DiTer: Diverse Terrain and Multi-Modal Dataset for Field Robot Navigation in Outdoor Environments

Seokhwan Jeong[1*] Hogyun Kim[1*] and Younggun Cho[1†]

Abstract—**Construction robots require autonomy in diverse environments to navigate and map their surroundings efficiently. However, the lack of diverse and comprehensive datasets hinders the evaluation and development of autonomous construction robots. To address this challenge, we present a multi-modal and diverse terrain dataset for the ground mapping of construction robots. The dataset includes various terrain types, such as sandy roads, vegetation, and sloping terrain. It comprises RGB-D cameras, thermal camera, light detection and ranging (LiDAR), inertial measurement unit (IMU), and global positioning system (GPS). We utilize a quadrupedal robot as a base platform to collect the dataset. The dataset and supplement materials are available at `https://sites.google.com/inha.edu/diter/datasets`.**

I. INTRODUCTION

In robot navigation, the characteristic of the terrain is essential since it greatly affects the safety and performance of robots. For example, if the ground is slippery, the robot's wheels or feet may become unstable, causing the robot to roll. Also, if the ground is sloped or has large differences in elevation, the robot may fall. In such hazardous situations, traversability estimation plays an important role in keeping the robot safe during navigation. Traversability is a concept that takes attributes of terrain nearby such as roughness and slope into account and then converts them into a binary value or probability score which indicates the robot can go or not even in an unstructured environment.

Unlike well-structured environments like paved roads, constructed and unstructured sites may include uneven terrain that threatened the platform. Although there are several datasets for rugged terrain environments [1], [2], we expect the need for supplementary ground-level data to robustly derive the traversability. To this end, the RGB-D camera in our experiment is deposited to face ground to solely contain information about nearby terrain. We deploy a quadrupedal robot to collect diverse terrain scenarios.

The main contributions of this paper are as follows:

- We propose a dataset that traverses multi terrain included HILL01, HILL02, FOREST, LAWN.
- The dataset is collected by multiple perceptual sensors including RGB-D, thermal camera, RGB, and LiDAR.
- The dataset contains motion measurements including IMU, GPS, and odometry of a quadrupedal robot.

[1*]S. Jeong [1*]H. Kim and [1†]Y. Cho is with Dept. the Electrical and Computer Engineering, Inha University, Incheon, South Korea [eric5709, hg.kim]@inha.edu, yg.cho@inha.ac.kr
(*) represents equal contribution.
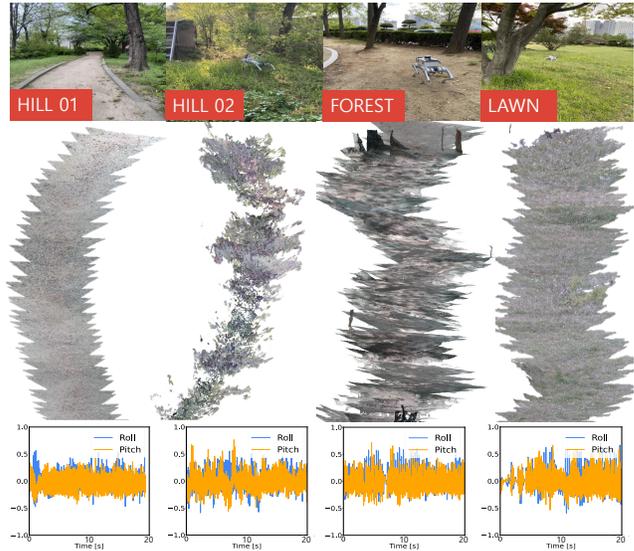
Fig. 1. Illustration of test environments and ground mapping results: HILL01, HILL02, FOREST, LAWN. (Top) The first row represents the environment of our dataset's sequence. (Middle) The second row is the accumulated depth camera's pointcloud along the trajectories from conventional LiDAR SLAM [3]. HILL02's pointcloud has lots of vacancies since the tall grass obscures the camera. (Bottom) The third row shows graphs that represent the roll and pitch in angular velocities [rad/s] from the IMU.

II. RELATED WORKS

In this section, we describe various terrain datasets and related research. In recent years, many datasets have become available online and are actively studied. However, most of the papers focus on urban environments, and the available datasets for field environments generally include single or similar terrains. In reality, various terrains with different properties coexist, such as paved and unpaved roads.

The urban environment is structurally simple and can assume a variety of settings, and traversability can be easily assessed or determined due to well-structured environments and man-made rules. However, challenges still exist, such as different characteristics among cities or GPS signal obstruction around high buildings. Many datasets have been developed for urban environments, including complex urban scenes [7] and multi-modal datasets for urban place recognition [8]. On the other hand, field environments can be more complex and diverse, as they often contain various terrain types and natural obstacles. While there are some datasets for field environments, most of them include single or similar terrains. Additionally, construction sites often include a combination of both urban and field environments.

