CEA Comité Español de Automática

XLV JORNADAS AUTOMÁTICA MÁLAGA 2024

UNIVERSIDADE DA CORUÑA

Jornadas de Automática

Perception Sensor Integration for Improved Environmental Reconstruction in Quadruped Robotics

Cruz, C.[a,*], del Cerro, J.[a], Barrientos, A.[a]

[a] *Centro de Automática y Robótica (CSIC-UPM) Universidad Politécnica de Madrid — Consejo Superior de Investigaciones Científicas, 28006 Madrid, Spain; christyan.cruz.ulloa@upm.es*

To cite this article: Cruz, C., del Cerro, J., Barrientos, A. 2024. Perception Sensor Integration for Improved Environmental Reconstruction in Quadruped Robotics.
Jornadas de Automática, 45. https://doi.org/10.17979/ja-cea.2024.45.10830

Abstract

Perception systems are fundamental in outdoor robotics, as their correct functionality is essential for tasks such as terrain identification, localization, navigation, and analysis of objects of interest. This is particularly relevant in search and rescue (SAR) robotics, where one current research focuses on the mobility and traversal of unstructured terrains (commonly resulting from natural disasters or attacks) using quadruped robots. 3D sensory systems, such as those based on 360-degree LiDAR, tend to create dead zones within a considerable radius relative to their placement (typically on the upper part of the robot), leaving the locomotion system without terrain information in those areas. This paper addresses the problem of eliminating these dead zones in the robot's direction of movement during the process of environment reconstruction using point clouds. To achieve this, a ROS-based method has been implemented to integrate "n" point clouds from different sensory sources into a single point cloud. The applicability of this method has been tested in generating elevation maps of the environment with different resolutions, using the quadruped robot ARTU-R (A1 Rescue Task UPM Robot) and short- and long-range RGB-D sensors, strategically placed on its lower front part. Additionally, the method has demonstrated real-time functionality and robustness concerning the issue of frame association in the fusion of information from decentralized sources. The code is available to the community in the authors' GitHub repository `https://github.com/Robcib-GIT/pcl_fusion`.

Keywords: Perception and sensing, Mobile robots, field robotic, Sensor integration and perception, Map building.

1. Introduction

The search and rescue robots field has experienced remarkable growth in recent years to improve the efficiency and safety of missions in hazardous environments resulting from natural or human-caused disasters. These advances seek to protect the lives of rescue brigades, conducting initial inspections in search of victims. One of the primary goals of this area of robotics is to provide support to these brigades during post-disaster interventions, safeguarding their lives while assisting in the detection of victims (Whitman et al., 2018; Wannous and Velasquez, 2017; Cruz Ulloa et al., 2021, 2023a; Ulloa et al., 2023).

Some real-world situations where these interventions have been utilized include events such as the attack on the Twin Towers (United States - 2001) (Blackburn et al., 2002), the Fukushima nuclear accident (Japan - 2011) (Eguchi et al., 2012), the Amatrice earthquake (Italy - 2016) (Kruijff et al., 2016), and the Mexico City earthquake (Mexico - 2017) (Whitman et al., 2018). Post-disaster events create environments with unstructured terrain that hinder rescue missions. In this context, quadruped robots have demonstrated great skill navigating these terrains, outperforming conventional robots (wheel, leg, track locomotion) Cruz Ulloa et al. (2023b); Cruz Ulloa (2024).

On the other hand, perception systems based on lidar or infrared projection commonly tend to leave blind spots that are not captured and, therefore, not included in subsequent reconstruction phases Benedek et al. (2021); Li and Ibanez-Guzman (2020); Li et al. (2024). These blind spots tend to encompass

*Corresponding author: christyan.cruz.ulloa@upm.es

a cone relative to the sensor's vertical axis and vary depending on the height of placement, leaving the robot without a field of vision within a radius (typically 1.5 [m]), which presents a problem for planning in that workspace, especially for legged robots. Figure 1 shows an environment reconstruction carried out using a front-facing RGB-D camera, distinguishing a radius of 1.7 [m] around the robot, over which no elevation information is available, while the 2D lidar system assumes it to be free space.

