Journal of Electronic Research and Application, 2025, Volume 9, Issue 6
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Autonomous Obstacle Avoidance Decision Mechanism of Intelligent Robot Based on Multimodal Perception

Jiaming Yan*

Sussex Artificial Intelligence Institute, Zhejiang Gongshang University, Hangzhou 310000, China

**Author to whom correspondence should be addressed.*

Abstract: Intelligent robots are increasingly being deployed across industries ranging from manufacturing to household applications and outdoor exploration. Their autonomous obstacle avoidance capabilities in complex environments have become a critical factor determining operational stability. Multimodal perception technology, which integrates visual, auditory, tactile, and LiDAR data, provides robots with comprehensive environmental awareness. By establishing efficient autonomous obstacle avoidance decision-making mechanisms based on this information, the system's adaptability to challenging scenarios can be significantly enhanced. This study investigates the integration of multimodal perception with autonomous obstacle avoidance decision-making, analyzing the acquisition and processing of perceptual information, core modules and logic of decision-making mechanisms, and proposing optimization strategies for specific scenarios. The research aims to provide theoretical references for advancing autonomous obstacle avoidance technology in intelligent robots, enabling safer and more flexible movement in diverse environments.

Keywords: Multimodal perception; Intelligent robot; Autonomous obstacle avoidance; Decision-making mechanism; Environmental cognition

Online publication: December 16, 2025

1. Introduction

With the rapid development of artificial intelligence (AI) and robotics, intelligent robots are increasingly moving from controlled, structured environments into complex and unstructured real-world settings, including household services, industrial inspection, and post-disaster rescue. Operating in such environments requires robots to perceive and respond to dynamic obstacles in real time, making autonomous obstacle avoidance a fundamental capability for ensuring task completion and preventing collisions. Traditional single-modal perception methods, such as vision-only or LiDAR-only sensing, are often hindered by environmental noise, limited detection range, and insufficient decision-making robustness. In contrast, multimodal perception integrates complementary information from multiple sensors, effectively overcoming the weaknesses of individual modalities and providing a richer, more reliable basis for decision-making. Consequently, advancing autonomous obstacle avoidance decision-making mechanisms grounded in multimodal perception is essential for improving environmental adaptability and operational safety in intelligent robots, and aligns with the growing industry demand for more intelligent, resilient, and dependable robotic systems.

2. The basic role of multimodal perception technology in obstacle avoidance of intelligent robots

2.1. Types and core roles of information acquisition in multimodal perception

Multimodal perception forms the foundation for robot environmental cognition, with its core mechanism involving diverse information acquisition through multiple sensors [1]. SLAM algorithms then transform this "dispersed information" into "spatialized environmental cognition". Visual sensors capture obstacle features to assist classification and provide semantic annotations for SLAM maps. LiDAR measures distances and positions to generate point clouds, enabling SLAM to construct high-precision geometric maps. Tactile sensors transmit pressure data to supplement the "physical attributes" of obstacles in SLAM maps. This dual-sensor synergy addresses the blind spots of single-sensor systems and SLAM's semantic limitations, allowing robots to grasp obstacle states and self-localization. It provides a "position-obstacle-attribute" tripartite framework for obstacle avoidance decision-making, forming the core of autonomous obstacle avoidance.

2.2. Sensor selection and scene adaptation for multimodal perception

Multimodal perception requires sensor selection based on specific scenarios and SLAM requirements. For indoor environments, prioritize high-definition cameras and short-range LiDAR; The former aids in constructing semantic maps for SLAM, while the latter provides high-precision distance data, making it suitable for confined spaces and reducing SLAM computational load. For outdoor industrial inspections, use anti-interference LiDAR and infrared sensors; The former maintains stable SLAM point clouds through dust penetration, while the latter supplements dynamic target positions. In post-disaster rescue operations, tactile and ultrasonic sensors are recommended; The former marks high-risk SLAM zones, while the latter identifies coordinates of hidden obstacles to mitigate risks and reduce SLAM deviations [2].

