ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Research on multimodal motion switching and cooperative control of wheel-legged robot on complex terrain

Kexu Yan[1,*] *and Shuozhou Niu*[1]

[1] School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China

SUMMARY: *With the continuous growth of the demand for complex terrain operations, the efficient perception, smooth switching and stable control of wheel-legged robots in unstructured environments have become the research focus. Focusing on the problem of multimodal motion switching and cooperative control of wheeled legged robots on complex terrain, this paper constructs a system modeling and kinematic analysis framework, fuses stereo vision, IMU, encoder, wheel speed meter and foot contact information, and designs a multimodal environment perception and feature fusion method. A motion mode switching mechanism based on switching demand function, dual-threshold hysteresis determination and smooth trajectory transition is proposed, and a wheel-leg cooperative control strategy for whole-body stability constraint is constructed. The experimental results show that the terrain recognition accuracy of the proposed method reaches 97.4%, and the reasoning time is 18.7 ms. The comprehensive passing rate on four types of complex terrain reaches 94.6%, and the average passing time is 12.8 s. Under the disturbance condition, the maximum attitude deviation is controlled within 6.4°, the recovery time is shortened to 1.8 s, and the task completion rate reaches 95.4%. The results show that the proposed method can effectively improve the continuous passing ability and operation robustness of the wheel-legged robot in complex terrain, and has practical significance for promoting the intelligent development of autonomous mobile equipment in complex environments.*

KEYWORDS: *Wheel-legged robot; Complex terrain; Motion mode switching; Cooperative control*

1 Introduction

Wheel-legged robots have both the mobile advantages of high-speed and low-energy consumption of wheeled mechanisms and good obstacle, ditch and terrain adaptation capabilities of legged mechanisms, showing high application value in mountain inspection, disaster rescue, field transportation, mining operations and other scenarios [1]. Compared with the single wheeled robot, it is not easily limited by steps, gravel, slope and pothole terrain. Compared with the pure legged robot, it can also maintain high maneuvering efficiency in relatively flat areas [2, 3]. Therefore, how to make the robot switch reasonably among wheeled, legged and their combination states according to environmental changes, and maintain attitude stability, trajectory continuity and execution efficiency during the switching process has become a key issue in current mobile robot research [4].

Under the condition of complex terrain, the motion state of robot often presents the

*632307030510@mails.cqjtu.edu.cn
https://doi.org/10.65102/is2026222
Published: 30 April, 2026

characteristics of strong coupling, strong nonlinearity and multi-constraint. Ground adhesion characteristics, slope fluctuation, obstacle distribution, contact impact and body center of gravity migration will simultaneously affect motion decision and control results [5]. We note that if a fixed mode or a single control law is still used, the robot is prone to problems such as terrain recognition lag, inaccurate switching timing, excessive body pitch fluctuation, and local path passing rate decrease. Especially when continuous slopes, loose road surfaces and unstructured obstacles are mixed, the traditional method of relying on single sensor judgment or preset rule base to complete mode switching is difficult to balance real-time performance, robustness and movement smoothness [6].

With the development of embedded computing platform, visual perception, inertial measurement, contact detection and intelligent control algorithm, the research of wheel-legged robot is shifting from single mechanism optimization to the integration of perception, decision and control. We believe that multimodal perception technology can fuse multi-source information such as vision, laser, inertial navigation, joint encoder and foot contact to provide a more complete representation of terrain category, passable area, body pose and movement risk. The motion switching mechanism needs to establish a unified mapping between environmental features, dynamic boundaries and task objectives, so that the robot can adaptively convert between rolling through, leg-raising crossing, and wheel-leg hybrid support modes. The cooperative control needs to balance the speed tracking, foot landing, fuselage stability and drive energy consumption, so as to ensure the continuous operation ability of the whole machine in complex scenes [7, 8].

Although the existing research has made some progress in wheel-leg mechanism design, terrain recognition, obstacle planning and stability control, there are still several prominent shortcomings. First, environmental perception and motion control are often processed in sections, and multi-source information does not form a unified state representation, which makes the perception results difficult to directly service the handover decision. Secondly, the research of mode switching pays more attention to whether it can switch, and pays less attention to trajectory continuity, impact suppression and attitude recovery before and after switching. Third, the control strategy is often established around a single working condition, and the overall robustness still has room for improvement in the face of continuous disturbance in complex terrain and multi-task switching [9]. Based on the above analysis, we believe that it is of clear engineering significance and methodological value to carry out the research on multimodal motion switching and cooperative control of wheel-legged robots on complex terrain.

Based on this, this paper constructs the system modeling and kinematics analysis framework of wheel-legged robot for complex terrain operation requirements, studies the multi-modal environmental perception and feature fusion method, designs the modal switching criterion and transition planning mechanism of wheel-legged motion, and further proposes the wheel-legged cooperative control strategy for whole-body stability constraints. We hope to improve the robot's ability to recognize complex terrain, the rationality of motion mode switching and the stability of cross-scene operation through the above research, and provide technical support for efficient autonomous mobility of wheel-legged robots in unstructured environments.

2 Methods and materials

2.1 System modeling and kinematic analysis of wheel-legged robot on complex terrain

In order to keep the wheel-legged robot's continuous traversing ability in steps, gravel slopes,

shallow trenches and irregular raised terrain, we first make a unified modeling of the overall mechanism configuration, coordinate system relationship and kinematic mapping. The research object adopts the wheel-leg composite chassis, and the whole machine is composed of the fuselage platform, the front and rear wheel-leg execution modules, the hub drive unit, the joint drive unit, the inertial measurement unit, the visual perception module and the airborne controller. The wheeled mechanism is responsible for the fast movement task of leveling the road surface, while the legged mechanism is responsible for crossing obstacles, lifting the body, and adjusting the contact posture. The two types of executive units work together under the same control framework, so as to give consideration to both movement efficiency and complex terrain adaptability.

As shown in Figure 1, the world coordinate system Σ_w, the body coordinate system Σ_b, the wheel end coordinate system Σ_{wi} and the foot end coordinate system Σ_{fi} are established. Here, Σ_w is used to describe the global position of the robot in the terrain environment, Σ_b is fixed to the body mass center, the xb axis points to the forward direction, the y_b axis points to the left of the body, and the z_b axis is vertical to the body upward. Each wheel-leg module is equipped with hip joint, knee joint and wheel-end drive DOF respectively, which is used to realize body height adjustment, support domain reconstruction and wheel-ground contact switching. In order to describe the state of the whole machine conveniently, the generalized coordinate vector is defined as follows.

$$q = [x, y, z, \phi, \theta, \psi, q_1, q_2, \dots, q_n]^T \tag{1}$$

where, x,y and z represent the position of the fuselage centroid in the global coordinate system, ϕ,θ and ψ are the roll Angle, pitch Angle and yaw Angle respectively, and q_n is the driving variable of each joint. This definition unifies the body pose and the wheel-leg joint state into the same state space, which is conducive to the subsequent terrain perception results being directly mapped to the mode switching and control solution process.

At the kinematic level, the wheel-legged robot has both wheel-type rolling constraints and leg-type joint chain constraints. The linear and angular velocity of the fuselage can be expressed as follows.

$$v_b = J_b(q)\dot{q} \tag{2}$$

where $v_b = [v_x, v_y, v_z, \omega_x, \omega_y, \omega_z]^T$ is the airframe speed state, and $J_b(q)$ is the Jacobian mapping matrix determined by the mechanism parameters and the current attitude. For the ith wheel-leg module, the position of its foot or wheel-end in the global coordinate system can be written as follows.

$$p_i = p_b + R(\phi, \theta, \psi)(l_i + f_i(q_i)) \tag{3}$$

where, p_b is the fuselage centroid position, R(ϕ,θ,ψ) is the fuselage attitude rotation matrix, l_i is the module mounting bias vector, and $f_i(q_i)$ is the terminal configuration function of the local joint chain. This formula is used to obtain the spatial position of the wheel end, foot end and terrain contact point under different motion modes, and is the basis for analyzing the obstacle height, leg lift amplitude and airframe ground clearance.

