

Collision-Free Humanoid Traversal in Cluttered Indoor Scenes



Fig. 1: Using a single generalist policy, our humanoid robot achieves collision-free traversal in cluttered indoor environments, including (a) detouring through narrow passages, (b) crouching under low-hanging obstacles, (c) squeezing through tight indoor spaces, and (d) hurdling over objects scattered on the floor.

Abstract—We study the problem of collision-free humanoid traversal in cluttered indoor scenes, such as hurdling over objects scattered on the floor, crouching under low-hanging obstacles, or squeezing through narrow passages. To achieve this goal, the humanoid needs to map its perception of surrounding obstacles to the corresponding traversal skills. However, due to the reward engineering bottleneck, perception–control gap and sim-to-real transfer challenge, directly learning such mappings is highly challenging. Therefore, we introduce Humanoid Potential Field (HumanoidPF), a unified representation that tightly bridges environmental perception with whole-body control. It induces dense, structured guidance to streamline reward engineering and provides compact, task-relevant and sim-to-real-robust perceptual observations. To enable generalizable traversal skills through diverse and highly challenging cluttered indoor scenes, we further propose a hybrid scene generation method, incorporating crops of realistic 3D indoor scenes and procedurally synthesized obstacles. We successfully transfer our policy to the real world and develop a teleoperation system that allows a user to command the humanoid to traverse in cluttered indoor scenes with just a single click. Extensive experiments are conducted in both simulation and the real world to validate the effectiveness of our method.

I. INTRODUCTION

Imagine that a domestic humanoid robot needs to frequently traverse between the bedroom, living room, and kitchen to perform household chores. A key challenge for the robot is to avoid collisions with the surrounding environment during movement, preventing potential damage to the robot itself or the environment. In cluttered indoor scenes, the

humanoid may need to hurdle over objects scattered on the floor, crouch under low-hanging obstacles, or squeeze through narrow passages. This requires the robot to perceive the environment and map obstacles with diverse spatial layouts and geometries to the corresponding traversal skills.

While legged locomotion in complex environments has seen remarkable advances for quadrupeds [1]–[13] and humanoids [14]–[23], existing works are often limited to obstacles with partial-space layouts and simple geometric types, as shown in Table I. When extending existing approaches to full-space traversal in cluttered indoor scenes, two fundamental challenges arise. (i) The reward engineering bottleneck: reinforcement learning methods typically rely on sparse collision penalties or manually designed constraints, which often provide insufficient guidance in challenging obstacle layouts and intricate geometries, leading to either overly conservative behaviors or unintended collisions. (ii) The perception–control gap and sim-to-real transfer challenge: high-dimensional visual representations such as voxel grids are computationally expensive for RL policies to ingest directly, and their domain gap between simulation and reality is profound due to extreme sensitivity to sensor noise.

To bridge these gaps, we introduce Humanoid Potential Field (**HumanoidPF**), a unified representation that tightly couples environmental perception with whole-body control. Inspired by classical Artificial Potential Fields [24] (APF), HumanoidPF encodes the environment as a continuous, dif-

ferentiable gradient field that induces “virtual forces” over the humanoid’s kinematic structure.

This representation provides three key advantages for perception and control. (i) Reward design streamlining: instead of relying on sparse collision penalties or manually tuned constraints, the potential field naturally induces a dense and structured reward that guides whole-body motion along collision-free trajectories, eliminating tedious tuning of collision-related reward terms. (ii) Reduction of the perception–control gap: humanoid–obstacle spatial relationships are encoded into a compact, task-relevant state representation, making it easier for policies to ingest and reason over traversal-relevant information. (iii) Improved sim-to-real robustness: unlike discrete voxel grids that are brittle to sensor noise, the continuous nature of potential fields acts as a low-pass perceptual filter, smoothing out perception artifacts and promoting consistent policy behavior in real-world environments.

To enable generalizable traversal skills across diverse and challenging cluttered indoor environments, we further propose a hybrid scene generation strategy. By augmenting crops of realistic 3D indoor datasets with procedurally synthesized “extreme” obstacles, we expose the robot to a curriculum of challenging clutter configurations that are rarely present in existing datasets, enabling it to acquire rich collision-avoidance experience and substantially improving robustness in near-collision and emergency scenarios.

In summary, we study the problem of collision-free humanoid traversal in cluttered indoor environments. Building upon our method, we further develop a teleoperation system, termed Click-and-Traverse (CAT), where user can simply click a goal to command the humanoid to safely traverse cluttered indoor environments. Extensive experiments in both simulation and realistic real-world indoor scenes validate the practical applicability of HumanoidPF and its strong generalization across diverse environments.

Our contributions are fourfold:

- We study the problem of collision-free humanoid traversal in cluttered indoor scenes, taking an important step toward applying humanoid robots in real-world home environments.
- We propose **HumanoidPF**, a novel representation that unifies perception and control. It streamlines RL reward design via automated collision-avoidance guidance, reduces the perception–control gap and enables robust sim-to-real transfer.
- We propose a hybrid scene generation strategy that exposes the policy to realistic, diverse and challenging cluttered scenarios, significantly improving robustness and generalization in complex indoor environments.
- We successfully transfer our policy to the real world and develop a convenient and useful teleoperation system to command the humanoid to traverse in cluttered indoor scenes.

Method	Spatial layout	Intricate geometry
PIM [18]	$S = \{g\}$	✗
HumanoidParkour [15]	$S = \{g\}$	✗
BeamDojo [21]	$S = \{g\}$	✗
Vb-com [19]	$S \subset \{g, l\}, S = 1$	✗
Gallant [23]	$S \subset \{g, l, o\}, S = 1$	✗
Ours	$S = \{g, l, o\}$	✓

TABLE I: Overall comparison with existing works. S : Spatial layout. g, l, o : Ground, lateral and overhead obstacles.

