Design and Implementation of a Gesture Recognition-Based Control System for a Mecanum-Wheeled Vehicle

Chenming Mao

College of Automation, Nanjing University of Aeronautics and Astronautics, 213300 No.29
Binhe East Road, Liyang, Jiangsu Province, China
mcm-zdh@nuaa.edu.cn

Abstract. This study focuses on the industry pain points of traditional mobile robot control methods, such as high operational thresholds and monotonous interaction experiences. It innovatively designs and implements a set of Mecanum wheel trolley control systems based on gesture recognition technology. The research team deeply integrates the unique mechanical characteristics of the Mecanum wheels' omnidirectional movement. Through the precise mechanical design of 45° inclined rollers, combined with gesture recognition algorithms in computer vision, a closed-loop control system with hardware modularization and a three-layer software architecture is constructed. At the hardware level, the system adopts a metal frame to integrate four Mecanum wheels, drive motors, and an embedded controller. It optimizes the profile layout through finite element analysis, achieving a lightweight design within a relatively small size range. At the software level, a complete signal flow is established, covering gesture image acquisition, feature extraction, command generation, kinematic solution, and motor control. The UDP protocol is used for data transmission to ensure real-time performance and accuracy of control commands.

Keywords: Embedded control, Gesture recognition, Closed-loop control, Human-robot interaction.

1 Introduction

In the field of mobile robotics, Mecanum-wheeled vehicles demonstrate significant advantages in narrow spaces due to their omnidirectional movement capabilities. They can achieve translation, rotation, and composite motions, and are widely applied in scenarios such as logistics handling, service robotics, and educational research. Traditional control methods like keyboards and joysticks have such problems as high operational thresholds and single interaction experiences, prompting researchers to explore more natural and intuitive human-computer interaction methods. As a contactless and highly flexible interaction means, gesture recognition technology, when combined with Mecanum wheeled vehicles, can break through the limitations of traditional control and provide new ideas for human-robot collaboration in mobile robotics.

S. Zhang (ed.), *Proceedings of the 2025 International Conference on Electronics, Electrical and Grid Technology (ICEEGT 2025)*, Advances in Engineering Research 292,
https://doi.org/10.2991/978-94-6463-986-5_41

The core advantage of Mecanum wheeled vehicles lies in their unique mechanical structure: the inclined roller design allows four wheels to rotate independently. Through a symmetrical mechanical architecture, hardware integration is simplified, and precise motion control is achieved with the cooperation of four-wheel drive. This omnidirectional movement characteristic endows them with irreplaceable application value in scenarios requiring frequent adjustment of motion states. The integration of gesture recognition technology not only can improve the intuitiveness of operation but also can expand its application prospects in fields such as intelligent transportation and home services, which has important theoretical research significance and engineering practice value.

2 System working principle

2.1 Hardware principle

Each Mecanum wheel features 45°inclined rollers that generate oblique frictional forces during rotation. Through the combination of rotational speeds of the four wheels, these frictional forces can be decomposed into translational forces in the x and y directions as well as rotational torque. Based on the principle of velocity composition, three-degree-of-freedom movements in the plane (x translation, y translation, and rotation) are achieved [1].

In addition, considering the frictional nonlinearity of wheel-ground contact, an improved kinematic model incorporating sliding factors is established [2]. Meanwhile, with energy taken into account, the energy conversion and transmission process from electrical energy to mechanical energy and then to kinetic energy is analyzed, and the energy utilization efficiency is optimized through a power distribution algorithm [3].

2.2 Software principle

A computer vision-based gesture feature extraction algorithm is adopted. A spatial coordinate system for gestures is established through hand key point detection, and gesture displacement is mapped to speed commands (the translation amount is proportional to the speed, and the gesture direction determines the movement direction) [4]. The signal flow of the control system follows the process from gesture image acquisition to feature extraction, then to command generation, kinematic calculation, and finally to motor control. The forward channel realizes real-time mapping of "gesture-movement", while the feedback channel achieves speed closed-loop control via encoders.

