하기보다, 검

출 유효도에 따라 우선순위를 부여하는 접근이 필요하다. 또한 복셀 선택 과정 자체를 하나의 전처리 최적화 문제로 다루는 관점이 필요하다.

본 논문에서는 이러한 한계를 해결하기 위하여 복셀 밀도, 반사강도 기반 활성화도, 주변 대비 밀도 변화, 곡률 근사치, 법선 변화, 센서 거리의 여섯 가지 기하학 특성을 이용한 적응적 복셀 샘플링 기법을 제안한다. 제안 방법은 hard voxelization 이후 더 큰 후보 복셀 집합을 구성한 뒤 각 후보 복셀의 기하학 점수를 계산하고, 상위 복셀만을 선택하여 기존 PV-RCNN에 입력한다. 또한 각 특성의 결합 가중치와 선택 임계값을 학습 가능하게 설계함으로써 휴리스틱 기반 선별의 경직성을 완화하였다.

본 논문에서 PV-RCNN을 주요 검증 백본으로 사용한 이유는 다음과 같다. PV-RCNN은 복셀 기반 sparse convolution과 포인트 기반 set abstraction을 결합한 point-voxel 구조로, 입력 복셀의 공간적 품질이 voxel feature와 keypoint feature 모두에 영향을 줄 수 있다. 따라서 입력 복셀 선별 전략의 효과를 분석하기에 적합하며, KITTI와 MMDetection3D 환경에서 재현성이 높아 기준 검출기로 사용하기에도 적절하다. 다만 제안 기법이 특정 모델에 종속되지 않음을 확인하기 위해 다른 복셀 기반 검출기에 대한 추가 비교 실험도 함께 수행하였다.

본 논문의 기여는 다음과 같다. 첫째, 검출에 유효한 복셀을 선별하기 위한 기하학 특성 기반 점수화 모델을 제안하였다. 둘째, 기존 검출기 구조를 크게 변경하지 않고 전처리 단계에 적응적 복셀 선택 모듈을 삽입하여 실용적인 통합 구조를 구현하였다. 셋째, KITTI 실험을 통해 후보 복셀 수 상한 변화에 따른 성능 추이를 분석하고, 여러 복셀 기반 검출기에 대한 확장 실험과 특성 중요도 분석을 수행하여 제안 기법의 효과를 검증하였다.

II. 본 론

1. 제안 기법의 개요

제안 기법의 전체 흐름은 그림 1과 같다. 입력 포인트 클라우드는 먼저 hard voxelization으로 변환한 뒤 기존 모델보다 큰 후보 복셀 집합을 먼저 생성하고, 각 후보 복셀에 대해 기하학 특성 기반 복셀 점수화를 수행한다. 이후 적응적 복셀 선택 전략을 통해 최종 보존 복셀 수 상한에 맞는 상위 복셀만 선택되어 복셀 기반 검출기의 sparse encoder로 전달되며, 제안 기법은 입

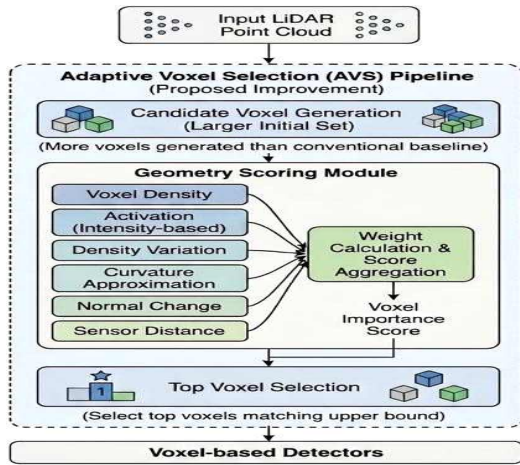


그림 1. 제안 기법의 전체 구조

Fig. 1. Overall pipeline of the proposed method.

력 전처리 단계의 품질 향상으로 해석할 수 있다.

2. 기하학 특성 기반 복셀 점수화

복셀 v_i 의 중요도는 여섯 가지 기하학 특성으로 정의하였다. 여기서 v_i 는 i 번째 복셀, n_i 는 해당 복셀 내부의 포인트 수, $N(i)$ 는 인접 복셀 집합, (x_i, y_i, z_i) 는 복셀 대표 좌표를 의미한다.