II. RELATED WORKS

A. Legged locomotion in complex environments

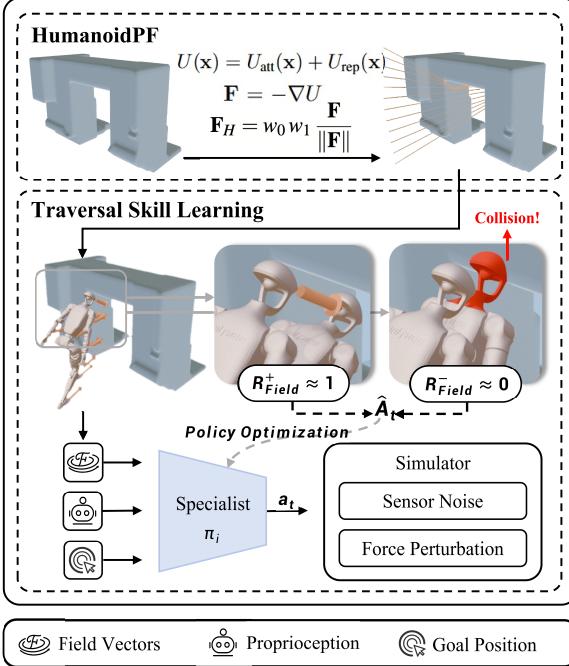
Legged robots are expected to perform stable locomotion in complex environments, including challenging terrains and obstacles. Quadruped robots have demonstrated robust parkour capabilities on highly challenging terrains [1]–[7] and confined or cluttered spaces [8]–[13]. Humanoids have also demonstrated the ability to navigate in height-constrained environments [14] and advanced locomotion skills against risky terrains or obstacles, such as stepping stairs, balance beams, and stepping stones [15]–[23].

However, existing works on humanoids often limited to obstacles with partial spatial layouts (e.g. terrains [15]–[22], or over-hanging obstacles [14]) and simple geometries (e.g. rectangular blocks [14]–[17], [20]–[22], or regular polyhedra [19], [23]). Notably, while Gallant [23] addresses ground, lateral, and overhead obstacle layouts in isolation, it does not consider scenarios in which these constraints coexist, and is limited to relatively simple obstacle geometries. In contrast, our method constructs HumanoidPF to operate in environments where ground, lateral, and overhead constraints are jointly present, while supporting highly intricate obstacle geometries. The comparison of existing works and our work is shown in Table I.

B. Artificial potential field for obstacle avoidance

Originally introduced in the late 1980s, the Artificial Potential Field [24] (APF) method generates a virtual force field to guide the motion of manipulators or mobile robots for obstacle avoidance. Inspired by physical analogies, the goal position is modeled as an attractive pole, while obstacles act as repulsive surfaces. Traditionally, APF has been widely used in 2D path planning for mobile robots [25]–[27] and robotic manipulators [28]–[30]. Only a few studies have combined model-based quadruped control with APF in limited forms. Some apply the APF by abstracting the quadruped as a single rigid body, guiding only the center of mass [31], [32]. Others extend it to a few selected joints to influence gait and posture [33], [34]. Nevertheless, these methods remain restricted to quadrupeds and fail to achieve coordinated whole-body obstacle avoidance, limiting their applicability to locomotion over simple terrains. In contrast, we propose HumanoidPF that serves as a soft motion prior, guiding whole-body, collision-free motion in complex 3D environments.

Training Specialist with HumanoidPF



Generalizing to Cluttered Indoor Scenes

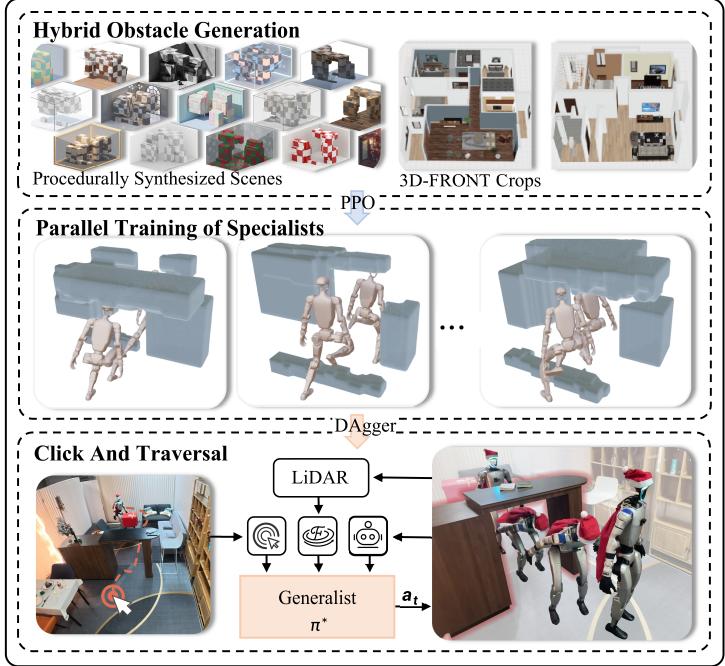


Fig. 2: **Overall pipeline.** We learn a visuomotor policy that maps diverse obstacle geometries and spatial layouts to corresponding traversal skills. **Left:** training of specialist policies using the proposed HumanoidPF. (*Top*) Construction of the HumanoidPF for humanoid whole-body control; (*Bottom*) its use in guiding humanoid traversal skill learning through field-based rewards and observations. **Right:** generalization and deployment pipeline. (*Top*) Hybrid obstacle generation for constructing diverse training environments; (*Middle*) parallel training of multiple specialists across varied obstacle configurations and difficulties; (*Bottom*) the final Click-and-Traverse deployment, where the learned generalist policy enables intuitive loco-navigation teleoperation through real-time location and mapping. Sections III-A and III-B provide detailed descriptions of the HumanoidPF for skill learning and generalization pipeline, respectively.

