Proceedings of CONF-MLA 2026 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2026.LD37220

Perception Technologies for Quadruped Robots on Unstructured Terrain

Ziheng Gan

NingboHD School, Ningbo, China
2301012@stu.hdschools.org

Abstract. This paper reviews the environmental perception technologies for quadruped robots that tackle the issue of uneven terrain, irregularly distributed challenges in obtaining ground physical properties directly, and the obstacles in obtaining the ground physical properties in unstructured terrain. The paper starts with an analysis of stereo vision and light detection and ranging (LiDAR) applications, followed by an explanation of how stereo vision and LiDAR systems are integrated. In terrains reconstruction, obstacles recognition and localization, discussing effects of Perception performance as affected by lighting, texture and bad weather. Second, it summarizes the contribution of multisensor fusion approaches to improve perception robustness and studies problems involving computation and observation conflicts. Furthermore, it explores the uses and drawbacks of force and tactile perception in ground contact Decisions in determination, slippage detection and ground hardness estimation. Finally, addressing problems like the lack of information from one sensor, lack of real time how fusion systems work, and how motion influences contact perception. In this paper, a hierarchical closed-loop perception system is proposed, which offers a reference for the autonomous operation of quadruped robots in relief, inspection, and field disaster operations.

Keywords: Quadruped robots, unstructured terrain, multisensor fusion, force and tactile perception

1. Introduction

With the continuous development of smart manufacturing and mobile robot technology, quadruped robots are more and more widely used in the field of manufacturing automation. Robots' growth in the global robotics market is strong. The quadruped robots are also playing an important role in various fields. As opposed to wheeled and tracked robots, quadruped robots have benefits including flexibility and flexibility. They have great opportunity in, for instance, disaster relief, cargo transport and security systems, to name a few. Currently, quadruped robots use a range of types of sensors to develop perception systems. The following conclusions can be drawn from these data: they assess road conditions and surrounding environments to perform tasks including route planning and posture adaption [1, 2].

In real world scenarios, there are many areas where the terrain is not structured, and may be quite complex like irregular terrain, rocky terrain, or differences in hardness of the ground. This puts very

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.LD37220

high requirement on the robot environmental perception ability. Mao's team has been investigating path planning for a quadruped robot in complicated environments [3]. Tang et al . surveyed localization, tracking and SLAM technologies for quadped robot perception from the target, robot and environment perspectives [4]. The AMCO was suggested by Elnoor et al. that adaptively links the visual semantics, body feedback and the current state of the robot to a possability cost map [5]. But there are still some aspects that are not investigated in the current research on the quadruped robots shortcomings. When they come across unknown and unstructured environment, robots have difficulty adapting their behavior accordingly. Moreover, robots can easily fall over on slippery or soft grounds, which greatly restricting their use of application skills.

To confront these problems, this paper will deeply investigate about perception technologies that can be used for quadruped robots. In the first place, will analyze three aspects, stereovision, LiDAR and multisensor. fusion – by focusing on comparison of their characteristics in obstacle recognition, localization, terrain mapping. Secondly, we will examine the roles of force and tactile perception in ground contact detection and the physical properties of the ground. Finally, we will provide a summary and outlook for quadruped robots. So that this paper can serve as a valuable reference for closed-loop perception systems for state estimation and contact verification in quadruped robots.

2. Visual perception technology

Visual perception technology can help quadruped robots reconstruct terrain and identify obstacles in unstructured environments by using optical sensors to obtain three-dimensional geometric information about the environment. Therefore, this section analyzes three modules: stereo vision, light detection and ranging (LiDAR), and multimodal fusion, as shown in Figure 1.

