

야지환경에서의 딥러닝 기반 영상 깊이 추정 데이터셋 구축을 위한 플랫폼 연구

하진수¹⁾ · 강현욱¹⁾ · 이종현¹⁾ · 박정희²⁾ · 이채현²⁾ · 김양곤²⁾ · 조기춘^{*1)}

건국대학교 스마트운행체공학과¹⁾ · LIG 넥스원 드론개발단²⁾

A Study of Deep-Learning Depth Estimation Dataset Construction Platform in Unstructured Outdoor Environment

Jinsu Ha¹⁾ · Hyunwook Kang¹⁾ · Jonghyun Lee¹⁾ · Junghee Park²⁾ · Chaehyun Lee²⁾ · Yanggon Kim²⁾ · Kichun Jo^{*1)}

¹⁾Department of Smart Vehicle Engineering, Konkuk University, Seoul 05029, Korea

²⁾Dept. Ground Control System S/W, LIG Nex1Co. Ltd, 21 Pangyo-ro 255beon-gil, Bundang-gu, Seongnam-si, Gyeonggi 13486, Korea

(Received 7 December 2023 / Revised 16 January 2024 / Accepted 24 January 2024)

Abstract : Accurate perception is critical for unmanned ground systems in unstructured outdoor environments. LiDAR sensors utilize laser beams to generate point-based 3D spatial information. However, they are vulnerable to external impact. Conversely, camera sensors, known for their affordability and durability, require indirect methods, such as stereo vision and deep learning, to generate spatial information. Creating datasets for deep learning in outdoor environments is challenging due to the absence of specialized data acquisition platforms and the time-consuming and costly process of generating manual annotations. To address these issues, this study proposes an outdoor data acquisition platform and a LiDAR-based technique to generate depth estimation datasets. The platform was employed to construct real outdoor datasets, and to conduct a qualitative evaluation of the constructed dataset. The dataset was then used to train and evaluate a depth-estimation network, validating the method's effectiveness. In conclusion, this study offers a comprehensive solution to acquire data in unstructured outdoor environments.

Key words : Unstructured outdoor environment(야지환경), Image depth estimation(영상 깊이 추정), Dataset construction platform(데이터셋 구축 플랫폼), Deep learning(딥러닝), Dataset(데이터셋)

1. 서론

지상무인체계란 전투 환경에서 사람의 개입 없이 스스로 주변 환경을 인지하고 판단하여 명령받은 임무를 수행하는 시스템이다. 지상무인체계의 동작을 위해서는 라이다(LiDAR), 카메라, 레이더, GNSS 등 다수의 인지 센서 정보를 처리하여 주변 환경을 인지해야 한다. 다양한 인지 센서가 사용 가능하며, 풍부한 주변 환경정보를 제공하는 라이다, 카메라 센서가 환경 인지에 활발하게 사용되고 있다.

라이다 센서는 발광부에서 빛을 방출한 후 물체에 반사되어 수광부로 돌아올 때까지의 시간(Time of flight)을 측정하여 정확한 거리 정보를 제공한다. 따라서 주변

환경에 대한 정확한 3차원 공간 정보를 제공할 수 있다. 이를 활용하여 객체 인지, 위치 추정, 주행 가능 영역 판단 등의 다양한 기능을 수행할 수 있으나, 라이다 센서의 높은 가격, 물리적 회전부로 인한 내구성 부족 등의 문제로 야지환경에서 주행하는 지상무인체계에 적용하기 어렵다.

카메라 센서의 경우, 고해상도의 풍부한 인지 정보를 제공하기 때문에 다양한 분야에서 사용되고 있다. 기존에는 이미지 평면에서의 객체 검출, 분할 연구들이 수행되었으나, 환경 인지 정보를 공간상에서 표현하기 위한 Monocular 3D detection,¹⁾ Monocular depth estimation^{2,3)} 등의 연구가 진행되고 있다. 단안 카메라를 기반으로 3D

*Corresponding author, E-mail: kichun@konkuk.ac.kr

[†]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.

인지 정보를 생성하는 연구들은 직접적으로 3차원 공간 정보를 직접 생성하는 대신 딥 러닝 네트워크를 활용하여 공간 정보를 간접적으로 추정한다. KITTI,⁴⁾ Cityscape⁵⁾ 등의 자율주행 오픈 데이터셋에서도 3차원 공간 정보를 학습하기 위한 여러 센서 데이터들을 공개하고 있다. 카메라 센서는 낮은 가격과 단순한 구조를 가지기 때문에 센서 고장 시에도 수리나 교체가 간편하며 충격과 진동에 강인하다. 따라서 야지 환경에서 주행하는 지상 무인체계로의 적용이 용이하다. 하지만 딥 러닝 네트워크 학습을 위해서는 대량의 학습 데이터셋이 필요하다.

딥 러닝 학습 데이터셋 구축에는 많은 시간과 비용이 필요하기 때문에 기 구축된 오픈 데이터셋을 활용할 수 있다. 그러나, 오픈 데이터셋은 차량이 일반적으로 주행하는 도로 환경에서 수집되어 지상무인체계 작동 환경과 차이가 존재한다. 또한 오픈 데이터셋은 일반적인 도로 주행 환경의 데이터 수집을 고려한 플랫폼을 사용하기에 야지 환경에 대한 시스템적 고려가 부족하다. 따라서 지상무인체계가 실제 주행하는 야지환경을 충분히 고려한 학습 데이터셋 생성이 필요하다.