	Motion	Environments	RGB-D for Ground	RGB	Thermal	LiDAR	GPS	IMU
RUGD [1]	Wheel Robot	Offroad	x	✓	x	✓	✓	✓
Rellis-3D [2]	Wheel Robot	Offroad	x	✓	x	✓	✓	✓
SubT [4]	Wheel Robot	Underground	x	✓	✓	✓	x	✓
Wild-Places [5]	Handheld	Jungle	x	✓	x	✓	x	✓
Hilti [6]	Handheld	In- / Outdoor	x	✓	x	✓	x	✓
Ours	Quadrupedal Robot	Diverse Terrain	✓	✓	✓	✓	✓	✓

In an environment that lacks clear rules, guidelines, or organization, there may be insufficient prescribed routines, explicit instructions, or established patterns of behavior. Wigness et al. [1] proposed pixel-wise semantic labeling datasets in off-road autonomous visual navigation. The datasets supported fine-grained terrain identification for path planning tasks. Because the RUGD dataset focuses on visual navigation, there is LiDAR Data, but it does not have point-wise labeling. To compensate for this, Jiang et al. [2] suggests a dataset that includes both semantic labels for visual navigation and pointcloud semantic labeling for LiDAR-based navigation. Using 3D information, they can recognize objects that cannot be found only in 2D, however, abstract information such as traversable/non-traversable cannot include. Recently, Knights et al. [5] introduced Wild-Places Datasets, a challenging large-scale dataset for LiDAR place recognition in unstructured, natural environments. Likewise, the Hilti SLAM challenge [6] and Sub-T challenge [4] have been held annually and contribute by providing datasets for various terrain.

However, for safe navigation, direct information on the ground in front of the robot is needed to determine the traversability. Also, information to overcome dark environments such as caves and nights cannot be ignored. Therefore, unlike other datasets as represented in Table I, we directly grasp the information about the ground through RGB-D Camera and use a thermal camera to secure robustness in dark places and plan for future work. We also utilize a quadrupedal robot to provide data on environments where wheeled robots are difficult to navigate.

III. SENSOR SETUP AND CALIBRATION

A. Sensor setup

Fig. 2 represents the robot platform and compact sensor configuration. Sensor specifications and each rostopic are configured as Table II. To capture visual measurements with different wavelengths, the RGB and thermal cameras are placed to form a forward-looking sensor system. The terrain-looking sensor system uses an RGB-D camera to obtain sufficient ground information. In addition, LiDAR enables the recognition of the surrounding environment regardless of angle. Finally, GPS, IMU, and built-in odometry are responsible for the robot's Navigation.
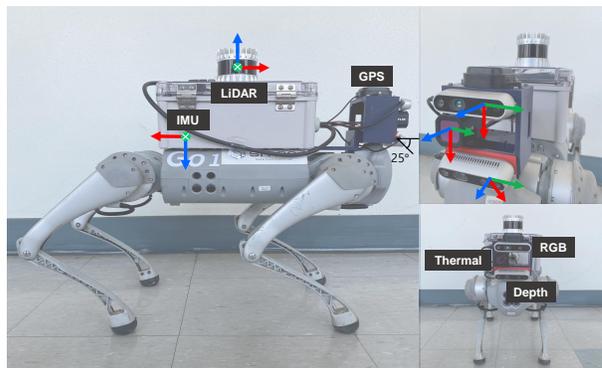
Fig. 2.   Robot and sensors used in DiTer Dataset

Hardware	Sensors	Specifications	Topic name
Intel NUC	RGB-D for Ground	Intel Realsense D435i	/cam_1/depth/color/points
			/cam_1/depth/camera_info
			/cam_1/color/image_raw
			/cam_1/color/camera_info
	RGB	Intel Realsense D435i	/cam_2/color/image_raw
			/cam_2/color/camera_info
	16bit−Thermal	FLIR Boson 640	/flir_boson/image_raw
			/flir_boson/camera_info
	LiDAR	Ouster OS1-32	/ouster/points
	GPS	Ublox C099-F9P with ANN-MB antenna	/ublox_gps/fix/
	9DOF−IMU	LORD Microstrain 3DM-GX5-25	/gx5/imu/data
Unitree-GO1	6DOF−IMU	Built in Quadrupedal robot	/dog_odom
			/imu_raw

B. Intrinsic Calibration

We utilize `ROS Camera Calibration` to camera intrinsic parameters. Fig. 3 represents the checkerboard for camera calibration. Also, we apply pixel value inversion to thermal images as described in Fig. 3 (b) and (c). Because