식 (1)은 각각 복셀 밀도와 활성도를 의미한다. 밀도는 포인트가 얼마나 집중되어 있는지를 나타내고, 활성도는 반사강도 기반 표면 특성을 반영한다. 여기서 d_i 는 i 번째 복셀의 밀도, a_i 는 활성도, $N_{\max}$ 는 하나의 복셀 안에 허용되는 최대 포인트 수, r_p 는 포인트 p 의 반사강도를 의미한다. 즉, 식 (1)은 복셀 내부 포인트 수와 반사강도 분포를 각각 정규화하여 두 개의 기본 기하학 특성으로 변환한 것이다. 밀도 특성은 전경 객체가 주변 배경보다 상대적으로 응집된 포인트 분포를 보이는 경우에 유용하며, 활성도는 표면 반사 특성의 차이를 통해 배경 구조와 객체 표면을 구분하는 데 도움을 준다. 두 특성은 단독으로 사용할 경우 편향될 수 있으나, 다른 구조적 특성과 결합될 때 보다 안정적인 구분력을 갖는다.

$$d_i = \frac{n_i}{N_{\max}}, \quad a_i = \frac{1}{n_i} \sum_{p \in v_i} |r_p| \quad (1)$$

식 (2)는 주변 대비 밀도 변화, 곡률 근사치, 법선 변화를 나타낸다. 이 세 항은 객체 경계와 구조 변화가 큰 영역에 높은 값을 부여하므로, 평탄한 배경 복셀보다 유의미한 후보를 선별하는 데 유리하다. 식 (2)에서

Δd_i 는 주변 복셀 대비 밀도 변화량, c_i 는 높이 차이를 이용한 곡률 근사치, nv_i 는 법선 변화량을 의미한다. 또한 v_j 는 인접 복셀 집합 $N(i)$ 에 포함된 복셀을 뜻하며, $n_i^T \cdot n_j$ 는 두 복셀 법선 벡터의 내적이다. 따라서 이 식은 복셀 주변의 구조 변화 정도를 수치화하는 역할을 한다.

$$\begin{aligned} \Delta d_i &= \left| d_i - \frac{1}{|N(i)|} \sum_{v_j \in N(i)} d_j \right|, \\ c_i &= \frac{1}{|N(i)|} \sum_{v_j \in N(i)} (z_i - z_j)^2, \\ nv_i &= \frac{1}{|N(i)|} \sum_{v_j \in N(i)} \frac{1 - n_i^T \cdot n_j}{2} \end{aligned} \quad (2)$$

구체적으로 Δd_i 는 현재 복셀의 밀도 d_i 와 인접 복셀들의 평균 밀도 차이를 절댓값으로 계산한 값이다. 따라서 현재 복셀이 주변과 유사한 밀도를 가지면 값이 작아지고, 객체 경계와 같이 주변과 상이한 분포를 가지면 값이 커진다. 한편 c_i 는 현재 복셀의 대표 높이 z_i 와 인접 복셀 높이 z_j 의 차이를 제곱한 뒤 평균한 값으로, 평탄한 지면에서는 작고 높이 변화가 큰 구조에서는 크게 나타난다. 법선 변화량 nv_i 는 인접 복셀 간 높이 기울기로부터 국소 법선을 근사한 뒤, 현재 복셀의 법선과 이웃 복셀 법선 사이의 방향 차이를 평균하여 계산한다. 즉, 표면 방향이 거의 일정한 영역에서는 nv_i 가 작고, 모서리나 외곽선처럼 표면 방향이 급격히 변하는 영역에서는 값이 커지므로, 객체 형상의 구조적 불연속성을 포착하는 데 유리하다. 특히 밀도 변화와 곡률 근사치는 지면이나 벽면과 같이 상대적으로 균일한 구조를 낮게 평가하고, 높이 변화와 표면 굴곡이 발생하는 복셀에 우선순위를 부여한다. 법선 변화는 인접 복셀 간 표면 방향 차이를 반영하므로 객체 외곽선이나 모서리 영역을 유지하는 데 중요한 역할을 한다.