III. METHOD

We study the problem of collision-free humanoid traversal in cluttered indoor scenes. Given a start position $\mathbf{s} \in \mathbb{R}^2$, a target position $\mathbf{g} \in \mathbb{R}^2$, and a set of indoor obstacles $\mathcal{O} = \{\mathcal{O}_i\}_{i=1}^N$, the humanoid needs to move from \mathbf{s} to \mathbf{g} without any collision with \mathcal{O} . To solve this problem, the humanoid needs to map its perception of surrounding obstacles to the corresponding traversal skills. Our method can be split into two parts. We will first introduce how to construct our **HumanoidPF** and utilize it to guide humanoid traversal skill learning in a single cluttered scene in Section III-A. We will further introduce how to generalize our policy to diverse and challenging indoor scenes with our proposed hybrid scene generation method in Section III-B. The overall pipeline of our method is shown in Figure 2.

A. HumanoidPF

To construct HumanoidPF, we first model the given target location \mathbf{g} as an attractive pole and obstacles \mathcal{O} as repulsive surfaces, creating the attractive field U_{att} and repulsive field U_{rep} respectively. These fields are blended to form the APP U , whose gradient $\mathbf{F} = -\nabla U$ generates a collision-free path toward the goal. But it's difficult to let every body part \mathbf{p}_k at Cartesian position \mathbf{x}_k to strictly follow its corresponding

field guidance $\mathbf{F}(\mathbf{x}_k)$. To this end, we calculate priority for different body part and weight the field vector to get final HumanoidPF \mathbf{F}_H . HumanoidPF \mathbf{F}_H then plays crucial role as policy observation and reward to learn traversal skills.

1) HumanoidPF Construction

We begin by constructing the attractive field U_{att} :

$$U_{\text{att}}(\mathbf{x}) = \eta \|\mathbf{x} - \mathbf{g}\|_{\text{geo}}, \quad (1)$$

where the geodesic distance $\|\mathbf{x} - \mathbf{g}\|_{\text{geo}}$ represents the shortest 3D path from position \mathbf{x} to the goal \mathbf{g} without intersecting obstacles, and η is a scaling factor. This formulation inherently accounts for obstacle geometry, providing safer guidance than a simple Euclidean distance.

Next, the repulsive field U_{rep} prevents collisions and is defined as:

$$U_{\text{rep}}(\mathbf{x}) = \begin{cases} \frac{1}{2} \xi \left(\frac{1}{d(\mathbf{x})} - \frac{1}{d_0} \right)^2, & d(\mathbf{x}) \leq d_0, \\ 0, & d(\mathbf{x}) > d_0, \end{cases} \quad (2)$$

where $d(\mathbf{x})$ is the signed distance, ξ is a gain coefficient, and d_0 defines the influence range of obstacles.

The final guidance field is the negative gradient of a combined potential,

$$\mathbf{F} = -\nabla U, \quad U(\mathbf{x}) = U_{\text{att}}(\mathbf{x}) + U_{\text{rep}}(\mathbf{x}), \quad (3)$$

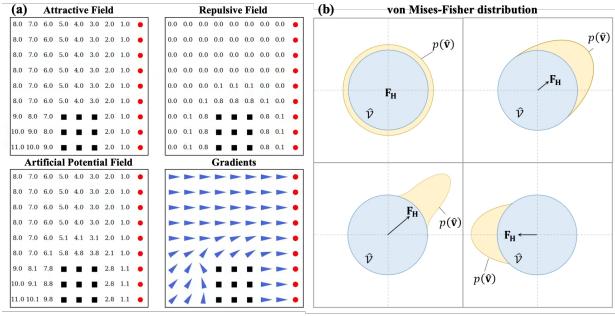


Fig. 3: (a) Construction of the APF and (b) motion prior induced by the HumanoidPF.

which is then queried at the locations of different body parts, yielding field vectors $\mathbf{F}(\mathbf{x}_k)$ for each body part \mathbf{p}_k . A 2D visualization of our APF is illustrated in Figure. 3 (a).

While APF method typically models robots as a single rigid body, its direct application to multi-jointed humanoid robots could lead to conflicts between body parts. For instance, when the robot faces an obstacle directly ahead, it must choose whether to avoid it by moving left or right. At this point, the potential fields on the left side of the body direct the robot toward the left path, while the fields on the right side suggest the opposite direction, creating a dilemma. To address this challenge, we propose a priority-weighting scheme that prioritizes the influence of certain body parts over others according to their contribution to the task.

Priority-weighting. Instead of treating all body parts equally, our priority-weighting scheme adjusts the influence of each body part based on its role in the overall motion.

To establish coherent global guidance, we assign a higher priority to the root body part (e.g., pelvis) since it plays a central role in maintaining stability and direction:

$$w_0(\mathbf{p}_{\text{root}}) = 1, \quad w_0(\mathbf{p}_{\text{others}}) = 0.5. \quad (4)$$

Furthermore, some body parts are more critical in avoiding obstacles, particularly those closer to potential collisions. To account for this, we define a dynamic collision-urgency weight based on the signed distance $d(\mathbf{x}_k)$ and the Cartesian velocity \mathbf{v}_k of body part \mathbf{p}_k , with a scaling factor λ :

$$w_1(\mathbf{p}_k) = \lambda \max(-\nabla d(\mathbf{x}_k) \cdot \mathbf{v}_k, 0.5) \exp(-d(\mathbf{x}_k)). \quad (5)$$

The resulting HumanoidPF is defined as $\mathbf{F}_H = w_0 w_1 \frac{\mathbf{F}}{\|\mathbf{F}\|}$. Body parts with higher priority receive a stronger potential field, thereby enforcing stricter adherence to the guidance, while lower-priority parts experience weaker fields, allowing for greater flexibility in motion.

2) Traversal skill learning with HumanoidPF

HumanoidPF for policy observation. To reduce the perception-control gap and enhance sim-to-real robustness, we leverage HumanoidPF to construct a compact, task-relevant visual observation. It's sampled at $K = 13$ body parts,

$$OBS_{Field} = \{\mathbf{F}_H(\mathbf{x}_k) \mid \mathbf{x}_k\}_{k=1}^K, \quad (6)$$

to encode humanoid–obstacle relationships and traversal-guiding information via forward-kinematics and self-localization. In contrast, most existing representations (e.g.,

voxel maps) encode the environment independently of the humanoid's morphology, hindering the policy from relating high-dimensional raw visual observations to traversal skills.

