Transforming a Quadruped into a Guide Robot for the Visually Impaired: Formalizing Wayfinding, Interaction Modeling, and Safety Mechanism

Anonymous Author(s)
Affiliation
Address
email

Figure 1: We develop a robot guide dog by formalizing the wayfinding task, developing an interaction model from the collected data (**Left**), and employing action shielding, to guide a user (**Right**).

Abstract: This paper explores the principles for transforming a quadrupedal robot into a guide robot for individuals with visual impairments. A guide robot has great potential to resolve the limited availability of guide animals that are accessible to only two to three percent of the potential blind or visually impaired (BVI) users. To build a successful guide robot, our paper explores three key topics: (1) formalizing the navigation mechanism of a guide dog and a human, (2) developing a data-driven model of their interaction, and (3) improving user safety. First, we formalize the wayfinding task of the human-guide robot team using Markov Decision Processes based on the literature and interviews. Then we collect real human-robot interaction data from three visually impaired and six sighted people and develop an interaction model called the "Delayed Harness" to effectively simulate the navigation behaviors of the team. Additionally, we introduce an action shielding mechanism to enhance user safety by predicting and filtering out dangerous actions. We evaluate the developed interaction model and the safety mechanism in simulation, which greatly reduce the prediction errors and the number of collisions, respectively. We also demonstrate the integrated system on a quadrupedal robot with a rigid harness, by guiding users over $100+$ m trajectories.

Keywords: Assistive Robot, Autonomous Navigation, Interaction Modeling

1 Introduction

Guide dogs play a crucial role in the lives of blind or visually impaired individuals (BVIs) by providing them with increased mobility and independence [1]. Unfortunately, limited availability is one of the known issues of guide animals due to the lengthy training process and their relatively short working lifespan. Consequently, only a small fraction, approximately two to three percent [2], of potential users are able to benefit from the assistance of guide dogs. Researchers have explored various types of assistive technologies for BVIs, including smart canes [3, 4, 5, 6, 7], belts [8, 9], glasses [10], audio/haptic systems [11, 12], and mobile robots [13, 14, 15, 16, 17]. Recently, the

Submitted to the 7th Conference on Robot Learning (CoRL 2023). Do not distribute.

development of affordable and capable quadrupedal robots has opened up opportunities for robot guide dogs [18, 19, 20, 21, 22], which possess notable qualities in terms of autonomy and mobility.

Developing robotic guiding systems for BVIs requires additional challenges beyond the common practices of the existing autonomous robot navigation systems [23, 24]. Unlike conventional formulations that simplify the problem as *Point-goal* or *Object-goal* navigation [25], navigation of the human-robot dyad often involves more complex and subtle bidirectional interactions, including high-level command generation by the user and intelligent disobedience by the robot. Furthermore, ensuring BVI user safety is critical during navigation, yet it is a challenging problem due to that the sensor configurations are robot-centric.

This paper discusses principles and practical solutions for transforming an autonomous robot into a guide robot for individuals with visual impairments. The topics that we explore include formalizing the wayfinding task of a human and guide dog team, developing a data-driven model of the human and the guide robot interaction, and improving the safety of the BVI user. First, we identify the mechanism of wayfinding in the literature [26, 27, 28] and formally define the guided navigation task using Markov Decision Processes (MDPs). Then we collect real human-robot interaction data of three BVI and six sighted users and develop a concise model, *Delayed Harness*, to better predict the navigation behaviors of the team in simulation. Finally, we introduce an action shielding mechanism [29] to improve the safety of the BVI user, which predicts the human position in the next step and filters out dangerous actions.

A mid-size quadrupedal robot, AlienGo [30], with a rigid harness is chosen as the evaluation platform. Initially, we evaluate the accuracy of the developed *Delayed Harness* interaction model and the improved safety of the learned policy in a Habitat [24] simulator. Then we deploy the developed system on a real AlienGo robot to demonstrate the effective long-range navigation of the human-robot team. Our contributions include: (1) We develop a formal definition of the wayfinding task of the human and guide dog team. (2) We collect human-robot interaction data, develop a concise interaction model, and open-source it to foster guide robot research in the robotics community. (3) We improve the safety of a BVI user by employing action shielding. (4) We demonstrate a complete system on hardware that can travel along trajectories for more than 100 meters.

