Uncertainty Adaptive Shared Control with Multimodal Intent Fusion for Human-Guided Quadruped Locomotion on Uneven Terrain

Likai Wu
Shandong Huayu University of Technology, Dezhou, Shandong, China

***Abstract*—Human-guided quadruped locomotion requires reliable intent inference and safe authority allocation under uncertain interaction and uneven terrain. This study proposes an uncertainty-adaptive shared-control framework with multimodal intent fusion for a Unitree Go2 quadruped. Interaction force, relative human–robot motion, operator-motion cues, and terrain context are fused through a temporal model to estimate desired motion and maneuver mode. Intent confidence is evaluated using posterior entropy and prediction variance, while locomotion risk is computed from traversability, slope, obstacle proximity, interaction load, and stability margin. The human authority factor is then adjusted online to blend the inferred human command with a robot-safe command. The shared command is realized through gait scheduling, terrain-aware foothold optimization, whole-body control, and interaction-compliant base regulation. Experiments across five guidance scenarios show that the proposed method improves intent accuracy, interaction load, slip resistance, stability margin, and task success compared with force-only guidance, fixed-authority shared control, and ablated variants.**

***Keywords*—Quadruped robot; human–robot interaction; shared autonomy; intent recognition; uncertainty estimation; terrain-aware locomotion; assistive robotics**

I. Introduction

Quadruped robots are increasingly used as mechatronic platforms for unstructured indoor and outdoor environments because they can negotiate stairs, curbs, discontinuous footholds, and uneven ground more flexibly than wheeled robots [1]–[5]. Recent progress in model predictive control, whole-body control, adaptive actuation, and learning-based locomotion has improved the robustness of legged robots on challenging terrain [6]–[12]. These advances have motivated assistive and collaborative applications, including robotic guide dogs, physical guidance, and human–robot load transportation [13]–[19].

However, human-guided quadruped locomotion remains difficult because the robot must solve two coupled problems at the same time. First, the robot must infer what the human intends from noisy and incomplete observations. Interaction force alone is informative, but it can also include involuntary disturbance, balance correction, hesitation, or conflict with robot motion. Second, the robot must decide how much control authority should be assigned to the human command and how much should be reserved for autonomous stabilization. A fixed authority ratio is insufficient because high intent confidence on flat ground should lead to responsive motion, whereas low confidence or high terrain risk should trigger stronger robot-side safety control.

This study addresses the coupled inference-and-control problem by proposing an uncertainty-adaptive shared-control framework with multimodal intent fusion. The central idea is to treat human-guided locomotion as a coupled inference and control problem rather than a simple force-to-velocity mapping. The robot estimates a posterior distribution over human locomotion intent, computes confidence from posterior entropy and temporal prediction variance, evaluates terrain and stability risk, and then blends the human-oriented command with a robot-safe command using an adaptive authority factor. The final command is implemented by terrain-aware gait scheduling, foothold optimization, whole-body control, and interaction-compliant base regulation. Section II reviews three closely related research directions: human-guided quadruped assistance, physical human–robot cooperation and shared control, and terrain-aware legged locomotion.

The main contributions are threefold. First, a multimodal temporal intent-fusion model is formulated for human-guided quadruped locomotion, combining interaction force, relative motion, operator-motion cues, and task context. Second, an uncertainty-adaptive shared-control law is developed to regulate human authority according to both intent confidence and locomotion risk. Third, a terrain-aware locomotion realization layer is integrated with the shared-control layer and evaluated through repeated trials across representative guidance scenarios.

II. Related Work

A. Human-Guided Quadrupeds and Assistive Navigation

Human-guided quadruped robots have been studied mainly for assistive navigation, where the platform must preserve stable locomotion while remaining responsive to physical guidance. Hamed et al. developed a hierarchical and safety-oriented control framework for cooperative locomotion between a robotic guide dog and a human [13]. Xiao et al. demonstrated leash-guided hybrid physical interaction for a robotic guide dog, showing that the physical connection can serve as an informative guidance interface rather than being treated solely as an external disturbance [14]. Chen et al. considered comfort-oriented quadruped guidance for visually impaired users and emphasized the importance of controlling interaction behavior in addition to achieving the navigation objective [15]. From a complementary perspective, Hwang et al. analyzed handler–guide-dog interaction patterns to inform the design of robotic companions, illustrating how physical and motion cues can convey guidance information during locomotion [16].

Shared-navigation studies further motivate adaptive allocation of control authority between a human and an autonomous robot. Kamikubo et al. investigated shared-control design for navigation robots used by blind and low-vision people, demonstrating the value of combining user input with autonomous assistance rather than assigning complete control to either side [20]. These studies establish physical guidance, interaction quality, and shared navigation as central elements of assistive mobility. However, human-guided locomotion on uneven terrain introduces a coupled problem that is less directly addressed in prior work: uncertainty in the inferred guidance intent can arise simultaneously with changes in terrain traversability and locomotion stability. A shared-control system for this setting must therefore determine not only the intended motion, but also how strongly that inferred intent should influence the robot when confidence or locomotion safety changes.

B. Physical Human–Robot Cooperation and Shared Control

Physical human–robot interaction provides a broader foundation for coordinating agents that exchange forces, motion constraints, and task-level commands. Darvish et al. proposed a hierarchical architecture for human–robot cooperation processes, connecting high-level cooperative objectives with lower-level robot behaviors [21]. In collaborative manipulation, de Cos and Dimarogonas developed adaptive cooperative control for human–robot load manipulation [17], while Culbertson et al. studied decentralized adaptive control for collaborative manipulation of rigid bodies [22]. Elwin et al. extended related principles to human–multirobot collaborative mobile manipulation, in which physical coupling and coordination must be maintained among a human and multiple mobile agents [23]. Collectively, these works show the importance of compliance, adaptation, and force regulation when a robot shares control of a physically coupled task.

Related coordination problems have also been studied for legged systems. Hamed et al. formulated distributed feedback controllers for stable cooperative locomotion of quadrupedal robots using virtual constraints [24], and Kim et al. investigated layered centralized and distributed control for cooperative locomotion of two quadrupedal robots [25]. Yang et al. considered collaborative navigation and manipulation of a cable-towed load by multiple quadrupedal robots [18], whereas Gu et al. studied adaptive interactive control of coupled human and quadruped robot load motion [19]. These approaches primarily regulate coupled dynamics, interaction forces, or coordination constraints after the cooperative objective has been specified. Human-guided quadruped locomotion additionally requires online interpretation of an operator's intended maneuver from noisy physical and motion cues. This makes uncertainty estimation directly relevant to shared-control allocation: when the inferred intent is ambiguous, the robot should rely less on the human-oriented command, whereas clear intent under safe conditions should permit more responsive guidance.

C. Terrain-Aware Legged Locomotion

The execution of a shared command ultimately depends on the ability of the quadruped to maintain dynamically feasible and terrain-compatible locomotion. Model-based legged-locomotion research has developed several complementary tools for this purpose. Hybrid zero-dynamics methods provide a formal basis for stabilizing periodic locomotion and optimizing dynamically consistent gaits [26], [27], while divergent-component-of-motion control offers an alternative formulation for regulating dynamic balance in three-dimensional walking [28]. Model predictive control has been widely used to incorporate future body motion, contact evolution, and dynamic constraints. Representative examples include variable-horizon MPC with swing-foot dynamics [29], representation-free MPC for dynamic quadruped motions [7], and MPC with explicit terrain insight for dynamic legged locomotion [6].