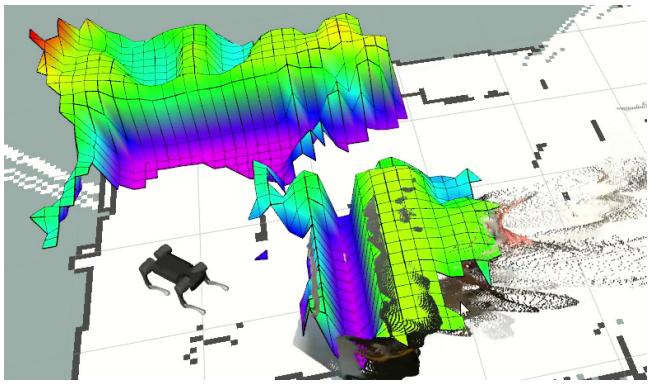


Figure 1: Reconstruction of an elevation map using a front-facing RGB-D sensor, with a blind spot around the robot.

This paper proposes a method to minimize blind spots in the perception system of legged robots. A method based on a Robot Operating System (ROS) has been implemented. It can receive multiple inputs (depth images and point clouds), preprocess them, integrate the "n" point clouds from different sources into a single cloud, and post-process them to associate them with the robot's reference system.

The method was tested using the quadruped robot ARTU-R (A1 Rescue Task UPM Robot), strategically equipped with short and long-range RGB-D cameras in its forward field of vision. Elevation maps were reconstructed using this perception data, evaluating different resolutions and the method's functionality. The main results demonstrated the minimization of blind spots and real-time operational capability.

This manuscript is structured as follows: Section 2 presents the proposed materials and methods, Section 3 discusses experiments and results, and Section 4 highlights the main conclusions and future work arising from this development.

2. Methodology

2.1. Materials

The quadruped robot ARTU-R from Unitree, model A1, was used for this development. It was equipped with two RGB-D sensors: the first one (A. Realsense D435i) has a working range of up to 10 meters depending on the environment, and the second one (B. Realsense SR305) is typically used for short-range applications with a range between 0.2 to 1.5 meters, as shown in Figure 2.

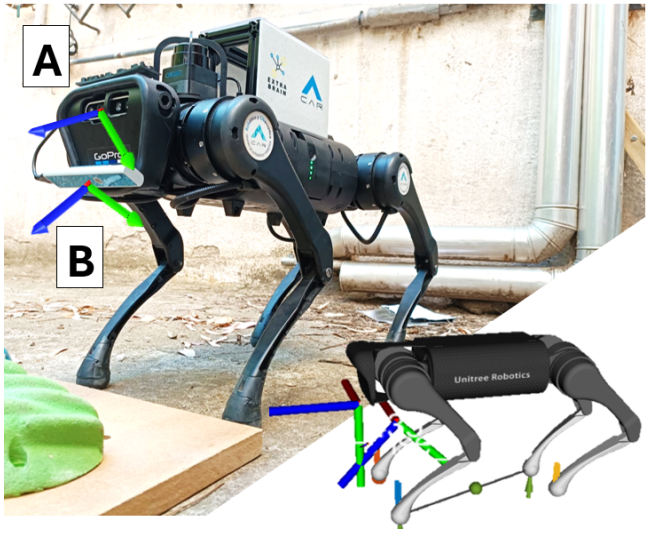


Figure 2: The implementation was developed using the quadruped robot ARTU-R, equipped with two RGB-D cameras mounted on its front-lower section to maximize the reconstructed area during movement.