2 Related Work

Assistive Guidance Systems for Visually Impaired People. Researchers have developed diverse assistive systems to enhance mobility for visually impaired people. Such assistive systems consist of essential elements, including sensors for perceiving the environment, computers for processing information, and various interfaces to instruct users. One common form is passive hand-held or wearable devices, such as canes [3, 4, 5, 6, 7], belts [8, 9, 11, 12], smart glasses [10], and smartphone applications [31, 32], which generally give audio or haptic feedback to inform users. On the other hand, there exist guide robots [33, 13, 14, 18, 19, 21] that are designed to lead users actively toward destinations. Since the earliest guide robot MELDOG [33], such systems are often built on top of wheeled robots for mobility [13, 14, 15, 34, 16, 17, 35, 36, 37]. Recently, legged robots are emerging in consideration of the ability to travel as real guide dogs [18, 19, 20, 21, 22, 38].

Human Modeling in Guide Robots. Developing guide robots necessitates modeling human-robot interaction because they communicate with users via physical interactions. For instance, Wang et al. [16] introduce a rotating rigid rod model that assumes the user is holding the end of the rotating rod at a fixed distance. On the other hand, some guide robots [18, 39, 21] communicate via a flexible leash, which investigates a mathematical model for capturing slack and taut modes. Likewise, interactions are typically modeled based on the interface of robots that have been developed [40, 34]. The most common interface for a real guide dog is a fixed harness [41, 42]. In this paper, we claim that a simple offset between the human and the robot is insufficient to capture bidirectional interactions and propose a novel mathematical model to improve the human-robot navigation experience.

Safety-aware Reinforcement Learning for Navigation. Autonomous robot navigation has gained significant attention in robotics and has been approached through both planning-based methods [43, 44] and learning-based techniques [45, 46, 47]. Please refer to the survey paper [48, 49] for further detailed references. Numerous algorithms have been developed to enhance the safety of navigation, including action shielding [29, 50], model predictive methods [51], multi-agent strategies [52], and near future considerations [53, 54]. Our work is also inspired by these prior works while being more customized toward the human-robot team navigation task.

3 Formalizing the Wayfinding Task

This section identifies how a human interacts with a guide dog and formulate the wayfinding task for a guide robot using Markov Decision Processes (MDPs). Our problem definition aims to develop more natural and comfortable user interfaces of guide robots for BVI users, compared to typical *Point-Goal* or *Objective-Goal* navigation problems [25].

Wayfinding with a Guide Dog. We first identify how guide dogs and human handlers communicate to work as a team by reviewing relevant background materials [41, 26, 27, 28] and conducting interviews with actual guide dog users. Typically, the human user is the one who generates high-level directional cues, such as "go straight" or "turn left", based on his or her own knowledge, while the guide dog handles local path following and collision avoidance based on visual perception. Unless the team routinely travels to the same destinations, the guide dog is not typically aware of the goal or path. Then, the guide dog sees and leads the human handler based on the given command. The guide dog needs to follow the virtual "travel line", a centerline of the hallway directed by the handler while adjusting a trajectory to ensure the safety of both the human and the dog itself. We refer to this type of human-dog navigation as *Wayfinding*, and it is our goal to replicate this interaction with a human-robot team.

In addition, a guide dog should understand user intention and compensate for any errors. We find two major types of human error: *Orientation error* and *Timing error*. For instance, *orientation error* can occur in environments such as a long hallway, where the user command is misaligned with the travel line and must be corrected. *Timing error* occurs when a human handler gives turn commands before or after actual turning point. For instance, if a turning command is given early while traveling down a hallway, the robot is expected to find the next available turn or stop if no turn is available. Please refer to Figure 6 and the supplemental video for illustrative examples.

