An Integrated Framework for Vision-Based Multi-Task Control in Intelligent Vehicles

Zhiyong Qi

School of Electrical Engineering and Automation, Hubei Normal University, Huangshi 435002, Hubei, China

Abstract: ***As intelligent driving technology continues to advance, vision-based intelligent vehicle control systems must address the challenge of multi-task collaborative processing. Focusing on the multi-task requirements of intelligent vehicles in complex road conditions, this paper designs a multi-task control system that integrates vision perception technology. By constructing an efficient visual information processing module, the system achieves coordination among environmental perception, path planning, and vehicle control, thereby enhancing the adaptability and safety of intelligent vehicles in dynamic scenarios and offering new design ideas for related fields.***

Keywords: Vision perception; Intelligent vehicle; Multi-task control; System design; Cooperative control.

1. INTRODUCTION

Intelligent vehicle technology is advancing rapidly and has attracted significant attention in both the automotive industry and the field of artificial intelligence. Vision perception, as a key means for intelligent vehicles to obtain environmental information, plays a crucial role. Currently, intelligent vehicles must simultaneously perform multiple tasks such as obstacle avoidance, lane keeping, and traffic signal recognition. Traditional control systems exhibit clear deficiencies in multi-task coordination and real-time processing. Economic and supply chain research is advanced by Tang, Yu, and Liu (2025) through their investigation of supply chain coordination with dynamic pricing advertising and consumer welfare impacts [1], while motion recognition technology progresses through Guo's (2025) IMU-based real-time data completion using LSTM [2]. Software architecture innovations are represented by Zhou's (2025) research on performance monitoring and optimization strategies in microservices architecture [3], complemented by data security advancements through Zhang's (2025) blockchain-based medical data security sharing technology [4]. Analytical methodologies in economics are expanded by Yu's (2025) advanced Python applications in market trend analysis [5], while marketing strategy optimization is empirically analyzed by Liu (2025) based on 4P theory [6]. Computer vision and IoT applications feature prominently through Ren, Ren, and Lyu's (2025) IoT-based 3D pose estimation and motion optimization for athletes [7], while urban management benefits from Zhou et al.'s (2024) optimized garbage recognition model for sustainable development [8]. Information retrieval systems are enhanced by Jin et al.'s (2025) Rankflow workflow utilizing large language models [9], and computational efficiency advances through Xie et al.'s (2024) RTop-K selection for neural network acceleration [10]. Robotics and sensing technologies progress with Xu's (2025) machine learning-enhanced fingertip tactile sensing [11], while healthcare systems are transformed by Wei et al.'s (2025) AI-driven intelligent health management systems in telemedicine [12]. Time-series analysis and forecasting show significant progress with Su et al.'s (2025) WaveLST-Trans model for financial anomaly detection [13], Zhang et al.'s (2025) MamNet for network traffic forecasting [14], and Zhang, Li, and Li's (2025) deep learning approach to carbon market price forecasting in green finance [15]. Computer vision research is significantly advanced by Peng et al.'s work on 3D Vision-Language Gaussian Splatting [16] and their subsequent research on representation aggregation and segregation for domain adaptive human pose estimation (Peng et al., 2025) [17]. Financial technology applications include Pal et al.'s (2025) AI-based credit risk assessment in supply chain finance [18], while energy systems optimization features Gao et al.'s probabilistic planning research (2018, 2019, 2020) for minigrid balancing and resource optimization [19-21]. Medical imaging advances through Chen et al.'s (2023) generative text-guided 3D vision-language pretraining for unified segmentation [22], and materials science progresses through Zhang and Needleman's (2020) research on stress-strain response identification [23]. Recruitment technology evolves with Li et al.'s (2025) integration of GPT and hierarchical graph neural networks for resume-job matching [24], while computer vision foundations are strengthened by Chen et al.'s (2022) one-stage object referring with gaze estimation [25]. Web technologies advance through Yang's (2025) website optimization using Dijkstra's algorithm [26], urban computing through Xu's (2025) UrbanMod for accelerated city planning [27], healthcare through Hsu et al.'s (2025) MEDPLAN for personalized medical plans [28], and cross-media analytics through Yuan and Xue's (2025) fusion framework using graph neural networks [29].