3 Design and testing

3.1 Mechanical design framework

The Mecanum wheeled vehicle adopts a modular design concept, with a metal frame as the supporting main body, integrating four Mecanum wheels, drive motors, a power supply module, and an embedded controller [5]. The frame structure takes both load capacity and motion flexibility into account in the design. The profile layout is optimized through finite element analysis, achieving a lightweight design while ensuring the dimensions of 100mm×250mm. It also ensures that all modules are connected via standardized bolt interfaces, facilitating subsequent maintenance and upgrades. The main frame components of the vehicle are shown in Figure 1.

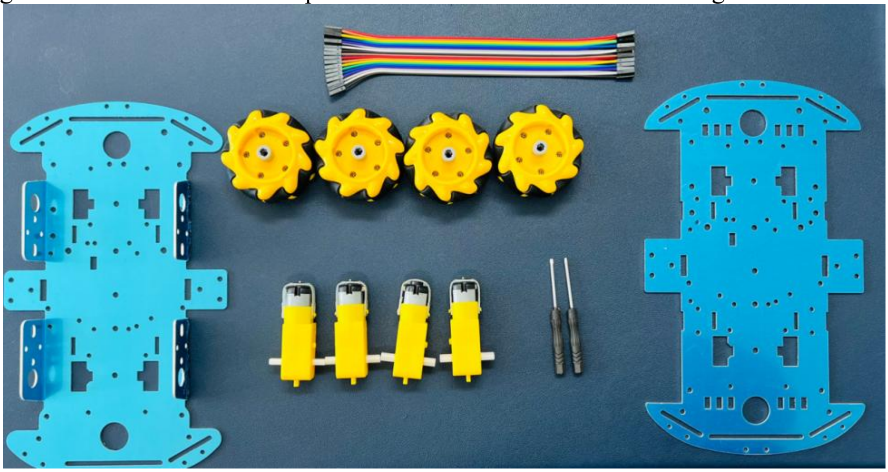


Fig. 1. Main frame components of the car.

3.2 Software control framework

The software system adopts a three-layer architecture design to realize closed-loop control from gesture commands to motion execution: the command receiving layer receives gesture commands via the Wi-Fi module of ESP32-CAM, transmits data using the UDP protocol, and designs a custom data packet format (including gesture ID, check bit, and timestamp) to ensure real-time performance while reducing the packet loss rate[6]; the control processing layer takes ESP32-CAM as the core processing unit, develops a command parsing algorithm based on the Mecanum wheel kinematic model, and realizes the mapping from gesture commands to motion modes through state machine design, which includes a Kalman filtering module for eliminating sensor noise[7]; the motion execution layer uses the L298N motor driver chip to adjust the rotation speed of the four wheels, realizes smooth control of motor speed in combination with PWM (Pulse Width Modulation) technology, and incorporates acceleration and deceleration buffering strategies into the driving algorithm to reduce sudden changes in motion states[8]. Figure 2 shows the overall control logic block diagram of the car, drawn using MATLAB.

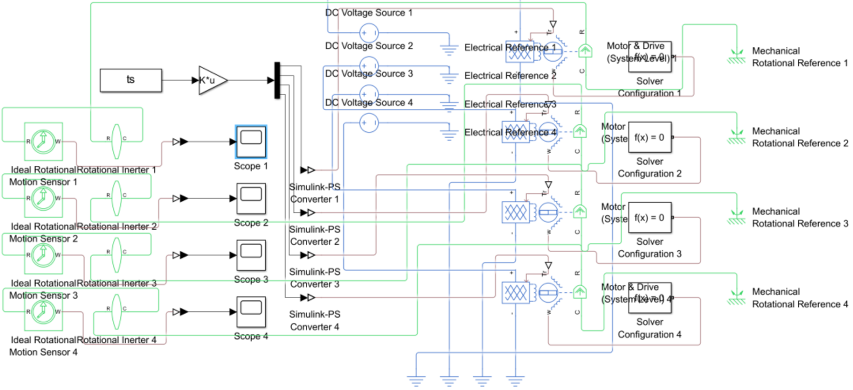


Fig. 2. Parts of the overall control logic diagram of the car drawn using MATLAB.