기존 영상 깊이 추정 연구에서는 스테레오 카메라에서의 시차 정보를 이용하여 깊이 이미지 정답 데이터를 생성하였다. 카메라를 기반으로 깊이 이미지를 생성할 경우, 낮은 거리 정확도로 인해 모델 학습 성능이 저하될 수 있으며, 추론된 거리 정보의 정확도를 신뢰하기 어렵다. 따라서 고 정밀 깊이 이미지 정답 데이터 생성이 필수적이며, 라이다-카메라 센서를 정확히 동기화, 캘리브레이션 하여 라이다 데이터를 이미지 평면으로 투영하는 방식으로 생성할 수 있다.

따라서 본 논문에서는 야지환경 깊이 추정 데이터셋 구축 플랫폼을 제안하고자 한다. 야지 환경이 충분히 고려된 플랫폼을 구성하고 이를 활용하여 깊이 추정 데이터셋을 구축함으로써 플랫폼의 유효성을 검증하고자 한다. 본 논문의 기여는 다음과 같다. 야지환경을 고려한 데이터 구축 플랫폼 구성 프로세스 제안하였으며, 제안하는 데이터 구축 플랫폼으로 생성한 데이터로 깊이 추정 네트워크를 학습하여 연구의 유효성과 타당성을 검증한다.

본문 구성은 다음과 같다. 2장에서는 본 연구와 관련이 있는 데이터 구축 플랫폼, 센서 간 시간 동기화 및 캘리브레이션, 깊이 추정 네트워크에 대한 사전 연구를 다룬다. 3장에서는 본 논문에서 제안하는 플랫폼 구성, 깊이 이미지 생성 프로세스에 대한 시스템 구조를 소개한다. 4장에서는 데이터 구축 플랫폼 구성 프로세스에 대해 하드웨어 구성, 시간 지연 파라미터 도출 및 센서 캘리브레이션 과정을 소개한다. 5장에서는 깊이 이미지 생

성 프로세스에 대해 시간 동기화, 라이다 투영, 깊이 이미지 생성 과정을 소개한다.

2. 관련 연구

관련 연구로는 데이터셋 취득 플랫폼, 센서 캘리브레이션, 딥 러닝 기반 깊이 추정 네트워크 연구를 소개하고자 한다.

2.1 데이터셋 취득 플랫폼

목표로 하는 데이터셋을 취득하기 위해서는 그에 맞는 플랫폼 구성이 중요하다. 많은 오픈 데이터셋 논문들은 원하는 형태의 데이터셋을 취득하기 위해 구성된 플랫폼에 대해 설명한다. NuScenes⁶⁾ 오픈 데이터셋은 Object detection, Semantic segmentation, Panoptic segmentation 데이터셋을 생성하기 위해 설치한 라이다, 카메라, 레이더, GPS 및 IMU 센서를 포함한 데이터 취득 플랫폼을 구축하였다. A2D2⁷⁾ 데이터셋 연구에서는 사각지대를 최소화한 데이터 취득을 위해 복수의 라이다 센서를 틸트하여 장착하였다. Waymo⁸⁾ 데이터셋 연구에서는 차량 주변 인지 가능 거리를 최대화하기 위해 5개의 라이다와 카메라 센서를 장착하였다.

2.2 센서 캘리브레이션

이기중 센서 데이터 융합을 위해서는 센서 간 정확한 캘리브레이션 파라미터가 필요하다. 정확한 캘리브레이션 파라미터를 구하기 위한 다양한 접근의 연구가 수행되고 있다. 이러한 이종 센서 간 캘리브레이션은 타겟의 존재 여부와 머신러닝 기법 적용에 따라 Target-based 캘리브레이션, Learning-based 캘리브레이션, Target-less 캘리브레이션 3가지로 나눌 수 있다.

Target-based 캘리브레이션 방법은 고전적인 센서 캘리브레이션 방법으로 특정한 타겟 객체를 통해 센서 간 외부 파라미터를 추정한다.

Beltrán 등⁹⁾은 자체 제작한 체커보드로 센서 사이의 6 자유도 파라미터를 도출하였다. Ruan과 Huber¹⁰⁾는 구(球)를 타겟으로 하여 센서 간 회전 행렬(Rotation matrix), 이동 행렬(Translation matrix)을 구하였다. Zhou 등¹¹⁾은 MATLAB calibration tool box를 이용하여 체커보드를 활용한 캘리브레이션을 진행하였다.

Learning-based 캘리브레이션 방법은 머신러닝 기법을 적용하여 센서 간 Extrinsic parameter를 추정한다. RegNet¹²⁾은 CNN을 활용하여 카메라와 라이다 센서 간 외부 파라미터를 연산하였다. RGGNet¹³⁾은 리만 기하학을 이용하여 새로운 Loss function을 제안하였다. LCCNet¹⁴⁾은 RGB

이미지와 포인트 클라우드, 깊이 이미지를 CNN의 입력으로 사용하여 캘리브레이션 성능을 향상하였다.

Target-less 캘리브레이션 방법은 별도의 타겟을 사용하지 않고 주변 환경 특징을 추출하여 외부 파라미터를 추정한다. Muñoz-Bañón 등¹⁵⁾은 카메라와 라이다 Edge 특징을 이용하여 캘리브레이션을 진행하고, Park 등¹⁶⁾은 각 센서의 Odometry를 이용하여 센서 간 외부 행렬을 구한다.

2.3 딥 러닝 기반 깊이 추정

단안 카메라 깊이 추정은 딥러닝 알고리즘을 활용하여 단안 카메라를 통해 취득된 이미지의 각 픽셀별 깊이를 추정한다. 학습된 딥러닝 알고리즘을 활용하면, 단안 카메라 이미지만으로도 깊이를 추정이 가능하다.