Foothold optimization further connects desired body motion with discrete contact placement. Xin et al. proposed online relative footstep optimization using discrete-time MPC and later combined robust footstep planning with LQR control for dynamic quadrupedal locomotion [30], [31]. These methods improve stability and contact feasibility by considering the relationship between body motion and future support locations. In parallel, learning-based approaches have substantially expanded quadruped robustness on irregular terrain. Hwangbo et al. demonstrated learned agile and dynamic motor skills for legged robots [9]; Lee et al. developed learning-based locomotion over challenging terrain [8]; and Miki et al. integrated terrain perception with learned locomotion for operation in unstructured outdoor environments [10]. DreamWaQ introduced implicit terrain imagination for robust quadrupedal locomotion through deep reinforcement learning [11], while Wang et al. investigated dynamic locomotion with perception-less terrain adaptation [12].

Model-based and learning-based locomotion methods therefore provide strong mechanisms for stability, terrain adaptation, and foothold feasibility, but the commanded motion is typically generated independently of a physically interacting human. In human-guided operation, the locomotion layer must instead execute a command whose reliability changes with the quality of the inferred human intent. The framework developed in this study connects these three research directions by fusing multimodal guidance cues, estimating intent confidence, combining that confidence with terrain and stability risk to allocate control authority online, and realizing the resulting shared command through terrain-aware gait and foothold control. This organization preserves responsiveness to physical guidance while maintaining robot-side stabilization when either the inferred intent or the terrain conditions become less reliable.

III. System Modeling and Multimodal Intent Fusion

Fig. 1 summarizes the proposed framework. The high-level layer performs multimodal feature encoding, temporal intent fusion, posterior confidence estimation, terrain-risk assessment, and authority allocation. The low-level layer realizes the resulting command through gait scheduling, foothold optimization, whole-body control, and interaction-compliant base regulation.

Let $\mathcal{W}$ and $\mathcal{B}$ denote the world and robot base frames. The quadruped state at time step k is

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{p}_k^\top & \mathbf{\Theta}_k^\top & \mathbf{v}_k^\top & \boldsymbol{\omega}_k^\top & \mathbf{p}_{f,1,k}^\top & \cdots & \mathbf{p}_{f,4,k}^\top \end{bmatrix}^\top, \quad (1)$$

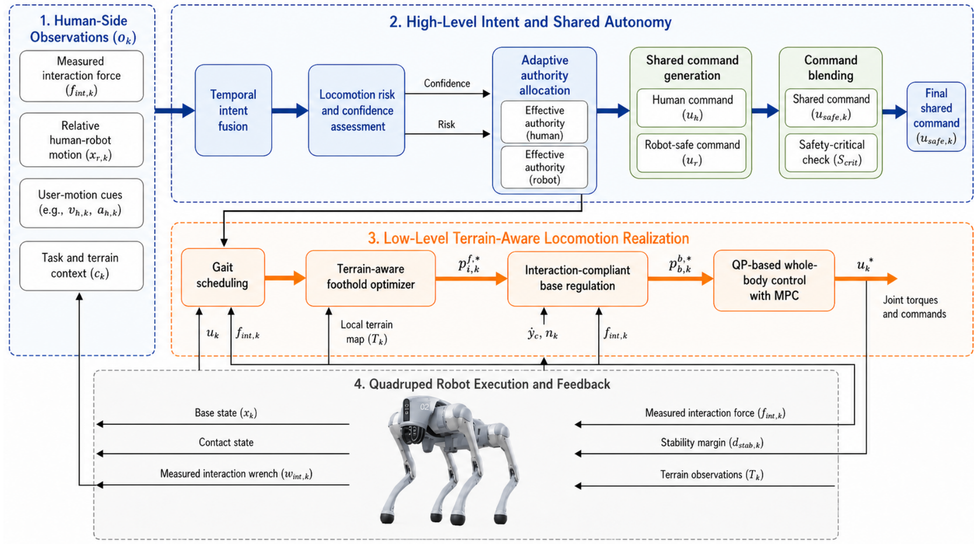

Fig. 1. Proposed uncertainty-adaptive shared-control framework. The framework explicitly connects multimodal intent inference, confidence and risk estimation, online authority allocation, and terrain-aware locomotion realization.

where, $\mathbf{p}_k$, $\mathbf{\Theta}_k$, $\mathbf{v}_k$, and $\boldsymbol{\omega}_k$ are the base position, orientation, linear velocity, and angular velocity, and $\mathbf{p}_{f,i,k}$ is the foot position of leg i. The commanded base motion is

$$\mathbf{u}_k = [v_{x,k}^{\text{cmd}},\ v_{y,k}^{\text{cmd}},\ \omega_{z,k}^{\text{cmd}}]^\top. \tag{2}$$

The physical interaction wrench is

$$\mathbf{w}_{\text{int},k} = [\mathbf{f}_{\text{int},k}^\top,\ \boldsymbol{\tau}_{\text{int},k}^\top]^\top, \tag{3}$$

and the relative human–robot state is

$$\mathbf{x}_{r,k} = [\mathbf{r}_k^\top,\ \dot{\mathbf{r}}_k^\top,\ \Delta\psi_k]^\top, \tag{4}$$

where, $\mathbf{r}_k$ is the planar relative displacement and $\Delta\psi_k$ is the heading difference.

The multimodal observation vector is constructed as

$$\mathbf{o}_k = [\mathbf{f}_{\text{int},k}^\top,\ \dot{\mathbf{f}}_{\text{int},k}^\top,\ \mathbf{x}_{r,k}^\top,\ \hat{\mathbf{v}}_{h,k}^\top,\ \hat{\mathbf{a}}_{h,k}^\top,\ \boldsymbol{\xi}_k^\top]^\top, \tag{5}$$

where $\hat{\mathbf{v}}_{h,k}$ and $\hat{\mathbf{a}}_{h,k}$ are estimated operator-motion cues, and $\boldsymbol{\xi}_k$ contains task-context and local-terrain features. The latent human intent is represented by

$$\mathbf{z}_k = [v_k^d,\ \omega_k^d,\ \boldsymbol{\mu}_k^\top]^\top, \tag{6}$$

where, v_k^d and ω_k^d are the desired speed and yaw rate, and $\boldsymbol{\mu}_k$ is the maneuver-mode probability vector.