2.3. Preliminary processing of multimodal perceptual information

After acquiring multimodal perceptual information, preliminary processing is required to address SLAM input demands by removing invalid data and standardizing formats, thereby providing high-quality data for fusion and mapping. The process begins with noise reduction for sensor raw data, where Gaussian filtering is applied to visual data to eliminate blurriness, while statistical filtering is used for LiDAR to eliminate outliers, ensuring SLAM data accuracy. Following that, data calibration involves unifying timestamps and coordinate systems, followed by "hand-eye calibration" to align sensor coordinates with SLAM's body coordinates, minimizing positioning errors. Finally, key obstacle information is filtered according to robot movement and SLAM mapping requirements. For example, information about obstacles ahead and on both sides during straight-line travel is retained to reduce SLAM interference and ensure the map focuses on the path area.

3. Multimodal perceptual information fusion methods and practices

3.1. The core significance of multimodal perceptual information fusion

Multimodal perception fusion serves as the critical bridge connecting perception, SLAM, and decision-making by integrating multimodal information with SLAM data to provide precise foundations for obstacle avoidance decisions. Both single-modal systems and standalone SLAM operations have inherent limitations, whereby visual systems are susceptible to lighting-induced SLAM semantic errors, LiDAR struggles to identify obstacle transmissibility, while standalone SLAM lacks crucial obstacle attributes. The fusion of multimodal and SLAM data addresses these deficiencies, as visual systems can identify "transmissible small stones," LiDAR measures "3-meter distance," and when combined with SLAM positioning, it enables robots to accurately determine locations, assess attributes, and avoid errors caused by partial information.

3.2. Common fusion methods of multimodal perceptual information

Multimodal perception information fusion, leveraging SLAM algorithm characteristics, adopts three-tiered approaches at data, feature, and decision-making levels to adaptively address obstacle avoidance and application scenarios. The data-level fusion directly integrates sensor raw data into SLAM systems, maximizing information retention to enhance map accuracy while requiring substantial processing power. This approach is particularly suitable for scenarios with limited sensors and small data volumes to prevent computational overload. The feature-level fusion first extracts modal-specific characteristics before integrating them with SLAM map features, optimizing positioning accuracy while reducing computational load and improving efficiency. This method proves ideal for industrial inspection robots and home service robots, balancing performance with obstacle avoidance requirements. The decision-level fusion enables independent decision-making across modalities before integrating SLAM map decisions, ensuring high reliability and robustness against sensor failures or positioning deviations. This approach is particularly suitable for high-security environments, guaranteeing effective obstacle avoidance decisions.

3.3. Major challenges of multimodal perceptual information fusion

Multimodal perception information fusion requires simultaneous resolution of collaborative issues with SLAM algorithms, facing three major challenges (data heterogeneity, real-time performance, and reliability). Data heterogeneity arises from differences in sensor data formats, dimensions, and incompatibility with SLAM data. This necessitates data conversion and semantic mapping to ensure multimodal data and SLAM results are integrated within a unified framework. The real-time challenge stems from dynamic environmental changes, requiring rapid fusion and SLAM map updates. This can be achieved by optimizing algorithms, adopting high-performance hardware, reducing processing and update latency, and preventing delayed obstacle avoidance [3]. The reliability challenge manifests through sensor failures, data anomalies, or SLAM positioning drift. This demands the introduction of redundant sensors and anomaly detection mechanisms. By leveraging data complementarity and correction, the fused information becomes reliable, thereby avoiding obstacle avoidance errors.

4. Construction of autonomous obstacle avoidance decision mechanism based on multimodal perception

4.1. Core logic framework of autonomous obstacle avoidance decision mechanism

The autonomous obstacle avoidance decision-making mechanism based on multimodal perception relies on SLAM positioning and map data as spatial support, with its core being the formulation of safe and efficient obstacle avoidance strategies by integrating environmental information with SLAM results. The mechanism follows a logical sequence of:

SLAM-assisted environmental cognition → decision analysis → action execution

Multimodal perception fusion and SLAM complete obstacle localization identification and status assessment to form spatial environmental cognition. Then, combining SLAM-provided robot position, velocity, and target path information, the feasibility of obstacle avoidance plans is evaluated, such as determining detour or human-avoidance solutions based on SLAM map data. Execution commands are finally sent while real-time updates to SLAM maps through multimodal perception are made. If obstacle states change, decisions are adjusted to ensure stable obstacle avoidance in a closed-loop system.