The arrangement of the sensors, with camera A mounted on the front and camera B inclined at an angle of 60 degrees relative to the robot's horizontal axis, has been fixed in place using mechanical couplings. This configuration (position and orientation relative to the centre of the robot) has been replicated using fixed frames in the robot's .xacro model, shown as a reference in the bottom right corner of Figure 2. The model and CAD files of the robot used in the simulation process to represent the robot's states are available in the authors' GitHub repository in Appendix A. The development and experimentation phase was conducted at the Centro de Automática y Robótica (UPM-CSIC) facilities (40°26'21.5" N 3°41'18.1" W)

2.2. Methods

Figure 3 illustrates the process of point cloud fusion from multiple inputs (depth sensors and lidar). Each depth sensor captures data ($Depth\_Scan\_1$, $Depth\_Scan\_2$, ..., $Depth\_Scan\_n$), which is processed using the $sensor\_msgs$ and cv_bridge libraries to transform image pixel coordinates into 3D point cloud coordinates ($Point\_cloud\_1$, $Point\_cloud\_2$, ..., $Point\_cloud\_n$). At this stage, point clouds directly from lidar sensors ($Point\_cloud\_n + 1$) can also be incorporated.

The point clouds are referenced to a common frame ($/tf\_referenced$) for alignment. Then, they are fused into a single point cloud ($Fused\_cloud$) through processes of resizing, downsampling, noise removal, rotation, and translation. Resizing ensures uniform point density; downsampling reduces computational load; noise removal improves data quality; rotation correctly aligns the clouds in three-dimensional space; and translation positions the clouds in the common reference frame.

The result is a fused point cloud that provides a detailed and accurate representation of the environment. It integrates the advantages of the different sensors used, particularly those of short—and long-range. These algorithms are available in Appendix A.

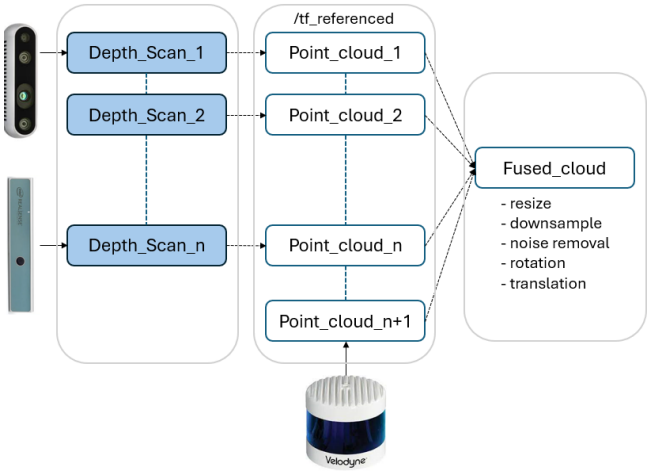


Figure 3: Schematic of the functionality of the implemented system for sensor integration from different sources.

In Figure 4, the upper left section shows the depth image from camera A with intensity variations; dark areas represent closer regions, while light areas indicate greater distances. The upper section shows the depth image from camera B. The respective lower sections display the local point clouds associated with the respective frames, coloured according to the vertical axis relative to the maximum height. The point cloud associated with camera B complements the first one, overlapping the area between the robot's legs and the beginning of that cloud.

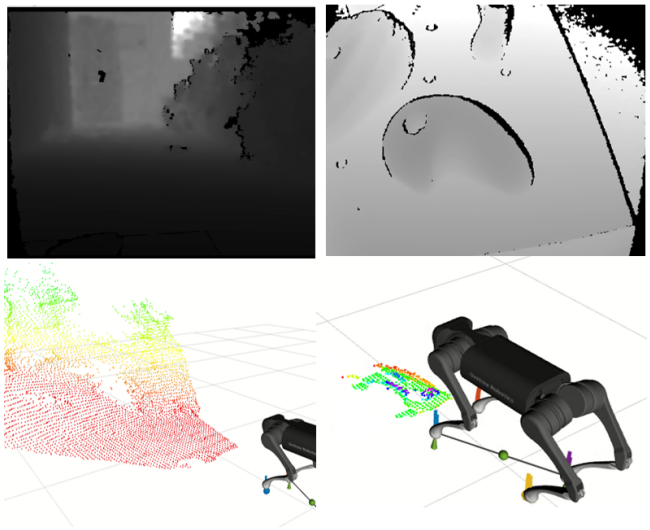


Figure 4: Partial reconstructions with each camera of ARTU-R are shown, including the depth images and the generated point clouds. (Left) The front camera has a range of up to 5 meters but features a blind spot within a 60 cm radius. (Right) The lower camera covers the blind spot of the front camera.