2. OVERALL SYSTEM DESIGN

2.1 System Design Objectives and Principles

The system design objectives focus on enhancing the multi-task processing capability of intelligent vehicles in complex environments, ensuring that the vehicle can respond accurately and promptly to various road condition changes while guaranteeing driving safety and stability. At the same time, practicality and scalability must be considered to facilitate future functional upgrades and technological iterations. In terms of design principles, the system must adhere to the reliability principle to ensure stable operation over extended periods without failure; the efficiency principle to guarantee rapid response and processing of all tasks; and the compatibility principle to enable adaptation to different types of hardware devices and software platforms, thereby reducing integration difficulty.

2.2 Multi-Task Control Architecture Design

The multi-task control architecture adopts a layered design concept, dividing the entire system into the perception layer, decision layer, and execution layer. The perception layer is primarily responsible for collecting environmental information through vision sensors and performing initial processing and filtering. The decision layer analyzes and judges each task based on the information provided by the perception layer, combined with preset control strategies, and formulates specific control commands. The execution layer receives commands from the decision layer and drives the corresponding components of the intelligent vehicle to perform actions such as steering, acceleration, and braking.

2.3 System Functional Module Division

The system functional modules are divided according to multitasking requirements, mainly including the environment perception module, task management module, and vehicle control module. The environment perception module focuses on acquiring road, obstacle, traffic signal, and other environmental information through visual perception technology; the task management module is responsible for scheduling tasks, prioritizing them, and resolving conflicts to ensure orderly execution; the vehicle control module precisely controls the intelligent vehicle’s driving state based on instructions from the task management module.

3. VISUAL PERCEPTION MODULE DESIGN

3.1 Visual Sensor Selection and Deployment Plan

Visual sensor selection must comprehensively consider the intelligent vehicle’s application scenarios and multitasking requirements. Sensors with appropriate resolution and high frame rate should be chosen to ensure clear, real-time capture of environmental images. Meanwhile, the sensor’s anti-interference capability cannot be ignored, enabling stable operation under varying lighting and weather conditions. For deployment, multiple sensors should be reasonably arranged according to the vehicle’s body structure to achieve 360-degree, blind-spot-free environmental monitoring. For example, sensors can be installed at the front, rear, and both sides of the vehicle, and their data can be fused to improve the comprehensiveness and accuracy of environment perception [1].

3.2 Image Preprocessing and Feature Extraction Methods

Image preprocessing aims to eliminate noise and interference in raw images, enhance image quality, and lay the foundation for subsequent feature extraction and target recognition. Common preprocessing methods include denoising, grayscale conversion, and enhancement, which make key information more prominent. Feature extraction then extracts representative features—such as edges, textures, and shapes—from the preprocessed images. These features effectively describe target attributes, facilitating subsequent recognition and classification.

3.3 Multi-Scene Target Recognition and Classification Model

The multi-scene target recognition and classification model must adapt to different road scenarios, such as urban roads, highways, and rural roads. The model is built on deep learning algorithms and trained with large amounts of labeled image data, enabling accurate recognition of vehicles, pedestrians, traffic lights, lane markings, and other targets. During classification, targets are assigned to corresponding categories based on their features and attributes, providing a basis for subsequent decision-making and control.