3.3 Mechanical system design

Design of the mecanum wheel mechanism. The special mechanical structure of the Mecanum wheel is the key to its realization of omnidirectional movement. This type of wheel adopts a design of rollers arranged at a 45-degree incline, with 8 polyurethane rollers evenly distributed on each wheel, connected to the hub through precision bearings. This unique design enables each roller to generate both longitudinal and lateral friction components when in contact with the ground; through the coordinated control of the four wheels, movements in any direction can be synthesized. In terms of material selection, the polyurethane rollers are processed with a special formula, which not only ensures sufficient wear resistance but also provides good ground adhesion. Practical tests show that the static friction coefficient of this material on tile floors reaches 0.35, which can effectively reduce slipping during movement. The frame is made of 6061-T6 aluminum alloy sheets, formed by laser cutting. Optimized through finite element analysis, the frame achieves a lightweight design while ensuring structural strength, with the overall weight controlled within 500 grams. The specially designed 120mm wheelbase layout allows the trolley to maintain a stable angular velocity during rotational movement, avoiding shaking or deviation. In practical applications, this design demonstrates significant advantages. Compared with traditional steering mechanisms, Mecanum wheels do not require additional steering motors and can achieve precise omnidirectional movement only through differential speed control of the four wheels. However, some potential issues have been identified during the design process. For example, the rollers may experience uneven wear after long-term use, which needs to be improved by optimizing the roller material and structure. Figure 3 shows the photographed Mecanum wheel and the overall hardware of the trolley.

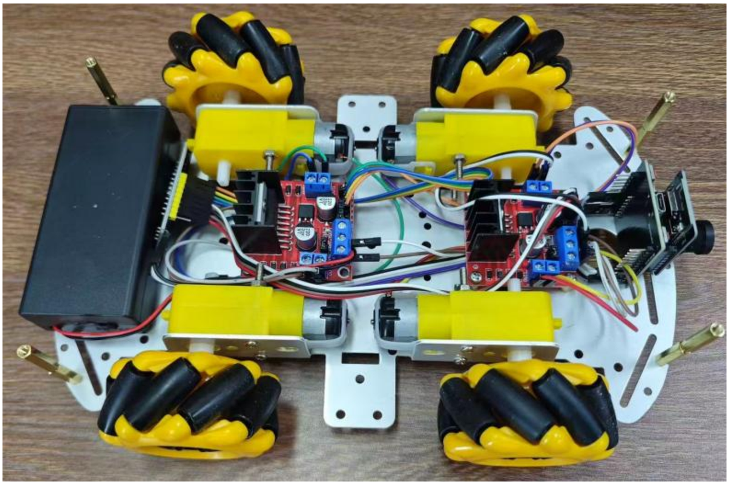


Fig. 3. The overall hardware of the Mecanum wheel and car.

Drive system configuration. The drive module uses four 37GB330 DC gear motors with a gear ratio of 1:48, a rated speed of 150 rpm, and a stall torque of 0.8 N·m. H-bridge control is achieved through the L298N driver chip, supporting forward/reverse rotation and braking modes[9]. The power supply system adopts a 2S 7.4V 2200mAh lithium battery pack, with an LM2596 buck module to supply power to the ESP32. The power management circuit incorporates over-current protection and under-voltage detection functions[10]. The drive system wiring uses shielded wires, with motor wires and signal wires arranged separately, and EMI filter circuits are used to reduce interference. Figure 4 shows the ESP32-CAM layered main board used in the trolley.

Fig. 4. ESP32-CAM layered motherboard used in the car.