When the robot is in the wheel-dominant mode, the wheel-ground non-slip rolling relation determines the coupling characteristics between the wheel-end line velocity and the angular velocity, which can be expressed as follows.

$$v_i = r_i \omega_i \tag{4}$$

Here, r_i is the wheel radius, ω_i is the angular velocity of the wheel end, and v_i is the tangential velocity of the wheel edge. This constraint ensures that the robot can pass quickly with low energy consumption in a relatively flat area. However, when the step, trench or contact height difference increases significantly, the wheel constraint will no longer meet the stable traffic requirements, and the support relationship needs to be reconstructed through the leg joint. In order to measure the static stability level of the whole machine on complex terrain, the support margin index is introduced:

$$S - \min_{k \in \Omega} d\left(\mathbf{p}_c, \mathcal{L}_k\right) \tag{5}$$

where p_c is the projection point of the body mass center on the support plane, $\mathcal{L}_k$ is the boundary of the support polygon, $d(\cdot)$ represents the directed distance from the point to the boundary, Ω is the set of support edges. The larger S is, the closer the centroid projection is to the inside of the support domain, and the stronger the anti-rollover and anti-pitch disturbance ability of the whole machine is.

Based on the above modeling, it can be seen that the motion of wheel-legged robots in complex terrain does not simply depend on a single executive unit, but is reflected in the joint evolution of body pose, joint chain motion, wheel-ground contact relationship and support stability. The system modeling and kinematic analysis not only provide a unified state interpretation framework for subsequent multimodal environmental perception results, but also lay the foundation for the modal switching criterion, transition trajectory generation and wheel-leg cooperative control law design.

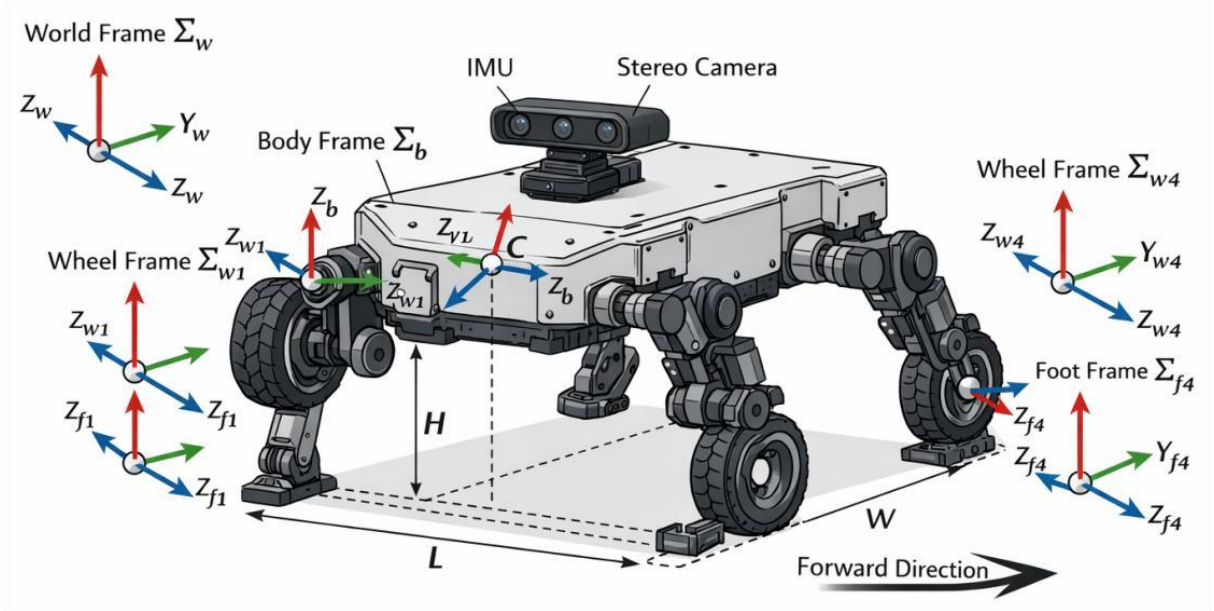

Figure 1: Definition diagram of the whole structure and coordinate system of the wheel-legged robot

2.2 Multi-modal environment perception and feature fusion method for complex terrain

The wheel-legged robot in complex terrain not only needs to identify whether there are steps, gravel, shallow trenches, slopes and loose areas in front of it, but also synchronously determines the change of body posture, wheel-ground attachment state and foot contact reliability. To this end, we construct a multi-modal environment perception system composed of stereo vision,

IMU, joint encoder, wheel speed meter and foot contact sensor. Through timing alignment, feature extraction and adaptive weighted fusion, a unified state representation for motion switching and cooperative control is formed.

In the data processing flow, firstly, the original data of each sensor is sampled synchronously and normalized. The visual branch extracts the geometric contour of the terrain, surface relief and obstacle edge information, the IMU branch outputs the pitch Angle, roll Angle and angular velocity changes, the encoder and wheel speed meter branches reflect the execution state and slip trend of the wheel leg, and the foot contact branch identifies whether the support is stable. The features extracted by the m-th sensor at time k are denoted as $f_m(k)$, then the initial set of multi-modal features can be expressed as follows.

$$\mathcal{F}(k) = \{\mathbf{f}_v(k), \mathbf{f}_i(k), \mathbf{f}_e(k), \mathbf{f}_c(k)\} \tag{6}$$

where, $f_v(k), f_i(k), f_e(k), f_c(k)$ correspond to visual, inertial, encoder wheel speed, and contact perception features, respectively. This representation puts the environment information and the state information of the ontology into a unified processing at the same time, and facilitates the subsequent establishment of the association between terrain and motion state.

Considering that the reliability of different sensors is not constant in complex working conditions, this paper introduces a quality assessment factor to construct an adaptive fusion weight. Let the quality score of the m-th mode be $s_m(k)$, then its fusion weight is defined as follows.

$$\alpha_m(k) = \frac{\exp(s_m(k))}{\sum_{j=1}^{M} \exp(s_j(k))} \tag{7}$$

Here, M is the number of modes, $\alpha_m(k) \in (0,1)$, and it satisfies $\sum_{m=1}^{M} \alpha_m(k) = 1$. When the visual occlusion or illumination mutation occurs, the quality score will decrease, and the corresponding weight will automatically decrease. When the body has a continuous impact, the contribution of IMU and contact information in the state determination will be increased accordingly. Based on this, the fusion feature vector can be obtained as follows.

$$F(k) = \sum_{m=1}^{M} \alpha_m(k) W_m f_m(k) \tag{8}$$

where, W_m is the feature mapping matrix of each modality. This formula can not only preserve the complementarity of multi-source information, but also suppress the interference of low-confidence input on the overall judgment result.

In order to serve the subsequent movement mode switching, this paper further defines the terrain traversability evaluation index. Combining the standard deviation of local elevation difference σ_h, slope Angle γ, obstacle density ρ_o and slip risk coefficient ξ_s, the difficulty index of complex terrain is constructed as follows.

$$D_t = \beta_1\sigma_h + \beta_2\gamma + \beta_3\rho_o + \beta_4\xi_s \tag{9}$$

Among them, $\beta_1, \beta_2, \beta_3, \beta_4$ are the normalized weight coefficients, and meet $\sum \beta_i = 1$. The larger D_t is, the less suitable the current area is to maintain the pure wheel rapid passage mode. After it exceeds the threshold, the system will increase the leg participation ratio and trigger the transition planning module. Therefore, the result of multimodal perception is no longer limited to terrain classification itself, but can be directly transformed into the criterion

variables required for handover decision.

As shown in Table 1, the multi-source sensor configuration and its output variables are uniformly defined in this paper. Through the above method, the system can simultaneously obtain key states such as terrain type, body posture, contact stability and slip risk, which provides reliable input for reasonable switching between wheeled and legged motion modes under complex terrain, and also lays a data foundation for subsequent trajectory correction and attitude compensation in wheel-leg cooperative control.

Table 1: Description of sensor configuration and fusion variables for multimodal sensing system

Sensor Type	Installation Position	Sampling Frequency / Hz	Main Output Variables	Feature Dimension	Main Function
Stereo Camera	Front upper side of the robot body	30	Terrain texture, obstacle edges, local height difference, traversable area	128	Identify steps, trenches, gravel, and slope contours
IMU	Near the center of mass of the robot body	200	Roll angle, pitch angle, angular velocity, linear acceleration	12	Perceive body attitude fluctuations and impact responses
Joint Encoder	Hip joints, knee joints, and wheel drive shafts	500	Joint angle, angular velocity, wheel rotational speed	16	Reflect wheel-leg execution state and joint coordination
Wheel Odometer	Left and right wheel ends	200	Wheel speed, displacement increment, slip estimation	8	Determine rolling state and adhesion variation
Foot Contact Sensor	Foot ends and contact points at the end of the wheel-leg modules	250	Contact flag, supporting force, contact duration	10	Identify support stability and foothold reliability
Fusion Output	Controller state layer	100	Terrain category, difficulty index D_t, attitude state, slip risk	20	Provide unified input for mode switching and coordinated wheel-leg control

2.3 Modal switching criterion and transition planning method for wheeled and legged motion

The wheel-legged robot is not suitable for maintaining a single movement mode for a long time under complex terrain conditions. When the ground is relatively flat, obstacles are sparse and adhesion is stable, wheeled motion can maintain high speed and low energy consumption. However, when the elevation difference is increased, the surface relief is enhanced or the wheel-ground slip is aggravated, the pure wheel passing will often cause the pitch fluctuation amplification, wheel end idling and even the body trapped. Based on this, we divide the motion modes into M_w (wheeled navigation mode), M_h (wheel-legged hybrid transition mode), M_l (leg-legged obstacle crossing mode) and M_r (recovery after passage mode), and construct a unified method around switching demand assessment, hysteresis judgment and transition trajectory generation, so that the robot can complete continuous and stable mode change in complex terrain.