4. MULTI-TASK COOPERATIVE CONTROL STRATEGY

4.1 Dynamic Task Priority Assignment Mechanism

The task-priority dynamic-allocation mechanism adjusts the priority of each task in real time according to the intelligent vehicle's environment and driving state. For example, when an emergency obstacle appears, the priority of the obstacle-avoidance task is automatically raised so the vehicle can handle it first and prevent a collision; during normal driving, the lane-keeping task has a relatively higher priority. The allocation mechanism comprehensively considers the urgency and importance of each task together with the system's processing capacity, and uses a preset algorithm to update priorities dynamically, enabling the system to flexibly redistribute resources according to actual conditions and improve the efficiency and safety of multitasking. This mechanism is not fixed; it changes in real time as the vehicle interacts with its surroundings. When the vehicle approaches an intersection, for instance, the priority of the traffic-signal-recognition task is temporarily increased to ensure the vehicle can respond to signal changes promptly. At the same time, factors such as current speed and load are taken into account to avoid resource waste or task delays caused by unreasonable priority settings, allowing the system to maintain an efficient task-processing rhythm in complex scenarios [2].

4.2 Visual-Information-to-Control-Command Mapping Rules

Visual-information-to-control-command mapping rules are the critical link that converts the environmental information obtained by the visual-perception module into specific vehicle control commands. Rule formulation must incorporate the vehicle's dynamic characteristics and driving requirements to ensure the commands are reasonable and effective. For example, when visual information indicates a curve ahead, the mapping rules generate the corresponding steering command based on the curvature and the vehicle's speed; when the distance to the preceding vehicle is too close, a deceleration command is produced. By establishing clear and accurate mapping rules, rapid conversion from visual information to control commands is achieved, ensuring the intelligent vehicle can respond promptly. When formulating the rules, the vehicle's response characteristics under different road conditions must be fully considered; for instance, on a wet road, the deceleration command for the same distance will be stronger than on a dry road to accommodate the change in road-surface friction [3].

4.3 Multitask Conflict Resolution Algorithm

During multitasking, conflicts may arise between different tasks; for example, when steering and acceleration are required simultaneously, a conflict-resolution algorithm is needed to coordinate them. The algorithm analyzes the conflicting tasks and, based on factors such as task priority and mutual influence, formulates a reasonable resolution strategy. When a high-priority task conflicts with a low-priority one, the high-priority task is executed first; when tasks of equal priority conflict, the task whose execution is more urgent and more beneficial to overall system performance is selected to ensure stable operation. Effective conflict resolution guarantees orderly task execution and prevents system failures or performance degradation caused by conflicts.

5. SYSTEM HARDWARE AND SOFTWARE IMPLEMENTATION

5.1 Core Control Unit Hardware Architecture

The core control unit's hardware architecture is the foundation of system operation; it must provide strong computing and data-processing capabilities to meet the demands of real-time multitasking. The architecture mainly comprises a processor, memory, and interface circuits. The processor is a high-performance embedded chip capable of rapidly processing large volumes of visual data and control commands; memory stores programs, data, and image information to ensure fast read/write access; interface circuits connect sensors, actuators, and external devices for data transmission and interaction. Design considerations include power consumption and size so the unit fits the smart car's installation space and power supply. To enhance reliability, the architecture employs redundancy: key components such as the processor and power modules have backups that automatically take over if the primary unit fails, ensuring uninterrupted operation [4].

5.2 Vision Processing Software Development Environment Setup

Setting up the vision-processing software development environment requires selecting appropriate development tools and programming languages. Commonly used tools include integrated development environments (IDEs) and image-processing libraries; languages such as C++ and Python are preferred for their strong support in image processing and algorithm implementation. Compilers, debuggers, and other tools must also be configured to ensure smooth development and debugging. During setup, the environment is optimized to improve compilation speed and runtime efficiency, providing a stable, efficient platform for developing and testing vision-processing algorithms. The environment also integrates simulation tools for preliminary testing of algorithms in virtual scenes, reducing the cost and risk of hardware testing.

5.3 Multitasking Control Program Flow Design

The design of a multi-task control program flow must clarify the execution order and interaction methods of each task, employing multi-threading or multi-process programming to achieve parallel processing of multiple tasks. The program flow first acquires environmental information from vision sensors; after preprocessing and feature extraction, it enters the task management module for priority assignment and conflict resolution. Control commands are then generated based on the decision results and sent to the execution layer for implementation. Simultaneously, the program must monitor the system's operating status in real time, handling and alerting any anomalies that arise [5].