Fig. 3(a)   **Fig. 3(b)**   **Fig. 3(c)**
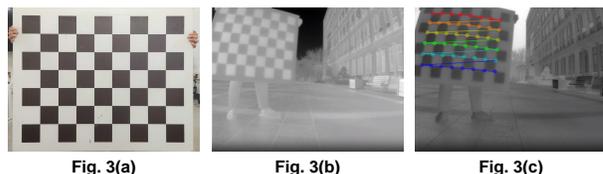
Fig. 3.   Calibration Images. (a) Our calibration board (b) Normal 8-bit thermal image (marker detection failure) (c) Inverted Thermal Image with marker detection.
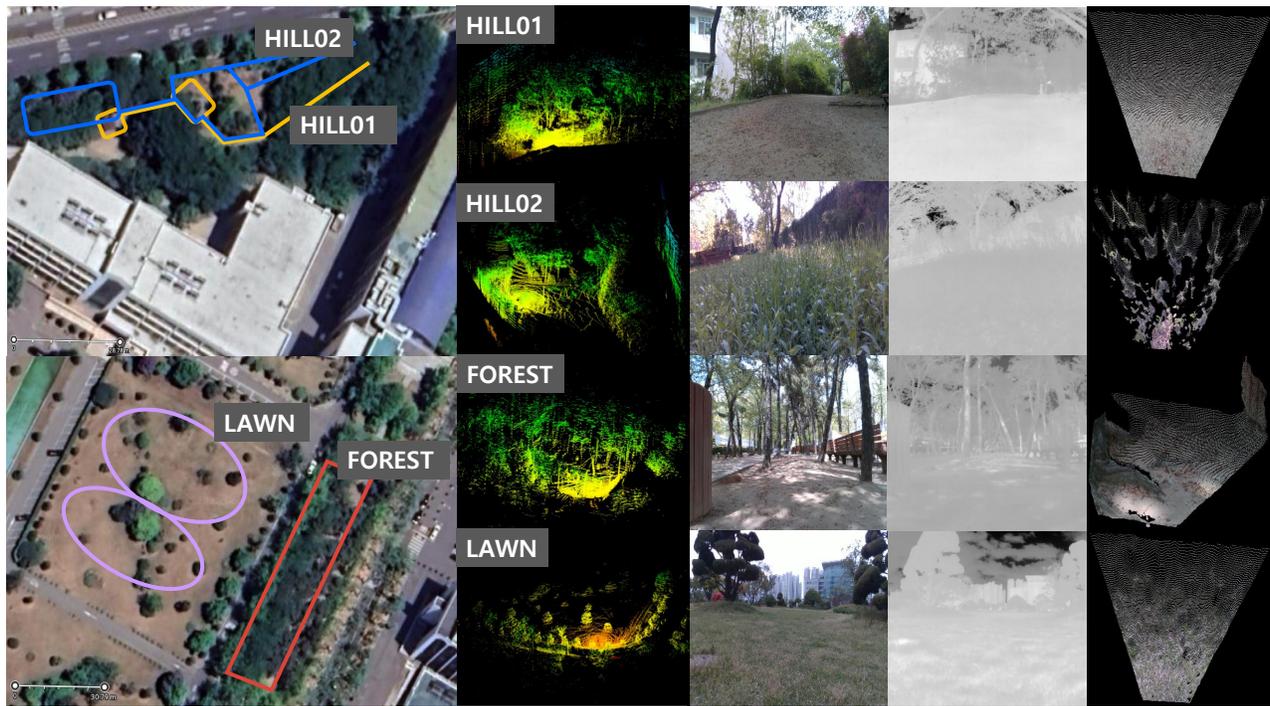
Fig. 4. Reference trajectory of each sequence from the satellite map. From left to right, the LiDAR submap, sample images of RGB, thermal, and pointcloud from the depth camera are shown.

the thermal values of the black and white pattern are opposed to the original pixel values, we inverted the pixel values in 8-bit thermal images.

C. Extrinsic Calibration

It is essential to know the transformations between sensors to utilize multiple sensor measurements. We first applied marker-based LiDAR-camera calibration [9]. As described in Fig. 3, the UV-printed checkerboard was used to capture co-visible measurements from cameras and LiDAR. We selected the nearest neighbor pairs to the LiDAR data for RGB and thermal cameras and used the measurements in the optimization process. For LiDAR-IMU calibration, we utilized the robust real-time LiDAR-inertial initialization method proposed in [10]. The projected LiDAR pointclouds are represented in Fig. 5. The figure confirms the results of
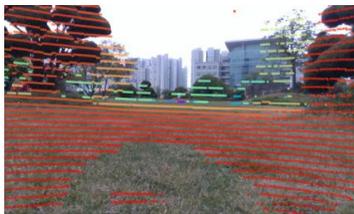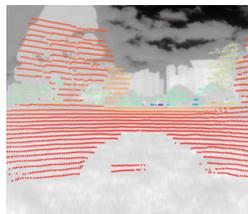
Fig. 5(a) **Fig. 5(b)**

Fig. 5. Pointcloud projection images. (a) and (b) are images of LiDAR points projected on the camera image and thermal camera image, respectively. Since the resolution and size of the images of each camera are different, we identify the different projected images.

extrinsic calibration.