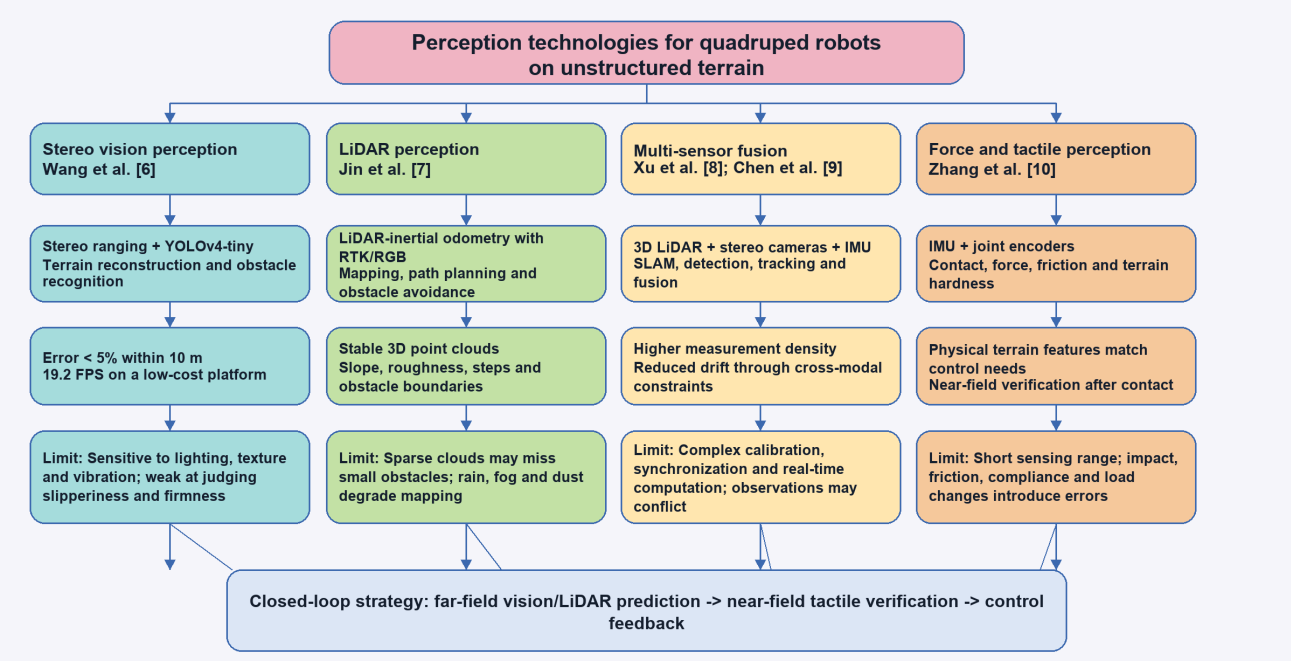


Figure 1. Perception technologies for quadruped robots on unstructured terrain. (picture credit: original)

DOI: 10.54254/2755-2721/2026.LD37220

2.1. Stereo vision perception

Stereo vision would determine the three-dimensional depth of the environment through stereo matching, combined with lightweight deep learning networks to extract semantic features of the scene, it performs terrain geometry reconstruction and material recognition. Therefore, stereo vision perception is a viable approach for implementing vision perception on low-cost quadruped platforms.

Wang et al. combined binocular stereo ranging with an improved YOLOv4-tiny object detection model for a quadruped robotic mobile platform. Distance to an object would be caculated by processing images from the left and right cameras and used a lightweight network to identify obstacle categories [6]. Experimental results showed that when the test distance was within 10 meters, the system's error remained below 5%, with an average processing speed of 19.2 FPS. Furthermore, the experimental platform for rangefinding and object detection was built using a stereo camera and an NVIDIA Xavier NX development board, demonstrating that the platform can handle both distance estimation and object recognition under limited computational resources.

While Wang's study managed to estimate the distance of objects and recognise obstacles by implementing stereo range finding and an enhanced YOLOv4-tiny object detection model, it also has some drawbacks – it fails to perform well in the absence of good lighting conditions or stable ground surfaces. In addition, using the robot in low lighting, on low-textured surfaces and steep terrain causes vibrations. Information gaps can easily be created by the content removed from the robot's body, while mechanical characteristics like slipperiness and firmness of surface and load-bearing capacity of the robot are difficult to be judged by the detection network. capacity.

