

3 차원 라이다 데이터의 딥러닝 적용을 위한 특징 추출 방법 분석

김동훈, 정성환, 김병준, 김선형, 김서정, 조연호, 박근호*
한국전자기술연구원

{clickmiss123, shjeong, jun0420, sh.kim, scott3554, geijin0821, *root}@keti.re.kr

Analysis of Feature Extraction Method for Applying 3D LiDAR Data to Deep Learning

Donghoon Kim, Shunghwan Jeong, Byoungjun Kim,
Seonhyeong Kim, Seojeong Kim, Yeonho Jo, Keunho Park*
Korea Electronics Technology Institute

요 약

본 논문에서는 3 차원 데이터를 딥러닝에 활용하기 위한 입력 특징 중 복셀 특징과 조감도 특징을 설명하고 각 특징별 장단점을 분석하였다. 결론적으로 복셀 특징의 경우 데이터를 최대한 많이 보존하므로 딥러닝 성능을 높이기 유리하고, 조감도 특성의 경우 데이터를 정규화 하므로 스케일 별로 추가학습 없이 활용 가능하다.

I. 서론

최근 수십 년 동안 레이저 스캐닝 및 스테레오 비전 이미지로 얻은 3 차원 데이터는 도시 매핑, 3D 모델링, 교통 모니터링, 토목 공학 및 산림 모니터링과 같은 광범위한 응용 분야에 사용되어 왔다[1]. 본 논문에서는 이러한 3 차원 데이터를 딥러닝에 활용하기 위한 입력 특징들을 비교 분석하였다.

II. 본론

본 논문에서는 3 차원 라이다 데이터를 딥러닝의 입력에 적용하기 위해 복셀(voxel) 특징[1,2]과 조감도(bird eye view) 특징[3,4]을 비교 분석하는 연구를 진행하였다. 이번 장에서는 각 특징의 개념을 설명하고 각 특징별 장단점을 분석한다.

2.1. 복셀 특징

3 차원 데이터의 경우 지평면을 기준으로 하는 x, y 축과 지평면에 수직인 z 축으로 이루어져 있다. 3 차원 데이터를 표현하는 방법인 복셀은 영상을 표현하는 픽셀(pixel)이 가지는 x, y 축에 대한 정보에 부피(volume)의 개념을 추가하여 z 축으로 확장시킨 공간정보를 포함한다. 복셀 특징을 구성할 때 실수(real number)의 위치값을 가지는 원본 데이터의 정량화를 수행한다. 그림 1 은 원통 모양의 원본 데이터를 각각 다른 스케일로 정량화한 복셀 특징을 생성하는 예를 나타낸다.

정량화를 위해 각각의 복셀에 3 차원 원본 데이터 값이 포함되면 1, 그렇지 않으면 0 으로 표현하는 단순한 방법이 있다. 하지만 이는 잡음(noise)에 취약하므로 그림 2 와 같이 각각의 복셀 안에 포함된 3 차원

데이터의 평균값이나 중앙값 등 사용자가 분석 환경에 맞게 정량화 방법을 선택할 수 있다. 스케일의 크기가 작으면 데이터가 세밀하지만 저장용량이 증가하고, 크기가 크면 데이터의 저장용량이 작지만 데이터가 세밀하지 못하다는 단점이 있으므로 스케일 크기의 적절한 선택이 필요하다.

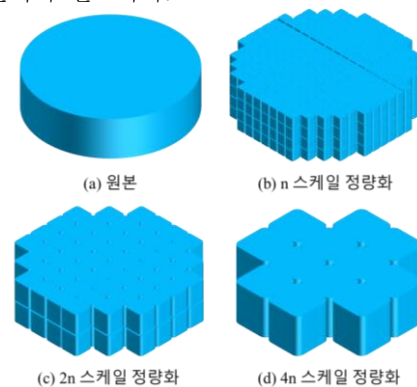


그림 1. 스케일 크기에 따른 복셀의 형태 예시

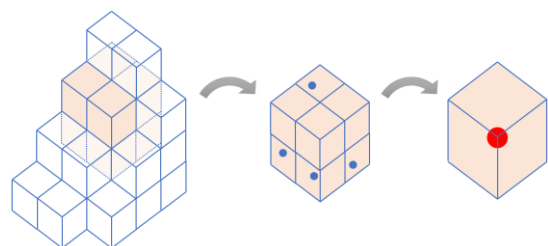


그림 2. 정규화를 위한 voxel 구성의 예시

2.2. 조감도 특징

조감도 특징은 그림 3 과 같이 공간상의 x, y 좌표에 대해 지정한 사각형 영역을 일정한 간격의 격자 형태로 표현함으로써 정량화를 수행한다. 조감도 특징은 강도 맵(intensity map), 높이 맵(height map), 밀도 맵(density map) 3 가지 특징 맵으로 구성된다.

- 강도 맵

라이다로부터 얻을 수 있는 3 차원 데이터 각각의 값은 피사체의 표면의 반사 정도에 따라 각각 다른 값의 강도(intensity)를 가진다. 이러한 특성으로부터 강도 맵은 각각의 격자 내부에 위치한 포인트 중 z 축상에서 가장 높은 위치에 있는 포인트의 반사 강도 값을 이용한다. 그림 3 에서 강도 맵은 6×6 크기의 행렬로 표현할 수 있으며 가장 높은 z 값을 가지는 지점을 빨간색으로 표현하였다. 빨간색 지점의 i 값이 강도 특징 맵의 값에 해당한다. 데이터가 없는 위치는 0 으로 표현한다.