HumanoidPF for observation further facilitates sim-to-real transfer. Voxel-based representations are often highly sensitive to sensor noise—such as missing LiDAR points or floating artifacts—leading to notable sim-to-real gaps. By contrast, our HumanoidPF contextualizes the influence of surrounding geometry into a continuous spatial field, akin to a low-pass perceptual filter that smooths out isolated local noise. As a result, local geometric perturbations has diminished impact on the resulting control signals, yielding more stable real-world performance. We empirically validate this property with experiments in Section IV-C).

HumanoidPF for policy reward. To streamline reward engineering, we employ HumanoidPF to induce dense, structured, and automated guidance that generalizes across diverse environments. At each time step, HumanoidPF encodes a distribution of preferred motion directions, and the policy is optimized to produce actions that align with this distribution, thus fostering safe and efficient collision-avoidance behaviors.

The von Mises–Fisher (vMF) distribution is used to model directional preferences $\mu(\mathbf{x})$ on the unit sphere and allow the strength of this preference to be controlled by a single concentration parameter $\kappa(\mathbf{x})$:

$$p(\hat{v} \mid \mu(\mathbf{x}), \kappa(\mathbf{x})) = C_d(\kappa) \exp(\kappa(\mathbf{x}) \mu(\mathbf{x})^\top \cdot \hat{v}), \quad (7)$$

where \hat{v} is the motion direction of a humanoid body part, $C_d(\kappa)$ is the normalization constant.

$\mu(\mathbf{x})$ and $\kappa(\mathbf{x})$ is directly derived from HumanoidPF:

$$\mu(\mathbf{x}) = \frac{\mathbf{F}_H(\mathbf{x})}{\|\mathbf{F}_H(\mathbf{x})\|}, \quad \kappa(\mathbf{x}) = \kappa_{\max} \|\mathbf{F}_H(\mathbf{x})\|. \quad (8)$$

Body parts with higher priority receive a field vector $\mathbf{F}_H(\mathbf{x})$ with larger magnitude; accordingly, the concentration parameter $\kappa(\mathbf{x})$ increases to encourage these parts to follow the guidance direction $\mu(\mathbf{x})$ more strictly, and vice versa for lower-priority parts. This priority-aware concentration design promotes coordinated whole-body motion while improving collision-avoidance behavior, as illustrated by a 2D visualization in Figure 3 (b).

During policy training, let the motion direction of the k -th body part be $\hat{v}_k = \mathbf{v}_k / \|\mathbf{v}_k\|$, the associated prior direction be $\mu_k = \mu(\mathbf{x}_k)$ and concentration parameter be $\kappa_k = \kappa(\mathbf{x}_k)$. Assuming independence among joints, the whole-body motion prior and log-likelihood reward are expressed as:

$$p(\hat{v}_{1:K} \mid \mathbf{x}_{1:K}, \mathbf{F}_H) = \prod_{k=1}^K C_d(\kappa_k) \exp(\kappa_k \mu_k^\top \cdot \hat{v}_k), \quad (9)$$

$$R_{Field} = \sum_{k=1}^K \left[\log C_d(\kappa_k) + \kappa_k \mu_k^\top \cdot \hat{v}_k \right]. \quad (10)$$

This reward formulation exhibits strong cross-scene generalization without requiring manual tuning, thereby enabling an automated training pipeline that scales effectively across diverse environments.

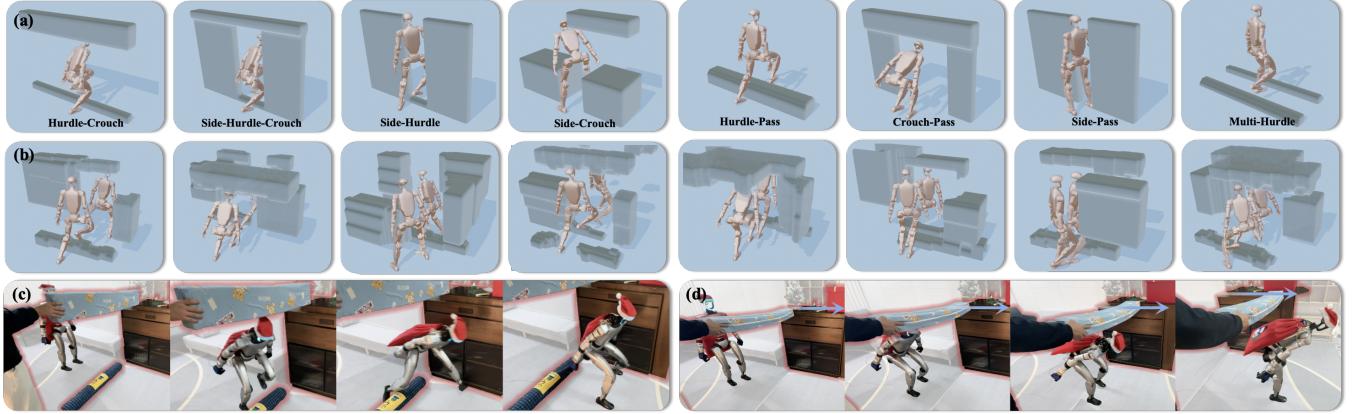


Fig. 4: Collision-free humanoid traversal in both simulation and the real world. (a) Humanoid traversal behaviors on eight representative test scene types; (b) traversal behaviors in procedurally generated cluttered environments; (c) real-world “hurdle-crouch” scenario, validating sim-to-real transfer in a cluttered indoor setting; (d) robustness under dynamic disturbances, where simple object movements (blue arrows) are introduced during traversal.

B. Generalizing to more scenes

For general practical use, the humanoid need to handle diverse scenes within a single unified policy. It need the policy trained with a sufficiently large and challenging indoor scene dataset to enable generalization and robustness. Therefore, we propose a hybrid scene generation method in Section III-B.1. It incorporates crops of realistic 3D indoor scenes for structural realism, and procedurally synthesized obstacles to enrich highly challenging clutter configurations that are rare in existing indoor datasets. In addition, we found that even with the assistance of our proposed **HumanoidPF**, directly learning a single policy that can handle all scenes remains very challenging due to the low sample efficiency of RL. Thus, we adopt a specialist-to-generalist strategy inspired by DAgger [35], which we will introduce in Section III-B.2.