In order to avoid triggering switching only by a single index, this paper synthetically combines terrain complexity, forward height difference, wheel-ground slip rate and support

margin to construct motion switching demand function:

$$\Gamma(k) = \mu_1 \bar{D}_t(k) + \mu_2 \bar{\Delta}h(k) + \mu_3 \bar{\sigma}_s(k) + \mu_4(1 - \bar{S}(k)) \tag{10}$$

where, $\bar{D}_t(k)$ is the normalized terrain difficulty index, $\bar{\Delta}h(k)$ is the local height difference in the forward direction of the robot, $\bar{\sigma}_s(k)$ is the normalized slip rate, $\bar{S}(k)$ is the normalized value of the support margin, μ_i is the weight coefficient and it can satisfy $\sum_{i=1}^{4} \mu_i = 1$. The larger $\Gamma(k)$ is, the stronger the demand of the environment for legged participation and active posture adjustment.

Considering the short-term fluctuations of sensor signals in complex terrain, it is easy to cause frequent jitter between wheel type and leg type if switching directly according to the instantaneous value. To this end, we also introduce a dual-threshold hysteresis mechanism and define the mode update rule as follows.

$$M(k+1) = \begin{cases} M_l, & \Gamma(k) \geq \tau_2 \text{and} \bar{\Gamma}_N(k) \geq \tau_2 \\ M_h, & \tau_1 \leq \Gamma(k) < \tau_2 \\ M_w, & \Gamma(k) < \tau_1 \\ M_r, & M_l \rightarrow M_w \text{during recovery} \end{cases} \tag{11}$$

where τ_1 and τ_2 are the low and high order switching thresholds, respectively, and $\tau_1<\tau_2$. Let $\bar{\Gamma}_N(k)$ denote the mean value in the last N sampling periods. Combined with the passing requirements of complex terrain test scenarios, τ_1 is set to 0.38 and τ_2 is set to 0.67 in this paper, so that the system can maintain efficient wheel-type passing under slightly undulated ground, and timely turn into leg dominant mode when significant obstacles or stability decreases. For the local height difference, when Δh exceeds 0.08 m, the system enters the wheel-leg mixing transition state. When Δh is further increased to more than 0.16 m, the leg support and lifting weight are increased. For wheel-ground slip rate, the wheel-speed output is suppressed when σ_s is greater than 0.12, and the leg-type dominant strategy is triggered when σ_s is greater than 0.22. When the support margin S drops below 0.20, then pose compensation and gait reconstruction are initiated to prevent the centroid projection from continuously approaching the support boundary.

In the transition planning stage, the smooth transition factor is constructed as follows.

$$\lambda(t) = 10\left(\frac{t}{T_s}\right)^3 - 15\left(\frac{t}{T_s}\right)^4 + 6\left(\frac{t}{T_s}\right)^5, \quad 0 \leq t \leq T_s \tag{12}$$

where, T_s is the duration of mode switching, and the first and second derivatives of $\lambda(t)$ at the starting and ending time are zero, which can effectively reduce the speed impact and attitude mutation at the switching moment. Considering the switching requirements of different terrain sections, the T_s is controlled in the interval of 0.6-1.2s in this paper, so that the robot can keep fast switching when the terrain changes slowly, and adopt a smoother transition mode in steps, gravel piles and discontinuous support areas. Based on this factor, the wheeled reference trajectory $r_w(t)$ and the legged reference trajectory $r_l(t)$ are as follows.

$$r_c(t) = (1 - \lambda(t))r_w(t) + \lambda(t)r_l(t) \tag{13}$$

Generate a uniform transition trajectory $r_c(t)$. In the wheel-leg mixing stage, the robot gradually rebuilds the support relationship through the front wheel deceleration, joint lifting, foot prepositioning and fuselage pitch compensation. In the recovery phase, the above process is reversed to make the robot return to the efficient rolling mode after passing the obstacle. In

this way, the body speed, joint displacement and contact force change can be kept continuous, and the wheel end slip or fall foot shock caused by switching too fast can be avoided. Figure 2 shows the state machine diagram of the motion mode switching.

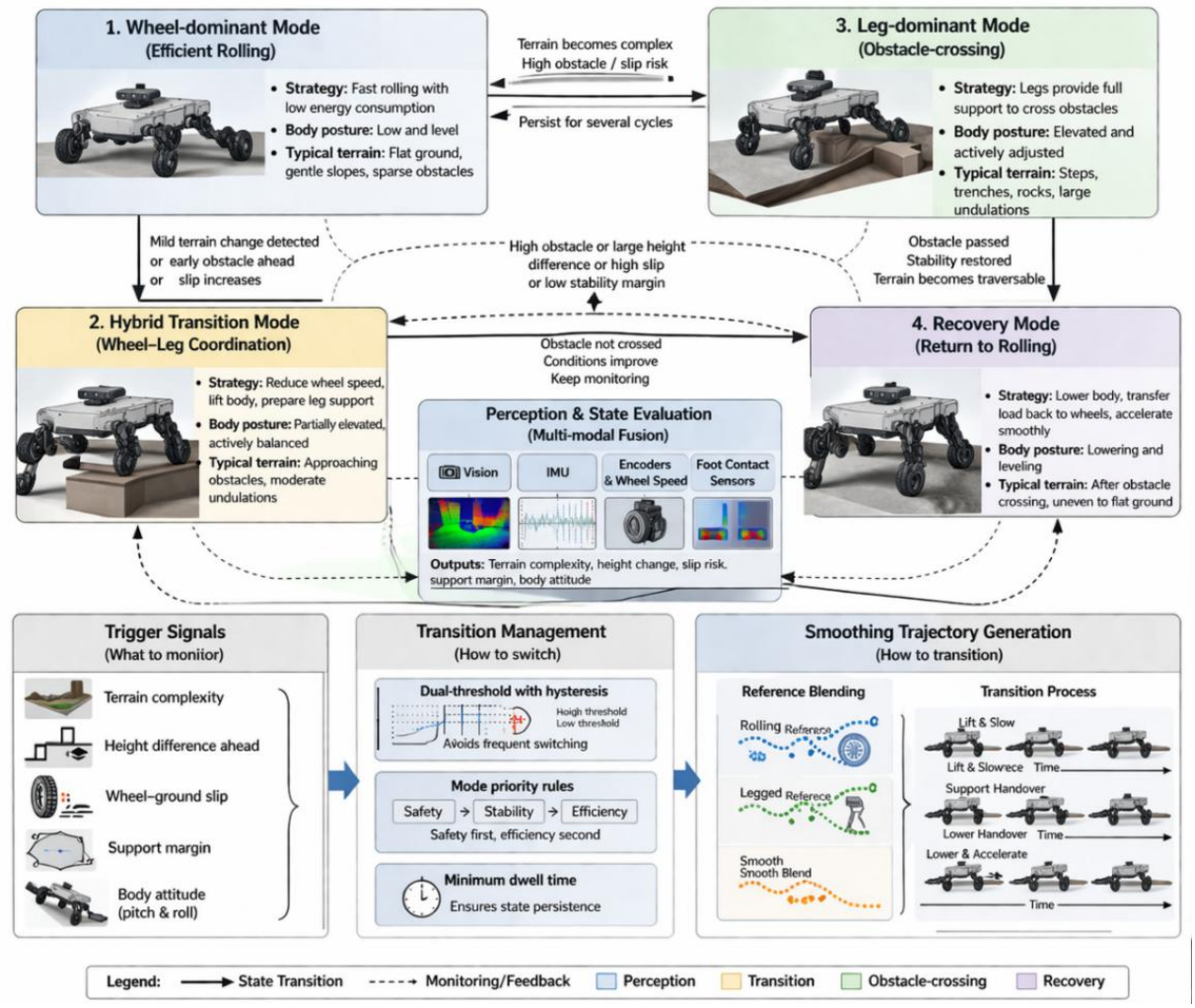

Figure 2: State machine of locomotion mode switching for the wheeled-legged robot

The motion-mode switching state machine constructed by us consists of four core states: the wheel-type cruise state is responsible for fast passage in flat areas, the wheel-leg hybrid state is responsible for modal buffering and support reconstruction, the leg-type obstacle crossing state is responsible for stable passage of steps, trenches and large undulating terrain, and the recovery state is responsible for center of gravity recovery and rolling recovery after obstacle crossing. The transition between each state is jointly triggered by switching demand function $\Gamma(k)$, forward height difference, slip rate and support margin, forming a switching mechanism driven by environmental change, constrained by state evaluation and connected by smooth trajectory. In general, the method logically realizes "first judge the demand, then switch hierarchically, and finally smooth transition", which provides a clear mode boundary for attitude tracking and torque distribution in the subsequent wheel-leg cooperative control.

