

# 자율주행의 정밀한 지도 생성을 위한 NeRF 기반 LiDAR SLAM 기술 동향

남민석, 이동훈, 김평주, 김재호\*  
세종대학교

{namminseok.sejong, donghoon.sejong, pyongjoo.sejong}@gmail.com,\*kimjh@sejong.ac.kr

## A Survey of NeRF-based LiDAR SLAM for Precise Mapping in Autonomous Driving

Minseok Nam, Donghoon Lee, Pyongjoo Kim, Jaeho Kim\*  
Sejong University

### 요약

자율주행 기술이 빠르게 발전함에 따라 복잡한 환경에서 정밀한 지도를 생성하는 것은 안전한 주행을 위해서 중요하다. 정밀한 지도 생성을 위한 이전 연구들은 격자, 복셀, 트리 구조와 같은 명시적인 표현 방식들이 사용되었다. NeRF 연구가 활발하게 진행되면서 고해상도를 표현 하기위해 NeRF 와 LiDAR SLAM 을 접목한 기술들이 연구가 되고 있다. 정밀한 지도 생성을 위한 NeRF 와 LiDAR SLAM 을 접목한 기술들에 대 최신 기술 동향을 살펴보고 각 기술들의 자세한 설명과 향후 SLAM 의 미래 연구 방향에 대해서 제시한다.

### I. 서론

SLAM(Simultaneous Localization And Mapping)은 카메라 및 LiDAR 와 같은 인지 센서를 통해 차량의 위치를 추정함과 동시에 주변 환경의 지도를 생성하는 기술이다. 자율주행 산업에서 SLAM 의 중요성이 점점 강조되면서, 안전한 주행을 보장하기 위해 높은 정확도를 제공하는 LiDAR SLAM 연구가 활발히 진행되고 있다. 그러나, LiDAR SLAM 에서는 채널과 센서의 거리에 따라서 포인트 클라우드가 희소한 영역이 발생하여 전체 환경을 충분히 표현하지 못하는 문제가 존재한다. 이전 연구들은 문제를 해결하고자 지도 표현으로 격자, 복셀, 트리 구조와 같은 명시적인 표현 방식들이 사용되었다[1]. 그러나 명시적 표현들은 고정된 해상도를 가지기 때문에 고해상도를 표현할 경우 상당한 메모리를 필요로 하며, 처리해야할 데이터 양에 따라 연산 비용이 급격히 높아진다는 문제를 가지고 있다. 반면에 저 해상도는 연산 비용을 절감할 수 있지만 실제 세계의 복잡한 환경을 충분히 반영하지 못해 자율주행 기술의 실용적 적용에 제약이 따른다. 이러한 명시적 표현 방식의 한계를 극복하기 위해 최근 연구들은 딥러닝을 통해 3D 환경을 직접 학습하는 암시적인 표현들에 대한 연구들이 진행되고 있으며, 대표적으로 NeRF(Neural Radiance Fields)[2]와 SLAM 을 접목한 기술들이 주목받고 있다[3]. NeRF 는 심층 신경망을 통해 복잡한 3D 환경의 색상과 밀도를 연속적인 함수로 모델링하여 효율적으로 표현할 수 있는 강점을 가진다. 이를 SLAM 에 적용함으로써, 차량의 위치를 추정하고 높은 정밀도를 가진 지도를 재구성할 수 있다. 또한 적은 양의 데이터로 복잡한 3D 환경을 효율적으로 표현할 수 있으며, 특징점 추출과 정합 과정이 없기 때문에 리소스를 줄일 수 있어, 높은 정밀도의 지도와 효율적인 SLAM 을 구현할 수 있다. 따라서 본 논문에서는

자율주행을 위해 정밀한 지도 생성을 위한 NeRF 와 LiDAR SLAM 을 접목한 기술들에 대한 최신 기술 동향을 살펴보고, 각 기술들의 자세한 설명과 향후 SLAM 의 미래 연구 방향에 대해서 제시한다.

### II. 본론

#### 2.1 NeRF 기반 LiDAR SLAM

NeRF 는 3D 환경을 연속적인 함수로 모델링하는 딥러닝 기반 암시적 표현 방식이다. NeRF-SLAM 은 NeRF 와 SLAM 기술을 통합을 처음 수행하였으며, 이미지를 통해 3D 환경을 암시적인 표현으로 학습하여 실시간으로 재구성하면서 높은 정확도의 SLAM 을 구현했다. 그러나 카메라의 내재적인 특성 때문에 조명 조건에 많은 영향을 받아 실내에서만 수행되는 한계가 있다. 실외에서도 높은 정확도의 SLAM 을 수행하기 위해 NeRF 와 LiDAR 를 결합한 방법들이 제안되고 있다. NeRF 기반 LiDAR SLAM 은 LiDAR 센서로부터 얻은 3D 점 구름 데이터를 학습하여 3D 환경의 색상과 밀도를 함수로 모델링하고 지도를 생성하고 위치 추정을 수행할 수 있다.