For implementation, the observation is divided into force, motion, and context branches:

$$\mathbf{h}_{f,k} = \phi_f([\mathbf{f}_{\text{int},k}^\top, \dot{\mathbf{f}}_{\text{int},k}^\top]^\top), \tag{7}$$

$$\mathbf{h}_{m,k} = \phi_m([\mathbf{x}_{r,k}^\top, \hat{\mathbf{v}}_{h,k}^\top, \hat{\mathbf{a}}_{h,k}^\top]^\top), \tag{8}$$

$$\mathbf{h}_{c,k} = \phi_c(\boldsymbol{\xi}_k), \tag{9}$$

where, ϕ_f, ϕ_m, and ϕ_c are lightweight multilayer encoders. The fused temporal feature is

$$\mathbf{h}_k = \text{GRU}([\mathbf{h}_{f,k}^\top, \mathbf{h}_{m,k}^\top, \mathbf{h}_{c,k}^\top]^\top, \mathbf{h}_{k-1}). \tag{10}$$

The intent posterior is then predicted as

$$\hat{v}_k^d = \mathbf{W}_v \mathbf{h}_k + b_v, \tag{11}$$

$$\hat{\omega}_k^d = \mathbf{W}_\omega \mathbf{h}_k + b_\omega, \tag{12}$$

$$\hat{\boldsymbol{\mu}}_k = \text{softmax}(\mathbf{W}_\mu \mathbf{h}_k + \mathbf{b}_\mu). \tag{13}$$

In the implemented model, each modality encoder contains two fully connected layers with 32 hidden units, the GRU uses a 16-step observation window and 64 hidden units, and the output head jointly optimizes an L_1 regression loss for v_k^d and ω_k^d and a cross-entropy loss for the maneuver class. The maneuver set contains forward, backward, leftward, rightward, and S-bend guidance. Desired-speed and yaw-rate labels are obtained from the time-aligned reference command and smoothed base trajectory, while ambiguous transition samples are retained for uncertainty estimation instead of being forced into a single deterministic class. All modality inputs are normalized using training-set statistics before temporal fusion. This implementation makes the modality fusion explicit and avoids treating physical interaction force as the only source of intent information.

The intent-confidence score is computed from the normalized maneuver entropy and the temporal variance of the regressed speed and yaw predictions:

$$H_{\mu,k} = -\sum_{j=1}^{M} \hat{\mu}_{k,j} \log(\hat{\mu}_{k,j}), \tag{14}$$

$$\sigma_{z,k}^2 = \mathrm{var}_{i\in\mathcal{E}}([\hat{v}_{k,i}^d, \hat{\omega}_{k,i}^d]), \tag{15}$$

$$\gamma_k = \mathrm{clip}\left(1 - \lambda_H \frac{H_{\mu,k}}{\log M} - \lambda_\sigma \sigma_{z,k}^2,\ 0,\ 1\right), \tag{16}$$

where, M is the number of maneuver classes and $\mathcal{E}$ denotes an ensemble or temporal dropout set used to estimate prediction variance.

IV. Uncertainty-Adaptive Shared Control

The shared-control layer produces two commands. The human-oriented command is obtained from the inferred intent:

$$\mathbf{u}_{h,k}^* = \Phi(\hat{v}_k^d, \hat{\omega}_k^d, \hat{\boldsymbol{\mu}}_k, \mathbf{x}_{r,k}), \tag{17}$$

where, $\Phi(\cdot)$ maps the intended speed, yaw rate, and maneuver mode to a feasible base command. The robot-safe command is computed through a constrained optimization problem:

$$\mathbf{u}_{r,k}^* = \arg\min_{\mathbf{u}} J_r(\mathbf{u}; \mathbf{x}_k, \mathbf{x}_{r,k}, \mathbf{m}_k), \tag{18}$$

with

$$J_r = w_p J_{\mathrm{prog}} + w_s J_{\mathrm{safe}} + w_f J_{\mathrm{force}} + w_t J_{\mathrm{terrain}} + w_c J_{\mathrm{smooth}}. \tag{19}$$

The terrain and stability descriptor is

$$\mathbf{m}_k = [\Delta h_k,\ s_k,\ \rho_k,\ d_{\mathrm{obs},k},\ d_{\mathrm{stab},k}]^\top, \tag{20}$$

where, Δh_k is terrain-height variation, s_k is local slope, ρ_k is traversability, $d_{\mathrm{obs},k}$ is obstacle clearance, and $d_{\mathrm{stab},k}$ is the support stability margin. The normalized locomotion risk is

$$\begin{aligned}\eta_k = \mathrm{clip}\big(&\omega_h \overline{\Delta h}_k + \omega_s \bar{s}_k + \omega_o(1 - \bar{d}_{\mathrm{obs},k}) \\ &+ \omega_f \|\mathbf{f}_{\mathrm{int},k}\|_2 / f_{\max} + \omega_d(1 - \bar{d}_{\mathrm{stab},k}), 0, 1\big).\end{aligned} \tag{21}$$

The risk weights were normalized to satisfy $\sum_i \omega_i = 1$; in the Go2 experiments, $[\omega_h, \omega_s, \omega_o, \omega_f, \omega_d] = [0.18, 0.18, 0.20, 0.22, 0.22]$. The weights are intentionally close in magnitude so that no single cue dominates the aggregate risk score. Interaction load and support margin receive the largest weights (0.22 each) because they are the two terms that most directly represent immediate physical interaction load and loss of support stability in the shared-control loop. Obstacle proximity is assigned an intermediate weight (0.20), whereas terrain-height variation and local slope are each assigned 0.18. This choice therefore introduces a modest safety bias while retaining contributions from all five normalized risk cues.

A. Sensitivity to Risk-Weight Variation

The sensitivity of the aggregate risk score to moderate weight variation was assessed by perturbing each nominal risk weight by either -10% or $+10\%$ and renormalizing the resulting vector to sum to one. All $2^5 = 32$ simultaneous corner combinations were evaluated. The largest ℓ_1 distance between a perturbed weight vector and the nominal vector was 0.0998 (Algorithm 1). Because the five normalized risk components lie in $[0, 1]$ and both weight vectors sum to one, the corresponding change in the unclipped aggregate risk score is bounded by

$$|\eta_k' - \eta_k| \le \frac{1}{2}\|\boldsymbol{\omega}' - \boldsymbol{\omega}\|_1 \le 0.0499. \tag{22}$$

Thus, simultaneous $\pm 10\%$ perturbations of the individual weights change the aggregate risk score by at most approximately 0.05 before clipping. The selected weighting scheme is therefore locally insensitive to moderate perturbations, although samples close to an authority-allocation or safety threshold can remain sensitive to small changes in the aggregate risk value.

The human-authority factor is designed to increase with intent confidence and decrease with locomotion risk:

$$\alpha_k = \mathrm{clip}(\alpha_{\min} + (\alpha_{\max} - \alpha_{\min})\gamma_k(1 - \eta_k), \alpha_{\min}, \alpha_{\max}). \tag{23}$$

To avoid abrupt switching, it is filtered as

$$\bar{\alpha}_k = \beta \bar{\alpha}_{k-1} + (1 - \beta)\alpha_k, \quad 0 < \beta < 1. \tag{24}$$

The disagreement between the human-oriented and robot-safe commands is

$$\delta_k = \|\mathbf{W}_u(\mathbf{u}_{h,k}^* - \mathbf{u}_{r,k}^*)\|_2, \tag{25}$$

and the effective authority is

$$\tilde{\alpha}_k = \bar{\alpha}_k \exp(-\lambda_\delta \delta_k). \tag{26}$$

The final shared command is

$$\mathbf{u}_k = \tilde{\alpha}_k \mathbf{u}_{h,k}^* + (1 - \tilde{\alpha}_k)\mathbf{u}_{r,k}^*. \tag{27}$$

In safety-critical states,

$$\mathcal{S}_{\mathrm{crit}} = \{\|\mathbf{f}_{\mathrm{int},k}\|_2 > f_{\mathrm{crit}} \text{ or } \eta_k > \eta_{\mathrm{crit}} \text{ or } d_{\mathrm{stab},k} < d_{\mathrm{crit}}\}, \tag{28}$$

the command is replaced by a conservative safe command $\mathbf{u}_{\mathrm{safe},k}$.