2.4 Wheel-leg cooperative control strategy and whole-body stability constraint design

After the completion of complex terrain perception, multi-modal state assessment and motion

mode switch decision, the robot still faces a more critical problem, that is, how to maintain the overall motion stability under the condition of wheel drive, leg support and body attitude adjustment at the same time. Different from single wheeled chassis, wheel-legged robot has obvious coupling relationship among wheel-end velocity, joint action, contact force distribution and body posture during obstacle crossing, trench crossing and slope crossing. If the control layer only deals with wheel speed tracking and leg joint control separately, it is easy to occur that the local execution is correct but the overall machine attitude is unstable. Based on this, we construct a wheel-leg cooperative control strategy for complex terrain passing tasks, which unifies upper-level pattern scheduling, trajectory generation, stability constraints and lower-level execution into the same control architecture to achieve whole-body dynamic stability under wheel-leg cooperative behavior.

The control architecture adopts a hierarchical design idea. According to the terrain perception results and mode switching state, the upper layer outputs the desired fuselage height, target pitch Angle, desired wheel speed and foot drop reference. The middle layer completes the adjustment of the fuselage attitude, the reconstruction of the support relationship and the coordination of the contact force. The bottom layer sends speed or torque commands to wheel end actuators and joint actuators respectively. In order to simultaneously take into account trajectory tracking, attitude maintenance and control smoothness, the whole body cooperative control objective function is defined as follows.

$$J_c = \frac{1}{2}\|W_p(p_b^d - p_b)\|^2 + \frac{1}{2}\|W_a(\eta^d - \eta)\|^2 + \frac{1}{2}\|W_u \Delta u\|^2 \tag{14}$$

where, p_b^d and p_b are the desired and actual fuselage positions, η^d and η represent the desired and actual attitude vectors, Δu is the input variation of adjacent control periods, and W_p, W_a and W_u are the weight matrices. On the one hand, this objective function constrains the body pose to be as close as possible to the desired reference, on the other hand, it suppresses the mutation of the control input, so as to reduce the impact of the wheel leg switching stage.

In order to ensure that the robot maintains a feasible stable state under non-flat support conditions, a zero moment point constraint is introduced to describe the whole body stability boundary. Let the equivalent zero moment point of the robot on the support plane be pz, then it satisfies:

$$p_z = \frac{\sum_{i=1}^{n} m_i (\bar{z}_i + g) p_i - \sum_{i=1}^{n} m_i \bar{p}_i z_i}{\sum_{i=1}^{n} m_i (\bar{z}_i + g)} \tag{15}$$

Here, m_i is the mass of each rigid body, p_i is the position of the corresponding particle in the support plane, and z_i is the vertical direction coordinate. During the control process, p_z is required to be always located inside the current support polygon to avoid an increased risk of rollover due to excessive fuselage pitch or too fast load transfer. For slope and step edge scenes, the system also needs to dynamically reduce the available support region to improve the conservative estimation of boundary uncertainty.

In the wheel-leg collaborative allocation layer, the wheel-speed reference, joint acceleration reference and contact force adjustment are uniformly written into the optimization variables, and the optimal execution amount of the current control cycle is obtained by constraint solving. The solution is given as follows.

$$\min_u \frac{1}{2}\|Au - b\|_Q^2 + \frac{1}{2}\|u\|_R^2 \tag{16}$$

$$s.t.\, u_{min} \le u \le u_{max}, \quad F_n \ge 0, \quad |F_t| \le \mu F_n \tag{17}$$

where, u contains wheel end velocity correction, joint driving torque and contact force distribution variables; Au=b describes the mapping relationship between desired motion and actual execution; Q and R are task error weight and control cost weight respectively; Fn and Ft represent normal contact force and tangential contact force respectively; μ is the ground friction coefficient. The optimization model enables the robot to strike a balance between speed efficiency, attitude stability and contact safety. When the terrain attachment is reduced or the support contact is reduced, the system automatically reduces the wheeled propulsion weight and increases the leg support proportion.

For the underlying joint and wheel-end execution, this paper uses a tracking control form based on dynamics compensation to convert the desired motion into joint torque output:

$$\tau = M(q)\ddot{q}_r + C(q,\dot{q})\dot{q} + G(q) - J_c^T\lambda \tag{18}$$

where M(q) is the system inertia matrix, $C(q,\dot{q})$ is the Coriaceous force and centrifugal force term, G(q) is the gravity term, $J_c^T\lambda$ is the contact constraint reaction force mapping term, and $\ddot{q}_r$ is the reference acceleration generated by the upper planning and optimal allocation. By introducing contact reaction force compensation, the robot can more stably maintain joint output consistency in the process of foot landing, wheel-end contact change and fuselage lifting, and avoid local joint overload or transient instability. Figure 3 shows the overall architecture of wheel-leg cooperative control.

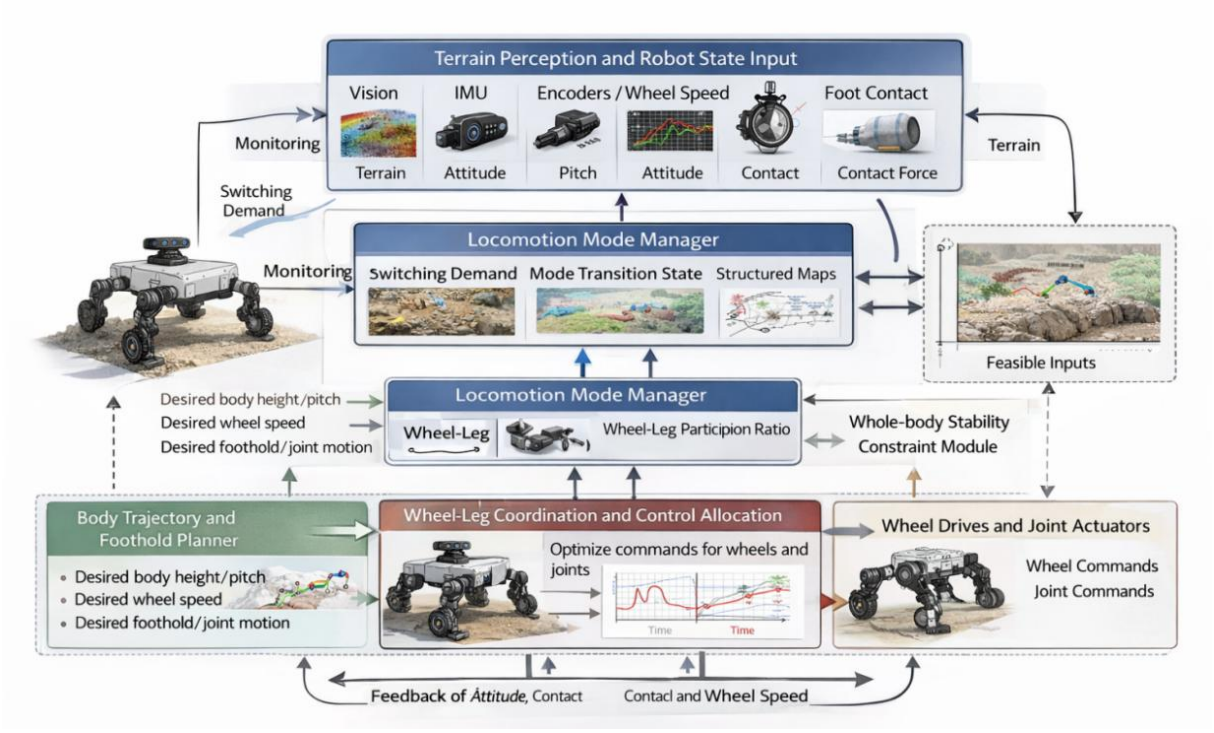

Figure 3: Overall architecture of wheel-leg coordinated control for the wheeled-legged robot.