4.2. Key functional modules of autonomous obstacle avoidance decision-making mechanism

The autonomous obstacle avoidance decision-making mechanism comprises three core modules: obstacle recognition and classification, path planning, and execution control, all requiring deep collaboration with SLAM algorithms. The recognition module utilizes SLAM coordinate data to determine precise obstacle locations, types, and statuses while tracking dynamic trajectories, providing "coordinate + attribute" data. The path planning module employs SLAM global map topology and algorithms like A* to generate obstacle-avoidance optimal routes that balance robot performance with obstacle coordinate accuracy. The execution control module converts paths into operational commands, using SLAM real-time positioning to correct deviations. For instance, detecting 0.5-meter positional offset triggers adjustments to motor speed and wheel direction, ensuring adherence to the planned trajectory.

4.3. Principles of priority setting for autonomous obstacle avoidance decision mechanism

The priority setting for autonomous obstacle avoidance decision-making mechanisms should be determined based on the spatial distribution of obstacles in the SLAM map, considering threat levels, environmental changes, and task requirements. Threat levels are assessed according to SLAM positioning distances, with larger and closer obstacles at 1-meter range being prioritized for avoidance. Environmental changes are referenced through SLAM's dynamic updates, where newly added dynamic obstacles have higher priority than existing static ones. Task requirements are aligned with the SLAM global path, where emergency rescue operations prioritize path efficiency, while routine inspections emphasize safety assurance. Priorities must be dynamically adjusted in real-time as the SLAM map updates, such as immediately elevating the priority of low-threat obstacles upon approach, to adapt to scenario-specific demands [4].

5. Application and optimization of autonomous obstacle avoidance decision mechanism based on multimodal perception

5.1. Scenario-based adaptation of autonomous obstacle avoidance decision-making mechanism

The autonomous obstacle avoidance decision-making mechanism based on multimodal perception requires adaptation to SLAM algorithm characteristics according to scene features and obstacle avoidance needs. In indoor home scenarios where static furniture dominates with occasional dynamic pedestrians, the mechanism optimizes SLAM semantic map construction by integrating multimodal perception-based pedestrian trajectory prediction data. It updates pedestrian coordinates in real-time within the SLAM map and plans short-distance, small-turn paths to avoid collisions. For outdoor industrial inspection scenarios with uneven roads, excessive dust, and variable lighting, the mechanism enhances LiDAR-infrared sensor fusion to provide interference-resistant data for SLAM, improving positioning accuracy while annotating road obstacle coordinates to optimize avoidance strategies. In post-disaster rescue scenarios where obstacles mainly consist of unstable ruins and rubble, the mechanism prioritizes tactile-ultrasound sensor coordination, annotates “high-risk zones” in the SLAM map, supplements hidden obstacle coordinates, and prioritizes avoiding these areas during path planning to ensure robot safety.

5.2. Performance optimization direction of autonomous obstacle avoidance decision mechanism

To enhance system performance, improvements should focus on three key areas, algorithm optimization, hardware upgrades, and data accumulation, to strengthen collaboration with SLAM algorithms. For instance:

(1) Algorithm optimization requires simultaneous enhancements in multimodal fusion and SLAM performance: Introducing deep learning techniques to improve multimodal information fusion accuracy, while optimizing SLAM algorithms to reduce positioning drift and enhance environmental recognition precision. When optimizing path planning algorithms, integrating SLAM map topology can shorten planning time and meet real-time obstacle avoidance requirements;

(2) Hardware upgrades should support multimodal perception and efficient SLAM operation: Selecting higher-precision sensors to improve multimodal data quality and SLAM map accuracy; adopting high-performance processors and edge computing devices to boost data processing, fusion, and SLAM computational capabilities while reducing latency [5];

(3) Data accumulation involves building a comprehensive dataset integrating “multimodal perception + SLAM + obstacle avoidance decision-making”: Collecting multimodal data from various scenarios, SLAM positioning and mapping data, and obstacle avoidance decision cases to train fusion algorithms and SLAM models, thereby improving their adaptability to different environments.