Problem Formulation. We define the wayfinding task for the robot as following high-level human directional cues: *going straight*, *turning left/right*, or *stop* while adjusting a detailed trajectory to avoid collisions. We use Markov Decision Processes (MDPs) to formalize this wayfinding task, which is defined as a tuple of the state space $\mathcal{S}$, action space $\mathcal{A}$, stochastic transition function $p(\mathbf{s}_{t+1}|\mathbf{s}_t, \mathbf{a}_t)$, reward function $r(\mathbf{s}_t, \mathbf{a}_t)$, and the distribution of initial states $\mathbf{s}_0 \sim \rho_0$. The state and action spaces can vary depending on the choice of the robot. For instance, one common choice is to construct the state using onboard sensors, such as RGB-D or lidar, while defining the action space as a set of possible navigation commands to the robot. We use depth images, lidar readings, and a relative location from the start location as a state and adopt discrete actions based on the recent autonomous navigation papers [46, 55, 47].

The reward function is critical to describe the desirable wayfinding behavior, which is defined as:

$$r_t = (d_t - d_{t-1}) - a\Delta \max\left(|\bar{\theta} - \theta_t| - b, 0\right) - c_t^{\text{collide}} - \lambda \tag{1}$$

where d_t is the human's travel distance from the start point, θ is the human's orientation, $\bar{\theta}$ is the target orientation (e.g., 0° for the straight command and 90° for the turn left command), a is a scalar weight, b is the error margin above which the corresponding term is activated, c_t^{collide} is a collision penalty, and λ is a time penalty. The first term encourages the team to travel as far as possible while minimizing the human's unnecessary movements, such as stepping backward, the

second term encourages matching the desired orientation (with some flexibility), and the third and fourth terms regularize unsafe and unnecessary behaviors, respectively.

Because our formulation rewards the team to travel long distances, it encourages the navigation policy to compensate for any orientation or timing errors, particularly in narrow hallways. Also note that all the terms are defined in a human-centric manner in order to provide a smooth trajectory for the visually impaired user, not for the robot itself.

4 Modeling Human-Guide Robot Interactions

In order to develop a successful guiding policy, it is important to simulate human-robot interactions accurately. Guide dogs wear a rigid harness, instead of a loose leash, that allows the human to feel both the guide dog's position and its orientation. We adopt such a fixed harness for our human-robot teams. While there are established human-robot models, such as a rotating rod [16], a slack/taut leash [18], and a geometric model [34], we require a new model specifically tailored to our guide dog robot with a harness. In this section, we propose a new interaction model, the *Delayed Harness*, and optimize the model parameters using real-world data.

Preliminary: Rigid Harness Model. The simplest model for describing rigid harness interactions is to use a fixed offset between the human and the robot. In this model, the human maintains their relative position to the robot, facing the same forward direction. Formally, when the robot position and orientation $(x_t^R, y_t^R, \theta_t^R)$ is given, the human state $(x_t^H, y_t^H, \theta_t^H)$ is computed by assuming a fixed distance d between the handler and the robot as follows:

$$x_t^H = x_t^R + d \cdot \cos\theta_t,\ y_t^H = y_t^R + d \cdot \sin\theta_t,\ \theta_t^H = \theta_t^R. \tag{2}$$

This model expects that the human and the robot rotate together, requiring extra space compared to leash, string, or rod connections.

Delayed Harness Model. A rigid harness model is not accurate because a human cannot exactly follow the robot in reality. As a result, a learned robot policy based on such a model may cause unnecessary movements or even collisions to a user.