As shown in Figure 3, the overall wheel-leg cooperative control architecture constructed in this paper consists of a perception input layer, a modal scheduling layer, a trajectory generation layer, a stability constraint and control allocation layer, and an execution feedback layer. The perceptual input layer provides terrain, pose and contact states. The mode scheduling layer determines the participation ratio of wheel type and leg type according to the current switching

state. The trajectory generation layer outputs the reference values of the body, wheel end and foot end. The control allocation layer solves the optimal control input by combining the whole body stability constraint. The feedback layer is executed to return the wheel speed, joint Angle and fuselage attitude in real time for closed-loop correction. On the whole, the strategy realizes the continuous transfer from environmental changes to control execution, so that the wheel-legged robot can not only complete the mode switch in complex terrain, but also maintain the stability of the body, reliable contact and continuous motion during the switch process.

3 Results

3.1 Performance analysis of multimodal perception and terrain recognition methods

In order to verify the effectiveness of the proposed multi-modal environment perception and terrain recognition method in complex scenes, we conduct comparative experiments on terrain classification accuracy, recognition stability and real-time reasoning ability. The experimental object covers typical complex terrain such as flat ground, gravel road, slope, step, trench and loose irregular surface. The input information is composed of vision, inertia, encoder and contact state. In order to highlight the role of multi-modal fusion strategy, this paper selects a single convolution classification model, a vision and inertia combination model, and a transformer-based terrain recognition model as comparison methods, and evaluates the proposed method under the same training set and test set conditions. The evaluation metrics include Accuracy, Precision, Recall, F1-score and Inference Time to fully reflect the performance of the model in terms of classification accuracy, class discrimination ability and deployment efficiency. Figure 4 shows the curve of terrain recognition accuracy and training rounds change.

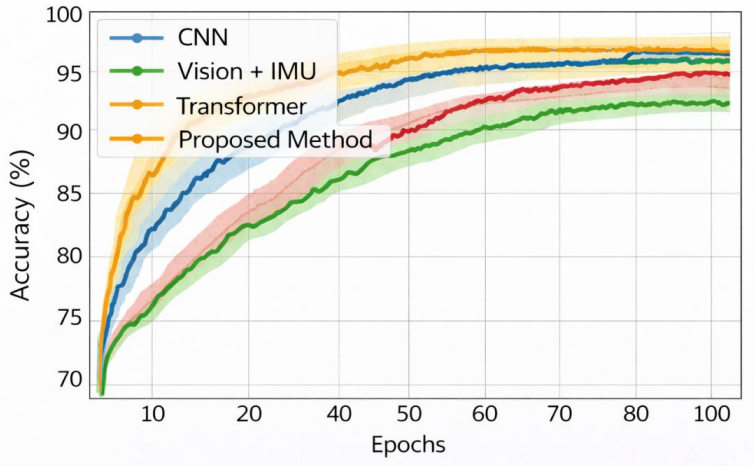

Figure 4: Accuracy curves versus training epochs for terrain classification methods.

As shown in Figure 4, with the continuous increase of training rounds, the recognition accuracy of all kinds of methods shows a trend of rapid rise at first and then gradually stable, but there are obvious differences in the convergence speed and final stable accuracy of different methods. The single CNN model had certain convergence ability in the early stage, but the improvement was limited in the middle and late stages, and it entered the platform area after about the 60th round. The combined model of vision and IMU is more stable than the single modal method, and the recognition ability of slope and gravel terrain is enhanced, but there are

still misjudgments on the complex terrain with stronger structure such as steps and trenches. The Transformer model relies on strong global feature expression ability and shows good fitting effect in the middle and late stages, but the inference overhead is relatively high. In contrast, the accuracy of the proposed method can form a faster improvement in the early stage of training, which is significantly higher than that of the other comparison methods after about 20 rounds, and maintains a smooth growth trend in the subsequent training process, indicating that the multi-modal feature fusion not only enhances the ability of the model to identify local geometric changes and body state disturbances. It also improves training stability in complex terrain classification tasks.

To further compare the comprehensive performance of each method on the test set, Table 2 shows the main performance comparison of different models.

Table 2: Comparison table of multimodal perception and terrain recognition performance

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Inference Time (ms)
CNN	87.2	85.9	84.4	85.1	15.2
Vision + IMU	88.8	87.5	86.9	87.2	17.4
Transformer	91.3	90.1	89.4	89.7	26.5
Proposed Method	97.4	96.8	96.1	96.4	18.7

From Table 2, we can see that the proposed method is at the optimal level in all indicators. Among them, the Accuracy reaches 97.4%, which is 10.2 percentage points higher than that of the CNN model, 8.6 percentage points higher than that of the vision and IMU combination model, and 6.1 percentage points higher than that of the Transformer model. The Precision, Recall and F1-score of the proposed method are 96.8%, 96.1% and 96.4%, respectively. The results show that the proposed method not only has a high overall classification accuracy, but also has a more balanced recognition ability among different terrain categories. In terms of real-time performance, the Inference Time of the proposed method is 18.7 ms, which is slightly higher than 15.2 ms of the basic CNN model, but still significantly lower than 26.5 ms of the Transformer model, which can meet the requirements of online terrain recognition and state update of wheel-legged robots in complex terrain.

In summary, the recognition accuracy of the proposed method stably reaches 97.4% at the 100th round of training, which is 6.1 percentage points higher than that of the suboptimal method. At the same time, the inference time is controlled within 18.7 ms, which indicates that the proposed method has achieved a good balance between recognition accuracy, convergence speed and online deployment feasibility. This indicates that the constructed multi-modal perception and terrain recognition mechanism can provide more reliable environmental input for subsequent motion mode switching and wheel-leg cooperative control.

3.2 Motion mode switching effect and passing performance analysis under complex terrain

In order to verify the practical effectiveness of the motion mode switching mechanism proposed in this paper in complex terrain, we further select four typical scenes such as step terrain, gravel slope, trench terrain and mixed obstacle terrain to conduct experiments. In the experiments, the proposed method is compared with the rule-driven switching method, the single threshold switching method and the adaptive switching method based on reinforcement learning. The coherence of mode switching, the success rate, the average passage time and the attitude fluctuation control effect of the robot in complex terrain are mainly investigated. Different from

the previous section, which focuses on the sensing recognition accuracy, this section pays more attention to whether the sensing results can be transformed into stable, smooth and efficient actual passing behavior after entering the decision-making layer.

In order to visually present the mode evolution process of the robot in complex terrain, Figure 5 shows the schematic diagram of the motion mode switching process of the wheel-legged robot in a typical obstacle scene.

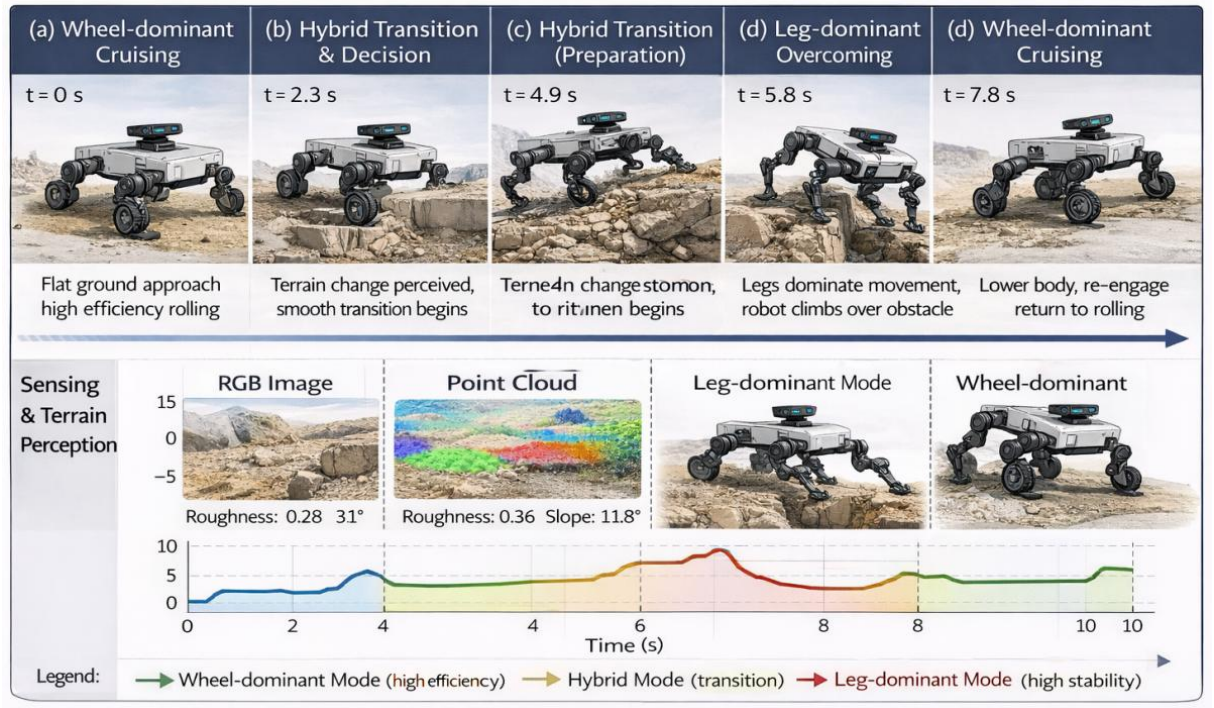

Figure 5: Schematic diagram of the motion mode switching process under complex terrain