Algorithm 1 Uncertainty-Adaptive Shared Control

Require: $\mathbf{o}_k$, $\mathbf{x}_k$, $\mathbf{x}_{r,k}$, local terrain map
Encode force, motion, and context features.
Update temporal feature $\mathbf{h}_k$ and predict $\hat{v}_k^d$, $\hat{\omega}_k^d$, $\hat{\boldsymbol{\mu}}_k$.
Compute intent confidence γ_k from entropy and prediction variance.
Compute terrain and stability risk η_k.
Generate $\mathbf{u}_{h,k}^*$ and $\mathbf{u}_{r,k}^*$.
Update $\tilde{\alpha}_k$ using confidence, risk, and command disagreement.
if $\mathcal{S}_{\text{crit}}$ is true **then**
  $\mathbf{u}_k \leftarrow \mathbf{u}_{\text{safe},k}$.
else
  $\mathbf{u}_k \leftarrow \tilde{\alpha}_k \mathbf{u}_{h,k}^* + (1-\tilde{\alpha}_k)\mathbf{u}_{r,k}^*$.
end if
return $\mathbf{u}_k$.

V. Terrain-Aware Locomotion Realization

The low-level controller converts $\mathbf{u}_k$ into contact timing, foothold locations, desired base motion, and joint torques. The contact state is

$$\boldsymbol{\chi}_k = [\chi_{1,k}, \chi_{2,k}, \chi_{3,k}, \chi_{4,k}]^\top, \quad \chi_{i,k} \in \{0,1\}. \tag{29}$$

The gait frequency, duty factor, and swing duration are adjusted as

$$f_{g,k} = \Gamma_f(\|\mathbf{u}_k\|, \eta_k), \tag{30}$$
$$\zeta_k = \Gamma_\zeta(\eta_k, \|\mathbf{f}_{\text{int},k}\|_2), \tag{31}$$
$$T_{s,k} = \Gamma_T(f_{g,k}, \zeta_k). \tag{32}$$

When terrain risk or interaction load increases, the gait scheduler increases the duty factor and reduces aggressive swing motions.

The local terrain map is

$$\mathcal{M}_k = \{h(x,y), s(x,y), \rho(x,y), o(x,y)\}, \tag{33}$$

where, h, s, ρ, and o denote local height, slope, traversability, and obstacle occupancy. The nominal foothold of leg i is

$$\mathbf{p}_{f,i,k}^{\text{nom}} = \mathbf{p}_{\text{hip},i,k} + \mathbf{R}(\psi_k)\mathbf{r}_i^{\text{ref}} + \mathbf{K}_v \mathbf{u}_k T_{\text{step},k}. \tag{34}$$

Interaction and risk biases are incorporated as

$$\mathbf{p}_{f,i,k}^{\text{bias}} = \mathbf{p}_{f,i,k}^{\text{nom}} + \mathbf{K}_{\text{int}}\mathbf{f}_{\text{int},k} + \mathbf{K}_\eta \nabla \eta_k. \tag{35}$$

The final foothold is selected from the candidate set $\mathcal{C}_{i,k}$:

$$\mathbf{p}_{f,i,k}^* = \arg\min_{\mathbf{p}\in\mathcal{C}_{i,k}} J_{f,i}(\mathbf{p}), \tag{36}$$

where,

$$J_{f,i} = \lambda_n \|\mathbf{p} - \mathbf{p}_{f,i,k}^{\text{bias}}\|_2^2 + \lambda_t J_{\text{terr}} + \lambda_s J_{\text{supp}} + \lambda_k J_{\text{kin}}. \tag{37}$$

Here, J_{terr} penalizes unsafe terrain, J_{supp} penalizes low support margin, and J_{kin} enforces kinematic feasibility.

For whole-body execution, the robot dynamics are

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{h}(\mathbf{q},\dot{\mathbf{q}}) = \mathbf{S}^\top \boldsymbol{\tau} + \mathbf{J}_c^\top \boldsymbol{\lambda} + \mathbf{J}_{\text{int}}^\top \mathbf{w}_{\text{int}}. \tag{38}$$

The controller solves

$$\min_{\ddot{\mathbf{q}},\boldsymbol{\lambda},\boldsymbol{\tau}} \omega_b J_b + \omega_c J_c + \omega_{\text{sw}} J_{\text{sw}} + \omega_\tau J_\tau + \omega_{\text{int}} J_{\text{intc}}, \tag{39}$$

subject to torque, friction-cone, and contact constraints. The QP is solved at the joint-command rate with the friction coefficient set according to the estimated terrain class; infeasible solutions fall back to a conservative trot command with reduced velocity limits. Interaction compliance is added through

$$\Delta\dot{\mathbf{x}}_{b,k} = \mathbf{K}_{a,k}\mathbf{f}_{\text{int},k} + \mathbf{D}_{a,k}\dot{\mathbf{f}}_{\text{int},k}, \quad \mathbf{K}_{a,k} = k_a \gamma_k (1-\eta_k)\mathbf{I}. \tag{40}$$

Thus, when the operator intent is inferred with high confidence and the terrain is safe, the robot behaves more compliantly; when risk increases, the compliance gain decreases and stabilization dominates.

VI. Experimental Protocol

A. Unitree Go2 Platform and Implementation

The framework was implemented on a Unitree Go2 quadruped with 12 actuated joints, onboard IMU, joint encoders, and the Unitree SDK communication interface. A handle-type physical guidance module was mounted near the robot trunk and equipped with a six-axis force/torque sensor to measure physical guidance interaction. A waist-mounted IMU provided operator-motion cues, and a forward-facing RGB-D camera produced a local height-map and obstacle map for terrain-risk estimation. All sensor streams were time synchronized in ROS 2. The shared-control layer ran at 50 Hz, the terrain and foothold layer ran at 100 Hz, and the joint-level execution loop ran at 500 Hz. Table I summarizes the implementation settings.