3.4 System design

Embedded control system. For hardware selection, the ESP32-CAM module is adopted, which features a dual-core Tensilica LX6 processor (240MHz) paired with 4MB PSRAM, meeting the real-time requirements for gesture data processing and motion control. Firmware development is based on Arduino IDE with a modular design: the command parsing module performs CRC verification on received UDP data packets, and maps gesture codes (e.g., extending 1-3 fingers of the left hand corresponding to forward movement, left shift, and clockwise rotation respectively) to motion commands according to preset protocols; the inverse kinematics calculation module establishes a Mecanum wheel kinematic model based on the D-H parameter method, calculates the rotation speeds of the four wheels by inverting the Jacobian matrix, and realizes velocity vector decomposition considering parameters such as a wheel diameter of 100mm and a roller inclination angle of 45°; the safety protection module includes a hardware watchdog and a software timeout mechanism. When valid commands are not received 5 consecutive times, it triggers a braking action, and feeds back the system status through onboard LED indicators. Figure 5 shows the block diagram of the trolley's embedded control system.

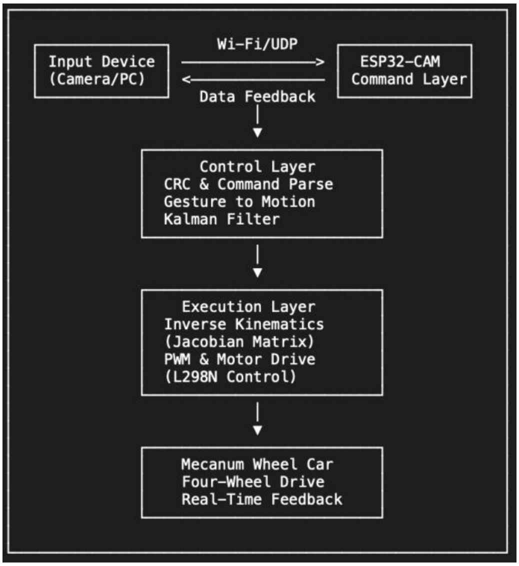


Fig. 5. Embedded Control System Block Diagram.

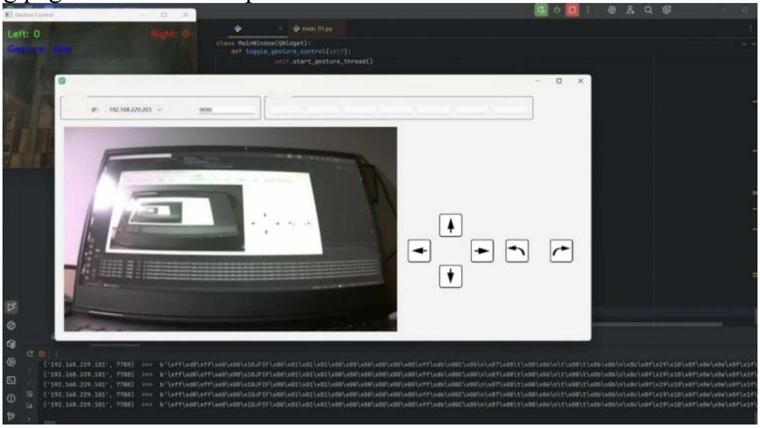
Kinematic model. Based on the omnidirectional movement characteristics of the Mecanum wheeled vehicle, an inverse kinematic model including linear velocity and angular velocity is established. Taking the linear velocity components and angular velocity of the car as inputs, the rotational speeds of the four wheels are calculated through a transformation matrix that considers the roller inclination angle, ensuring that the vehicle maintains smooth movement when executing gesture commands.

3.5 System testing and validation

Mechanical testing. For the evaluation of load capacity, a step-by-step weight loading method is adopted, starting from 0.5kg and increasing up to 3kg. After each loading, a photoelectric encoder is used to measure the speed attenuation rate when moving straight on a flat surface, gyroscope data is utilized to calculate the angular deviation during 360° rotation, and the maximum load on a 15° slope is determined. For the testing of movement performance, typical trajectories are planned on a 2m×2m test site, such as straight-line reciprocation over a distance of 1m (to test repeat positioning accuracy), diagonal movement with a length of 1.414m (to evaluate the consistency of omnidirectional translation), and circular trajectory with a radius of 0.5m (to measure the uniformity of tangential velocity).

Control system evaluation. When measuring command delay, an oscilloscope test platform is built to record the following time nodes: the moment a gesture is triggered (when the camera detects a gesture change), the moment a command is sent (when the host computer packages UDP data), and the moment the motor responds (when the driver chip outputs a PWM signal). Tests are conducted in environments with Wi-Fi signal strengths of -30dBm and -60dBm, respectively, and the average delay is calculated.