The solution is to introduce online parameter calibration, image deblurring and parallax confidence assessment; to combine short range depth (or laser) measurements with low confidence areas; This allows, and continually optimises, visual judgements based on the nature of the slip and the state of support after the foot has hit the ground, which minimises the likelihood of a single misjudgement having a direct impact on gait control

2.2. LiDAR perception

The core value of LiDAR lies in its minimal susceptibility to ambient light. By providing a stable 3D point cloud, the system can calculate slope, roughness, step height, clearance, and obstacle boundaries to plan paths. Jin et al. used a tightly coupled LiDAR-inertial odometer as the basis for constructing a global geometric map when designing a guide-dog quadruped robot [7]. Furthermore, this study incorporated RTK factors and RGB information to supplement absolute localization and map representation, and combined cost maps, an improved A* algorithm, and TEB to perform global path planning and local obstacle avoidance.

In the aforementioned research, when the quadruped robot encountered small obstacles, the sparse point cloud failed to adequately represent them. Furthermore, in adverse weather conditions —such as rain, fog, or dust—three-dimensional point clouds may produce abnormal echoes, making continuous mapping computationally challenging.

To address this, ground-based adaptive segmentation, echo quality filtering, and hierarchical local maps can be employed to retain only geometric features relevant to safe foot placement; simultaneously, IMU constraints can be used to stabilize short-term motion, thereby preventing sudden changes in localization caused by single-match results.

DOI: 10.54254/2755-2721/2026.LD37220

2.3. Multi-sensor fusion perception

A major method to increase the robustness of the perception in unstructured terrain is multimodal fusion, since failure situations differ among the various types of sensors. Xu et al. combined 3D LiDAR, stereo cameras, and an IMU into a multi-sensor SLAM system; and extended the system by adding map building, object detection, and tracking, to form a single environmental perception system [8]. In contrast, Chen et al. integrated point clouds, visual features and the IMU pre-integration in a factor graph framework, and reduced the influence of the misalignment in one loop with a vision-lidar dual-loop approach [9]. For both, it is seen that fusion would not only boost the density of measurements, but also constrain drift due to cross-modal constraints.

The two studies discussed above are multimodal fusion studies, but the compromises of such an approach are predominantly in the form of system complexity and real time resource demands. As the number of sensors increases, calibration and maintenance costs increase, as does the complexity of time synchronization and the quantity of data bandwidth and if various sensors are impacted, then there can be conflicting data.

One possible solution to problems of high computational load and observation conflict in multi-sensor systems is to employ a multi-frequency architecture, where high frequency signals (e.g., the IMU and joint states) are used for short-term stabilization and low frequency signals (e.g., vision and LiDAR) are used to correct the map. Furthermore, the time alignment, credibility assessment should be carried out for all the sensors for balancing real time performance and robustness.

3. Force and tactile perception

External geometric observations are augmented with force and tactile perception, which cannot be replaced. Data from various devices can be correlated to calculate the time of ground contact, the magnitude of the normal support force, the nature of tangential friction forces, the point of contact with the ground and ground compliance. These quantities can sometimes provide information on the state of the road surface, such as softness of the ground, or subsided road. Zhang et al. used an IMU and joint encoders to create a terrain classification method. In contrast, this study uses physical terrain features (roughness, slippage, hardness, slope, etc.) to describe terrain, which better matches practical control requirements than semantic categories such as grass or gravel [10].

The drawbacks of the above works are short sensing range and dependenceof contact on the motion state of the robot and the mechanical structure errors. The typical robot can only verify surface properties after contact; however, impact, joint friction, the compliance of the robot and load changes can influence the calculation of these properties. Additionally, normal impacts that occur while moving fast can be mistaken for collisions and/or slippage. In order to resolve these problems, it is necessary to build a contact model as mentioned above based on the combining of joint states, IMU data and local terrain prediction. If the system senses slip or sink, it should slow or change its posture or choose different foot placement. Thus, tactile feedback can modify visual information in a feedback loop – a closed process that goes from far-field prediction to near-field verification.

4. Summary of perception technologies

Table 1 summarizes the main limitations, and suggestions for the perception technologies discussed above.