SHINE-Mapping[4]은 환경에 대한 정보를 신경망에 암시적으로 저장함으로써 비교적 컴팩트한 표현을 사용하는 방법을 제안하였다. 이 방식은 포인트 클라우드를 옥트리 노드 코너에 저장되며 포인트 클라우드 이진 교차 엔트로피 손실을 사용하여 최적화한다. 다음 방법을 이용하여 최적화를 진행하면 데이터가 불균등하게 분포될 수 있는 대규모 야외 환경에서 모델의 일반화를 돕는 희소 데이터 영역에서 특징 표현을 효과적으로 최적화를 해준다. 매핑 과정에서 지역 특징에 부호가 있는 거리 값으로 디코딩 하기 위해 다중 퍼셉트론을 사용한다. 또 이진 교차 엔트로피와

<표 1> NeRF 기반한 LiDAR SLAM 개요

방법	venue	장면 인코딩	기하학적 표현	특징
SHINE-Mapping	ICRA 2023	Feature Octree + MLP	SDF	재앙적 망각에 효과적 대응, 전통적인 방식과 비교해서 메모리 효율적인 3D 재구성을 제공
NF-Atlas	RA-L 2023	Feature Octree + MLP	SDF	다중 볼륨 기능을 이용한 다양한 해상도를 효과적으로 관리, 효율적인 메모리 관리
NeRF-LOAM	ICCV 2023	Voxel Octree + MLP	SDF	최소 데이터 효율적 처리, 신경 서명 거리 함수를 통해 지면과 비지면 점을 분류하고, 지면의 정확한 SDF를 계산하여 Z축 표류를 감소

Eikonal 항을 결합한 새로운 손실 함수를 사용하여 기하학적으로 일관된 매핑 결과를 생성할 수 있도록 한다.

NF-Atlas[5]는 최소 볼륨 표현과 암시적 신경 표현의 장점을 결합하여 대규모 맵을 다중 볼륨 신경 기능 필드로 표현하는 것을 제안했다. 기존 NeRF의 기반한 LiDAR SLAM은 옥트리 또는 해시 테이블과 같은 희박하고 이산화된 볼륨 표현을 채택하여 상대적으로 낮은 메모리 비용으로 빠른 매핑을 달성한다. 하지만 기존의 방법으로 재구성된 맵은 잡음이 많고 빈공간이 포함할 수 있다. 매핑 영역을 여러 개의 SDF(Signed Distance Fields)를 인코딩하는 희박한 다중 규모 신경 기능인 서브맵으로 나누고, 각 서브맵을 신경 특징 볼륨으로 표현하며, 이를 포즈 그래프로 연결한다. 각 신경 특징 볼륨을 포즈 그래프의 노드로 취급하여 지역적으로는 강성을 유지하면서 전역적으로는 탄력적인 매핑을 가능하게 했다.

NeRF-LOAM[6]은 소규모 환경에서만 재구성이 가능했던 문제를 대규모 환경에서도 재구성이 가능하도록 확장하였으며, 뉴런 오도메트리, 뉴런 매핑 및 매쉬 재구성의 세 가지 모듈로 구성된 새로운 NeRF 기반 LiDAR 주행 거리 측정 및 매핑 접근 방식을 제안하였다. 뉴런 오도메트리는 LiDAR 데이터를 증분 처리하고 SDF(Signed Distance Fields) 오류를 최소화하기 위해 오도메트리와 복셀 임베딩을 최적화한다. 뉴런 매핑은 사진 할당 없이 동적 복셀 임베딩 생성 전략을 사용하여 매핑 정확도를 향상시킨다. 매쉬 재구성은 키 스캔(Key Scans)이라 불리는 LiDAR 스캔의 중요한 시점들을 저장한 뒤 신경망을 통해 각 포인트의 거리 값을 계산하여 생성하는 SDF를 이용해 현재의 맵 상태를 반영한다. 3D 공간을 격자로 나누고, 각 격자의 꼭짓점에서의 SDF 값에 따라 표면을 구성하는 삼각형들을 결정하는 '마칭 큐브(Marching Cubes)' 알고리즘[7]을 사용하여 SDF 값들로부터 3D 지도를 생성한다.