B. Operator-in-the-Loop Trial Protocol

The hardware evaluation followed a standardized operator-in-the-loop protocol using a fixed pool of 10 laboratory operators. After a brief familiarization with the handle-guidance procedure, the operators executed predefined forward, backward, rightward, leftward, and S-bend guidance maneuvers. During each trial, the handle-mounted force/torque sensor and waist-mounted IMU described in Section VI-A supplied synchronized interaction and operator-motion signals together with the robot state. Route geometry, velocity limits, termination criteria, and sensor-processing settings were kept consistent

TABLE I. UNITREE GO2 EXPERIMENTAL AND CONTROLLER SETTINGS

Item	Setting
Robot platform	Unitree Go2 quadruped, 12 actuated DOFs
State feedback	Onboard IMU and joint encoders, 500 Hz
Interaction sensing	Handle-mounted 6-axis force/torque sensor, 100 Hz; 8 Hz low-pass filtering
Operator-motion cues	Waist IMU velocity and acceleration estimates, 100 Hz
Terrain perception	Front RGB-D camera and local height map, 10 Hz
Computation	ROS 2 workstation with Unitree SDK interface
Intent model	Three encoders, 16-step GRU, 64 hidden units
Shared-control rate	50 Hz; authority bounds $\alpha_{\min} = 0.20$, $\alpha_{\max} = 0.90$
Foothold/gait update	100 Hz; QP-based whole-body command update at 500 Hz
Thresholds	$\eta_{\text{crit}} = 0.82$, $f_{\text{crit}} = 38$ N, $d_{\text{crit}} = 2.5$ cm
Trials	30 trials per scenario and method; 150 trials per method
Terrain types	Asphalt path, leaf-covered soil, rocky/mossy uneven ground, grass, and simulated rough/block terrains

across the compared methods. The evaluation focused on controller-level measures of intent inference, interaction load, locomotion stability, and task execution.

C. Scenarios, Baselines, and Failure Criteria

The evaluation used five representative guidance scenarios: forward guidance, backward guidance, rightward guidance, leftward guidance, and S-bend guidance. Each method was tested using the same route length, terrain layout, velocity limits, and termination conditions. Hardware trials were conducted under real sensing, contact, and outdoor lighting conditions. Separate simulation environments were used for controller debugging and parameter pre-tuning and were not included in the repeated-trial hardware statistics. A trial was counted as failed if the robot violated the stability-margin threshold for more than 0.5 s, exceeded the force-safety threshold for more than 0.3 s, slipped continuously for more than three consecutive support phases, contacted an obstacle, or failed to complete the route within the maximum allowed time.

Five methods were compared. Force-Only Guidance (FOG) maps interaction force directly to the velocity command. Fixed-Authority Shared Control (FASC) blends human and robot commands using a constant authority factor. Multimodal Shared Control without Uncertainty Adaptation (MSC-w/o-UA) uses multimodal intent inference but keeps authority fixed. Uncertainty-Adaptive Shared Control without Terrain-Aware Foothold Adaptation (UASC-w/o-TFA) keeps the proposed high-level shared control but uses nominal foothold tracking. The proposed full framework uses multimodal intent fusion, uncertainty-adaptive authority allocation, and terrain-aware foothold optimization.

D. Metrics and Statistical Analysis

Intent understanding was evaluated by desired-speed error e_v, desired-yaw-rate error e_ω, and maneuver accuracy Acc_μ. Cooperative consistency was evaluated by relative-state error e_{rel}. Interaction load was evaluated by average force f_{avg}, peak force f_{peak}, and command jerk j_{cmd}. Locomotion robustness was evaluated by stability margin m_{stab} and slip rate r_{slip}, while task performance was evaluated by completion time and success rate. Continuous trial-level metrics are reported as mean $\pm$ standard deviation unless otherwise specified; slip and success are reported as rates over 150 trials per method. Statistical significance against the proposed method was evaluated using a two-sided Wilcoxon signed-rank test, with $p < 0.05$ considered significant.

VII. RESULTS AND ANALYSIS

A. Overall Quantitative Comparison

Table II reports the repeated-trial comparison. The proposed method achieved the lowest speed and yaw-rate estimation errors, the highest maneuver classification accuracy, the lowest interaction force, and the highest success rate. Compared with FOG, the proposed method reduced the speed estimation error from 0.184 m/s to 0.066 m/s, reduced the peak interaction force from 34.8 N to 17.9 N, reduced the slip rate from 12.6% to 4.3%, and increased the success rate from 76.7% to 96.7%. The improvements in e_v, f_{peak}, m_{stab}, and completion time were significant against all baselines under the Wilcoxon test ($p < 0.05$). Compared with FASC, the improvement indicates that fixed authority is insufficient when terrain risk and intent confidence vary over time.

The normalized metric summary in Fig. 3 further shows that the proposed method has balanced improvements across intent accuracy, interaction load, slip resistance, and task success. The advantage over UASC-w/o-TFA is especially visible in locomotion robustness, suggesting that uncertainty-aware authority allocation must be matched with terrain-aware foothold realization.

B. Base-Velocity Response and Authority Adaptation

Fig. 4 compares the forward base-velocity response. FOG and FASC have slower convergence because their commands depend on either raw force mapping or a fixed authority factor. MSC-w/o-UA converges faster, but it cannot reduce human authority when uncertainty increases. UASC-w/o-TFA improves high-level blending but still exhibits larger gait-cycle oscillations due to nominal foothold tracking. The proposed method reaches the target velocity fastest and has the smallest post-transition oscillation.

Fig. 5 illustrates how confidence and risk jointly adjust human authority. During rough terrain and obstacle-zone intervals, the risk score increases and intent confidence decreases, causing α_k to move from human-dominant control toward robot-stabilized control. When the robot returns to safer terrain, α_k increases again, allowing more responsive human guidance. This adaptation preserves responsive human guidance on safer terrain while shifting authority toward robot-side stabilization as uncertainty or terrain risk increases.

C. Multiscenario Guidance Results

Fig. 2 shows representative traces from the five guidance scenarios. The dominant interaction-force component is consistent with the intended motion direction, while the base velocity follows the same macroscopic trend without directly transmitting high-frequency force noise into body motion. In the S-bend scenario, the force direction changes continuously, and the resulting base response remains bounded and directionally