To this end, we develop a more flexible model named *Delayed Harness*. This model is based on the observation that a robot's action causes the change of the relative location, but that a human tends to gradually recover the default relative location (Figure 2). Let us define the robot and human states as $\mathbf{x}_t^H = [x_t^H, y_t^H, \theta_t^H]^T$ and $\mathbf{x}_t^R = [x_t^R, y_t^R, \theta_t^R]^T$, which gives the offset $\mathbf{o}_t = \mathbf{x}_t^H - \mathbf{x}_t^R$. Once the robot takes an action $\mathbf{a}_t$, it first changes its position: $\mathbf{x}_{t+1}^R = \mathbf{x}_t^R + \mathbf{a}_t$. Now, we have a temporary offset $\hat{\mathbf{o}}_{t+1} = \mathbf{x}_t^H - \mathbf{x}_{t+1}^R$ (black dashed line in Figure 2). We gradually interpolate this temporary offset toward the default offset, $\bar{\mathbf{o}}$ (green line), to obtain the corrected offset in the next time: $\mathbf{o}_{t+1} = \alpha\hat{\mathbf{o}} + (1-\alpha)\bar{\mathbf{o}}$. Finally, we compute the next human position $\mathbf{x}_{t+1}^H = \mathbf{x}_{t+1}^R + \mathbf{o}_{t+1}$. Note that we use simple arithmetic operators for illustration, but in reality, they should be handled as transformation operations.

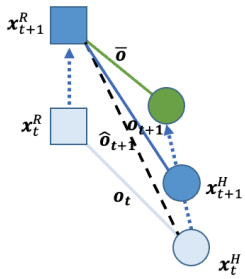

Figure 2: Delayed Harness Model.

Data Collection and Model Fitting. Our *Delayed Harness* model requires four parameters: the default offset $\bar{\mathbf{o}} = [\Delta x, \Delta y, \Delta\theta]^T$ and the decaying parameter α. We determine these parameters by fitting them to the collected trajectories. We recruit three BVI users and six sighted people, and ask them to walk with a manually-controlled robot guide dog over five trajectories described in the work of Nanavati et al. [34]. The sighted individuals each wore a eye mask while walking with the robot. Please refer to the supplemental material for more details of the data collection process. The collected data is provided as supplemental material (`interaction-data.zip`).

5 Improving Safety via Action Shielding

We further introduce an action-shielding approach [29] to improve safety during human-robot navigation. Specifically, our focus is on enhancing the safety of the BVI user based on data from sensors arranged in a robot-centric layout. Inspired by a guide dog user who is also using a cane, our action shielding predicts the changes in the human and robot positions according to the possible actions and filters out unsafe actions. Particularly, we designed the shielding mechanism based on lidar, which has relatively better accuracy (± 5 cm) than a depth camera (5 % up to 15 m).

The implementation of such action shielding requires two subroutines: (1) predicting possible collisions in the next time step, and (2) suppressing potentially risky actions.

Computation of Shielding Zone. To identify the shielding zone, we compute the occupied areas of both agents in the current and next time steps, which are approximated as circular shapes. While we have four collision primitives, the region between the occupied areas of the human and robot agents should be empty because the two agents are physically connected through a harness. In addition, the zone between the current and the next time steps should also be collision-free to account for continuous movements. Therefore, we compute a convex hull of all the collision shapes from both agents at two timesteps to compute the shielding zone (green line in Figure 3). Once computed, we sample rays from the origin to determine the distance thresholds for lidar. Refer to the Appendix 8.2 for more details.

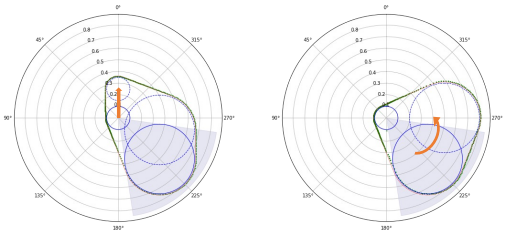

Figure 3: Examples of Shielding Zones.

Suppressing Unsafe Actions. When shielding is activated, we can suppress unsafe actions by adjusting the action probability distribution by a *suppression factor* β. A suppression factor of zero means assigning zero probability to the unsafe actions to force the policy to select safe actions only. It is also possible to employ a small positive value to encourage exploration during training.

The action shielding technique can be utilized during both training and testing. During training, we anticipate the agent will learn a safer policy, especially when sensing is not perfect due to sensor noises or blind spots. Alternatively, we can combine action shielding with any pre-trained policy to improve its safety at the evaluation stage. We examine how the learned policy performs under different action suppression factors in ideal and realistic sensing conditions in Section 6.3.