3. Experiments and Results

3.1. Variation of functionality parameters.

For the experimental phase, optimal parameters for the functionality of the combined cloud processing for environment reconstruction using elevation maps have been determined following the development of Fankhauser et al. (2018). These maps are generated from an input of point clouds and an associated odometry frame from ARTU-R. Figure 5 shows the result of the reconstruction of the test area in front of the robot based on different parameters, aiming to balance real-time processing and the maximization of the test area reconstruction (covered area). This area, composed of a platform and anchored modules, was placed in front of the robot in the previously blind spot. The experimentally determined parameters were an update frequency of 10 Hz, a spatial resolution of 0.02 m between points, and the downsampling value depending on the number of points; in this case, a 25% reduction was sufficient.

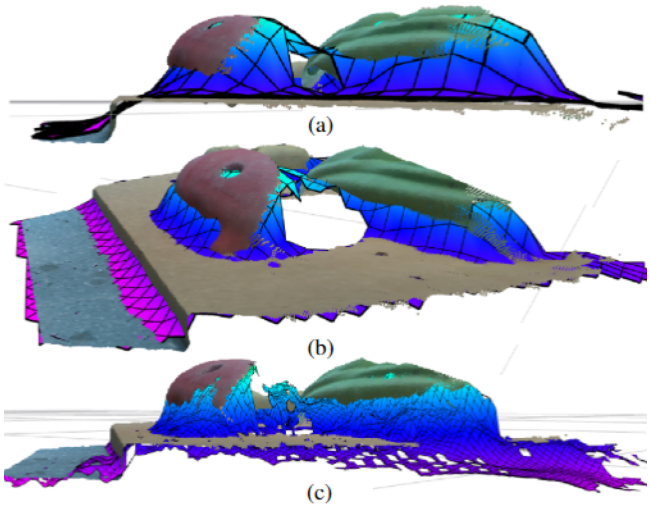


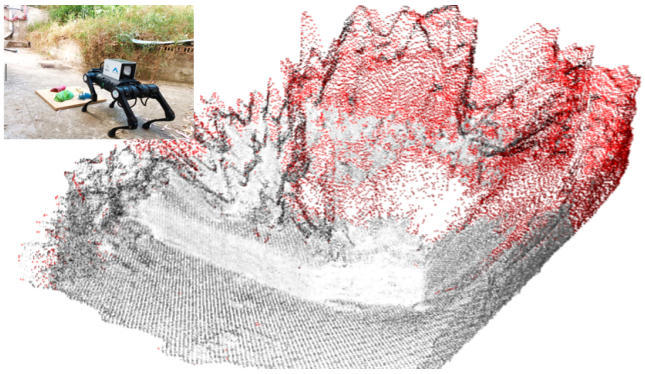
Figure 5: Reconstruction of the test area confirmed by elements anchored to a 0.5x0.5 meter panel. (a) 5 cm mesh. (b) 2 cm mesh. (c) 0.5 cm mesh.

Figure 5 highlights the different types of resolutions superimposed on the test area. In case (a), the geometric resolution of details is lost, but the reconstructed area is maximized (with the robot stationary). Case (c) maximizes the details, but the computational processing increases significantly. Case (b) balances processing speed and reconstruction details, making it the best option.