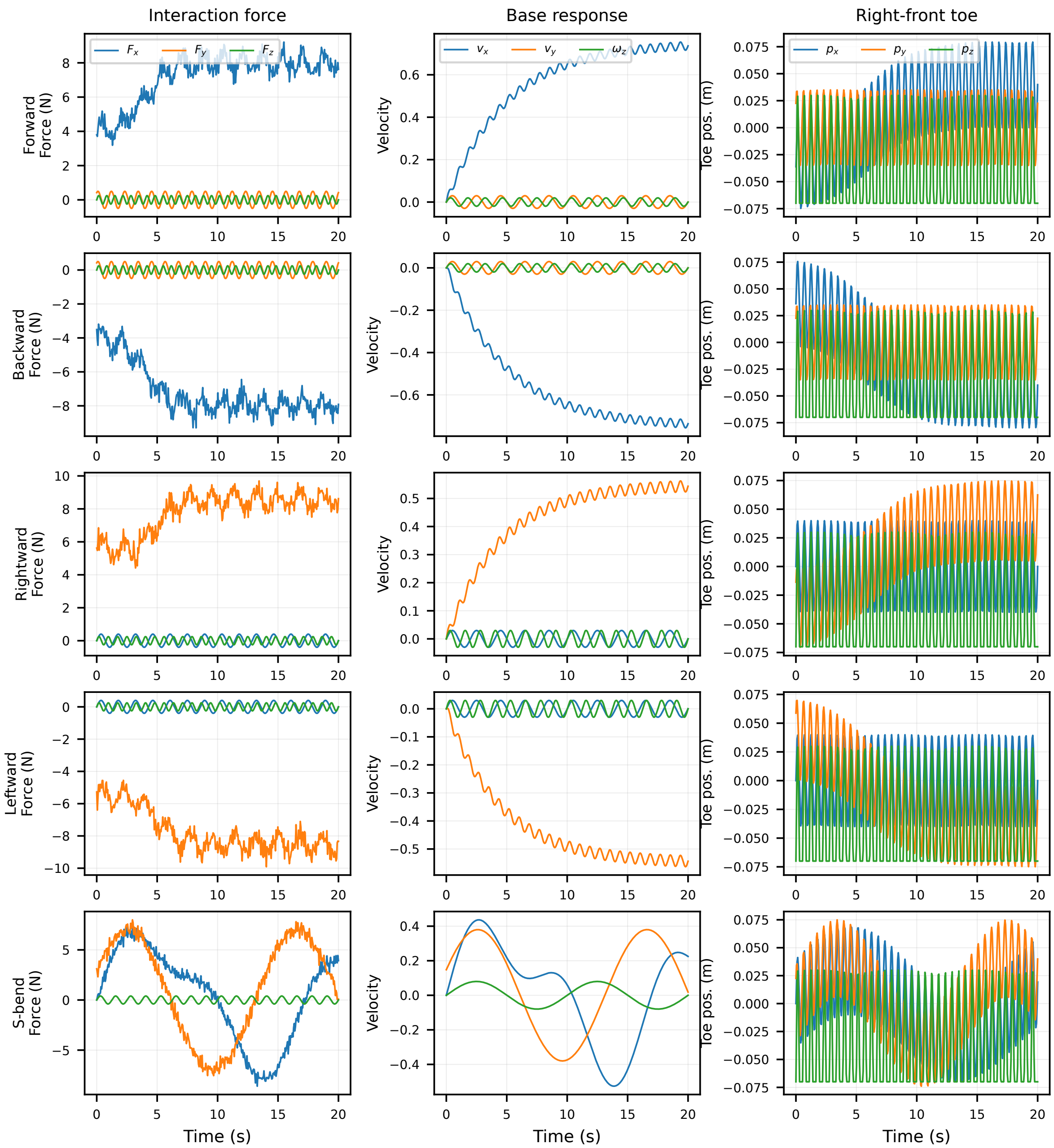

Fig. 2. Representative multiscenario guidance results. Rows correspond to forward, backward, rightward, leftward, and S-bend guidance. Columns show interaction force, base response, and right-front toe trajectory.

consistent. The right-front toe trajectories remain periodic and bounded, indicating that foothold adaptation maintains stable stepping under different commanded directions.

D. Ablation Study

Table III summarizes the ablation study. Adding motion cues to force guidance improves maneuver accuracy from 79.3% to 86.5%, showing that physical interaction alone is not sufficient for robust intent inference. Adding context further improves accuracy and success rate. Uncertainty adaptation mainly reduces peak force and increases success rate because the robot can reduce human authority when intent is ambiguous or terrain becomes risky. The full framework obtains the best slip rate and stability-related performance because the high-level shared command is physically realized through terrain-aware foothold optimization.

Fig. 6 visualizes the same ablation trend. The improvement from force-only guidance to force-plus-motion fusion mainly increases intent accuracy, while the addition of task context improves task success. The uncertainty-adaptive module reduces peak force by suppressing overreaction to ambiguous input, and the terrain-aware foothold module further lowers slip rate by improving physical feasibility during direction changes and rough-terrain traversal.

TABLE II. Repeated-Trial Performance Comparison Across Five Guidance Scenarios

Method	e_v (m/s)	e_ω (rad/s)	Acc_μ (%)	e_{rel} (m)	f_{avg} (N)	f_{peak} (N)	j_{cmd} (m/s^2)	m_{stab} (cm)	r_{slip} (%)	r_{succ} (%)
FOG	0.184 ± 0.041	0.217 ± 0.052	79.3 ± 4.7	0.198 ± 0.055	11.6 ± 2.5	34.8 ± 6.1	0.312 ± 0.071	3.8 ± 1.0	12.6	76.7
FASC	0.143 ± 0.035	0.181 ± 0.046	82.1 ± 4.1	0.163 ± 0.048	10.1 ± 2.1	28.9 ± 5.4	0.241 ± 0.060	4.6 ± 1.1	9.8	83.3
MSC-w/o-UA	0.103 ± 0.028	0.124 ± 0.035	90.8 ± 3.2	0.129 ± 0.037	8.4 ± 1.8	24.1 ± 4.2	0.186 ± 0.045	5.2 ± 1.0	7.1	90.0
UASC-w/o-TFA	0.094 ± 0.023	0.115 ± 0.030	91.5 ± 2.8	0.116 ± 0.034	8.0 ± 1.6	22.7 ± 3.9	0.179 ± 0.043	4.8 ± 0.9	8.4	86.7
Ours	$\mathbf{0.066 \pm 0.018}$	$\mathbf{0.081 \pm 0.023}$	$\mathbf{95.6 \pm 2.1}$	$\mathbf{0.086 \pm 0.025}$	$\mathbf{6.7 \pm 1.4}$	$\mathbf{17.9 \pm 3.1}$	$\mathbf{0.121 \pm 0.032}$	$\mathbf{6.7 \pm 1.2}$	**4.3**	**96.7**

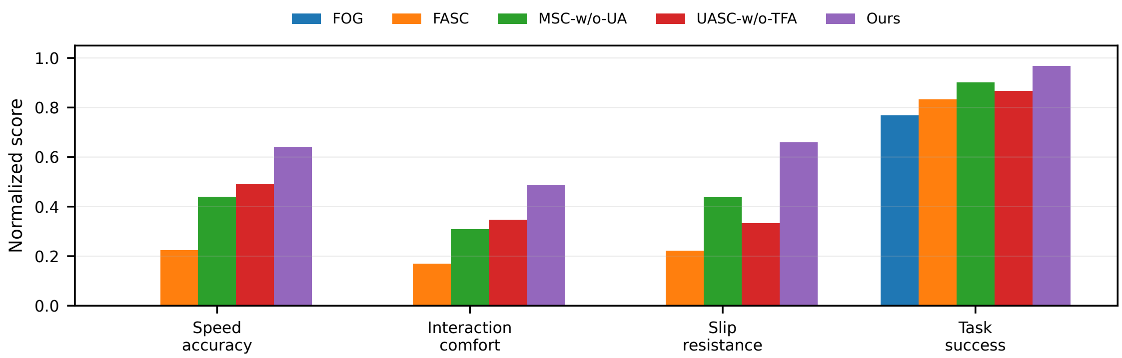

Fig. 3. Normalized comparison of key performance categories. Speed accuracy, interaction load, and slip resistance are computed from inverse-normalized desired-speed error, peak interaction force, and slip rate, respectively, while task success is the measured success rate scaled to $[0, 1]$. A larger score indicates better performance in all categories.