6 Experiments

We conducted experiments to address the following questions: (1) Is the proposed delayed harness model more accurate than other models? (2) How does the interaction model impact performance? (3) Does action shielding effectively decrease collisions in noisy environments? (4) When combined, can the proposed robot guide dog navigate to the destination with a human user?

6.1 Training Details

Our policies were trained in the Habitat simulator [24] with the Matterport3D dataset [56] using DD-PPO [46]. The training typically took 10 million steps until convergence, which roughly corresponds to two to three days. An observation space consists of all the sensor information, including depth images, 2D lidar reading, and the relative location from the starting position, followed by a one-hot encoded high-level directional cue provided by the user. For the low-level robot action space, we choose ten discrete actions {stop, forward, turn 10° left/right, sidestep left/right, diagonal steps toward the front left/right} inspired by the recent navigation literature [57]. We particularly included side and diagonal steps because they are effective for guiding the user to escape confined regions.

Table 1: Comparison of Different Interaction Models

Model	Fixed-unopt-ind	Fixed-opt-ind	RR-opt-ind	DH-opt-ind (ours)	DH-opt-all
RMSE	136.2	102.8	431.4	**86.3**	123.1

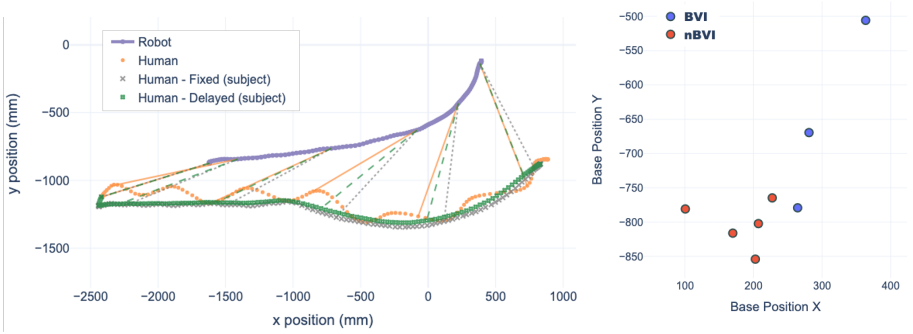

Figure 4: (**Left**) Different trajectories of interaction models optimized using the collected motion capture data on each individual subject. (**Right**) Each subject's base positions derived from the optimized results of the *delayed harness* model.

6.2 Comparison of Interaction Models

We first compared our *delayed harness* model against two baselines: a *fixed harness* and a *rotating rod* [16]. By default, the parameters are optimized for each individual subject, which gives us three models: *Fixed-opt-ind*, *RR-opt-ind* (rotating rod), and *DH-opt-ind* (delayed harness). In addition, we included the unoptimized setting, *Fixed-unopt-ind*, which uses the starting parameter to estimate the default offset $\bar{\mathbf{o}}$ and assumes zero decay $\alpha = 0$. We also investigated the *DH-opt-all* model, which optimizes a single set of parameters for all the subjects. For all the optimized models, we fitted the parameters over ten trajectories and validated them over five trajectories. The results indicated that the *DH-opt-ind* model exhibited the highest accuracy, followed by *Fixed-opt-ind*, *DH-opt-all*, *Fixed-unopt-ind*, and *RR-opt-ind* (Table 1). Figure 4 left presents the example of the different trajectories between *Fixed Harness* and *Delayed Harness* models. For the same time step, *Delayed Harness* (connected via green line) is closer to the actual human location (orange line) compared to *Fixed Harness* (grey line) by capturing the delayed response of the human user. It is also worth mentioning that a low accuracy of a rotating rod model was due to its different interface than a harness: therefore, we did not further investigate more complex rod-based models, such as a geometric model proposed by Nanavati et al. [34].

The accuracy difference between *DH-opt-ind* and *DH-opt-all* highlights the importance of models customized for users. We observed that subjects have different default offsets (Fig. 4 right) or responsiveness to direction changes. Therefore, we decided to train policies for each individual, assuming that it is possible to collect user-specific data at onboarding. In the future, it will be interesting to investigate preference optimization techniques [58] to improve robot guide dog behaviors.