깊이 추정을 위한 알고리즘 BTS(From Big To Small)¹⁷⁾은 인코딩 단계에서 Stride convolution과 Spatial pooling layer를 사용하고 디코딩 단계에서 풍부한 깊이 특징 추출을 위해 새로운 방식의 Local planar guidance layer를 활용하여 깊이를 추론한다.

NeWCRFs¹⁸⁾는 이미지를 각각의 창으로 나누고 각 창 별로 FC-CRF 최적화를 수행하여 시간 복잡성을 줄인다. 그래프간 관계를 감지하기 위해 Multi-head attention을 적용하고 Vision transformer를 인코더, FC-CRF를 디코더로 활용하여 최적의 깊이를 추론한다.

깊이 추정 알고리즘의 학습, 평가를 위해서는 카메라 이미지와 이와 쌍을 이루는 깊이 이미지가 필요하다. KITTI Depth Prediction Benchmark¹⁹⁾에서 제공하는 깊이 추정 데이터셋은 널리 사용되는 딥러닝 알고리즘 학습, 검증용 데이터셋으로 도시, 주거지역, 도로의 동시에 취득한 카메라 이미지와 깊이 이미지를 제공한다.

3. 시스템 구조

제안하는 시스템 아키텍처는 Fig. 1에서 확인할 수 있으며 플랫폼 구축 프로세스와 깊이 이미지 생성 프로세스로 구성된다.

플랫폼 구축 프로세스는 데이터 취득을 위한 차량 플랫폼을 구축하는 과정이다. 이는 센서 간 정적 시간 보상과 센서 내·외부 파라미터 캘리브레이션 과정을 포함한다. 플랫폼 구축 프로세스는 데이터셋 구축에 앞서 수행되며, 해당 과정을 통해 도출한 정적 시간 지연 상수와 센서 내·외부 파라미터는 5장의 깊이 이미지 생성 프로세스에 사용된다.

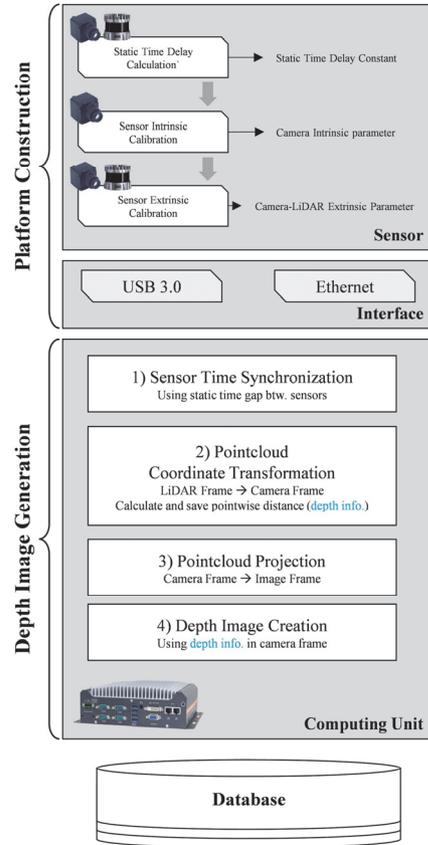


Fig. 1 Overall system architecture for data construction platform and depth image creation process

깊이 이미지 생성 프로세스는 라이다 포인트 클라우드와 카메라 이미지 데이터 스트림을 센서 인터페이스를 통해 수신하고, 두 센서의 시간 정보 비교를 통해 시간 동기화 된 카메라-라이다 데이터 쌍을 생성한다. 동기화 된 데이터에 대해 캘리브레이션 파라미터를 활용하여 라이다 포인트 클라우드를 이미지 평면에 투영하고, 깊이 추정 이미지 GT를 생성한다. 생성된 깊이 이미지는 대응되는 RGB 이미지와 함께 데이터베이스에 저장한다.

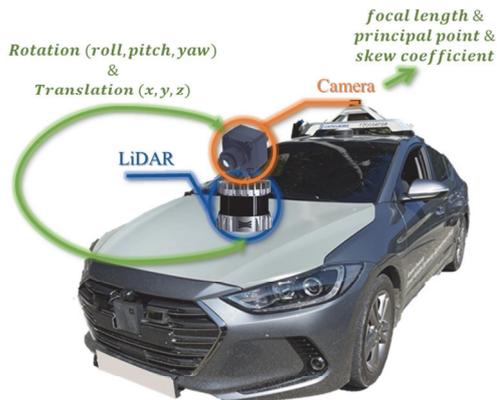
4. 야지환경을 고려한 데이터 취득 플랫폼 구성

4장에서는 야지환경을 고려한 데이터 취득 플랫폼의 구성에 대해 다루고자 한다. 일반 도로 환경과 비교할 때, 야지환경은 비정형 노면으로 구성되어 컴퓨팅 및 센서 작동에 악조건을 제공하기 때문에 이를 고려한 플랫폼을 구성하였다. 또한, 이기종 센서 데이터 융합을 위해 센서 간 시간 지연 파라미터와 센서 내·외부 파라미터를 도출하였다.

4.1 장비 구성

컴퓨팅 시스템은 야외 환경의 충격, 먼지 등의 악조건을 고려하여 Neousys社의 Rugged Embedded 제품군 Nuvo-7501 모델을 사용하였다. Rugged Embedded PC는 물리적인 냉각 팬을 사용하지 않고 자연적으로 냉각이 가능한 외장 케이스를 적용하여 강한 외부 충격과 수분, 먼지에서도 높은 내구성을 보장할 수 있다. 또한 데이터 수집 차량 트렁크에 마운트 랙을 구성하고 Embedded PC를 흔들리지 않도록 고정하였다.