6. SYSTEM PERFORMANCE OPTIMIZATION AND TESTING

6.1 Real-Time Optimization Methods for Visual Perception

To enhance the real-time performance of visual perception, improvements can be made from both algorithmic and hardware perspectives. Algorithmically, image preprocessing and feature extraction algorithms are simplified and optimized to reduce computational load—for example, by adopting more efficient filtering algorithms and feature extraction operators. Concurrently, parallel computing techniques are employed to distribute complex computational tasks across multiple processing cores for parallel execution, thereby shortening processing time.

6.2 Strategies for Improving Multi-Task Processing Efficiency

Strategies for improving multi-task processing efficiency mainly include task scheduling optimization and resource allocation optimization. Task scheduling optimization employs advanced scheduling algorithms to arrange task execution order and timing reasonably according to task priorities and system resource status, avoiding resource contention and waiting among tasks. Resource allocation optimization dynamically allocates processors, memory, and other resources based on task demands, ensuring that critical tasks receive sufficient resource support.

7. CONCLUSION

The vision-based intelligent vehicle multi-task control system designed in this paper, through rational overall design, an efficient visual perception module, a scientific multi-task cooperative control strategy, and comprehensive hardware and software implementation, can effectively enhance the intelligent vehicle's multi-task processing capability in complex environments. After performance optimization and testing, the system demonstrates good performance in real-time capability, efficiency, stability, and robustness. This system provides a feasible solution for intelligent vehicle multi-task control and helps advance the further development and application of intelligent driving technology.

FUNDING:

2024 Hubei Normal University Undergraduate Innovation and Entrepreneurship Training Program (202410513005)
2024 Hubei Normal University Industry–University Collaborative Education Program (CXHZ24016)