밀도 변화, 곡률 근사치, 법선 변화는 모두 객체 경계 부근에서 높은 값을 가질 수 있으나, 세 특성이 포착하는 물리적 의미는 서로 다르다. 밀도 변화는 현재 복셀이 주변보다 얼마나 다른 점유 분포를 갖는지를 나타내므로 객체와 배경의 분포 대비에 민감하다. 곡률 근사치는 인접 복셀 간 높이 변화의 제곱 평균을 이용하므로 차량 지붕, 보행자 상체, 자전거 프레임과 같이 국소적인 형상 굴곡을 강조한다. 반면 법선 변화는 표면 방향의 불연속성을 측정하므로 포인트 수가 많지 않더라도 외곽선이나 수직 구조가 뚜렷한 영역을 보존하는 데

기여한다. 따라서 세 특성은 동일한 edge 영역을 중복 검출하는 항이 아니라, 점유 대비, 높이 변화, 표면 방향이라는 서로 다른 관점에서 구조적 불연속성을 보완적으로 설명한다.

식 (3)에서 r_i 는 센서 중심과의 거리를 반영하는 거리 사전 정보이며, s_i 는 정규화된 기하학 특성들의 가중 합으로 계산되는 최종 점수이다. 이때 가중치의 합을 1로 제한하여 각 특성의 상대적 기여도를 해석 가능하게 하였다. 여기서 r_i 는 센서 중심으로부터의 거리 보정값, $R_{\max}$ 는 수평면 기준 최대 거리, $\hat{f}_{k,i}$ 는 k 번째 정규화 특성값, w_k 는 해당 특성의 가중치이다. 즉, 식 (3)은 서로 범위가 다른 기하학 특성을 동일 척도에서 결합하여 하나의 최종 중요도 점수로 환산하는 과정으로 이해할 수 있다. 거리 사전 정보는 원거리 영역의 희소성으로 인해 발생할 수 있는 불안정한 선택을 완화하기 위한 보조 특성이다. 다만 본 연구의 목적은 원거리 복셀을 일괄적으로 제거하는 것이 아니라, 동일 거리대 내에서 구조적으로 더 효과적인 복셀에 우선순위를 부여하는 것이므로 다른 특성과 결합한 상대적 평가 방식으로 사용하였다.

$$r_i = 1 - \frac{\| (x_i, y_i) \|_2}{R_{\max}}, \quad s_i = \sum_{k=1}^6 w_k \hat{f}_{k,i}, \quad \sum_{k=1}^6 w_k = 1 \quad (3)$$

그림 2는 복셀 점수 산출 과정을 예시적으로 나타낸다. 그림에서 볼 수 있듯이, 객체의 윤곽이나 표면 변화

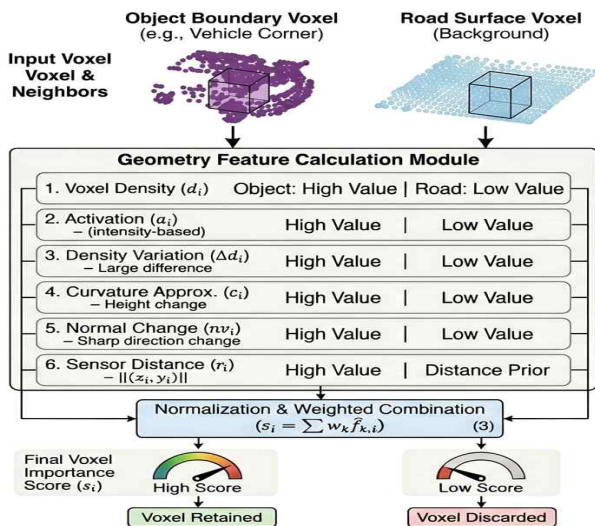


그림 2. 복셀 점수화 과정

Fig. 2. Geometric voxel scoring process.

가 집중된 복셀은 여섯 개 특성 중 여러 항에서 동시에 높은 값을 가지며, 결과적으로 최종 점수 역시 상승하게 된다. 반면 평탄한 지면이나 단조로운 배경 복셀은 일부 특성 값이 낮아 상위 후보에서 제외될 가능성이 커진다.

3. 적응적 복셀 선택 전략

제안 방법은 최종적으로 sparse encoder에 입력되는 최종 보존 복셀 수 상한은 기준 모델과 동일하게 유지하고, 후보 탐색 범위에 해당하는 후보 복셀 수 상한만 확장한다. 즉, 본 연구의 핵심은 복셀 수 자체를 증가시키는 것이 아니라 동일한 최종 예산 내에서 더 유의미한 복셀을 선택하는 데 있다.