It can be seen from Figure 5 that the robot mainly maintains the wheeled cruise state when approaching the flat area, the body height is low, and the wheel end undertakes the main driving task. When the vision and contact information identify the edge of the front step or the local elevation difference continues to rise, the system first enters the wheel-leg mixing transition stage, and completes the support region reconstruction by reducing the wheel speed, lifting the fuselage and adjusting the foot preparation drop point in advance. In the formal obstacle crossing stage, the leg support is dominant, the body pitch Angle is actively constrained within a small range, and the wheel end is more responsible for auxiliary traction and attitude synergy. After the obstacle is passed, the robot enters the recovery phase to gradually drop the center of gravity and switch back to the efficient rolling mode.

In order to further compare the overall passing ability of different methods in complex terrain, Figure 6 shows the comparison results of passing rate and average passing time of each method in the comprehensive scene test.

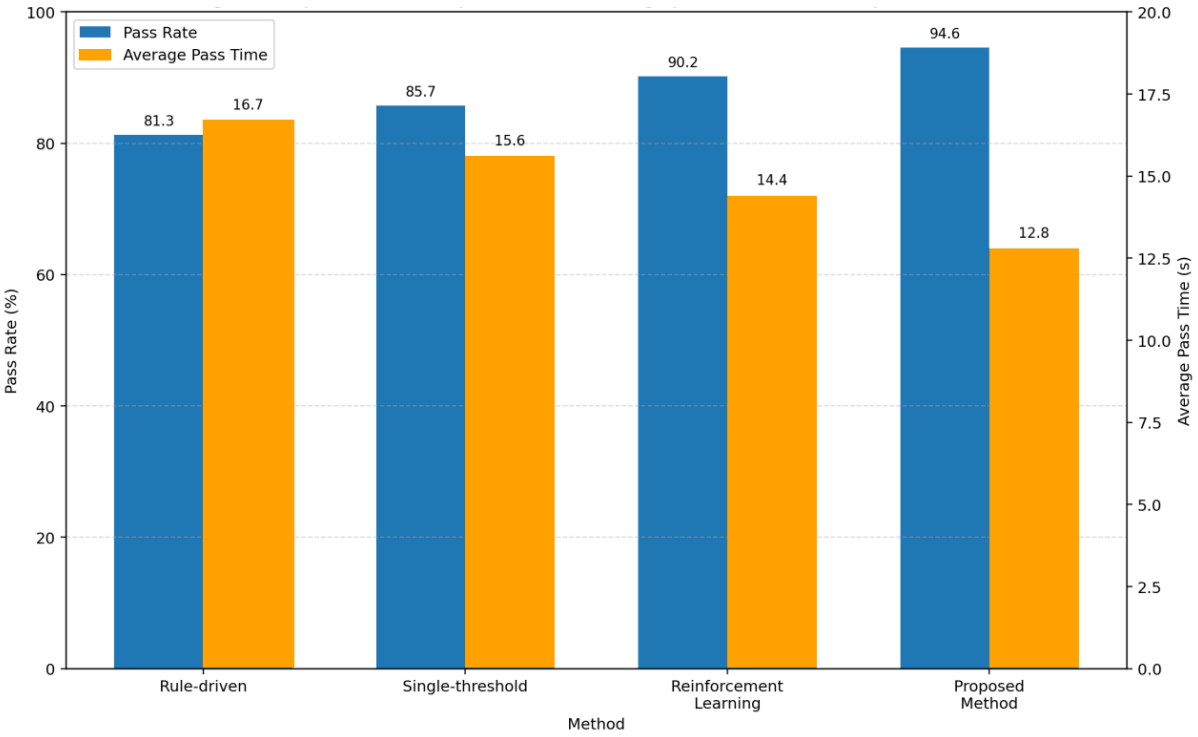

Figure 6: Comparison of passing rate and average passing time of different methods in complex terrain

As can be seen from Figure 6, the comprehensive pass rate of the proposed method on four complex terraforms reaches 94.6%, which is significantly higher than that of the rule-driven method (81.3%), the single threshold method (85.7%) and the reinforcement learning method (90.2%). In terms of the average passage time, the proposed method is 12.8 s, which is 3.9 s, 2.8 s and 1.6 s shorter than the rule-driven method, the single threshold method and the reinforcement learning method, respectively. The main reason for this difference is that the rule-driven method has a stiff response to the terrain change, and it is easy to have switching lag in the step edge and the trench transition area. Although the single threshold method can complete the trigger quickly, it has the problem of frequent switching and is easy to cause local jitter. Reinforcement learning method has strong adaptability, but it still has policy fluctuations when the scene changes drastically. In contrast, the proposed method combines multimodal state assessment, hysteresis determination and smooth trajectory transition to form a better balance between switching timing and passage efficiency.

Table 3 further statistics the motion switching and passing performance of the proposed method in different complex terrain scenes.

Table 3: Statistical table of motion switching and passing performance in different complex terrain scenes

Terrain Scenario	Average Traversal Time / s	Success Rate / %	Average Number of Mode Switches	Maximum Pitch Angle / (°)	Average Number of Slips
Step Terrain	11.9	96.2	2.0	7.8	0.6
Gravel Slope Terrain	12.7	93.8	2.3	8.2	1.3
Trench Terrain	13.4	92.5	2.5	8.6	0.9
Mixed Obstacle Terrain	15.2	91.7	3.1	9.4	1.5

According to the results in Table 3, in the step terrain, the average robot passage time is 11.9 s, the success rate reaches 96.2%, the average number of switching is 2.0, and the maximum pitch Angle is 7.8°. In the gravel slope scene, although the surface attachment condition is poor, the system still maintains a success rate of 93.8%, and the average number of slips is controlled at 1.3. In the trench terrain, due to the more sensitive support switching, the maximum pitch Angle increased to 8.6°, but the overall passage time was still controlled at 13.4 s. In the mixed obstacle scene, the robot needs to continuously complete multiple switches of wheel, hybrid and leg. The average passing time is 15.2 s, and the success rate is 91.7%, which shows a good ability to adapt to continuous scenes. On the whole, with the increase of terrain complexity, the number of switching, passage time, maximum pitch Angle and average slip increase, but the change range of each index is generally controllable, indicating that the transition planning and cooperative control mechanism designed in this paper can still maintain a relatively stable attitude adjustment ability and terrain passage performance in continuous obstacle environment.

3.3 Wheel-leg cooperative control results and system robustness analysis

In order to further verify the stability of the wheel-leg cooperative control strategy constructed in this paper under complex terrain and disturbance conditions, we select the traditional PID control, the wheel-leg decoupling control and the proposed method for comparative experiments. During the experiment, small lateral disturbances and contact changes were added to the continuous tasks of step crossing, gravel slope crossing and obstacle recovery, so as to investigate the differences of different control strategies in body attitude maintenance, recovery speed, slip suppression and task completion reliability. Different from the previous two sections, which focus on perception and mode switching, this section focuses on analyzing whether the control layer can stably implement the upper-level planning results into the overall machine motion.The attitude response curve of the fuselage under cooperative control is shown in Figure 7.