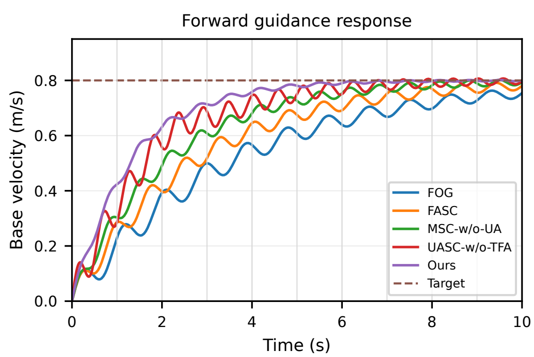

Fig. 4. Representative forward-guidance base-velocity response. Vertical grid lines correspond to gait-cycle intervals.

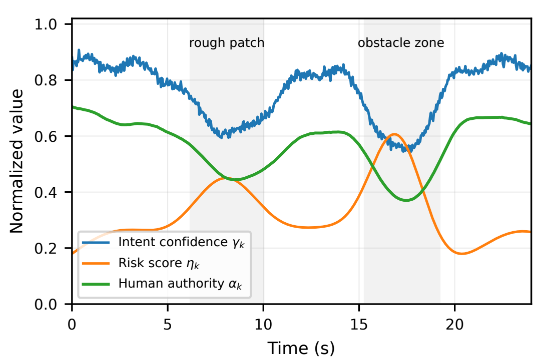

Fig. 5. Representative adaptation of intent confidence, terrain-risk score, and human authority factor. Higher risk or lower confidence reduces human authority and increases robot-side stabilization.

TABLE III. Ablation Study of the Main Components

Variant	Acc_μ (%)	Success (%)	Peak force (N)	Slip (%)
Force only	79.3	76.7	34.8	12.6
Force + motion	86.5	84.4	28.6	9.7
Force + motion + context	90.8	90.0	24.1	7.1
+ uncertainty adaptation	93.1	93.3	20.8	6.0
Full framework	**95.6**	**96.7**	**17.9**	**4.3**

E. Repeated-Trial Statistical Analysis

The distribution-level results in Fig. 7 further support the repeated-trial comparison. Across 150 trials, the proposed method has the lowest median desired-speed estimation error and peak interaction force, the largest stability margin, and the shortest completion time. The narrower interquartile range also indicates better repeatability under scenario changes. The distribution-level comparison complements the representative trajectories by showing that the performance advantage persists across guidance directions and terrain conditions.

F. Discussion

The results indicate that the proposed framework improves performance for three reasons. First, multimodal temporal fusion improves intent inference by combining force, relative motion, operator-motion cues, and context, rather than relying on a single noisy channel. Second, the confidence- and risk-aware authority factor prevents the robot from over-following uncertain human input in unsafe situations while preserving responsiveness when the intent is clear. Third, terrain-aware foothold optimization reduces the mismatch between high-level shared commands and low-level physical feasibility.

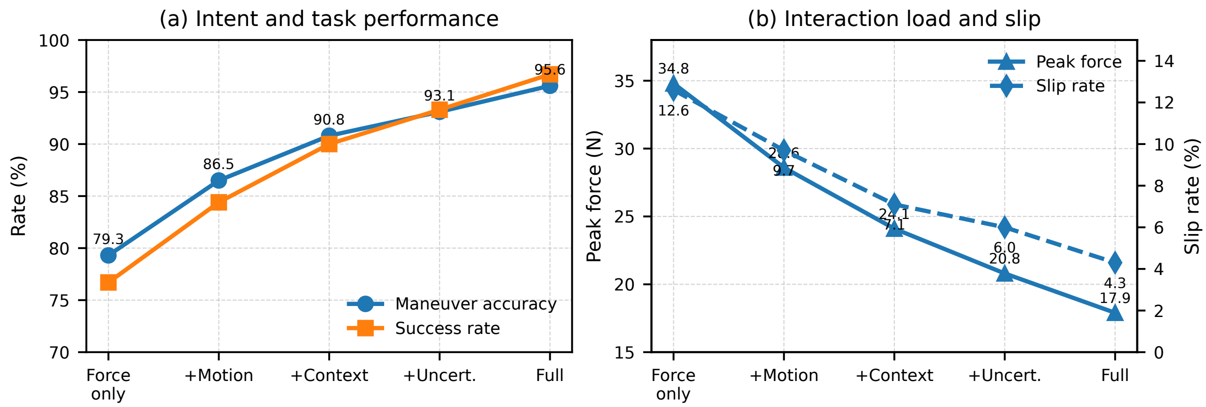

Fig. 6. Ablation trend of the proposed modules. Intent and task performance improve as motion cues, contextual information, uncertainty adaptation, and terrain-aware foothold optimization are added, while peak interaction force and slip rate decrease.

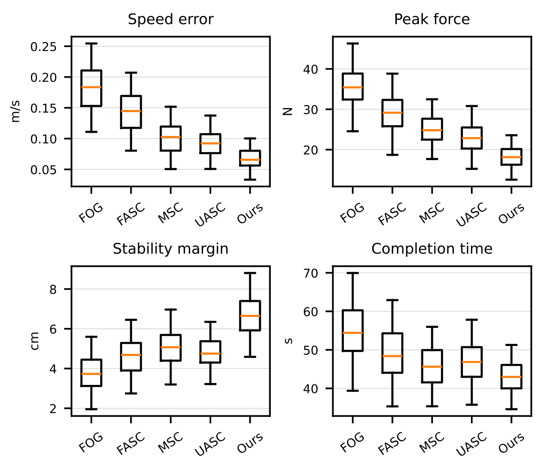

Fig. 7. Repeated-trial statistical comparison over 150 trials for each method. Box plots show median, interquartile range, and whisker range.

Therefore, the main benefit is not obtained from one isolated module, but from the integration of intent inference, authority adaptation, and terrain-aware locomotion realization.

G. Limitations and Generalization

The present study has several limitations. Hardware validation was conducted on a single Unitree Go2 platform, so the transferability of the sensing layout, controller gains, safety thresholds, and authority-allocation parameters to quadrupeds with different mass distributions, actuator limits, or contact dynamics remains to be established. The evaluation also considered five directional guidance scenarios—forward, backward, rightward, leftward, and S-bend guidance—and therefore does not cover long-horizon navigation, crowded environments, complex obstacle negotiation, or unconstrained guidance patterns that may arise in practical deployment. In addition, simulation was used for controller debugging and parameter pre-tuning before the hardware trials, which may introduce parameter choices specific to the simulation-to-hardware workflow. The standardized operator protocol was intended to isolate controller-level effects and does not characterize the full variability of real-world human guidance. Future studies will evaluate additional quadruped platforms, more diverse terrains and navigation tasks, broader interaction conditions, cross-platform parameter sensitivity, and personalized adaptation.

VIII. Conclusion

This study presented an uncertainty-adaptive shared-control framework with multimodal intent fusion for human-guided locomotion of a Unitree Go2 on uneven terrain. The method estimates human locomotion intent from interaction force, relative motion, operator-motion cues, and contextual information; computes confidence from posterior entropy and prediction variance; evaluates terrain and stability risk; and adjusts human authority online. The shared command is implemented by gait scheduling, terrain-aware foothold optimization, whole-body control, and interaction-compliant base regulation. Repeated-trial results across five guidance scenarios show improvements in intent accuracy, interaction load, slip resistance, stability margin, and task success rate compared with force-only, fixed-authority, and ablated shared-control methods. Validation to date is limited to one hardware platform, five representative guidance scenarios, and simulation-assisted controller pre-tuning. Future work will extend the framework to additional quadruped platforms, long-horizon navigation under broader interaction conditions, richer terrain perception, and personalized adaptation models.