식 (4)에서 M 은 voxelization 이후 생성된 전체 복셀 수, B_{cand} 는 후보 복셀 수 상한, B_{voxel} 은 최종 보존 복셀 수 상한, C 는 실제 후보 복셀 수, K 는 최종 선택 복셀 수를 의미한다. 후보 복셀 수가 지나치게 작으면 유효 후보의 다양성이 부족해지고, 반대로 지나치게 크면 배경 복셀이 증가하여 선택 이점이 감소할 수 있다. 계산 절차를 순서대로 정리하면 다음과 같다. 먼저 hard voxelization 이후 생성된 전체 복셀 수 M 과 후보 복셀 수 상한 B_{cand} 를 비교하여 실제 후보 복셀 수 C 를 결정한다. 이후 후보 집합에 포함된 각 복셀의 점수 s_i 를 계산하고, 이를 내림차순으로 정렬한 뒤 상위 K 개를 선택한다. 여기서 K 는 실제 후보 복셀 수 C 와 최종 보존 복셀 수 상한 B_{voxel} 중 더 작은 값으로 정해지므로, 최종적으로는 기준 모델과 동일한 복셀 예산 안에서 가장 효과적인 복셀만 유지하게 된다.

$$C = \min(M, B_{cand}), K = \min(C, B_{voxel}), \quad I_{keep} = TopK((s_i)_{i=1}^C, K_{keep}) \quad (4)$$

이와 같은 설계는 계산량 증가를 최소화하면서도 복셀 선택의 자유도를 높이는 장점을 가진다. 즉, 최종적으로 네트워크가 처리하는 복셀 수는 기준 모델과 동일하지만, 후보 집합 단계에서 보다 많은 대안을 탐색한 뒤 선별적으로 입력을 구성하므로 sparse encoder가 접하는 입력 품질을 개선할 수 있다. 다시 말해 제안 기법은 더 많은 복셀을 쓰는 방법이 아니라 같은 수의 복셀을 더 잘 고르는 방법에 가깝다. 이러한 관점은 연산 자원이 제한된 실제 시스템에서 특히 중요하며, 동일한 자원 제약 하에서 검출 성능을 개선하려는 목적과도 부합한다.

4. 복셀 기반 검출기 통합 및 학습 파라미터

제안 샘플러는 MMDetection3D의 데이터 전처리 단계에 삽입되어 hard voxelization 직후에 동작하도록 구현하였다. 따라서 PV-RCNN의 복셀 인코더, sparse encoder, ROI 기반 검출 헤드는 그대로 유지되며, 입력 복셀 집합만 적응적으로 재구성된다.

식 (5)는 가중치와 선택 임계값의 학습 방식을 나타낸다. 점수 가중치는 softmax로 정규화하여 안정적인 학습하고, 임계값은 sigmoid 매핑을 통해 사전 정의한 범위 안에서 조정되도록 설계하였다. 여기서 w_k 는 k 번째 기하학 특성의 결합 가중치, θ_k 는 학습 대상 파라미터, τ 는 선택 임계값, τ_{low} 와 τ_{high} 는 임계값의 하한과 상한, γ 는 임계값을 조정하는 학습 파라미터를 의미한다. 이와 같이 파라미터를 직접 학습함으로써 데이터셋 특성에 맞는 복셀 선택 기준을 자동으로 형성할 수 있다.

$$w_k = \frac{\exp(\theta_k)}{\sum_{m=1}^k \exp(\theta_m)}, \tau = \tau_{low} + (\tau_{high} - \tau_{low})\sigma(\gamma) \quad (5)$$