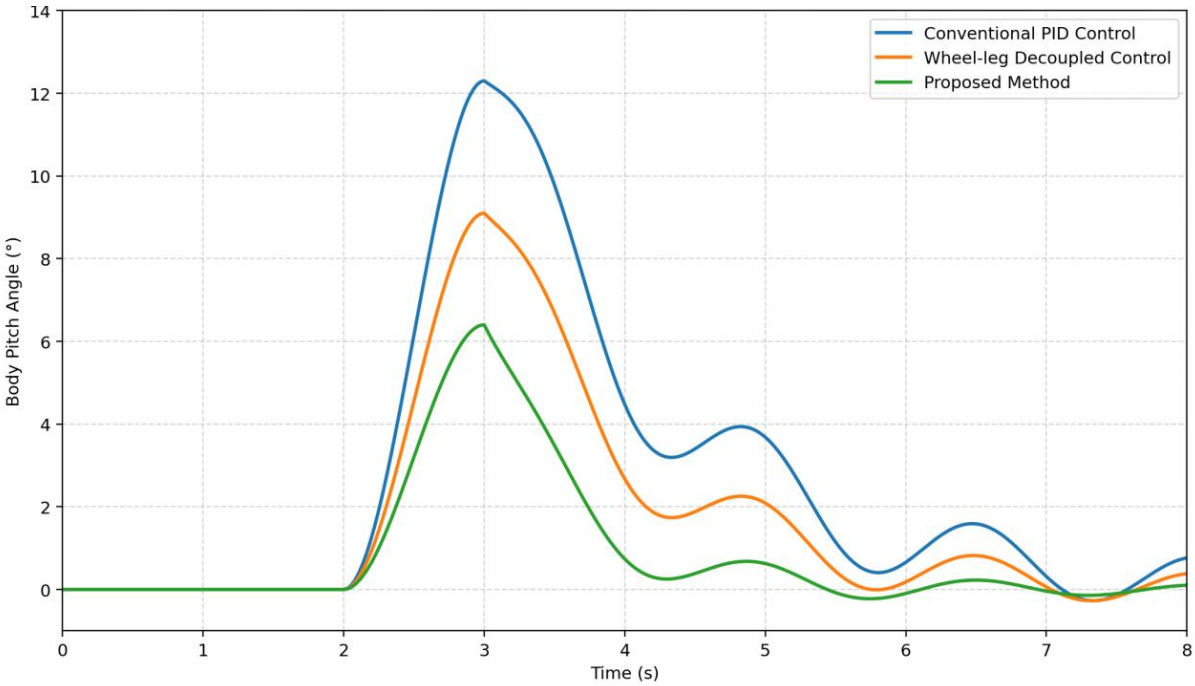

Figure 7: Fuselage attitude response curve under cooperative control

As shown in Figure 7, under the effect of cooperative control, although the pitch Angle and roll Angle of the fuselage fluctuate for a short time in the obstacle crossing stage, the overall response is relatively stable. At the moment of the transition between the obstacle contact and the support, the method proposed in this paper will appear a certain amplitude of attitude deviation, and then converge rapidly. The peak of the pitch Angle is controlled within 6.4°, and the peak of the roll Angle is 4.9°, and it returns to the stable interval within 1.8 s. In contrast, the traditional PID control lacks unified coordination between wheel speed regulation and leg support, and is prone to attitude oscillation in the obstacle crossing stage. The peak pitch reaches 12.3°, and the recovery time is extended to 3.8 s. Although decoupling control is more stable than PID, the peak value of pitch is still 9.1° and the peak value of roll is 6.7° because it fails to fully consider the wheel-ground contact change and joint load redistribution. This indicates that the proposed method can more effectively balance the relationship between wheel-end drive and leg support under the attitude closed-loop constraint, thereby reducing the body impact during obstacle crossing.

To further quantify the system performance of different control strategies, Table 4 shows the comparison results of stability and robustness indexes.

Table 4: Comparison table of system stability and robustness under different control strategies

Control Strategy	Average Attitude Error / (°)	Maximum Attitude Deviation / (°)	Recovery Time / s	Average Slip Rate / %	Energy Consumption / J	Task Completion Rate / %
Traditional PID Control	4.8	12.3	3.8	13.7	149.3	83.7
Wheel-Leg Decoupled Control	3.1	9.1	3.1	9.4	136.9	89.6
Proposed Method	1.7	6.4	1.8	6.2	128.6	95.4

It can be seen that the average attitude error of the proposed method is 1.7°, and the maximum attitude deviation is 6.4°, which are significantly lower than 4.8° and 12.3° of the traditional PID control, and also better than 3.1° and 9.1° of the decoupling control. In terms of recovery time, the proposed method is only 1.8 s, which is 2.0 s shorter than the traditional PID and 1.3 s shorter than the decoupled control. At the same time, the average slip rate of the proposed method in complex terrain is controlled at 6.2%, and the task completion rate reaches 95.4%, which indicates that the proposed method can not only quickly restore the stability of the airframe when the disturbance occurs, but also maintain a high reliability of terrain passage. In terms of energy consumption index, the proposed method is 128.6J, which is slightly lower than 136.9J of decoupling control and 149.3J of traditional PID, indicating that the wheel-leg cooperative control does not sacrifice efficiency for stability, but achieves a good compromise between stability and energy consumption.

It can be seen from the results in Figure 7 and Table 4 that the proposed method still maintains strong control stability under the combined effect of complex terrain and external disturbance: the maximum attitude deviation is controlled within 6.4°, the recovery time is shortened to 1.8 s, and the task completion rate reaches 95.4%, which is 11.7 percentage points higher than that of the traditional PID. This shows that the proposed wheel-leg cooperative control strategy can effectively enhance the system robustness and provide stable support for continuous motion in complex terrain.

4 Discussion

From the experimental results, we believe that the core difficulty of the wheel-legged robot in complex terrain is not only whether it can identify obstacles or whether it has two kinds of movement capabilities of wheel-leg, but also whether the perception, switching and control can form a continuous closed loop. At present, many methods perform well on a single module, for example, some have high recognition accuracy, some have strong obstacle crossing ability, and some controllers have fast response in local conditions. However, once the system is in the scene where the gravel slope, trench edge and mixed obstacles continuously appear, the system performance is often degraded because of the information transmission is not timely, the switching conditions are too rigid, or the control output is not coordinated. The key reason why our method can achieve better results in terrain recognition accuracy, passing efficiency and attitude stability at the same time is that multi-modal perception, hysteresis switching decision and smooth transition control are integrated into a unified framework, so that environmental changes can be more stably converted into motion decision and execution instructions.

From the perception level, our proposed multi-modal information fusion method significantly improves the reliability of complex terrain recognition. In the experiment, the recognition accuracy of the proposed method reaches 97.4%, which is significantly higher than that of the comparison methods, indicating that after the combination of visual information with IMU, wheel speed, and contact state, the robot no longer only depends on the surface appearance to judge the road condition, but can understand the terrain risk by combining its own posture and support changes. We believe that such a state expression is more in line with the working characteristics of wheel-legged robots in real scenes, because the impact on their motion safety is not a single geometric obstacle, but the combined effect of terrain relief, attachment conditions and body force changes. After the quality of the sensing results is improved, the trigger basis of mode switching is more stable, which provides a prerequisite for the subsequent improvement of the pass rate.

From the perspective of handover, the advantage of our method is not only faster handover, but more stable handover. The comprehensive pass rate reaches 94.6%, and the average pass time is 12.8 s, which shows that the system can better balance safety and efficiency in complex terrain. Although the single threshold method can be triggered quickly, it has the problems of local jitter and frequent round-trip switching. The reinforcement learning method has adaptive advantages in some scenes, but it still has strategy fluctuations when facing continuous and complex terrain. In contrast, by setting the intermediate transition mode, there is a buffer link between the wheeled cruise and the legged obstacle crossing, and then with the smooth trajectory generation, the speed, the fuselage height and the support relationship can be gradually changed, so the impact of the switching process is effectively weakened. This result indicates that we prefer to understand the motion mode switching in complex terrain as a continuous transition process driven by the environment, rather than a simple discrete state jump.

From the control level, the wheel-leg cooperative control strategy we constructed improves the robustness of the system significantly. The experimental results show that the maximum attitude deviation is controlled within 6.4°, the recovery time is shortened to 1.8 s, and the task completion rate reaches 95.4%, which indicates that the cooperative control not only improves the attitude maintaining ability of the airframe, but also enhances the recovery speed of the system under disturbance and contact change. We believe that the reason is that this control strategy does not simply regulate wheel speed and joint action separately, but unifies the body attitude, support allocation and wheel leg actuation into the control objective. In this way, when the topography mutation or load transfer occurs, the robot can more reasonably allocate the role

ratio of wheel end propulsion and leg support, so as to avoid the problem of effective local control but insufficient global stability.

Of course, we should also note that there are still some limitations in this study. Although the current experimental scene covers a variety of typical complex terrain, it is still mainly a controlled environment. The adaptability to high uncertainty scenes such as muddy, slippery and soft surfaces needs to be further verified. Some threshold parameters and weights still depend on experimental tuning, and it may be necessary to introduce a stronger online self-adjustment mechanism after the scene span is further expanded. In addition, this paper mainly focuses on the stability and control effect, and the discussion of energy consumption optimization and mechanism wear during long-term operation is not deep enough. In the future, we will continue to carry out research on higher frequency state prediction, online learning mechanism with stronger generalization ability, and real machine verification in complex field environments, so as to further improve the autonomous operation ability of wheel-legged robots in real unstructured scenes.