Indeed, selecting the appropriate interaction model is critical to achieving the best human-robot team navigation. Table 2 presents the performance of the learned policies using various interaction models at training and testing times. Both policies with the fixed or delayed harness models showed their best performance when evaluated with the corresponding model. However, if the model is altered, the performance drops, which is approximately 25 % lower than the original performance. Therefore, we suggest to train a policy with a more accurate interaction model, such as *Delayed Harness*.

Table 2: Rewards with Different Training and Testing Interaction Models.

		Test	
		Fixed	Delay
train	Fixed	5.96	4.27
	Delay	4.00	5.04

Sensor Config	Supp. Factor during Train	Supp. Factor during Test	Collision-Free Ep. Ratio ↑	Avg. Collisions Per Ep. ↓
Ideal	0.0	0.0	(diverges)	
	0.1	0.0	**0.96**	**0.04**
	0.5	0.0	0.84	1.46
	0.9	0.0	0.88	1.80
	1.0	0.0	0.92	1.40
(no shielding)	1.0	1.0	0.32	8.20
Noisy	0.0	0.0	(diverges)	
	0.1	0.0	0.68	4.12
	0.5	0.0	**0.84**	**1.48**
	0.9	0.0	0.64	2.52
	1.0	0.0	0.72	3.16
(no shielding)	1.0	1.0	0.52	3.96

Table 3: Performance of Action Shielding in Ideal and Noisy Environments.

6.3 Effectiveness of Action Shielding

The main objective of our action shielding mechanism is to improve the safety of the human-robot team in unseen environments with noisy sensor readings. To this end, we compared a policy with action shielding against a vanilla RL policy in both ideal and noisy simulated environments. Additionally, we varied an action suppression factor β during training as well, which has a large affect on the performance. For noisy environments, we employed a Gaussian noise for the lidar and a Redwood [59] noise with an intensity of 1.0 for the depth camera.

In Table 3, we present the collision-free episode ratio and the average number of collisions per episode as metrics for evaluating the safety of the policies. Overall, it is evident that activating action shielding ($\beta < 1.0$) during evaluation effectively reduces the occurrence of collisions. However, the desirable training configurations differed in ideal and noisy environments. In an ideal setting, the performance is not very sensitive to the choice of the suppression factor. It is even very beneficial to turn on action shielding only at the testing stage, which shows a great performance margin against no shielding (collision-free episode ratios: 0.92 vs 0.32). In a noisy environment, training a policy with mild action shielding ($\beta \sim 0.5$) was more effective for adapting a navigation strategy to cope with action shielding. In any configuration, zero action supression β during training results in divergence of learning, which is likely due to insufficient exploration.

Action shielding was effective even in real environments with noisy sensor readings. We presented two distinct scenarios where the vanilla policy (without action shielding) failed to generalize to novel situations, such as encountering curtains or large boxes that were unseen during training. Conversely, the policy with action shielding successfully navigated around these obstacles by predicting the user and robot trajectories. For more details, please refer to the supplemental video.

6.4 Real-world Experiments

We deployed the proposed robot guide dog system on the hardware of AlienGo [30] by Unitree (Figure 5). This quadruped robot was equipped with a Zed depth camera, a Slamtec RPLIDAR lidar, and an Intel t265 tracking camera. The robot leads the user via a harness while taking verbal commands as directional cues. The policy was trained and deployed with action shielding with action supression of 0.5 and 0.0, respectively, based on Table 3.

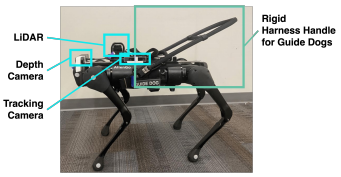

Figure 5: Hardware Platform

General Navigation. We assigned a sighted user wearing an eye mask the task of completing three indoor routes with four directional cues: `forward`, `left`, `right`, and `stop`. Route A (146 m)

is a curve path that can be traversed by a single forward cue. Route B (130 m) and C (121 m) required the user to issue multiple turns based on their own localization. Refer to Appendix 8.3 and the supplemental video for the details of the routes. In our experiments, the user successfully accomplished all the tasks without collisions. In Route A, a single command of `forward` was sufficient to complete all the tasks, even in a very long curved corridor. In Route B and C, the user sometimes issued turning commands early or late, but the robot was able to compensate for the timing errors. In Route C, a robot intelligently adjusted the trajectory along the hallway without additional cues.Where there were obstacles ahead, the robot guide dog led the user to avoid them while staying in the travel line. Occasionally, the human gets close to obstacles or walls and the robot prevented human collisions by actively using action shielding, which encouraged sidestepping as a means of avoidance.