- 높이 맵

다양한 응용을 위해서 사용자가 관심영역을 설정하여 높이의 값을 정규화 할 수 있다. 높이 맵은 강도 맵에서 강도의 특성을 이용한 것과 유사하게 높이 특성 그 자체를 활용한 특징으로 사용자가 지정한 최대 높이의 임계값 th_{max_z} 와 최소 높이 임계값 th_{min_z} 를 이용하여 행렬 격자 안에서 가장 높은 z 값을 정규화하며 수식은 다음과 같다.

$$height_map_{(x,y)} = \frac{z_{(x,y)}}{th_{max_z} - th_{min_z}} \quad (1)$$

- 밀도 맵

밀도 맵은 행렬 격자에 포함할 수 있는 가장 많은 수의 3 차원 데이터의 임계값 th_{max_num} 을 설정하고, 각각의 행렬 격자에 포함된 3 차원 데이터의 개수를 측정하여 적은 수의 데이터가 포함될수록 0 에 가까운 값을, 많은 수의 데이터가 포함될수록 1 에 가까운 값을 출력하도록 정규화한 행렬로 수식은 다음과 같다.

$$density_map_{(x,y)} = \frac{z_{(x,y)}}{th_{max_num}} \quad (2)$$

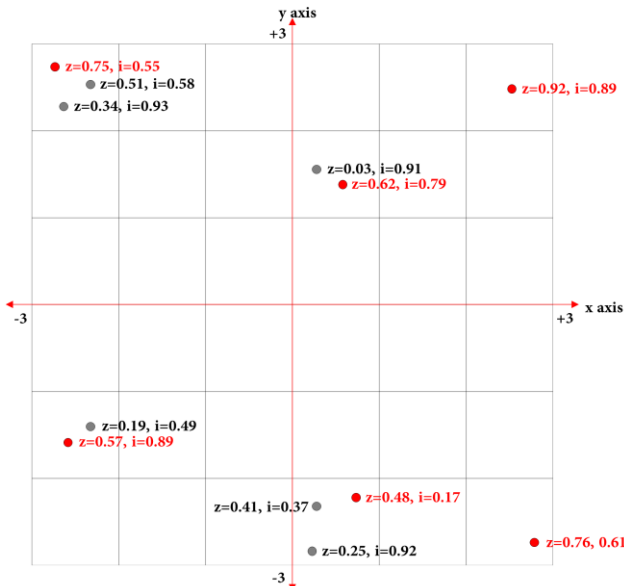


그림 3.3 3차원 데이터의 조감도 특징 값 생성 예시

3 가지 특징 맵을 동일 x, y 축 범위, 동일 격자 간격으로 구성하면 동일한 크기의 특징 맵이 만들어진다. 이러한 동일한 크기의 특징 맵을 겹치면 3 차원 특징을 구성할 수 있으며, 이를 입력으로 사용함으로써 기존의 RGB 영상을 활용하는 동일한 구조의 딥러닝 모델을 적용한 학습을 수행할 수 있다.

2.3. 특징 추출 방법 별 장단점

복셀 특징의 경우 데이터 원본 3 차원 데이터를 특정 스케일 크기의 복셀 안에 포함시켜 정규화하기 때문에 원본 데이터와 비교하면 데이터의 손실이 발생하지만 조감도 특징과 비교했을 때 z 축으로의 특징 손실이 비교적 적기 때문에 딥러닝 모델의 입력으로 사용했을 때 높은 성능을 가진 모델을 학습한다. 하지만 3 채널을 가지는 조감도 특징의 경우 x, y 축의 크기가 다른 특징 데이터가 입력으로 사용되어도 컨볼루션(convolution)이라는 딥러닝의 특성을 이용하여 활용성이 높지만 복셀 특징의 경우 활용성이 떨어진다는 단점이 있다.

III. 결론

본 논문에서는 3 차원 라이다 데이터를 딥러닝 입력에 적용하기 위한 방법론인 복셀 특징과 조감도 특징을 설명하고 장단점을 분석하였다. 복셀 특징의 경우 데이터를 최대한 많이 보존하므로 딥러닝 성능을 높이기 유리하고, 조감도 특징의 경우 데이터를 정규화 하므로 스케일 별로 추가학습 없이 활용 가능하다.

ACKNOWLEDGMENT

이 논문은 2021 년도 정부(경찰청)의 재원으로 과학기술혁신센터의 지원을 받아 수행된 연구임 (No.092021C28S02000, 협력적 교통제어전략 도입을 위한 교통정보 음영구간 정보생성 및 운영관리 기술개발)

참 고 문 헌

- [1] Xu Y., Tong X., Stilla, U. " Voxel-based representation of 3D point clouds: Methods, applications, and its potential use in the construction industry," Automation in Construction, Vol. 126, 2021.
- [2] Cheung G. K. M., Kanade T., Bouguet J.-Y., Holler M., " A real time system for robust 3D voxel reconstruction of human motions," Proceedings IEEE Conference on Computer Vision and Pattern Recognition, pp. 714-720, 2000.
- [3] Chen X., Ma H., Wan J., Li B., Xia T. " Multi-View 3D Object Detection Network for Autonomous Driving," Computer Vision and Pattern Recognition, pp. 1907-1915, Jun. 2017.
- [4] Huang J., Huang G., Zhu Z., Ye Y., Du D. " BEVDet: High-Performance Multi-Camera 3D Object Detection in Bird-Eye-View," Computer Vision and Pattern Recognition, pp. 1-19, Jun. 2022.