1) Hybrid scene generation

We observe that highly challenging obstacle layouts constitute merely a long-tail subset in most existing datasets [36]–[38], since typical indoor scenes featuring orderly object arrangements and clearly separated walkable regions. Simply scaling up the dataset does not alleviate this problem. Therefore, we propose a hybrid scene generation scheme that augments realistic 3D indoor datasets with procedurally synthesized “extreme” obstacles.

Crops of realistic 3D indoor scenes. For generalization to realistic indoor environments, we adopt the 3D-FRONT [36] dataset, containing structurally realistic scenes with large-scale high-quality furniture objects. We selectively crop and filter scene blocks for specialist and generalist training.

Specifically, we first project all furniture onto the ground and erode the resulting planar walkable regions with a radius of 0.1 m to account for clearance. Within the remaining walkable regions, we randomly sample a start location and crop a 5 m × 5 m block centered at this position. During training, goal location will be randomly sampled on a circle with a radius of 2 m around the start.

We initially train specialist policies on all such cropped

scenes and subsequently identify scenes with low traversal success rates. Scenes that are empirically found to be non-traversable are manually filtered out.

Procedurally generated obstacles. To supplement crops of 3D-FRONT with more challenging and cluttered environments, we procedurally generate obstacles that impose simultaneous ground, lateral, and overhead constraints, deliberately targeting highly restrictive scenarios. Specifically, we place boxes with varying positions, dimensions, and orientations that may extend upward from the floor, descend from the ceiling, and be placed in close proximity to form narrow traversable passages.

To break structural regularity and enhance geometric realism, we apply random SO(3) rotations and 2D Perlin noise to each box. The resulting artifacts, such as spiky surfaces or non-manifold regions, are mitigated via 3D morphological closing and opening at the voxel level before mesh conversion. Visualizations of the robot traversing generated obstacles are shown in Figure 2 and Figure 4 (b).

To support curriculum learning during policy training, we use a layout-agnostic difficulty factor to control obstacle complexity, such as the number and size of boxes. As the difficulty increases, the policy progressively acquires robust traversal skills under increasingly challenging configurations.

2) Specialists to Generalist

We construct an automated and parallel specialist-to-generalist training pipeline. We first conduct large-scale training of specialist policies across diverse scenes via PPO [39]. The reward derived from HumanoidPF is scene-general, enabling scalable training on 139 cropped 3D-FRONT scenes and 216 procedurally generated scenes. Each specialist is trained with 32,768 parallel environments and 5,000,000 episodes, with start and goal locations randomly sampled for each episode.

Subsequently, a generalist policy is distilled using DAgger [35] with varying obstacle configurations. During this process, dedicated specialist policies provide expert actions

	Hurdle-Crouch		Side-Hurdle-Crouch		Side-Hurdle		Side-Crouch	
	SR(%) \uparrow	DE(m) \downarrow						
Humanoid Parkour	33.3 \pm 6.1	1.16 \pm 0.63	0.4 \pm 0.3	1.49 \pm 0.04	45.1 \pm 4.5	0.62 \pm 0.39	64.4 \pm 19.3	0.56 \pm 0.16
ASTraverse	28.1 \pm 10.4	1.11 \pm 0.78	0.5 \pm 0.5	1.06 \pm 0.39	37.1 \pm 3.1	0.54 \pm 0.32	56.0 \pm 9.9	0.48 \pm 0.05
Ours	93.9 \pm 2.7	0.08 \pm 0.16	86.6 \pm 5.2	0.2 \pm 0.32	95.4 \pm 3.9	0.06 \pm 0.34	96.9 \pm 2.1	0.05 \pm 0.09
Ours w/o <i>OBSField</i>	77.8 \pm 5.4	0.33 \pm 0.23	53.7 \pm 9.9	0.59 \pm 0.08	60.4 \pm 9.6	0.53 \pm 0.68	90.1 \pm 5.3	0.19 \pm 0.35
Ours w/o <i>RField</i>	21.9 \pm 15.8	1.27 \pm 0.71	0.0 \pm 0.0	1.57 \pm 0.003	71.4 \pm 9.9	0.5 \pm 0.34	80.3 \pm 15.4	0.23 \pm 0.06
<hr/>								
	Hurdle-Pass		Crouch-Pass		Side-Pass		Multi-Hurdle	
	SR(%) \uparrow	DE(m) \downarrow						
Humanoid Parkour	84.3 \pm 8.0	0.32 \pm 0.18	48.1 \pm 3.3	1.34 \pm 0.18	41.3 \pm 2.5	0.91 \pm 0.34	88.7 \pm 2.6	0.23 \pm 0.35
ASTraverse	75.9 \pm 6.8	0.66 \pm 0.3	41.3 \pm 5.3	0.9 \pm 1.04	55.2 \pm 8.5	0.78 \pm 0.87	82.1 \pm 8.7	0.26 \pm 0.43
Ours	96.9 \pm 5.5	0.06 \pm 0.15	97.5 \pm 4.3	0.05 \pm 0.09	97.3 \pm 3.2	0.04 \pm 0.15	95.0 \pm 4.9	0.06 \pm 0.1
Ours w/o <i>OBSField</i>	92.3 \pm 4.9	0.1 \pm 0.15	96.8 \pm 5.0	0.07 \pm 0.3	95.2 \pm 4.4	0.07 \pm 0.22	90.5 \pm 3.5	0.09 \pm 0.11
Ours w/o <i>RField</i>	90.7 \pm 7.5	0.12 \pm 0.37	95.9 \pm 5.0	0.08 \pm 0.24	28.0 \pm 18.3	1.07 \pm 0.56	88.3 \pm 14.1	0.23 \pm 0.5

TABLE II: **Validation of HumanoidPF for skill learning.** To better characterize the performance of our method under diverse obstacle layouts, we design 8 distinct scene types for evaluation and ablation studies.

conditioned on current obstacle, enabling the generalist to acquire strong generalization capability across diverse scenarios.