Funding

This work was supported by the Dezhou Intelligent Equipment Research and Development Center.

References

[1] M. H. Raibert, "Legged robots," *Communications of the ACM*, vol. 29, no. 6, pp. 499–514, jun 1986.

[2] H. W. Park, P. M. Wensing, and S. Kim, "High-speed bounding with the mit cheetah 2: Control design and experiments," *The International Journal of Robotics Research*, vol. 36, no. 2, pp. 167–192, 2017.

[3] G. Bledt, M. J. Powell, B. Katz, J. D. Carlo, P. M. Wensing, and S. Kim, "Mit cheetah 3: Design and control of a robust, dynamic quadruped robot," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 2245–2252.

[4] Z. Li, Z. Li, L. M. Tam, and Q. Xu, “Design and development of a versatile quadruped climbing robot with obstacle-overcoming and manipulation capabilities,” *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 3, pp. 1649–1661, jun 2023.

[5] B. Jin, S. Ye, J. Su, and J. Luo, “Unknown payload adaptive control for quadruped locomotion with proprioceptive linear legs,” *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 4, pp. 1891–1899, aug 2022.

[6] O. Villarreal, V. Barasuol, P. M. Wensing, D. G. Caldwell, and C. Semini, “Mpc-based controller with terrain insight for dynamic legged locomotion,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 2436–2442.

[7] Y. Ding, A. Pandala, C. Li, Y.-H. Shin, and H.-W. Park, “Representation-free model predictive control for dynamic motions in quadrupeds,” *IEEE Transactions on Robotics*, vol. 37, no. 4, pp. 1154–1171, aug 2021.

[8] J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, “Learning quadrupedal locomotion over challenging terrain,” *Science Robotics*, vol. 5, no. 47, p. eabc5986, 2020.

[9] J. Hwangbo *et al.*, “Learning agile and dynamic motor skills for legged robots,” *Science Robotics*, vol. 4, no. 26, p. eaau5872, 2019.

[10] T. Miki, J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, “Learning robust perceptive locomotion for quadrupedal robots in the wild,” *Science Robotics*, vol. 7, no. 62, p. eabk2822, 2022.

[11] I. M. A. Nahrendra, B. Yu, and H. Myung, “Dreamwaq: Learning robust quadrupedal locomotion with implicit terrain imagination via deep reinforcement learning,” in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 5078–5084.

[12] L. Wang *et al.*, “Design and dynamic locomotion control of quadruped robot with perception-less terrain adaptation,” *Cyborg and Bionic Systems*, vol. 2022, p. 9816495, 2022.

[13] K. A. Hamed, V. R. Kamidi, W.-L. Ma, A. Leonessa, and A. D. Ames, “Hierarchical and safe motion control for cooperative locomotion of robotic guide dogs and humans: A hybrid systems approach,” *IEEE Robotics and Automation Letters*, vol. 5, no. 1, pp. 56–63, jan 2020.

[14] A. Xiao, W. Tong, L. Yang, J. Zeng, Z. Li, and K. Sreenath, “Robotic guide dog: Leading a human with leash-guided hybrid physical interaction,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 11 470–11 476.

[15] Y. Chen *et al.*, “Quadruped guidance robot for the visually impaired: A comfort-based approach,” in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 12 078–12 084.

[16] H. Hwang, H.-T. Jung, N. A. Giudice, J. Biswas, S. I. Lee, and D. Kim, “Towards robotic companions: Understanding handler-guide dog interactions for informed guide dog robot design,” in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 2024, cHI 2024.

[17] C. R. de Cos and D. V. Dimarogonas, “Adaptive cooperative control for human-robot load manipulation,” *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 5623–5630, apr 2022.

[18] C. Yang *et al.*, “Collaborative navigation and manipulation of a cable-towed load by multiple quadrupedal robots,” *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10 041–10 048, aug 2022.

[19] S. Gu, F. Meng, B. Liu, X. Chen, Z. Yu, and Q. Huang, “Adaptive interactive control of human and quadruped robot load motion,” *IEEE/ASME Transactions on Mechatronics*, 2024, to be published.

[20] R. Kamikubo *et al.*, “Designing shared control for navigation robots with blind and low vision people,” in *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction*, 2025.

[21] K. Darvish, E. Simetti, F. Mastrogiovanni, and G. Casalino, “A hierarchical architecture for human–robot cooperation processes,” *IEEE Transactions on Robotics*, vol. 37, no. 2, pp. 567–586, apr 2021.

[22] P. Culbertson, J.-J. Slotine, and M. Schwager, “Decentralized adaptive control for collaborative manipulation of rigid bodies,” *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1906–1920, jun 2021.

[23] M. L. Elwin, B. Strong, R. A. Freeman, and K. M. Lynch, “Human-multirobot collaborative mobile manipulation: The omnid mocobots,” *IEEE Robotics and Automation Letters*, vol. 8, no. 1, pp. 376–383, feb 2023.

[24] K. A. Hamed, V. R. Kamidi, A. Pandala, W.-L. Ma, and A. D. Ames, “Distributed feedback controllers for stable cooperative locomotion of quadrupedal robots: A virtual constraint approach,” in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 5314–5321.

[25] J. Kim, R. T. Fawcett, V. R. Kamidi, A. D. Ames, and K. A. Hamed, “Layered control for cooperative locomotion of two quadrupedal robots: Centralized and distributed approaches,” *IEEE Transactions on Robotics*, vol. 39, no. 6, pp. 4728–4748, dec 2023.

[26] A. D. Ames, K. Galloway, K. Sreenath, and J. W. Grizzle, “Rapidly exponentially stabilizing control lyapunov functions and hybrid zero dynamics,” *IEEE Transactions on Automatic Control*, vol. 59, no. 4, pp. 876–891, aug 2014.

[27] A. Hereid, C. M. Hubicki, E. A. Cousineau, and A. D. Ames, “Dynamic humanoid locomotion: A scalable formulation for hzd gait optimization,” *IEEE Transactions on Robotics*, vol. 34, no. 2, pp. 370–387, apr 2018.

[28] J. Englsberger, C. Ott, and A. Albu-Sch”affer, “Three-dimensional bipedal walking control based on divergent component of motion,” *IEEE Transactions on Robotics*, vol. 31, no. 2, pp. 355–368, apr 2015.

[29] E. Daneshmand, M. Khadiv, F. Grimminger, and L. Righetti, “Variable horizon mpc with swing foot dynamics for bipedal walking control,” *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2349–2356, apr 2021.

[30] S. Xin, R. Orsolino, and N. Tsagarakis, “Online relative footstep optimization for legged robots dynamic walking using discrete-time model predictive control,” in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 513–520.

[31] G. Xin *et al.*, “Robust footstep planning and lqr control for dynamic quadrupedal locomotion,” *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4488–4495, mar 2021.