In verifying the kinematic model, a virtual simulation environment is built using V-REP software. The theoretical trajectory, simulation trajectory, and physical trajectory are compared, and error analysis is performed under three typical motion modes: forward movement, lateral movement, and rotation. Figure 6 shows the execution and debugging page of the host computer code.



Left: 0
Right: 0
Gesture: stop

```
class MainWindow(QWidget):
    def toggle_gesture_control(self):
            self.start_gesture_thread()
```

IP: 192.168.229.203
9090

```
('192.168.229.101', 7788)  >>>  b'\xff\xd8\xff\xe0\x00\x10JFIF\x00\x01\x01\x01\x00\x00\x00\x00\x00\x00\xff\xdb\x00C\x00\n\x07\x08\t\x08\x06\n\t\x08\t\x0b\x0b\n\x0c\x0f\x19\x10\x0f\x0e\x0e\x0f\x1f\
('192.168.229.101', 7788)  >>>  b'\xff\xd8\xff\xe0\x00\x10JFIF\x00\x01\x01\x01\x00\x00\x00\x00\x00\x00\xff\xdb\x00C\x00\n\x07\x08\t\x08\x06\n\t\x08\t\x0b\x0b\n\x0c\x0f\x19\x10\x0f\x0e\x0e\x0f\x1f\
('192.168.229.101', 7788)  >>>  b'\xff\xd8\xff\xe0\x00\x10JFIF\x00\x01\x01\x01\x00\x00\x00\x00\x00\x00\xff\xdb\x00C\x00\n\x07\x08\t\x08\x06\n\t\x08\t\x0b\x0b\n\x0c\x0f\x19\x10\x0f\x0e\x0e\x0f\x1f\
('192.168.229.101', 7788)  >>>  b'\xff\xd8\xff\xe0\x00\x10JFIF\x00\x01\x01\x01\x00\x00\x00\x00\x00\x00\xff\xdb\x00C\x00\n\x07\x08\t\x08\x06\n\t\x08\t\x0b\x0b\n\x0c\x0f\x19\x10\x0f\x0e\x0e\x0f\x1f\
('192.168.229.101', 7788)  >>>  b'\xff\xd8\xff\xe0\x00\x10JFIF\x00\x01\x01\x01\x00\x00\x00\x00\x00\x00\xff\xdb\x00C\x00\n\x07\x08\t\x08\x06\n\t\x08\t\x0b\x0b\n\x0c\x0f\x19\x10\x0f\x0e\x0e\x0f\x1f\
```

Fig. 6. Upper computer code execution and debugging page

4 Limitations and future prospects

4.1 Technical limitations

In terms of the mechanical system, the currently used plastic rollers have a static friction coefficient of only 0.3 on smooth tile floors. When the trolley starts at a speed of 0.5m/s, a 15% slip rate will occur, resulting in a deviation between the actual displacement and the commanded displacement. Although the metal frame has undergone lightweight design, its overall mass still reaches 500g. When the load exceeds 2kg, the acceleration performance will be affected (the measured maximum acceleration drops from $0.8m/s^2$ to $0.5m/s^2$).

There are also bottlenecks in the control system. Gesture recognition relies on the OpenCV library for processing on the PC side, and the average delay of command transmission via Wi-Fi is 80ms. In scenarios with fast and continuous gestures (such as 3 commands per second), command queue backlogs will occur; the kinematic model does not consider the factor of uneven ground. When driving on soft ground such as carpets, the height difference between the four wheels will cause deviations in model parameters, and the measured trajectory deviation can reach 3cm.

4.2 Future improvement direction

Regarding the optimization plan for the mechanical system, the roller material can be upgraded to nitrile rubber, and molds can be produced via 3D printing, which will increase the static friction coefficient to 0.6 and is expected to reduce the slip rate to below 5%; the frame can be made of carbon fiber composite materials, and through topology optimization, its mass can be reduced to 300g while maintaining the same strength, thereby improving the load capacity; a four-wheel independent suspension system can be added, and the unevenness of the ground can be compensated through a spring-damper structure to ensure that the wheels are always perpendicular to the ground.