센서의 경우, 야외 환경 데이터 수집 과정에서 물, 먼지와 같은 악천후 조건에서의 동작을 보장하기 위해 IP69K 방수, 방진 사양을 지원하는 모델로 선정하였다.



(a) Overall platform configuration



(b) LiDAR and camera sensor mounting



(c) trunk & mounting rack

Fig. 2 Result of data construction platform configuration for unstructured outdoor environment

라이다 센서는 Ouster OS1-128ch, 카메라 센서는 CanLab CLCC-G-01C 모델을 사용하였다.

장비 장착 시에도 센서 별 시야각을 모델링하고 센서 간 시야 각 간섭을 고려하여 장착 위치를 선정하였다. 또한, 주행 중 라이다-카메라 센서 간 외부 캘리브레이션 파라미터의 변화를 최소화하기 위해 라이다 센서 위에 카메라 센서를 Rigid하게 장착하였다. 또한 차량 루프 대신 보닛에 센서 모듈을 장착함으로써 주행 중에도 센서를 확인할 수 있도록 하였다.

플랫폼 구성 결과는 Fig. 2에서 확인할 수 있으며, 각 센서 별 상세 사양은 Table 1에서 확인할 수 있다.

Table 1 Computing unit, LiDAR, camera sensor specification

Model	Neousys Nuvo-7501
CPU	Intel® i7-9700E (4.4Ghz, 35W TDP)
Memory	2x DDR4 16GB
I/O	2x GbE, 4x USB3.0 4x COM ports
Storage	2.5" SSD 1TB
Power	12V 160W DC
Environmental	MIL-STD-810G

(a) computing unit

Model	Ouster OS1-128 Rev.7
Vertical FOV	-22.5° ~ 22.5°
Vertical Resolution	0.351°
Horizontal FOV	360°
Horizontal Resolution	512, 1024, 2048
# of Layers	128
Framerate	10 or 20 Hz

(b) scanning LiDAR

Model	CanLab CLCC-G-01C
Sensor Type	CMOS
Horizontal FOV	98°
Resolution	1920 x 1080
Frame Rate	30hz
Dynamic range	120dB

(c) camera

4.2 센서 간 시간 동기화를 위한 정적 시간지연 예측

본 절에서는 깊이 이미지 생성의 바탕이 되는 센서 융합을 위한 센서 간 정적 시간 지연 분석과 센서 내·외부 파라미터 캘리브레이션에 대해 다룬다.

센서 시간 동기화의 경우, 동기화 방법 및 대상 센서에 따라 다양한 연구²⁰⁻²²⁾가 진행되었다. 주로 GPS PPS(Pulse Per Second), PTP(Precision Time Protocol)등을 활용하며, 밀리 초 단위의 정확도로 센서 간 시간을 동기화한다.

본 연구에서는 카메라-라이다 2개의 인지 센서에 대한 동기화를 목적으로 한다. 이 경우, 라이다 센서가 지원하는 Phase lock 기능과 카메라 센서가 지원하는 Hardware trigger 기능을 활용하는 것이 일반적이나, 본 시스템의 동작 환경인 야외환경은 다양한 조도 조건으로 카메라 Exposure time이 일정하지 않고, 주행 중 물리적 충격으로 라이다 센서의 회전 Phase가 변화할 수 있기 때문에 이러한 방법으로는 정확한 시간 동기화를 수행하기 어렵다. 따라서 본 연구에서는 두 센서 사이의 정적 시간 지연 파라미터를 정확하게 추정한 뒤, 센서 데이터 시간 정보를 보상함으로써 환경에 강인한 센서 시간 동기화를 수행하고자 한다.

정적 시간 지연 파라미터 예측 프로세스는 Fig. 3에서 확인할 수 있다. 이를 위해 고속 주행, 고속 선회 주행 데이터를 취득하고, 이미지에 포인트 클라우드 데이터를 투영하여 객체 형상 정보가 정확히 일치할 수 있는 센서 간 시간 지연 파라미터를 휴리스틱하게 추정한다. 만약 카메라 데이터가 라이다 데이터보다 빠른 경우, 이미지 상의 객체는 포인트 클라우드 상의 객체보다 먼저 움직이고, 라이다 데이터가 더 빠른 경우에는 포인트 클라우드 상의 객체가 먼저 움직인다. 도출한 정적 시간 지연 파라미터는 시간 동기화 모듈에서 사용되며, 더 긴 시간 지연을 가지는 센서 데이터의 시간 지연을 보상함으로써 두 센서가 정확한 시간 동기화가 가능하도록 한다.

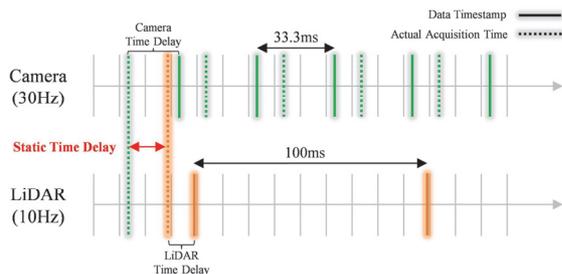


Fig. 3 Static time delay compensation process for multi-sensor synchronization

4.3 센서 내·외부 파라미터 캘리브레이션

캘리브레이션 단계는 크게 카메라 내부 캘리브레이션, 카메라-라이다 외부 캘리브레이션 두 단계로 나뉜다.