식 (5)의 첫 번째 항은 원시 파라미터 θ_k 를 softmax 함수에 통과시켜 각 가중치가 0 이상이며 전체 합이 1이 되도록 만든 것이다. 따라서 특정 특성의 기여도를 과도하게 증폭시키지 않으면서도 여섯 개 기하학 특성의 상대적 중요도를 안정적으로 학습할 수 있다. 두 번째 항은 자유 변수 γ 를 sigmoid 함수에 통과시킨 뒤 사전에 정한 구간 $[\tau_{low}, \tau_{high}]$ 로 선형 사상하여 임계값 τ 를 얻는 방식이다. 이와 같은 범위 제한은 임계값이 비현실적인 값으로 발산하는 것을 방지하고, 학습 과정에서 선택 강도를 점진적으로 조정할 수 있게 한다. 이러한 구성은 특정 기하학 특성에 과도하게 의존하는 문제를 줄이고, 데이터셋 및 검출기 특성에 따라 중요한 기하학 정보가 자동으로 강조되도록 한다. 또한 구현 측면에서 기존 검출기의 백본과 검출 헤드를 변경하지 않으므로 다른 복셀 기반 검출기에도 비교적 쉽게 확장할 수 있는 장점이 있다.

다만 본 연구의 샘플러는 클래스별로 별도의 선택 기준을 두기보다는, 여섯 특성의 정규화 및 학습 가능한 결합 가중치를 통해 데이터셋 전반에서 유효한 공통 기하학 기준을 형성하도록 설계하였다. 즉, 차량, 보행자, 자전거가 서로 다른 크기와 형상을 갖더라도 밀도, 반

사강도, 국소 구조 변화, 거리 특성의 조합을 통해 각 객체에 유효한 복셀을 상대적으로 높은 점수로 선택하도록 한다. 클래스별 adaptive sampling은 향후 연구에서 확장할 수 있는 방향으로 남겨둔다.

III 실험

1. 실험 환경

모든 실험은 Ubuntu 22.04 OS 워크스테이션을 사용하여 수행되었으며 하드웨어 사양은 Intel Xeon(R) Silver 4215R(3.20 GHz) CPU, 128GB RAM, Nvidia RTX 3090 GPU 2개(24GB VRAM)와 같다. 개발 환경은 CUDA 12.1, cuDNN 8.9.2, pytorch 2.1.1, mmdet3d 1.4.0 등의 라이브러리가 사용되었다.

PV-RCNN에서는 최종 보존 복셀 수 상한을 고정된 상태에서 후보 복셀 수 상한만 변화시키며 제안 기법의 동작 특성을 분석하였다. 반면 SECOND^[7], PointPillars^[8], Part-A2^[9] 확장 실험에서는 후보 복셀 수 상한을 (48K, 100K)으로 고정하여, 제안 기법의 일반화 가능성을 확인하는 데 초점을 맞추었다. 본 논문에서 제시하는 AP11 및 추론 시간 수치는 모두 5회 반복 결과의 평균값을 기준으로 정리하였으며, 논문 내 표와 본문에서는 평균값만을 제시하였다.

2. 데이터셋 및 평가지표

데이터셋은 KITTI 3D object detection benchmark^[10]를 사용하였다. 평가 대상 클래스는 Car, Pedestrian, Cyclist의 3종이며, 평가 지표는 KITTI의 3D AP11을 채택하였다. 본 논문에서는 특히 moderate 난이도에서의 Overall 3D AP11을 주요 성능 지표로 사용하였다. Overall 3D AP11은 easy, moderate, hard 각 난이도에서 클래스별 3D AP11을 평균한 지표이며, 여러 IoU 임계치를 평균한 값과는 다르다. KITTI 환경에서는 일반적으로 moderate 성능이 대표 지표로 활용되므로, 본 논문에서도 moderate 값을 중심으로 분석하였다. 실용성 분석을 위하여 프레임당 평균 end-to-end 추론 시간도 함께 측정하였다. 시간 측정은 warm-up 20프레임 이후 validation 데이터의 3749 프레임에 대하여 수행하였으며, 입력 포인트 클라우드의 voxelization, 제안한 복셀 선별 과정, 검출기 추론 및 후처리를 모두 포함하고 dataloader I/O 시간은 제외하였다.

3. 실험 결과 및 고찰

가. 정량적 평가

먼저 표 1과 같이 PV-RCNN에서 후보 복셀 수 상한 변화에 따른 Overall 3D AP11 moderate 평균과 평균 추론 시간을 함께 비교하였다. 후보 복셀 수 상한이 지나치게 작으면 객체 관련 후보 복셀 확보가 제한되고, 반대로 지나치게 크면 배경 후보가 증가하여 선별 효율이 저하될 수 있다. 5회 반복 평균 기준으로는 (48K, 100K) 설정이 가장 우수한 moderate 성능을 보였다.