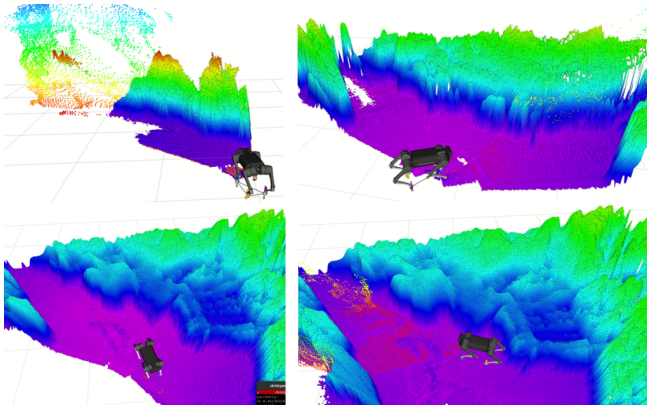
3.2. Application to the reconstruction of environments

Once the operating parameters were defined, the reconstruction of an outdoor environment was carried out. This environment, characterized by vegetation and medium-sized walls, had been previously reconstructed by the authors using the RTAB-Map method, as detailed in Cruz Ulloa et al. (2024). Figure 6(a) shows the result of the reconstruction with RTAB-Map in grey and provides a perspective of the environment in the upper left corner.

On the other hand, figure 6(b) shows the reconstruction of the same environment (in false colour according to the vertical axis) using the sensory input method proposed in this article. Four phases of the robot's movements to complete the reconstruction are shown, completing a 360-degree turn to capture ninety per cent of the information. Table 1 presents the significant improvements in reconstruction with the proposed method; for this evaluation, ten reconstruction tests were carried out.



(a)



(b)

Figure 6: (a) Reconstruction of an outdoor environment developed by the authors in Cruz Ulloa et al. (2024). (b) Stages of environment reconstruction with the implemented method.

Table 1: Comparative evaluation of parameters for environment reconstruction, related to the development shown in Figure 6.

	Single sensory input	**Strategic sensory combination**
Reconstruction time	56 s	13 s
Required robot maneuverability	High	Low
Spatial resolution	0.02 m	0.02 m
Complexion index (A Captured/A Total)	89.3 %	98.2 %

3.3. Relevance of the Proposed Method

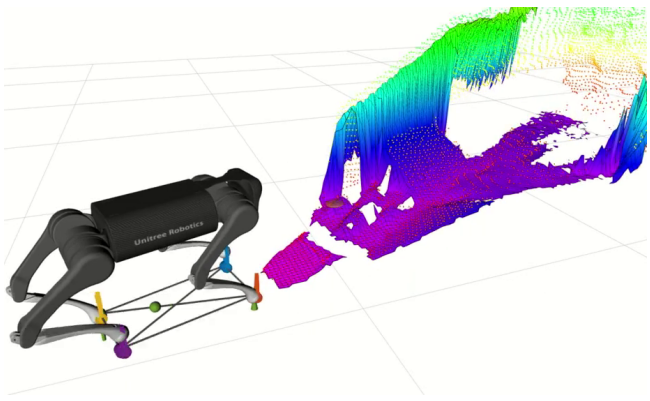


Figure 7: Initial state of capturing environmental information before movement.

One of the main advantages of the implemented method is the rapid acquisition of spatial depth information for subsequent trajectory estimation. In Figure 7, the robot is shown in a static state before movement, with its four legs forming a support polygon. Here, a projection of the elevation map is evident, extending from the legs to the first three meters, where a wall is located.

The visible field is initially reconstructed from the robot's perspective in a static state 30 centimetres above the ground. However, as the robot advances, terrain details are refined using the short-range sensor. The method for generating the elevation map relies on the odometry provided by the robot's IMU, allowing the map to be refined with a higher degree of detail. Therefore, if the robot experiences oscillations and sharp turns during movement, the generation of the map is not affected.

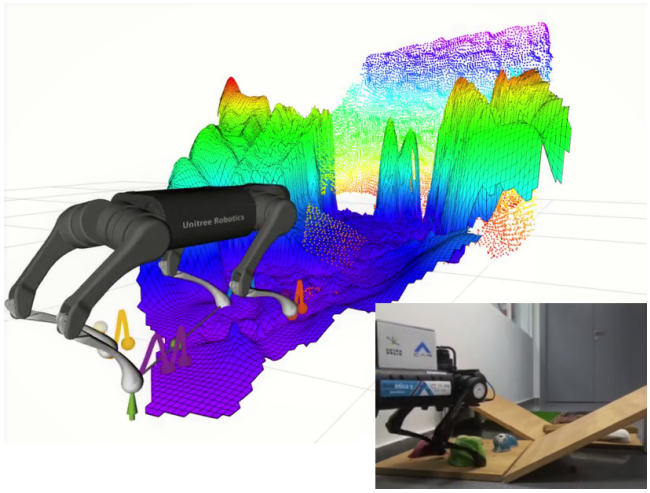


Figure 8: Highly detailed elevation map, corresponding to the robot's progress over unstructured terrain.