카메라 캘리브레이션은 카메라 센서 내부 파라미터인 초점거리, 주점, 비대칭 계수와 렌즈 왜곡 파라미터인 방사, 접선 왜곡을 구하는 과정을 의미한다. 카메라 캘리브레이션을 위해 MATLAB camera calibrator toolbox²³⁾를 사용하였다. 캘리브레이션 과정은 체커보드 패턴 검출 단계, 카메라 파라미터 추정 단계, 왜곡 보정 결과 확인 단계로 수행된다. 첫 단계는 이미지 내에서 체커보드를 검출하는 과정으로, 사전 정의된 크기의 체커보드를 카메라 센서로 촬영한 뒤 OpenCV 라이브러리를 이용하여 체커보드의 코너 픽셀 좌표를 검출한다. 다음 단계에서는 체커보드 좌표와 크기 정보를 이용하여 카메라 내부 파라미터를 추정한다. 마지막 왜곡 보정 결과 확인 단계에서는 도출한 카메라 내부 파라미터를 이용하여 원본 이미지 데이터의 렌즈 왜곡을 보정하며, 이를 시각화하여 캘리브레이션 결과를 검증한다.

라이다-카메라 캘리브레이션은 카메라 센서와 라이다 센서 간의 6자유도의 이동 행렬과 회전 행렬을 도출한다. Target-based 캘리브레이션은 라이다-카메라 외부 파라미터 추출 성능이 다른 방법들에 비해 우수하지만 주행 중 센서 외부 파라미터가 변경될 경우 추가적인 캘리브레이션 단계가 필요하다는 단점을 가진다. 하지만 야외 환경이라는 특수 주행 환경에 강인한 센서 플랫폼을 구성하였기 때문에 주행 중 라이다-카메라 외부 파라미터의 변화가 현저히 적다. 따라서 주행 중 센서 외부 파라미터 추정이 필요하지 않기 때문에 가장 높은 성능을 가지는 Target-based 캘리브레이션 알고리즘을 적용하였다.

먼저 라이다 센서에서는 RANSAC(RANdom SAMple Consensus) 알고리즘을 적용하여 체커보드의 평면을 탐색하고, 3D 코너 포인트를 추정한다. 카메라 센서에서는 내부 파라미터인 초점 거리와 식 (1)을 적용하여 3D 코너 포인트를 추출한다.

$$checkerboard\ depth = \frac{checker\ board\ length * focal\ length}{checkerboard\ pixel\ length} \quad (1)$$

도출한 센서 별 체커보드 Pose를 동일 좌표계로 일치 시킴으로써 두 센서 간 6 자유도의 회전, 병진 행렬을 추정한다.

5. 깊이 추정 이미지 데이터 생성

깊이 이미지 GT 생성 프로세스는 4장의 시간 동기화

와 캘리브레이션 결과를 기반으로 수행된다. 먼저 시간 동기화 결과를 이용하여 라이다-카메라 데이터 쌍을 생성하며, 캘리브레이션 결과를 이용하여 라이다 포인트 클라우드를 이미지 평면으로 투영한다. 투영된 포인트의 공간 정보를 이용하여 이미지에 픽셀 별 깊이 정보를 생성한다. 각 프로세스에 대한 상세한 과정은 5.1 ~ 5.4 절에서 다룬다.

5.1 센서 시간 동기화 및 카메라-라이다 쌍 생성

라이다, 카메라 센서 중 더 높은 주기를 가지는 이미지 데이터를 기준으로 동기화를 수행한다. 이미지 데이터가 수신되었을 경우, 이미지 큐에 데이터를 저장하고 라이다 데이터가 수신될 경우, 이미지 큐에서 동일 시간 정보를 가지는 데이터를 탐색한다. 시간 정보 관리에는 4.2절에서 도출한 센서 간 정적 시간 지연 파라미터를 반영하여 물리적으로 동일한 시간의 센서 데이터를 탐색할 수 있도록 한다.

5.2 카메라-라이다 좌표 변환

카메라 센서 기준의 거리 정보를 획득하기 위해 라이다-카메라 좌표계를 정렬하며, 카메라 좌표 기반의 거리 정보를 획득한다. 4.3절에서 도출한 센서 간 외부 파라미터를 이용하며, 라이다 포인트 클라우드를 카메라 좌표계로 이동한다. 이동한 카메라 좌표계 기준으로 라이다 포인트 클라우드 별 거리(Depth)정보를 저장한다.

5.3 라이다 포인트 투영

4.2절에서 도출한 카메라 내부 파라미터를 이용하여 카메라 좌표계 상의 라이다 포인트 클라우드를 이미지 좌표계로 투영한다. 이 과정에서 각 포인트와 대응되는 이미지 픽셀 정보를 획득한다.

5.4 깊이 이미지 생성

5.2절에서 저장했던 카메라 좌표계 기준의 포인트 별 깊이 정보와 5.3절에서 라이다 투영을 통해 생성한 포인트-픽셀 대응 관계를 이용하여 깊이 이미지를 생성한다.

6. 실험 설계 및 결과

6.1 실험 환경 소개

실험을 위해 본문에서 제안한 데이터 구축 플랫폼 구성 프로세스를 따라 데이터 수집 차량을 구성하였다. 야지환경 데이터는 강원도 평창의 신기봉산로, 강원도 홍천 울전리 밤바치길에서 비포장도로, 산길을 주행하며 취득하였다. 데이터셋의 유효성 검증을 위한 깊이 추정 네트워크 학습에는 NVIDIA RTX 3090 GPU를 활용하였다.