표 1에서 확인되듯이 기준 모델이 가장 빠른 추론 속도를 보였으며, 제안 기법은 후보 복셀 선별 과정이 추가되므로 평균 추론 시간이 다소 증가하였다. 그럼에도 불구하고 5회 평균 기준 증가폭은 약 1.48~2.48ms/frame 범위에 머물렀으며, 후보 복셀 수 상한의 선택에 따라 성능 향상과 시간 오버헤드의 절충 관계가 달라짐을 확인할 수 있었다.

특히 (48K, 100K) 설정은 평균 추론 시간이 81.74ms/frame으로 기준 모델 대비 2.26ms/frame 증가하는 데 그쳤으나, Overall 3D AP11 moderate는 72.15로 기준 모델 대비 1.59 향상되어 정확도 측면에서 가장 우수한 결과를 보였다. 한편 (42K, 90K) 설정은 제안 기법 가운데 시간 증가폭이 가장 작게 나타나, 후보 복셀 수 상한이 정확도와 추론 효율의 균형을 결정하는 핵심 변수임을 보여준다.

표 2는 제안 기법의 일반화 가능성을 확인하기 위해 PV-RCNN, SECOND, PointPillars, Part-A2의 Overall 3D AP11 moderate 평균 결과를 통합 비교한 것이며, 모두 후보 복셀 수 상한 (48K, 100K) 조건의 결과를 기준으로 정리하였다. 표 2에서 보듯이 제안 기법은 네 검출기 모두에서 positive gain을 보였다. 특히

표 1. PV-RCNN에서의 후보 복셀 수 상한에 따른 Overall 3D AP11 moderate 및 평균 추론 시간 비교

Table 1. Comparison of mean Overall 3D AP11 moderate and average inference time according to candidate voxel limit on PV-RCNN.

Method	Candidate voxel limit	Overall 3D AP11 Moderate ↑	Average inference time (ms/frame) ↓
Original PV-RCNN	(16k, 40k)	70.56	79.48
Ours	(30k, 70k)	69.85	81.79
Ours	(36k 80k)	71.77	81.95
Ours	(42k, 90k)	71.11	80.95
Ours	(48k, 100k)	72.15	81.74

표 2. 복셀 기반 검출기의 Overall 3D AP11 moderate 비교

Table 2. Comparison of Overall 3D AP11 moderate of voxel-based detectors.

Model	Proposed candidate voxel limit	Original ↑	Ours ↑
SECOND	(48K, 100K)	63.34	66.52
PointPillars	(48K, 100K)	63.58	64.29
Part-A2	(48K, 100K)	67.76	68.10
PV-RCNN	(48K, 100K)	70.56	72.15

SECOND에서는 3.18의 가장 큰 향상이 확인되었고, PV-RCNN, PointPillars, Part-A2에서도 각각 1.59, 0.71, 0.34의 개선이 나타났다. 이는 입력 복셀 선택 전략의 개선이 특정 백본이나 검출 헤드에 국한되지 않고 복셀 기반 검출기 전반에 유효할 수 있음을 시사한다.

나. 특성 중요도 분석

제한한 여섯 가지 기하학 특성의 상대적 기여도를 분석하기 위하여 테스트 시점 feature masking 실험을 수행하였다. 학습이 완료된 PV-RCNN 제안 모델의 체크포인트는 고정하고, 평가 단계에서 하나의 특성만 비활성화한 뒤 나머지 특성 가중치를 재정규화하였다. 따라서 표 3의 성능 하락폭은 해당 특성이 최종 복셀 선택 점수에 기여하는 민감도를 나타낸다.

표 3에서 센서 거리 특성을 제거한 경우 Overall 3D AP11 moderate가 71.34로 감소하여 가장 큰 하락폭 (-0.95)을 보였다. 그 다음으로 밀도 특성 제거 시 -0.70의 하락이 나타났다. 이는 거리와 밀도가 KITTI 주행 장면에서 포인트 희소성 및 객체, 배경 구분에 핵심적인 역할을 수행함을 의미한다. 곡률 근사치와 밀도

표 3. 테스트 시점 feature masking 기반 기하학 특성 중요도 분석

Table 3. Geometry feature importance analysis based on test-time feature masking.