Orientation and Timing Errors. As explained in Section 3, the guide robot must effectively handle cues that are provided with some timing or orientation errors caused by human users. We showcase the robustness of the developed robot guide dog by displaying the trajectories involving human errors. Figure 6(a) displays *orientation errors* on a forward cue, where the robot corrected the initial orientation error and followed the direction of the hallway to travel as long as possible. Figure 6(b) illustrates the turning with *timing errors*, where the robot took diagonal stepping or 120° turning to correct early or late issued commands, respectively.

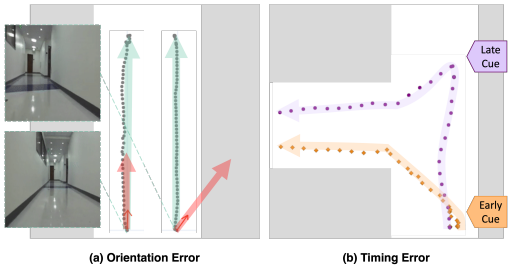

Figure 6: Trajectories with Two Type of Human Errors: Orientation Error (a) and Timing Error (b).

7 Conclusion and Limitation

This paper presents three topics and potential solutions for transforming a standard quadrupedal robot into a robot guide dog to assist blind or visually impaired individuals. First, we study the navigation patterns of human-guide dog pairs and establish a formal wayfinding task using Markov Decision Processes (MDPs). Then we propose a concise interaction model called the *Delayed Harness* to effectively represent the interaction between a human and a robot guide dog, which leads to more accurate behavior prediction and better navigation performance. The parameters of this interaction model are optimized with respect to the real interaction data of three visually impaired and six sighted subjects, which is provided as supplemental material to this manuscript. Finally, we introduce an action shielding mechanism to improve the safety of the human user, inspired by a guide dog user who is also utilizing a cane along with a guide dog. We demonstrate that the integrated robot guide dog can navigate with a user with an blindfold over multiple 100+m long trajectories. We hope that the topics and techniques discussed in this study will inspire further research in the development of guide robots for blind and visually impaired individuals.

Limitation. The proposed solutions in the paper have room for improvement, and we plan to investigate the following research directions. Firstly, while we have explored only a limited set of basic verbal cues, real guide dog users employ a wider range of commands, such as "find an elevator" or "follow the person", combined with non-verbal cues. It will be beneficial to formalize a more expanded set of human-guide dog interactions. Additionally, although our *delayed harness* model effectively captures typical navigation behaviors, it struggles with out-of-distribution situations like users pausing to chat with someone. In the future, we will expand the provided dataset to encompass such diverse scenarios. The developed action shielding mechanism is effective but not perfect. For instance, it can fail even in simulation due to blind spots of the lidar sensor. To resolve this issue, we will explore alternative sensors and safety mechanisms to improve the safety of the human and the robot. Finally, we plan to evaluate the system on a larger population of blind or visually impaired users to gather feedback for further improvements.

References

[1] L. Whitmarsh. The benefits of guide dog ownership. *Visual impairment research*, 7(1):27–42, 2005.

[2] Guiding Eyes for the Blind. https://www.guidingeyes.org/about/faqs/, 2022.

[3] J. Borenstein and I. Ulrich. The guidecane-a computerized travel aid for the active guidance of blind pedestrians. In *Proceedings of International Conference on Robotics and Automation*, volume 2, pages 1283–1288. IEEE, 1997.

[4] S.-J. Kang, Y. Ho, and I. H. Moon. Development of an intelligent guide-stick for the blind. In *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*, volume 4, pages 3208–3213. IEEE, 2001.