6.2 시간 동기화 및 센서 캘리브레이션 결과 검증

본 실험에서는 플랫폼 구축 과정에서 진행한 시간 동기화 및 센서 캘리브레이션 결과를 정성적으로 평가한다. 시간 동기화는 고 선회 주행 상황에서의 정적 상수 시

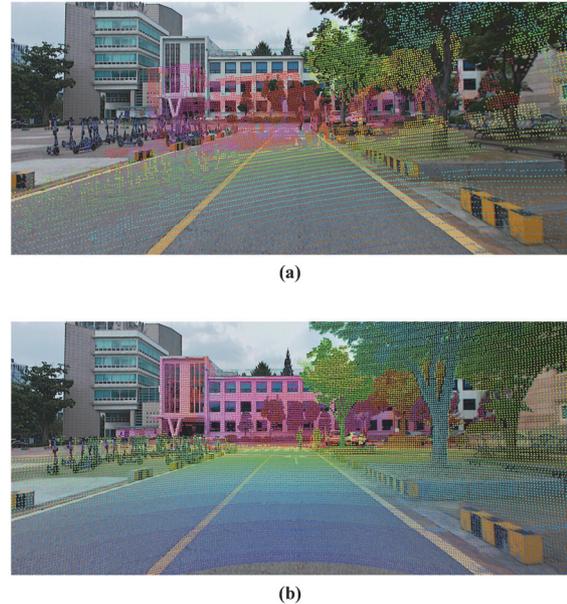


Fig. 4 Qualitative comparison of pointcloud projection of inaccurate calibration parameter and accurate calibration parameter

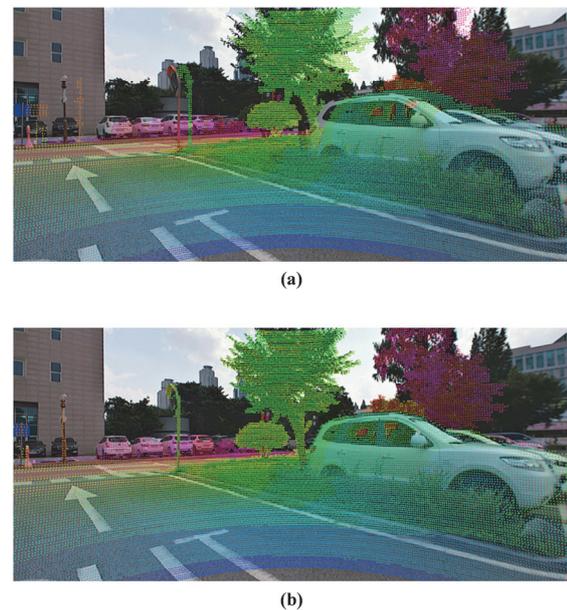


Fig. 5 Qualitative comparison of pointcloud projection of (a) Inaccurate time synchronization and (b) accurate time synchronization

간 보상 적용 여부에 따른 투영 결과를 비교하여 알고리즘의 유효성을 검증하였다. 정적 상수 시간 보상 진행 유무에 따른 정성적 평가 결과는 Fig. 4에서 확인할 수 있으며, 올바른 시간 보상 파라미터가 적용되었을 경우, 두 센서에 검출된 객체의 외곽선이 서로 일치함을 확인할 수 있다.

센서 캘리브레이션 결과 검증은 정지 상황에서의 이미지 왜곡 보정, 포인트 클라우드 투영을 통해 확인하였다. 정확한 카메라 내부 파라미터를 사용하였을 경우, 이미지 외곽에서 왜곡 보정이 잘 이루어진 것을 확인할 수 있고, 정확한 라이다-카메라 외부 파라미터를 찾은 경우 포인트 클라우드 투영이 잘 이루어졌음을 확인할 수 있다. 정성적 평가 결과는 Fig. 5를 통해 확인할 수 있다.

6.3 라이다 데이터 활용 깊이 이미지 생성 정성적 평가

본 실험에서는 직접 구축한 데이터셋과 Cityscapes 오픈 데이터셋의 깊이 이미지를 비교하여, 라이다 센서를 활용한 깊이 정보 생성 방법이 효과적임을 정성적으로 검증한다.

정성적 평가 결과는 Fig. 6에서 확인할 수 있다. 직접 구축한 데이터셋의 특징을 살펴보면, 원거리 객체의 깊이 정보를 뚜렷하게 확인할 수 있다. 또한 스테레오 비전 알고리즘의 한계로 근거리 객체의 표면 정보가 희소하게 표현된 반면, 직접 구축한 데이터에서는 객체의 표면

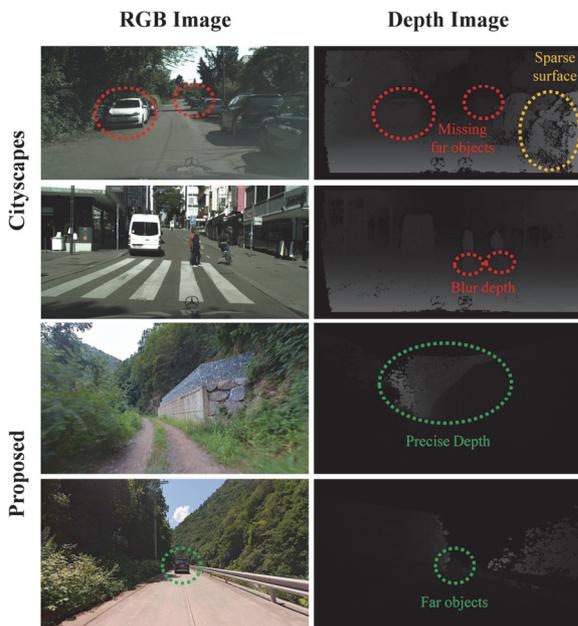


Fig. 6 Qualitative evaluation between cityscapes and custom dataset

정보가 정확하게 표현되었다. Cityscapes 데이터셋의 경우, 근거리 객체와 바닥 사이의 깊이 정보가 뚜렷하게 구분되지 않고 흐릿하게 표현되었다.