To learn robust and stable traversal skills, both specialist and generalist policies are trained with sensor noise and force perturbations to simulate realistic collision-avoidance conditions. In addition, a curriculum with progressively increasing scene difficulty is employed to gradually enhance traversal performance and explore the limits of the policy’s obstacle-avoidance capability.

IV. EXPERIMENT

In this section, we provide extensive experimental results in both the MuJoCo [40] simulator and the real-world deployment. We use the 29-DoF Unitree G1 humanoid robot in all our experiments. The experiments aim to address the following three questions:

- **Q1:** Can **HumanoidPF** improve the performance of traversal in cluttered indoor scenes compared to existing methods?
- **Q2:** Can our hybrid scene generation method help the policy generalize to unseen and challenging scenes?
- **Q3:** Can the HumanoidPF for observation *OBSField* help the sim-to-real transfer?

A. Validation of HumanoidPF for skill learning

To address **Q1** (*Can HumanoidPF improve the performance of traversal in cluttered indoor scenes compared to existing methods?*), we compare the performance of our method against existing ones on traversing cluttered scenes.

Experiment Setting. To systematically analyze performance under different obstacle layouts and geometric configurations, we design eight distinct types of cluttered scenes for evaluation. All scenes used in this experiment are manually generated, each type exhibiting distinct characteristics and collectively covering a broad range of challenging obstacle configurations, with 10 scenes per type. We train and evaluate all methods on these generated scenes to compare their

ability to traverse cluttered environments. Representative visualizations of the robot traversing each scene type are shown in Figure 4 (a).

Experiment Metrics. We use the following metrics to evaluate the performance on traversing cluttered scenes:

- **Success Rate (SR, %)** records the percentage of successful traversal trials. A trial is considered successful if the robot reaches within 0.1 meters of the target location in 5 seconds without colliding with any obstacle.
- **Distance Error (DE, m)** is the averaged closest horizontal distance between the humanoid root and the target location in a traversal.

Baselines. We choose the following methods as baselines and adapt them to fit our problem setting for a fair comparison:

- **Humanoid Parkour** [15]: Humanoid Parkour was originally proposed to handle the terrain below the feet using depth image. For a fair comparison, we replace the depth image with voxel map to deal with overhanging obstacles.
- **ASTraversal** [13]: ASTraversal was originally proposed for quadrupeds. We re-implement its core multi-layer elevation maps and obstacle-avoidance reward for our humanoid framework.
- **Ours w/o *OBSField*:** We replace the HumanoidPF in observation with multi-layer elevation maps as used in ASTraversal [13]. We adopt this baseline to evaluate the effectiveness of HumanoidPF for observation.
- **Ours w/o *RField*:** We remove the HumanoidPF-guided reward. Instead, we use a basic collision-penalty reward commonly used in collision-avoidance tasks [15]. We adopt this baseline to validate HumanoidPF for reward.

Experiment Results. Results are summarized in Table II. Our method consistently achieves the best performance across all types of test cases, demonstrating its strong traversal capability in cluttered scenes. In contrast, both Humanoid Parkour and ASTraversal only achieve competitive performance in relatively simple terrains. Nevertheless,

our approach exhibits low variance over multiple trials, indicating high robustness and stability of the learned policy.

B. Validation of hybrid scene generation

To address **Q2** (*Can our hybrid scene generation method help the policy generalize to unseen and challenging scenes?*), We evaluate and compare the zero-shot scene generalization ability of the traversal policies trained with different scene datasets.

Experiment Setting. We collect 30 artist-designed indoor scenes that were not included in the training dataset. These scenes are categorized according to obstacle density into 15 easy and 15 hard cases. We test the policies trained with different datasets in these unseen scenes and compare their zero-shot transfer performance.

Experiment Metrics. We use the **Success Rate (SR, %)** metric as mentioned in Section IV-A.

Baselines. We choose the following methods as baselines:

- **Base:** We use 3D-FRONT scenes only.
- **Base + Syn-Partial-Easy:** **Base** combined with a subset of moderate-difficulty (≤ 0.6) synthesized scenes.
- **Base + Syn-Full-Easy:** **Base** combined with the full set of moderate-difficulty (≤ 0.6) synthesized scenes.
- **Base + Syn-Full-Hard:** **Base** combined with the full set of high-difficulty ($0.4 \sim 1.0$) synthesized scenes.

Experiment Results The results shown in Table III confirm that procedural obstacle generation is crucial to enhance generalization, as it systematically scales both the volume and the difficulty of training data. Despite significant performance improvement from the expansion of obstacle diversity, saturation in gains can be observed once common layout patterns are sufficiently learned. The key breakthrough emerges from training in scenes with novel complexities and tighter constraints, which induce the policy to develop the superior capabilities necessary for complex terrains.

Group	Easy (SR%)	Hard (SR%)
Base	62.0 ± 23.4	1.2 ± 2.9
Base + Syn-Partial-Easy	78.6 ± 12.2	12.3 ± 10.3
Base + Syn-Full-Easy	83.1 ± 19.0	26.4 ± 5.8
Base + Syn-Full-Hard	95.2 ± 6.1	66.7 ± 17.9

TABLE III: Validation of hybrid scene generation. Mean success rate (%) on easy and hard subsets for the four experiment groups (mean \pm std).

C. Validation of HumanoidPF for real world transfer

To address **Q3** (*Can our method be successfully transferred to the real world?*), we construct a cluttered indoor scene in the real world and deploy our method on the real robot to test its traversal capability.