<표 1>는 NeRF 기반한 SLAM 방법들의 구조와 기술적 특징들로 구성 되어있다. SHINE-Mapping과 NF-Atlas는 특징 옥트리를 가지고 있지만 NeRF-LOAM은 복셀 옥트리를 가지고 있다. 기하학적 표현은 모두 SDF를 사용한다. SHINE-Mapping은 점진적 학습 기능 사용하기 때문에 재앙적 망각을 효과적으로 방지하는 것을 알 수 있고 NF-Atlas는 다수의 서브맵을 포즈 그래프로 관리함으로써 각각의 지역적 특성을 독립적으로 최적화하는 장점을 가지고 있다. NeRF-LOAM은 데이터가 불균일하게 분포된 환경에서도 효과적으로 작동하고 신경 암시적 표현을 사용하여 매우 세밀하고 정확한 매핑을 한다. 하지만 세가지 모든 방법들은 정밀한 지도 생성을 위해 복잡한 구조로 이뤄져 있다. 이는 실시간 성능에 대한 제한이 생긴다 보다는 이전 격자, 복셀, 트리와 비교해서 메모리와 시간을 절감할 수 있었으나 실시간 적용단계에는 제한이 있다.

### III. 결론

본 연구에서는 NeRF와 LiDAR SLAM을 결합한 최신 기술 세가지를 탐구하고, 이에 대한 다양한 방법들을

소개했다. NeRF를 활용한 LiDAR SLAM 기술은 고해상도의 지도를 생성하고 차량의 위치를 정확하게 추정할 수 있는 잠재력을 제공한다. 또한, 이러한 방법들은 기존의 SLAM 기술들과 달리 필요한 특징점 추출과 정합 과정이 없어 시간적, 연산적 효율성을 갖추고 있다. NeRF를 SLAM에 적용하는 다양한 방법들, 예를 들어 SHINE-Mapping, NF-Atlas, NeRF-LOAM 등은 각각의 특징을 가지고 있으며, 실외 환경에서의 적용 가능성과 성능 향상을 위한 연구가 진행 중이다. 대규모 환경에서의 재구성 가능성을 고려하여 확장된 기능을 제공하여 격자, 복셀 트리와 비교해서 메모리와 시간을 절감은 가능하나 시스템의 복잡성으로 인해 실시간 적용 단계에서는 제한이 존재한다. 추후 연구에서 실시간성을 보장하는 연구가 필요하다.

### ACKNOWLEDGMENT

본 연구는 2024년도 산업통상자원부 및 산업기술평가관리원 (KEIT) 연구비 지원에 의한 연구임 (RS-2022-00154678)

### 참고 문헌

- [1] Yang, X., Li, H., Zhai, H., Ming, Y., Liu, Y., & Zhang, G. (2022, October). Vox-fusion: Dense tracking and mapping with voxel-based neural implicit representation. In 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (pp. 499-507). IEEE.
- [2] Mildenhall, Ben, et al. "Nerf: Representing scenes as neural radiance fields for view synthesis." Communications of the ACM 65.1 (2021): 99-106.
- [3] Rosinol, Antoni, John J. Leonard, and Luca Carlone. "Nerf-slam: Real-time dense monocular slam with neural radiance fields." 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2023.
- [4] X. Zhong, Y. Pan, J. Behley, and C. Stachniss. SHINE-Mapping: Large-Scale 3D Mapping Using Sparse Hierarchical Implicit Neural Representations. In Proc. of the IEEE Intl. Conf. on Robotics & Automation (ICRA), 2023.
- [5] Yu, Xuan, et al. "Nf-atlas: Multi-volume neural feature fields for large scale LiDAR mapping." IEEE Robotics and Automation Letters (2023).
- [6] Junyuan Deng, Qi Wu, Xieyuanli Chen, Songpengcheng Xia, Zhen Sun, Guoqing Liu, Wenxian Yu, and Ling Pei. Nerf-loam: Neural implicit representation for large-scale incremental LiDAR odometry and mapping. In Proc. IEEE Int. Conf. Comput. Vis., 2023.
- [7] Lorensen, William E., and Harvey E. Cline. "Marching cubes: A high resolution 3D surface construction algorithm." Seminal graphics: pioneering efforts that shaped the field. 1998. 347-353