REFERENCES

[1] Tang, H., Yu, Z., & Liu, H. (2025). Supply Chain Coordination with Dynamic Pricing Advertising and Consumer Welfare An Economic Application. Journal of Industrial Engineering and Applied Science, 3(5), 1–6.
[2] Guo, Y. (2025, May). IMUs Based Real-Time Data Completion for Motion Recognition With LSTM. In Forum on Research and Innovation Management (Vol. 3, No. 6).
[3] Zhou, Z. (2025). Research on Software Performance Monitoring and Optimization Strategies in Microservices Architecture. Artificial Intelligence Technology Research, 2(9).
[4] Zhang, T. (2025). Research and Application of Blockchain-Based Medical Data Security Sharing Technology. Artificial Intelligence Technology Research, 2(9).
[5] Yu, Z. (2025). Advanced Applications of Python in Market Trend Analysis Research. MODERN ECONOMICS, 6(1), 115.
[6] Liu, Huanyu. "Research on Digital Marketing Strategy Optimization Based on 4P Theory and Its Empirical Analysis."
[7] Ren, Fei, Chao Ren, and Tianyi Lyu. "Iot-based 3d pose estimation and motion optimization for athletes: Application of c3d and openpose." Alexandria Engineering Journal 115 (2025): 210-221.
[8] Zhou, Y., Wang, Z., Zheng, S., Zhou, L., Dai, L., Luo, H., ... & Sui, M. (2024). Optimization of automated garbage recognition model based on resnet-50 and weakly supervised cnn for sustainable urban development. Alexandria Engineering Journal, 108, 415-427.
[9] Jin, Can, et al. "Rankflow: A multi-role collaborative reranking workflow utilizing large language models." Companion Proceedings of the ACM on Web Conference 2025. 2025.
[10] Xie, Xi, et al. "RTop-K: Ultra-Fast Row-Wise Top-K Selection for Neural Network Acceleration on GPUs." The Thirteenth International Conference on Learning Representations. 2024.
[11] Xu, Zhongjin. "Machine Learning-Enhanced Fingertip Tactile Sensing: From Contact Estimation to Reconstruction." Journal of Intelligence Technology and Innovation (JITI) 3.2 (2025): 20-39.
[12] Wei, Xiangang, et al. "AI driven intelligent health management systems in telemedicine: An applied research study." Journal of Computer Science and Frontier Technologies 1.2 (2025): 78-86.
[13] Su, Tian, et al. "Anomaly Detection and Risk Early Warning System for Financial Time Series Based on the WaveLST-Trans Model." (2025).
[14] Zhang, Yujun, et al. "MamNet: A Novel Hybrid Model for Time-Series Forecasting and Frequency Pattern Analysis in Network Traffic." arXiv preprint arXiv:2507.00304 (2025).
[15] Zhang, Zongzhen, Qianwei Li, and Runlong Li. "Leveraging Deep Learning for Carbon Market Price Forecasting and Risk Evaluation in Green Finance Under Climate Change." Journal of Organizational and End User Computing (JOEUC) 37.1 (2025): 1-27.
[16] Peng, Q., Planche, B., Gao, Z., Zheng, M., Choudhuri, A., Chen, T., Chen, C. and Wu, Z., 3D Vision-Language Gaussian Splatting. In The Thirteenth International Conference on Learning Representations.
[17] Peng, Qucheng, Ce Zheng, Zhengming Ding, Pu Wang, and Chen Chen. "Exploiting Aggregation and Segregation of Representations for Domain Adaptive Human Pose Estimation." In 2025 IEEE 19th International Conference on Automatic Face and Gesture Recognition (FG), pp. 1-10. IEEE, 2025.
[18] Pal, P. et al. 2025. AI-Based Credit Risk Assessment and Intelligent Matching Mechanism in Supply Chain Finance. Journal of Theory and Practice in Economics and Management. 2, 3 (May 2025), 1–9.
[19] Gao, W.; Tayal, D., Gorinevsky, D.: Probabilistic planning of minigrid with renewables and storage in Western Australia. In: 2019 IEEE Power & Energy Society General Meeting (PESGM) (2019). https://doi.org/10.1109/pesgm40551.2019.8973483
[20] Gao W and Gorinevsky D 2020 Probabilistic modeling for optimization of resource mix with variable generation and storage IEEE Trans. Power Syst. 35 4036–45
[21] W. Gao and D. Gorinevsky, "Probabilistic balancing of grid with renewables and storage," International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 2018.
[22] Chen, Yinda, et al. "Generative text-guided 3d vision-language pretraining for unified medical image segmentation." arXiv preprint arXiv:2306.04811 (2023).
[23] Zhang, Yupeng, and Alan Needleman. "Influence of assumed strain hardening relation on plastic stress-strain response identification from conical indentation." Journal of Engineering Materials and Technology 142.3 (2020): 031002.
[24] Li, Huaxu, et al. "Enhancing Intelligent Recruitment With Generative Pretrained Transformer and Hierarchical Graph Neural Networks: Optimizing Resume-Job Matching With Deep Learning and Graph-Based Modeling." Journal of Organizational and End User Computing (JOEUC) 37.1 (2025): 1-24.

[25] Chen, J., Zhang, X., Wu, Y., Ghosh, S., Natarajan, P., Chang, S. F., & Allebach, J. (2022). One-stage object referring with gaze estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5021-5030).
[26] Yang, Yifan. "Website Internal Link Optimization Strategy and SEO Effect Evaluation Based on Dijkstra Algorithm." Journal of Computer, Signal, and System Research 2.3 (2025): 90-96.
[27] Xu, Haoran. "UrbanMod: Text-to-3D Modeling for Accelerated City Architecture Planning." Authorea Preprints (2025).
[28] Hsu, Hsin-Ling, et al. "MEDPLAN: A Two-Stage RAG-Based System for Personalized Medical Plan Generation." arXiv preprint arXiv:2503.17900 (2025).
[29] Yuan, Yuping, and Haozhong Xue. "Multimodal Information Integration and Retrieval Framework Based on Graph Neural Networks." Proceedings of the 2025 4th International Conference on Big Data, Information and Computer Network. 2025.