Experiment Setting. For real-world deployment, we utilize the Livox Mid-360 LiDAR on the G1 robot, which provides both IMU measurements and raw point clouds. To obtain HumanoidPF, we first employ a SLAM system based on Fast-LIO2 [41] and OctoMap [42] to acquire surrounding

Real-world Exteroception	Crouch- Pass	Hurdle- Pass	Side- Pass	Crouch- Hurdle
Voxel Grids	1/5	3/5	2/5	2/5
Elevation Maps	3/5	3/5	1/5	2/5
HumanoidPF (Ours)	4/5	5/5	5/5	4/5

TABLE IV: Validation of HumanoidPF for real world transfer. Success rate on four challenging real-world scenes.

occupancy map. To ensure a clean and physically consistent field representation, the occupancy map is further post-processed through denoising, dilation, and erosion operations, producing a smooth HumanoidPF according to Section III-A.1. Finally, the learned policy queries the updated HumanoidPF at 50 Hz, enabling real-time gradient-based motion control. Qualitative results are shown in Figure 1 and Figure 4.

Experiment Metrics. We use the **Success Rate (SR)** metric as mentioned in Section IV-A.

Baselines. We distill the same specialists to two different visual representation as baselines:

- **Voxel Grids:** Distilling to voxel-based visuomotor policy.
- **Multi-layer Elevation Maps:** Distilling to two-layer-elevation visuomotor policy.

Experiment Results. Results are summarized in Table IV. Our HumanoidPF consistently outperforms voxel-based approaches across diverse environments. For example, on hurdle terrains, the voxel-based method often exhibits instability because it is sensitive to fine-grained geometric details of obstacles and to noise in task-irrelevant regions. In contrast, HumanoidPF remains largely unaffected by such disturbances and shows behavior that aligns more closely with simulation, further confirming its robustness in sim-to-real transfer.

V. CLICK-AND-TRAVERSE: TELEOPERATION SYSTEM

For ease of operation, we build a teleoperated locomotion system, termed Click-and-Traverse (CAT), on top of the learned generalist policy. The user simply clicks a desired goal location on a 2D grid map constructed by SLAM, and the humanoid autonomously traverses to the target while avoiding collisions. This interface eliminates the need for labor-intensive control modalities such as joysticks or motion capture, providing a lightweight, efficient, and highly automated teleoperation solution.

VI. CONCLUSION

In this work, we study the problem of collision-free humanoid traversal in cluttered indoor environments by investigating how to better map the robot’s perception of surrounding obstacles to corresponding traversal skills. While this work demonstrates strong potential for practical usage of humanoid robots in real-world domestic settings, several current limitations remain: 1) our HumanoidPF does not readily extend to exploiting contact interactions, such as leaning on obstacles or stepping onto support surfaces; 2) although our method performs well across a wide range of environments, generalization to unseen, unstructured layouts and extremely cluttered scenes remains challenging.