Through simulation testing with this dataset, scenarios like SLAM positioning deviations and sensor failures can be simulated to preemptively identify and address system limitations.

5.3. Future challenges and coping ideas of autonomous obstacle avoidance decision-making mechanism

The system must address core challenges in SLAM collaboration during development to establish clear directions for improvement as follows:

(1) Insufficient performance of perception and SLAM in extreme environments: In conditions like heavy rain and dense fog, sensor data distortion causes SLAM positioning errors and map construction inaccuracies. The solution involves developing specialized sensors resistant to extreme environments while optimizing multimodal fusion and SLAM coordination algorithms to mitigate environmental impacts;

(2) SLAM and decision-making coordination issues in multi-robot obstacle avoidance: When multiple robots operate in the same area, conflicts in SLAM maps and contradictory obstacle avoidance decisions may occur. The solution requires introducing multi-robot collaborative SLAM algorithms that share SLAM maps and positioning data, along with establishing a collaborative decision-making module to formulate

obstacle avoidance strategies based on unified SLAM maps;

(3) Limited adaptability to unknown obstacles: While the system and SLAM perform well with known obstacles, SLAM struggles to accurately annotate unknown obstacles, leading to decreased decision accuracy.

The solution involves applying transfer learning and online learning techniques to enable the system to rapidly learn characteristics of unknown obstacles and update SLAM semantic annotation rules, thereby enhancing adaptability to unfamiliar obstacles.

6. Conclusion

The intelligent autonomous obstacle avoidance decision-making mechanism based on multimodal perception integrates multiple sensing inputs to overcome the limitations of single-modal perception, providing reliable technical support for robots to autonomously navigate complex environments. This study examines four key aspects: the foundational role of multimodal perception technology, information fusion methods, decision-making mechanism design, and application optimization. It clarifies that multimodal perception forms the basis of decision-making systems, information fusion serves as the critical link, while rational decision modules and priority settings ensure efficient operation. The advancement of this mechanism not only enhances the environmental adaptability and operational safety of intelligent robots but also expands their applications in increasingly complex scenarios. Looking ahead, with continuous developments in sensing technologies and AI algorithms, the multimodal-based autonomous obstacle avoidance mechanism will evolve toward higher precision, faster response capabilities, and stronger adaptability. This progress will further facilitate the maturation and widespread adoption of intelligent robotics technology, providing robust support for intelligent upgrades across various industries.

Disclosure statement

The author declares no conflict of interest.

References

[1] Gu X, Li X, Liu J, 2025, Research on Autonomous Obstacle Avoidance Control for Archive Robots Based on Improved Machine Learning. Journal of Mechanical Manufacturing and Automation, 54(4): 248–253.

[2] Wang L, Zou H, Du Y, et al., 2023, Autonomous Obstacle Avoidance Technology Based on Deep Reinforcement Learning. Journal of Microcomputer Applications, 39(5): 47–50.

[3] Niu S, Han P, 2021, Design of Autonomous Collision Avoidance System for Intelligent Patrol Robot. Electronic Design Engineering, 29(4): 155–158.

[4] Ma M, Zhou H, Zhang H, et al., 2025, Design of an Autonomous Obstacle-Avoidance Path Planning Algorithm for Mobile Robots. Science and Technology Innovation and Application, 15(11): 32–36.

[5] Wang X, Gu H, 2025, Autonomous Obstacle Avoidance of Mobile Robots Based on Fuzzy Control. Computer and Digital Engineering, 53(2): 451–454.