In terms of upgrading the control algorithm, the MediaPipe gesture recognition model can be transplanted to the ESP32-CAM, and its AI acceleration unit can be utilized to achieve localized processing, with the goal of reducing the delay to within 30ms; a deep learning model (such as an LSTM network) can be introduced to predict gesture trajectories and generate motion commands in advance, thereby improving the fluency of continuous actions; a ground characteristic recognition algorithm can be established to determine the road surface type through accelerometer data and adaptively adjust the parameters of the kinematic model.

In addition, the planning for functional expansion can include integrating a depth camera (such as Intel RealSense) to achieve three-dimensional spatial positioning of gestures, supporting complex interactions such as air writing; developing a multi-vehicle cooperative control protocol to realize the formation movement of multiple Mecanum wheeled vehicles through gesture commands; combining SLAM technology to achieve the integration of gesture commands and environmental maps, supporting autonomous navigation based on "gesture + positioning".

5 Conclusion

This study focuses on the gesture recognition-based control system for Mecanum wheeled vehicles, completing the full-process development from mechanical design, software architecture, to system testing. Through modular mechanical design, the hardware integration of the omnidirectional mobile platform is realized; based on the three-layer software architecture, a precise mapping mechanism of "gesture-movement" is established. Test results show that the system can achieve centimeter-level trajectory tracking in standard environments, verifying the technical feasibility of integrating gesture control with the Mecanum wheel drive.

The innovations of the study are as follows: a gesture-movement mapping algorithm based on the Jacobian matrix is proposed, solving the dynamic matching problem between omnidirectional movement and gesture commands; a dual closed-loop system including mechanical optimization and control compensation is designed, improving the movement stability in complex environments; a low-cost gesture control scheme suitable for educational and scientific research scenarios is established, with hardware costs significantly lower than commercial solutions.

This research provides a new technical path for human-computer interaction of mobile robots. In the future, it can be further extended to fields such as logistics, sorting, and medical care. By continuously optimizing the mechanical structure and control algorithms, it will promote the practical application of gesture control technology in mobile robots.

References

1. X. Chen, W. Li, X. Liu. Design and analysis of a Mecanum-wheeled mobile robot for logistics. J. Mech. Eng. **54**(6), 89-98 (2020).
2. S. Liu, X. Zhao, W. Sun. Maneuverability comparison between Mecanum and traditional wheeled robots. Robot. Auton. Syst. **142**, 103765 (2021).
3. Y. Wang, H. Zhang. Adaptive control of Mecanum-wheeled mobile robots for trajectory tracking in uncertain environments. IEEE Trans. Control Syst. Technol. **30**(5), 1899-1906 (2022).
4. J. Smith, R. Johnson. Dynamic modeling and traction optimization for Mecanum wheel systems. Int. J. Robot. Res. **39**(7), 845-862 (2020).
5. L. Rodriguez, T. Garcia. Real-time control architecture for Mecanum-wheeled robots using wireless communication protocols. J. Control Eng. Appl. Inform. **23**(2), 45-56 (2021).
6. Y. Liu, Q. Wang. Embedded system design for omnidirectional mobile robots with Mecanum wheels. Embedded Syst. Lett. **13**(2), 101-104 (2021).
7. J. Chen, Y. Yang. Trajectory planning for Mecanum-wheeled mobile robots in complex environments. Int. J. Adv. Robot. Syst. **17**(4), 1-12 (2020).
8. A. Wilson, D. Brown. Kinematic control optimization for Mecanum-wheeled robots with application to precision positioning. J. Intell. Robotic Syst. **106**(3), 1-14 (2022).
9. M. Garcia, S. Lopez. A comparative study of gesture recognition techniques for robotic control interfaces. Robot. Auton. Syst. **132**, 103546 (2020).
10. M. Zhou, Z. Hu. Optimization of power management for Mecanum-wheeled mobile robots. Energy Convers. Manag. **242**, 114473 (2021).