REFERENCES

- [1] R. Grandia, F. Jenelten, S. Yang, F. Farshidian, and M. Hutter, “Perceptive locomotion through nonlinear model-predictive control,” *IEEE Transactions on Robotics*, vol. 39, no. 5, pp. 3402–3421, 2023.
- [2] 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.
- [3] C. Zhang, N. Rudin, D. Hoeller, and M. Hutter, “Learning agile locomotion on risky terrains,” in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 11 864–11 871.
- [4] X. Cheng, K. Shi, A. Agarwal, and D. Pathak, “Extreme parkour with legged robots,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 11 443–11 450.
- [5] 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.
- [6] J. Lee, M. Bjelonic, A. Reske, L. Wellhausen, T. Miki, and M. Hutter, “Learning robust autonomous navigation and locomotion for wheeled-legged robots,” *Science Robotics*, vol. 9, no. 89, p. ead19641, 2024.
- [7] N. Rudin, D. Hoeller, M. Bjelonic, and M. Hutter, “Advanced skills by learning locomotion and local navigation end-to-end,” in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 2497–2503.
- [8] Z. Wang, T. Ma, Y. Jia, X. Yang, J. Zhou, W. Ouyang, Q. Zhang, and J. Liang, “Omni-perception: Omnidirectional collision avoidance for legged locomotion in dynamic environments,” *arXiv preprint arXiv:2505.19214*, 2025.
- [9] D. Hoeller, N. Rudin, D. Sako, and M. Hutter, “Anymal parkour: Learning agile navigation for quadrupedal robots,” *Science Robotics*, vol. 9, no. 88, p. eadi7566, 2024.
- [10] T. Miki, J. Lee, L. Wellhausen, and M. Hutter, “Learning to walk in confined spaces using 3d representation,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 8649–8656.
- [11] Z. Zhuang, Z. Fu, J. Wang, C. Atkeson, S. Schwertfeger, C. Finn, and H. Zhao, “Robot parkour learning,” *arXiv preprint arXiv:2309.05665*, 2023.
- [12] N. Rudin, J. He, J. Aurand, and M. Hutter, “Parkour in the wild: Learning a general and extensible agile locomotion policy using multi-expert distillation and rl fine-tuning,” *arXiv preprint arXiv:2505.11164*, 2025.
- [13] Y. Chen, J. Ma, Z. Luo, Y. Han, Y. Dong, B. Xu, and P. Lu, “Learning autonomous and safe quadruped traversal of complex terrains using multi-layer elevation maps,” *IEEE Robotics and Automation Letters*, 2025.
- [14] Z. Li, J. Zeng, S. Chen, and K. Sreenath, “Autonomous navigation of underactuated bipedal robots in height-constrained environments,” *The International Journal of Robotics Research*, vol. 42, no. 8, pp. 565–585, 2023.
- [15] Z. Zhuang, S. Yao, and H. Zhao, “Humanoid parkour learning,” *arXiv preprint arXiv:2406.10759*, 2024.
- [16] J. He, C. Zhang, F. Jenelten, R. Grandia, M. Bächer, and M. Hutter, “Attention-based map encoding for learning generalized legged locomotion,” *Science Robotics*, vol. 10, no. 105, p. eadv3604, 2025.
- [17] J. Sun, G. Han, P. Sun, W. Zhao, J. Cao, J. Wang, Y. Guo, and Q. Zhang, “Dpl: Depth-only perceptive humanoid locomotion via realistic depth synthesis and cross-attention terrain reconstruction,” *arXiv preprint arXiv:2510.07152*, 2025.
- [18] J. Long, J. Ren, M. Shi, Z. Wang, T. Huang, P. Luo, and J. Pang, “Learning humanoid locomotion with perceptive internal model,” in *2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2025, pp. 9997–10 003.
- [19] J. Ren, T. Huang, H. Wang, Z. Wang, Q. Ben, J. Long, Y. Yang, J. Pang, and P. Luo, “Vb-com: Learning vision-blind composite humanoid locomotion against deficient perception,” *arXiv preprint arXiv:2502.14814*, 2025.
- [20] T. Huang, R. Xu, Y. Wang, W. Gao, and S. Zhang, “Traversing narrow paths: A two-stage reinforcement learning framework for robust and safe humanoid walking,” *arXiv preprint arXiv:2508.20661*, 2025.
- [21] H. Wang, Z. Wang, J. Ren, Q. Ben, T. Huang, W. Zhang, and J. Pang, “Beamdojo: Learning agile humanoid locomotion on sparse footholds,” *arXiv preprint arXiv:2502.10363*, 2025.
- [22] W. Sun, B. Cao, L. Chen, Y. Su, Y. Liu, Z. Xie, and H. Liu, “Learning perceptive humanoid locomotion over challenging terrain,” *arXiv preprint arXiv:2503.00692*, 2025.
- [23] Q. Ben, B. Xu, K. Li, F. Jia, W. Zhang, J. Wang, J. Wang, D. Lin, and J. Pang, “Gallant: Voxel grid-based humanoid locomotion and local-navigation across 3d constrained terrains,” *arXiv preprint arXiv:2511.14625*, 2025.
- [24] O. Khatib, “Real-time obstacle avoidance for manipulators and mobile robots,” *The international journal of robotics research*, vol. 5, no. 1, pp. 90–98, 1986.
- [25] Z. Wu, J. Dai, B. Jiang, and H. R. Karimi, “Robot path planning based on artificial potential field with deterministic annealing,” *ISA transactions*, vol. 138, pp. 74–87, 2023.
- [26] Y. K. Hwang, N. Ahuja, et al., “A potential field approach to path planning,” *IEEE transactions on robotics and automation*, vol. 8, no. 1, pp. 23–32, 1992.
- [27] S. S. Ge and Y. J. Cui, “Dynamic motion planning for mobile robots using potential field method,” *Autonomous robots*, vol. 13, no. 3, pp. 207–222, 2002.
- [28] Y. Gong, R. Laha, and L. Figueiredo, “Geopf: Infusing geometry into potential fields for reactive planning in non-trivial environments,” *arXiv preprint arXiv:2505.19688*, 2025.
- [29] M. Ginesi, D. Meli, A. Roberti, N. Sansonetto, and P. Fiorini, “Dynamic movement primitives: Volumetric obstacle avoidance using dynamic potential functions,” *Journal of Intelligent & Robotic Systems*, vol. 101, no. 4, p. 79, 2021.
- [30] Y. Chen, L. Chen, J. Ding, and Y. Liu, “Research on real-time obstacle avoidance motion planning of industrial robotic arm based on artificial potential field method in joint space,” *Applied Sciences*, vol. 13, no. 12, p. 6973, 2023.
- [31] H. Igarashi, T. Machida, F. Harashima, and M. Kakikura, “Free gait for quadruped robots with posture control,” in *9th IEEE International Workshop on Advanced Motion Control, 2006*. IEEE, 2006, pp. 433–438.
- [32] H. Igarashi and M. Kakikura, “Path and posture planning for walking robots by artificial potential field method,” in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, vol. 3. IEEE, 2004, pp. 2165–2170.
- [33] F. Zhuo, W. Jia, S. Ma, J. Yuan, and Y. Sun, “A variable artificial potential field method for gait generation of quadruped robot,” in *2021 IEEE International Conference on Mechatronics and Automation (ICMA)*. IEEE, 2021, pp. 1176–1181.
- [34] Y. A. Voeurn and M. I. Raza, “Motion planning of quadruped robot using potential field,” in *2021 International Conference on Nonlinearity, Information and Robotics (NIR)*. IEEE, 2021, pp. 1–6.
- [35] S. Ross, G. Gordon, and D. Bagnell, “A reduction of imitation learning and structured prediction to no-regret online learning,” in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2011, pp. 627–635.
- [36] H. Fu, B. Cai, L. Gao, L.-X. Zhang, J. Wang, C. Li, Q. Zeng, C. Sun, R. Jia, B. Zhao, et al., “3d-front: 3d furnished rooms with layouts and semantics,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 10 933–10 942.
- [37] M. Deitke, E. VanderBilt, A. Herrasti, L. Weihs, K. Ehsani, J. Salvador, W. Han, E. Kolve, A. Kembhavi, and R. Mottaghi, “Procthor: Large-scale embodied ai using procedural generation,” *Advances in Neural Information Processing Systems*, vol. 35, pp. 5982–5994, 2022.
- [38] S. K. Ramakrishnan, A. Gokaslan, E. Wijmans, O. Maksymets, A. Clegg, J. Turner, E. Undersander, W. Galuba, A. Westbury, A. X. Chang, et al., “Habitat-matterport 3d dataset (hm3d): 1000 large-scale 3d environments for embodied ai,” *arXiv preprint arXiv:2109.08238*, 2021.
- [39] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” *arXiv preprint arXiv:1707.06347*, 2017.
- [40] E. Todorov, T. Erez, and Y. Tassa, “Mujoco: A physics engine for model-based control,” in *2012 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 2012, pp. 5026–5033.
- [41] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, “Fast-llo2: Fast direct lidar-inertial odometry,” 2021. [Online]. Available: <https://arxiv.org/abs/2107.06829>
- [42] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, “OctoMap: An efficient probabilistic 3D mapping framework based on octrees,” *Autonomous Robots*, 2013, software available at <https://octomap.github.io>. [Online]. Available: <https://octomap.github.io>