종합적으로, 직접 구축한 데이터셋은 원거리 객체에 대해 명확한 깊이 정보를 생성하였고, 객체를 뚜렷하게 구분할 수 있는 충분한 특정 정보를 제공한다. 따라서 라이다 센서를 활용한 깊이 정보 생성 방법이 효과적임을 확인할 수 있다.

6.4 구축한 야지환경 깊이 추정 데이터셋 학습 및 성능 평가

데이터셋 구축 플랫폼과 이를 활용하여 생성한 데이터셋의 유효성을 검증하기 위해 취득된 데이터와 KITTI 데이터셋을 활용하여 오픈 소스 딥 러닝 네트워크 NeWCRFs를 학습, 평가하였다. 다양한 환경에서 취득된 데이터에 대한 검증을 위해 일반 도로 환경 및 야지 환경, 비포장 도로 데이터 또한 평가에 사용하였다. 플랫폼을 통해 취득된 총 11,307개의 데이터를 7:2:1로 분할하여 각각 학습, 검증, 테스트에 활용하였다. Fig.7은 학습된 딥러닝 알고리즘의 테스트 데이터에 대한 입력과 예측 데이터를 나타낸다. 첫번째 열은 입력데이터, 두번째 열은 출력 데이터를 나타낸다. 각 행은 도로 상황으로 분류하였다. Table 2는 학습한 딥 러닝 네트워크의 성능을 비교한 결과를 확인할 수 있다. 모델이 생성한 예측 결과와 정답 이미지의 차이를 SILog(Scale Invariant Logarithmic Error)와 RMSE(Root Mean Square Error)를 통해 정량적으로 분석하였다.

예측 이미지에서 밝은 영역은 데이터 취득 플랫폼으로부터 먼 거리로 예측된 영역, 어두운 영역은 플랫폼으

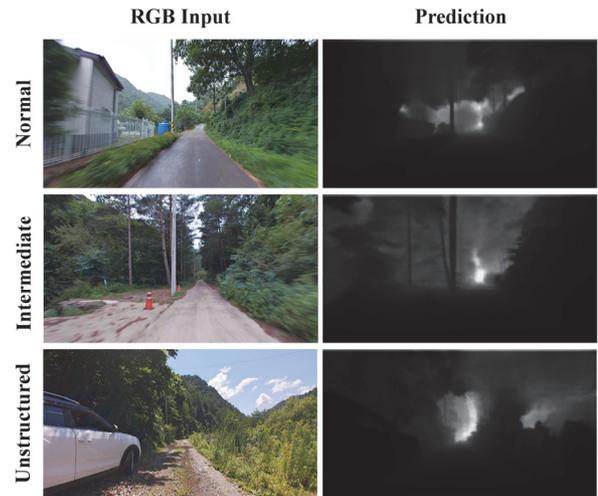


Fig. 7 Depth prediction result

Table 2 Depth estimation quantitative result

Trained dataset	Tested dataset	SILog ↓	RMSE ↓
Kitti	Kitti	8.30	17.03
Proposed	Kitti	27.34	18.08
Proposed	Proposed	33.44	1.86

로부터 가까운 거리로 예측된 영역이다. 이미지 내 전봇대 차량, 나무 등에 대한 위치에 대해 깊이를 잘 예측한 것을 확인할 수 있다. Table 2를 통해 제안하는 데이터 구축 플랫폼으로 생성한 데이터셋의 정량적 평가 결과를 확인할 수 있다. 1, 2열에서는 KITTI와 제안하는 데이터셋으로 평가할 경우, SILog 지표가 유사하기 때문에 새로 구축한 데이터셋의 유효성을 확인할 있으며, 3열의 RMSE 지표를 통해 딥러닝 네트워크 학습의 유효성을 확인할 수 있다.

7. 결론

이 연구의 결론을 요약하면 다음과 같다. 첫째, 야외 환경에서 주행하는 지상 무인 시스템을 위한 딥 러닝 기반 인지 네트워크 학습을 위한 데이터 구축 플랫폼을 제안하였다. 둘째, 제안된 플랫폼을 활용하여 이기종 센서 데이터를 취득, 융합하여 어노테이션 없이 고정밀 깊이 추정 데이터셋을 자동으로 생성할 수 있다. 셋째, 오픈소스 깊이 추정 네트워크의 학습과 성능 평가를 통해 제안하는 시스템의 유효성을 검증하였다.

또한 본 연구에서 제안한 데이터 취득 플랫폼을 활용하여 추가적인 야외 환경 데이터셋 생성, 그에 적합한 딥러닝 네트워크 개발과 관련된 추후 연구가 가능하다. 제안하는 깊이 추정 데이터 생성 프로세스를 활용하여 야외 환경 깊이 추정 데이터셋을 구축하고 딥 러닝 네트워크를 학습함으로써 라이다 센서 없이 카메라 센서만을 사용하는 야외 환경 주행 환경 인지 등의 추후 연구가 가능하다.

후 기

이 논문은 2022년도 정부(방위사업청)의 재원으로 국방기술평화연구소의 지원을 받아 수행된 연구임(No. KRIT-CT-22-011-00, 라이다-영상 신호 융합 기술).