Figure 8 shows the robot's progress over a test scenario composed of a series of planks with ramps and debris. Compared to Figure 7, a higher level of detail in the elevation map reconstruction is evident. In the first instance (initial information capture state), this would be useful for generating a general plan to approach the terrain in a certain direction. Meanwhile, the adjustments with the second sensor would allow corrections (local planner) to be made to the initial global plan.

4. Conclusions and Future Work

This work has developed and implemented a method for minimizing occluded areas in robots' fields of vision. It highlights its functionality and applicability to legged robots by fusing information from sensors strategically placed in the robot's forward field of view, using short—and long-range RGB-D cameras.

The proposed method has demonstrated real-time functionality and robustness in decentralized systems. It avoids timing latencies and synchronization issues in the association of frames when handling point clouds from multiple sources.

The pre and post-processing phases of point clouds have facilitated their applicability to the generation of elevation

maps in environment reconstruction, thanks to the management and handling of scales, noise, and parameters of the clouds from their acquisition.

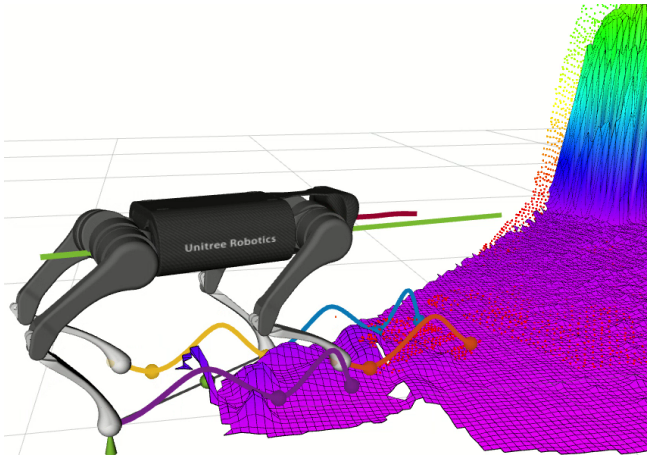


Figure 9: Displacement of the quadruped robot over terrain with a 2-2 alternating gait pattern, where sinusoidal waves of different colours represent the trajectories followed by the tip of each leg.

In future work, the integration of elevation maps with the robot's gait planning system is proposed, incorporating techniques based on BLNN (Brain-Like Neural Networks) to help predict movements for positioning the robot's leg at each moment, as well as optimizing routes through these maps to navigate terrain more efficiently in terms of both safety and time. Figure 9 demonstrates the current functionality of the planner on flat terrain (the elevation map is overlaid for visual purposes only), where different colours (blue, red, purple, and yellow) of the sinusoidal waves indicate the trajectory that each leg will follow with each step.

Funding

This research has been possible thanks to the financing of "Proyecto CollaborativE Search And Rescue robots (CESAR)" (PID2022-142129OB-I00) founded by MCIN/AEI/10.13039/501100011033 and "ERDF A way of making Europe".

Appendix A.

GitHub Repository: `https://github.com/Robcib-GIT/pcl_fusion`

References

Benedek, C., Majdik, A., Nagy, B., Rozsa, Z., Sziranyi, T., 2021. Positioning and perception in lidar point clouds. Digital Signal Processing 119, 103193.
URL: `https://www.sciencedirect.com/science/article/pii/S1051200421002323`
DOI: `https://doi.org/10.1016/j.dsp.2021.103193`

Blackburn, M. R., Everett, H. R., Laird, R. T., 8 2002. After action report to the jointprogram office: Center for the robotic assisted search and rescue (crasar) related efforts at the world trade center. Tech. rep., SPACE AND NAVAL WARFARE SYSTEMS CENTER SAN DIEGO CA.