DOI: 10.54254/2755-2721/2026.LD37220

Table 1. Summary of perception technologies

Perception aspect	Limitation	Suggestion
Visual perception	Affected by lighting, surface texture, and vehicle vibration; depth voids are more likely to occur in low-light or low-texture environments; it is difficult to determine whether the ground is slippery, soft or hard, and its load-bearing capacity.	Perform online calibration using external references, image deblurring, and parallax confidence assessment; fuse short-range depth or laser measurements in low-confidence areas, and use foot-level feedback to correct visual judgments.
Visual perception	Sparse point clouds do not adequately represent small obstacles; rain, fog, and dust can cause abnormal echoes; and continuous mapping places a heavy computational burden on the system.	By employing ground-adaptive segmentation, echo quality filtering, and hierarchical local maps, only features relevant to foot placement safety are retained; short-term motion is constrained using an IMU.
Visual perception	The system architecture is complex, and calibration and maintenance costs are high; there are significant challenges related to time synchronization, data bandwidth, and real-time computing demands; and observational data may conflict with one another.	The system employs a multi-frequency fusion architecture, using a high-frequency IMU and joint states to maintain short-term stability, low-frequency visual and laser data to correct the map, and confidence-weighted processing to resolve conflicts.
Force and tactile perception	The sensing range is short, and the surface properties can typically only be determined after contact is made; the results are easily affected by impacts, joint friction, structural compliance, and changes in load.	A contact model is established based on joint state, IMU data, and local terrain predictions; when slippage or sinking is detected, the speed, posture, or foot placement is adjusted, and the results are fed back to the vision system.

5. Conclusions

In this paper, perception requirements of the quadruped robots in the unstructured environment are discussed and stereo vision, LiDAR, multisensor fusion, and force and tactile perception are reviewed. It further discusses problems like lighting and texture effects on vision, LiDAR point cloud degradation in bad weather, the computational requirements for multimodal fusion, the sensitivity of contact perception to dynamic errors, etc., and provides an overview of representative work.

For the three problems that faced in this essay, which is about the environment's affect to the perception, the calculate burden and the influence to perception that caused by the motion state. Based on the review analysis, the development focus of quadruped robot perception system should be transferred from increasing the number of sensors to building a credible and usable perception model of terrain conditions and promote the autonomous operation capability of quadruped robots in the long term in disaster relief, inspection and field.

DOI: 10.54254/2755-2721/2026.LD37220

References

[1] Fu, Z., Kumar, A., Agarwal, A., et al. (2022). Coupling vision and proprioception for navigation of legged robots. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 17273–17283. New Orleans, Louisiana. IEEE, USA.

[2] Ulloa, C. C., Sánchez, L., Cerro, J. D., et al. (2023). Deep learning vision system for quadruped robot gait pattern regulation. *Biomimetics, 8*(3), 289.

[3] Mao, Y. (2024). *A study on path planning for quadruped robots in complex terrain*. Master's thesis, East China Normal University. Shanghai, China.

[4] Tang, Z., Zhang, H., Liu, Z., et al. (2026). A review of intelligent perception technologies for autonomous follow-along quadruped robots. *Military Industry Automation, 45*(4), 93–101.

[5] Elnoor, M., Weerakoon, K., Sathyamoorthy, A. J., et al. (2024). AMCO: Adaptive multimodal coupling of vision and proprioception for quadruped robot navigation in outdoor environment. *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Abu Dhabi, United Arab Emirates. IEEE, USA.

[6] Wang, X., Hu, N., Zhang, P., et al. (2020). Research on environmental perception of quadruped robots based on stereo vision. *Firepower and Command and Control, 50*(7), 192–198.

[7] Jin, H. (2024). *Research on navigation and obstacle avoidance algorithms for guide dogs*. Master's thesis, Shandong Jiaotong University. Jinan, China.

[8] Xu, Y. (2024). *Research and implementation of environmental perception algorithms for quadruped robots based on multisensor fusion*. Master's thesis, Qilu University of Technology. Jinan, China.

[9] Chen, Z. (2025). *Research on real-time localization algorithms for quadruped robots based on vision-laser fusion*. Master's thesis, Beijing Jiaotong University. Beijing, China.

[10] Zhang, Y., Huang, B., Hong, M., et al. (2025). A terrain classification method for quadruped robots with proprioception. *Electronics, 14*(6), 1231.