Masked feature	Overall 3D AP11 Moderate ↑	Delta ↑
None	72.29	0
Distance	71.34	-0.95
Density	71.59	-0.7
Curvature	71.84	-0.45
Density change	71.84	-0.45
Activation	71.91	-0.38

변화는 각각 -0.45, -0.45의 유사한 하락폭을 보여, 객체 경계와 국소 구조 변화 보존에 상호보완적으로 기여하는 것으로 해석된다.

다. 정성적 평가

정성적 평가는 동일한 KITTI validation 장면에서 기준 모델과 제안 방법의 검출 결과를 비교하는 방식으로 수행하였다. 작은 보행자, 원거리 객체, 차량 경계 영역과 같이 복셀 선택의 영향이 큰 사례를 중심으로 비교하면 제안 방법의 장점을 보다 명확하게 확인할 수 있다. 그림 3은 동일 장면에서 기준 모델과 제안 방법의 검출 결과를 비교한 예를 나타낸다. 제안 방법은 보행자와 원거리 차량의 경계 부근에서 더 안정적인 3D bounding box를 형성하며, 특히 희소한 포인트 분포를 갖는 객체 주변에서 누락을 줄이는 경향을 보인다. 이러한 시각적 비교는 정량적 성능 향상이 실제 장면에서도 유효함을 뒷받침한다.

제안 방법은 차량의 외곽선과 보행자의 직립 구조가 반영된 복셀을 보다 안정적으로 유지함으로써 검출 박스의 위치와 크기를 개선하는 경향을 보였다. 반면 배경 구조가 복잡하거나 객체 포인트 수가 극단적으로 적은 경우에는 여전히 오검출 또는 누락이 발생할 수 있었다. 이러한 정성적 결과는 정량적 성능 향상과도 일관된 방향을 보인다. 즉, 제안 기법은 단순히 수치상 평균을 높이는 것에 그치지 않고, 실제 검출 장면에서 객체 경계와 위치 정합성을 개선하는 방식으로 작동함을 확인할 수 있다.

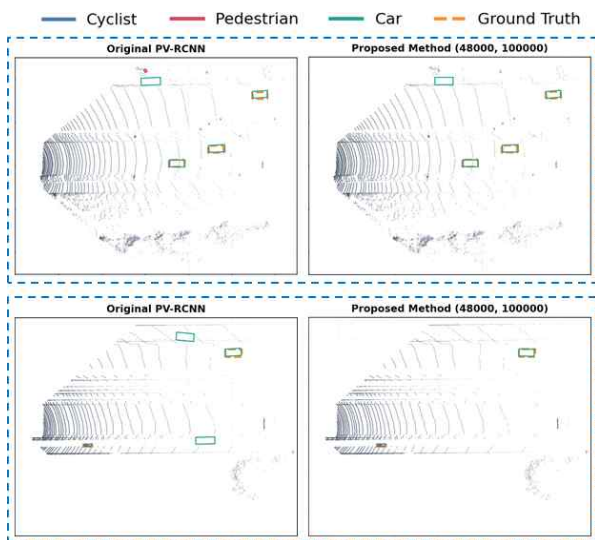


그림 3. 정성적 검출 결과 비교

Fig. 3. Qualitative comparison of detection results.

IV. 결 론

본 논문에서는 LiDAR 기반 3차원 객체 검출을 위한 기하학 특성 기반 적응형 복셀 샘플링 기법을 제안하였다. 제안 방법은 복셀 밀도, 활성도, 밀도 변화, 곡률 근사치, 법선 변화, 거리 특성을 이용하여 각 복셀의 중요도를 정량화하고, 후보 복셀 집합으로부터 검출에 유효한 상위 복셀만을 선별하여 기존 PV-RCNN의 입력으로 전달하도록 설계되었다.