References

1) S. Sim and D. Kum, "Image Sequence Based 3D Object Detector for Autonomous Driving," KSAE Annual Conference Proceedings, pp.474-475, 2021.

2) S. Jung, J. Jang, D. Kang and J. Kim, "Monocular Camera-Based Object Detection and 3D Point Cloud Creation," KSAE Annual Conference Proceedings, pp.532-536, 2021.

3) D. Lim and J. Yoo, "Improved Mono-SLAM Algorithm Using Depth Estimation and Segmentation Based on Deep Learning," Transactions of KSAE, Vol.31, No.8, pp.619-627, 2023.

4) A. Geiger, P. Lenz, C. Stiller and R. Urtasun, "Vision Meets Robotics: The Kitti Dataset," International Journal of Robotics Research, Vol.32, No.11, pp.1231-1237, 2023.

5) M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth and B. Schiele, "The Cityscapes Dataset for Semantic Urban Scene Understanding," Proceedings of the IEEE Conference On Computer Vision and Pattern Recognition, pp.3213-3223, 2016.

6) H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan and O. Beijbom, "Nuscenes: A Multimodal Dataset for Autonomous Driving," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.11621-11631, 2020.

7) J. Geyer, Y. Kassahun, M. Mahmudi, X. Ricou, R. Durgesh, A. S. Chung and L. Hauswald, "A2D2: Audi Autonomous Driving Dataset," arXiv preprint arXiv:2004.06320, 2020.

8) P. Sun, H. Kretzschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. Caine, V. Vasudevan, W. Han, J. Ngiam, H. Zhao, A. Timofeev, S. Ettinger, M. Krivokon, A. Gao, A. Joshi, Y. Zhang, J. Shlens, Z. Chen and D. Anguelov, "Scalability in Perception for Autonomous Driving: Waymo Open Dataset," Proceedings of the IEEE/CVF Conference On Computer Vision and Pattern Recognition, pp.2446-2454, 2020.

9) J. Beltrán, C. Guindel, A. de la Escalera and F. García, "Automatic Extrinsic Calibration Method for Lidar and Camera Sensor Setups," IEEE Transactions on Intelligent Transportation Systems, Vol.23, No.10, pp.17677-17689, 2022.

10) M. Ruan and D. Huber, "Calibration of 3D Sensors Using a Spherical Target," 2nd International Conference on 3D Vision, Vol.1, pp187-193, 2014.

11) L. Zhou, Z. Li and M. Kaess, "Automatic Extrinsic Calibration of a Camera and a 3D Lidar Using Line and Plane Correspondences," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) pp.5562-5569, 2018.

- 12) N. Schneider, F. Piewak, C. Stiller and U. Franke, "RegNet: Multimodal Sensor Registration Using Deep Neural Networks," IEEE Intelligent Vehicles Symposium(IV), pp.1803-1810, 2017.
- 13) K. Yuan, Z. Guo and Z. J. Wang, "RGGNet: Tolerance Aware Lidar-Camera Online Calibration with Geometric Deep Learning and Generative Model," IEEE Robotics and Automation Letters, Vol.5, No.4, pp.6956-6963, 2020.
- 14) X. Lv, B. Wang, Z. Dou, D. Ye and S. Wang, "LCCNet: LiDAR and Camera Self-Calibration Using Cost Volume Network," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.2894-2901, 2021.
- 15) M. Á. Muñoz-Bañón, F. A. Candelas and F. Torres, "Targetless Camera-Lidar Calibration in Unstructured Environments," IEEE Access, Vol.8, pp.143692-143705, 2020.
- 16) C. Park, P. Moghadam, S. Kim, S. Sridharan and C. Fookes, "Spatiotemporal Camera-Lidar Calibration: a Targetless and Structureless Approach," IEEE Robotics and Automation Letters, Vol.5, No.2, pp. 1556-1563, 2020.
- 17) L. Andraghetti, P. Myriokefalitakis, P. L. Dovesi, B. Luque, M. Poggi, A. Pieropan and S. Mattoccia, "Enhancing Self-Supervised Monocular Depth Estimation with Traditional Visual Odometry," International Conference on 3D Vision(3DV), pp.424-433. 2019.
- 18) W. Yuan, X. Gu, Z. Dai, S. Zhu and P. Tan, "New CRFs: Neural Window Fully-Connected CRFs for Monocular Depth Estimation," arXiv preprint arXiv:2203.01502, 2022.
- 19) A. Geiger, The KITTI Depth Prediction Benchmark, https://www.cvlibs.net/datasets/kitti/eval_depth.php?benchmark=depth_prediction, 2012.
- 20) Y. K. Abdelmohsen, Camera-LIDAR Detection Fusion, Bachelor Degree, German University in Cairo, New Cairo City, Egypt, 2020.
- 21) H. Sommer, R. Khanna, I. Gilitschenski, Z. Taylor, R. Siegwart and J. Nieto, "A Low-Cost System for High-Rate, High-Accuracy Temporal Calibration for Lidars and Cameras," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.2219-2226, 2017.
- 22) M. Faizullin, A. Kornilova and G. Ferrer, "Open-Source Lidar Time Synchronization System by Mimicking GNSS-Clock," IEEE International Symposium on Precision Clock Synchronization for Measurement, Control, and Communication(ISPCS), pp.1-5, 2022.
- 23) F. Azra, D. Jurić and D. Osmanković, "The Procedure of a Camera Calibration using Camera Calibration Toolbox for MATLAB," Proceedings of the 35th International Convention MIPRO, IEEE, 2012.