5 Conclusion

In this paper, a multi-modal motion switching and cooperative control method is proposed to solve the problems of perception instability, uneven mode switching and insufficient control robustness of wheel-legged robots in complex terrain. Starting from the whole machine modeling and kinematic analysis, the vision, inertia, wheel speed and contact information are unified into a multi-modal environment perception framework, and the continuous switching of wheel, hybrid and leg motions is realized through hysteresis determination and smooth transition planning. On this basis, a wheel-leg cooperative control and whole-body stability constraint model is further constructed to enhance the attitude retention ability of the system under the conditions of obstacle crossing, disturbance and support change. Experiments show that the proposed method is superior to the comparison strategies in terms of recognition accuracy, passing efficiency and control stability. The terrain recognition accuracy reaches 97.4%, the comprehensive pass rate reaches 94.6%, the maximum attitude deviation is reduced to 6.4°, and the task completion rate is improved to 95.4%. In general, the proposed method achieves the integration of perception, decision making and control, and provides feasible technical support for the application of wheel-legged robots in rescue inspection, field transportation and autonomous operation of complex landforms.

References

[1] Chen Z, Wang S, Wang J, Xu K, Lei T, Zhang H, Wang X, Liu D, Si J. Control strategy of stable walking for a hexapod wheel-legged robot[J]. ISA Transactions, 2021, 108: 367-380. DOI: 10.1016/j.isatra.2020.08.033.

[2] Cui L, Wang S, Zhang J, Zhang D, Lai J, Zheng Y, Zhang Z, Jiang Z P. Learning-Based Balance Control of Wheel-Legged Robots[J]. IEEE Robotics and Automation Letters, 2021, 6(4): 7667-7674. DOI: 10.1109/LRA.2021.3100269.

[3] Chen H Y, Wang T H, Ho K C, Ko C Y, Lin P C. Development of a novel leg-wheel module with fast transformation and leaping capability[J]. Mechanism and Machine Theory, 2021, 163: 104348. DOI: 10.1016/j.mechmachtheory.2021.104348.

[4] Chen Z, Li J, Wang S, Wang J, Ma L. Flexible gait transition for six wheel-legged robot

with unstructured terrains[J]. Robotics and Autonomous Systems, 2022, 150: 103989. DOI: 10.1016/j.robot.2021.103989.

[5] He J, Sun Y, Yang L, Gao F. Model Predictive Control of a Novel Wheeled-Legged Planetary Rover for Trajectory Tracking[J]. Sensors, 2022, 22(11): 4164. DOI: 10.3390/s22114164.

[6] Miki T, Lee J, Hwangbo J, Wellhausen L, Koltun V, Hutter M. Learning robust perceptive locomotion for quadrupedal robots in the wild[J]. Science Robotics, 2022, 7(62): eabk2822. DOI: 10.1126/scirobotics.abk2822.

[7] Shen Y, Chen G, Li Z, Wei N, Lu H, Meng Q, Guo S. Cooperative control strategy of wheel-legged robot based on attitude balance[J]. Robotica, 2023, 41(2): 566-586. DOI: 10.1017/S0263574722001436.

[8] Takahashi N, Nonaka K. Model Predictive Leg Configuration Control for Leg/Wheel Mobile Robots that Adapts to Changes in Ground Level[J]. Journal of Robotics and Mechatronics, 2023, 35(1): 160-170. DOI: 10.20965/JRM.2023.P0160.

[9] Qin D, Zhang G, Zhu Z, Zeng X, Cao J. An online terrain classification framework for legged robots based on acoustic signals[J]. Biomimetic Intelligence and Robotics, 2023, 3: 100091. DOI: 10.1016/j.birob.2023.100091.

[10] Cao J, Zhang J, Wang T, Meng J, Li S, Li M. Mechanism design and dynamic switching modal control of the wheel-legged separation quadruped robot[J]. Robotica, 2024, 42(3): 660-683. DOI: 10.1017/S0263574723001686.

[11] Zhu B, He J, Sun J. Kinematic modeling and hybrid motion planning for wheeled-legged rovers to traverse challenging terrains[J]. Robotica, 2024, 42(1): 153-178. DOI: 10.1017/S0263574723001406.

[12] Mao N, Chen J, Spyrakos-Papastavridis E, Dai J S. Dynamic modeling of wheeled biped robot and controller design for reducing chassis tilt angle[J]. Robotica, 2024, 42(8): 2713-2741. DOI: 10.1017/S0263574724001061.

[13] Liang L, Xu Y, Han L, Liu Y. System design and control of the sphere-wheel-legged robot[J]. Robotica, 2024, 42(10): 3363-3379. DOI: 10.1017/S0263574724001401.

[14] Korayem M H, Nikseresht N, Asadi A H. Vision-based dynamic modeling of wheeled-legged robot considering slippage using Gibbs-Appell formulation[J]. Robotica, 2024, 42(11): 3784-3802. DOI: 10.1017/S0263574724001644.

[15] Mohamed S, Im Y, Shin H, Kim Y, Shin B. Development of a novel quadruped hybrid wheeled-legged mobile robot with telescopic legs[J]. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2024, 238(22): 10785-10797. DOI: 10.1177/09544062241271624.

[16] Wang D, Fang B, Zheng J. Design and research of deformable wheel-legged robot based on origami mechanisms[J]. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2024, 238(17): 8769-8784. DOI: 10.1177/09544062241241417.

[17] Ren X, Liu H, Qin Y, Han L, Nie S, Xie J. Trajectory optimization and stability control for a novel unmanned intelligent wheel-legged vehicles on unstructured terrains[J]. Robotics and Autonomous Systems, 2024, 182: 104818. DOI: 10.1016/j.robot.2024.104818.

[18] Khan R A I, Zhang C, Pan Y, Zhang A, Li R, Zhao X, Shang H. Hierarchical optimum control of a novel wheel-legged quadruped[J]. Robotics and Autonomous Systems, 2024, 180: 104775. DOI: 10.1016/j.robot.2024.104775.

[19] Lee J, Bjelonic M, Reske A, Wellhausen L, Miki T, Hutter M. Learning robust autonomous navigation and locomotion for wheeled-legged robots[J]. Science Robotics, 2024, 9(89): eadi9641. DOI: 10.1126/scirobotics.adi9641.

[20] Zhang A, Zhou R, Zhang T, Zheng J, Chen S. Balance Control Method for Bipedal Wheel-Legged Robots Based on Friction Feedforward Linear Quadratic Regulator[J]. Sensors, 2025, 25(4): 1056. DOI: 10.3390/s25041056.

[21] Wang B, Xin Y, Chen C, Song Z, Sun B, Guo T. Whole-Body Control with Uneven Terrain Adaptability Strategy for Wheeled-Bipedal Robots[J]. Electronics, 2025, 14(1): 198. DOI: 10.3390/electronics14010198.

[22] Gao J, Fan R, Yang H, Pang H, Tian H. Research on Motion Control Method of Wheel-Legged Robot in Unstructured Terrain Based on Improved Central Pattern Generator (CPG) and Biological Reflex Mechanism[J]. Applied Sciences, 2025, 15(15): 8715. DOI: 10.3390/app15158715.

[23] Xu K, Wang S, Shi L, Li J, Yue B. Horizon-stability control for wheel-legged robot driving over unknow, rough terrain[J]. Mechanism and Machine Theory, 2025, 205: 105887. DOI: 10.1016/j.mechmachtheory.2024.105887.

[24] Pan Z, Liu S, Zhou S, Li B, Niu Z, Wang R. Collaborative control of the knee-wheeled wheel-legged robot: Considering actuators characteristics[J]. Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, 2025, 239(2): 272-289. DOI: 10.1177/09596518241272801.

[25] Zhao D, Liu J, Yang P, Cui T, Wu D, Zhang L. Modeling and stability control of steering and reconfiguration motion for wheel-legged metamorphic robot[J]. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2025, 239(11): 4256-4272. DOI: 10.1177/09544062251319073.

[26] Chen Z, Huang J, Wang S, Wang J, Xu Y. Whole-body stability control of foot walking for wheel-legged robot on unstructured terrain[J]. Robotic Intelligence and Automation, 2025, 45(2): 292-303. DOI: 10.1108/RIA-10-2024-0209.

[27] Mei X, Wei Y, Guo C, Zhang X. Experimental research of wheel-legged robot crossing obstacles[J]. Robotica, 2025, 43(7): 2459-2479. DOI: 10.1017/S0263574725000499.