IV. DATASET

All sequences are obtained within the outdoor sites on the campus. Information and trajectory for each sample route can be found in Fig. 4. Table III summarizes the names of our datasets and brief descriptions, etc. A detailed description of each sequence is as follows:

1) Hill: HILL sequences are environments composed of a gravel road and tall grass on either side of the road. Each sequence takes two laps around the same path.
   - HILL01 is targeted mainly on the gravel road.
   - HILL02 focuses on poor environments including under-canopy.
2) Forest: FOREST is a terrain where trees are sporadically planted. People are frequently appearing in this sequence. Also, the FOREST sequence is designed to make a loop by driving around the same path twice.
3) Lawn: LAWN is obtained in a rectangular park with hills scattered. In this sequence, the robot travels around with the shape of an infinite symbol (∞) for two laps using the stones in the center as a base point.

V. DISSCUSION & FUTURE WORKS

We proposed DiTer, a diverse terrain dataset leveraging ground details. Our sequences contain various challenges such as loss of data caused by tall grass. We utilize an RGB-D camera as well as thermal, LiDAR, and IMU to additionally interpret the robot's state in rugged terrain. We hope that our dataset support research targeting terrain where the wheeled platform is difficult to traverse.

TABLE III
SUMMARY OF OUR DATASET

Number	Sequence Name	Terrain	Description	Length	Duration
1	HILL01	Pebble	Gravel Road	351m	790s
2	HILL02	Pebble and Lush vegetation	Switching between the two environments	446m	1070s
3	FOREST	Land with scattered trees	Traverse in a square	350m	746s
4	LAWN	Lawn with hills	Drawing infinite symbol	501m	991s

In this paper, we also provide the reference trajectories by applying the well-known LiDAR SLAM method (LIO-SAM [11]). Because LiDAR-based mapping is also popular research for autonomous construction robots, we plan to evaluate various LiDAR-inertial SLAM approaches on our datasets. Also, we will extend the dataset to include multi-session measurements by capturing the sequences at night time.

VI. ACKNOWLEDGMENT

REFERENCES

[1] M. Wigness, S. Eum, J. G. Rogers, D. Han, and H. Kwon, "A rugd dataset for autonomous navigation and visual perception in unstructured outdoor environments," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 5000–5007.

[2] P. Jiang, P. Osteen, M. Wigness, and S. Saripalli, "Rellis-3d dataset: Data, benchmarks and analysis," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 1110–1116.

[3] J. Zhang and S. Singh, "LOAM: lidar odometry and mapping in real-time," in *Proc. Robot.: Science & Sys. Conf.*, 2014.

[4] J. G. Rogers, J. M. Gregory, J. Fink, and E. Stump, "Test your slam! the subt-tunnel dataset and metric for mapping," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 955–961.

[5] J. Knights, K. Vidanapathirana, M. Ramezani, S. Sridharan, C. Fookes, and P. Moghadam, "Wild-places: A large-scale dataset for lidar place recognition in unstructured natural environments," *arXiv preprint arXiv:2211.12732*, 2022.

[6] M. Helmberger, K. Morin, B. Berner, N. Kumar, G. Cioffi, and D. Scaramuzza, "The hilti slam challenge dataset," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7518–7525, 2022.

[7] J. Jeong, Y. Cho, Y.-S. Shin, H. Roh, and A. Kim, "Complex urban dataset with multi-level sensors from highly diverse urban environments," *The International Journal of Robotics Research*, vol. 38, no. 6, pp. 642–657, 2019.

[8] G. Kim, Y. S. Park, Y. Cho, J. Jeong, and A. Kim, "Mulran: Multimodal range dataset for urban place recognition," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 6246–6253.

[9] D. Tsai, S. Worrall, M. Shan, A. Lohr, and E. Nebot, "Optimising the selection of samples for robust lidar camera calibration," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2021, pp. 2631–2638.

[10] F. Zhu, Y. Ren, and F. Zhang, "Robust real-time lidar-inertial initialization," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 3948–3955.

[11] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," in *2020 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2020, pp. 5135–5142.