KITTI 실험의 평균 결과에 따르면, 후보 복셀 수 상한을 (48K, 100K)으로 설정한 경우 기준 모델 대비 Overall 3D AP11 moderate가 1.59 향상되었다. 또한 프레임당 평균 end-to-end 추론 시간은 기준 모델 79.48ms/frame, 제안 방법 81.74ms/frame으로 측정되어, 약 2.26ms/frame의 제한적인 시간 증가만으로 성능 향상을 달성할 수 있음을 확인하였다. 추가로 SECOND, PointPillars, Part-A2에 대한 확장 실험에서도 Overall 3D AP11 moderate가 각각 3.18, 0.71, 0.34 향상되어, 제안 기법이 특정 검출기 구조에 국한되지 않고 다른 복셀 기반 검출기에도 적용 가능함을 확인하였다.

종합하면, 본 연구는 복셀 기반 3차원 객체 검출에서 입력 전처리 단계가 독립적인 최적화 대상이 될 수 있음을 실험적으로 확인하였다. 그리고 후보 복셀 수 상한의 적절한 설정이 성능 향상에 중요한 역할을 한다는 점을 보임으로써, 복셀 선택 전략의 설계가 단순한 전처리 조정이 아니라 검출 성능과 추론 효율 간의 균형을 좌우하는 핵심 요소가 될 수 있음을 확인하였다. 또한 feature masking 분석을 통해 센서 거리와 복셀 밀도가 성능에 가장 큰 영향을 미치는 특성임을 확인하였고, 곡률 근사치와 밀도 변화 역시 객체 경계 보존에 상호보완적으로 기여함을 확인하였다. 이는 제안한 샘플러가 단일 밀도 기준이 아니라 복수의 기하학 특성을 결합하여 입력 복셀의 중요도를 평가한다는 점을 뒷받침한다.

향후에는 다중 스케일 기하학 특성의 도입, 보다 정교한 복셀 관계 모델링, 그리고 다른 복셀 기반 검출기로의 확장을 통해 제안 기법의 일반화 가능성과 적용 범위를 추가로 검증할 필요가 있다.

REFERENCES

- [1] Y. Zhou and O. Tuzel, "VoxelNet: End-to-end learning for point cloud based

- 3D object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Salt Lake City, UT, USA, pp. 4490-4499, 2018.
- [2] S. Shi, C. Guo, L. Jiang, Z. Wang, J. Shi, X. Wang, and H. Li, "PV-RCNN: Point-voxel feature set abstraction for 3D object detection," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Seattle, WA, USA, pp. 10529-10538, 2020.
- [3] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, pp. 652-660, 2017.
- [4] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep hierarchical feature learning on point sets in a metric space," in Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 30, pp. 5099-5108, 2017.
- [5] G. Zhang, L. Fan, C. He, Z. Lei, Z. Zhang, and L. Zhang, "Voxel Mamba: Group-free state space models for point cloud based 3D object detection," in Adv. Neural Inf. Process. Syst. (NeurIPS), Vancouver, Canada, 2024.
- [6] Z. Leng, P. Sun, T. He, D. Anguelov, and M. Tan, "PVTransformer: Point-to-voxel transformer for scalable 3D object detection," arXiv preprint arXiv:2405.02811, 2024.
- [7] Y. Yan, Y. Mao, and B. Li, "SECOND: Sparsely embedded convolutional detection," Sensors, vol. 18, no. 10, Art. no. 3337, 2018.
- [8] A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, and O. Beijbom, "PointPillars: Fast encoders for object detection from point clouds," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Long Beach, CA, USA, pp. 12697-12705, 2019.
- [9] S. Shi, Z. Wang, J. Shi, X. Wang, and H. Li, "From points to parts: 3D object detection from point cloud with part-aware and part-aggregation network," IEEE Trans. Pattern Anal. Mach. Intell., vol. 43, no. 8, pp. 2647-2664, 2021.
- [10] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Providence, RI, USA, pp. 3354-3361, 2012.

— 저 자 소 개 —



라 승 탁(학생회원)
2017년 국립한밭대학교
전자공학과 학사 졸업.
2019년 국립한밭대학교
전자공학과 석사 졸업.
2022년~현재 국립한밭대학교
전자공학과 박사과정.

<주관심분야: 자율주행, 딥러닝>



이 승 호(평생회원)
1986년 한양대학교 전자공학과
학사 졸업.
1989년 한양대학교 전자공학과
석사 졸업.
1994년 한양대학교 전자공학과
박사 졸업.

1994년~현재 국립한밭대학교 전자공학과 교수
<주관심분야: 영상신호처리, 딥러닝, AR>