Cruz Ulloa, C., Marzo 2024. Quadrupedal robots in search and rescue : Perception and teleoperation, no Publicado.
URL: `https://oa.upm.es/81769/`
DOI: `10.20868/UPM.thesis.81769`

Cruz Ulloa, C., Garcia, M., del Cerro, J., Barrientos, A., 2023a. Deep learning for victims detection from virtual and real search and rescue environments. In: Tardioli, D., Matellán, V., Heredia, G., Silva, M. F., Marques, L. (Eds.), ROBOT2022: Fifth Iberian Robotics Conference. Springer International Publishing, Cham, pp. 3–13.

Cruz Ulloa, C., Prieto Sánchez, G., Barrientos, A., Del Cerro, J., 2021. Autonomous thermal vision robotic system for victims recognition in search and rescue missions. Sensors 21 (21).
URL: `https://www.mdpi.com/1424-8220/21/21/7346`
DOI: `10.3390/s21217346`

Cruz Ulloa, C., Sánchez, L., Del Cerro, J., Barrientos, A., 2023b. Deep learning vision system for quadruped robot gait pattern regulation. Biomimetics 8 (3).
URL: `https://www.mdpi.com/2313-7673/8/3/289`
DOI: `10.3390/biomimetics8030289`

Cruz Ulloa, C., Álvarez, J., del Cerro, J., Barrientos, A., 2024. Vision-based collaborative robots for exploration in uneven terrains. Mechatronics 100, 103184.
URL: `https://www.sciencedirect.com/science/article/pii/S0957415824000497`
DOI: `https://doi.org/10.1016/j.mechatronics.2024.103184`

Eguchi, R., KenElwood, Lee, E. K., Greene, M., 2012. The 2010 canterbury and 2011 christchurch new zealand earthquakes and the 2011 tohoku japan earthquake. Tech. rep., Earthquake Engineering Research Institute.

Fankhauser, P., Bloesch, M., Hutter, M., 2018. Probabilistic terrain mapping for mobile robots with uncertain localization. IEEE Robotics and Automation Letters (RA-L) 3 (4), 3019–3026.
DOI: `10.1109/LRA.2018.2849506`

Kruijff, I., Freda, L., Gianni, M., Ntouskos, V., Hlavac, V., Kubelka, V., Zimmermann, E., Surmann, H., Dulic, K., Rottner, W., Gissi, E., Oct 2016. Deployment of ground and aerial robots in earthquake-struck amatrice in italy (brief report). In: 2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR). pp. 278–279.
DOI: `10.1109/SSRR.2016.7784314`

Li, Y., Ibanez-Guzman, J., 2020. Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems. IEEE Signal Processing Magazine 37 (4), 50–61.
DOI: `10.1109/MSP.2020.2973615`

Li, Y., Kong, L., Hu, H., Xu, X., Huang, X., 2024. Optimizing lidar placements for robust driving perception in adverse conditions.

Ulloa, C. C., Llerena, G. T., Barrientos, A., del Cerro, J., 2023. Autonomous 3d thermal mapping of disaster environments for victims detection. In: Robot Operating System (ROS) The Complete Reference (Volume 7). Springer, pp. 83–117.

Wannous, C., Velasquez, G., 2017. United nations office for disaster risk reduction (unisdr)—unisdr's contribution to science and technology for disaster risk reduction and the role of the international consortium on landslides (icl). In: Sassa, K., Mikoš, M., Yin, Y. (Eds.), Advancing Culture of Living with Landslides. Springer International Publishing, Cham, pp. 109–115.

Whitman, J., Zevallos, N., Travers, M., Choset, H., 2018. Snake robot urban search after the 2017 mexico city earthquake. In: 2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR). pp. 1–6.
DOI: `10.1109/SSRR.2018.8468633`