[5] K. Yatani, N. Banovic, and K. Truong. Spacesense: representing geographical information to visually impaired people using spatial tactile feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 415–424, 2012.

[6] S. Kher Chaitrali, A. Dabhade Yogita, K. Kadam Snehal, D. Dhamdhere Swati, and V. Deshpande Aarti. An intelligent walking stick for the blind. *International Journal of Engineering Research and General Science*, 3(1):1057–1062, 2015.

[7] R. Velázquez, E. Pissaloux, P. Rodrigo, M. Carrasco, N. I. Giannoccaro, and A. Lay-Ekuakille. An outdoor navigation system for blind pedestrians using gps and tactile-foot feedback. *Applied Sciences*, 8(4):578, 2018.

[8] S. Shoval, J. Borenstein, and Y. Koren. The navbelt-a computerized travel aid for the blind based on mobile robotics technology. *IEEE Transactions on Biomedical Engineering*, 45(11): 1376–1386, 1998.

[9] H.-C. Wang, R. K. Katzschmann, S. Teng, B. Araki, L. Giarré, and D. Rus. Enabling independent navigation for visually impaired people through a wearable vision-based feedback system. In *2017 IEEE international conference on robotics and automation (ICRA)*, pages 6533–6540. IEEE, 2017.

[10] W.-J. Chang, J.-P. Su, L.-B. Chen, M.-C. Chen, C.-H. Hsu, C.-H. Yang, C.-Y. Sie, and C.-H. Chuang. An ai edge computing based wearable assistive device for visually impaired people zebra-crossing walking. In *2020 IEEE International Conference on Consumer Electronics (ICCE)*, pages 1–2, 2020. doi:10.1109/ICCE46568.2020.9043132.

[11] J. Wilson, B. N. Walker, J. Lindsay, C. Cambias, and F. Dellaert. Swan: System for wearable audio navigation. In *2007 11th IEEE international symposium on wearable computers*, pages 91–98. IEEE, 2007.

[12] B. N. Walkera and J. Wilsonc. Swan 2.0: Research and development on a new system for wearable audio navigation.

[13] Y. Wei, X. Kou, and M. C. Lee. Smart rope and vision based guide-dog robot system for the visually impaired self-walking in urban system. In *2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pages 698–703, 2013. doi:10.1109/AIM.2013.6584174.

[14] T.-K. Chuang, N.-C. Lin, J.-S. Chen, C.-H. Hung, Y.-W. Huang, C. Teng, H. Huang, L.-F. Yu, L. Giarré, and H.-C. Wang. Deep trail-following robotic guide dog in pedestrian environments for people who are blind and visually impaired-learning from virtual and real worlds. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5849–5855. IEEE, 2018.

[15] D. R. Bruno, M. H. De Assis, and F. S. Osorio. Development of a mobile robot: Robotic guide dog for aid of visual disabilities in urban environments. *Proceedings - 2019 Latin American Robotics Symposium, 2019 Brazilian Symposium on Robotics and 2019 Workshop on Robotics in Education, LARS/SBR/WRE 2019*, pages 104–108, 2019. doi:10.1109/LARS-SBR-WRE48964.2019.00026.

[16] L. Wang, J. Zhao, and L. Zhang. Navdog: robotic navigation guide dog via model predictive control and human-robot modeling. In *Proceedings of the 36th Annual ACM Symposium on Applied Computing*, pages 815–818, 2021.

[17] V. Ranganeni, M. Sinclair, E. Ofek, A. Miller, J. Campbell, A. Kolobov, and E. Cutrell. Exploring levels of control for a navigation assistant for blind travelers. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '23, page 4–12, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450399647. doi:10.1145/3568162.3578630. URL https://doi.org/10.1145/3568162.3578630.

[18] 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)*, pages 11470–11476. IEEE, 2021.

[19] A. Sivacoumare, S. M. S, S. Satheesh, T. Athul, M. V, and T. Vinopraba. Ai quadruped robot assistant for the visually impaired. In *IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society*, pages 1–5, 2021. doi:10.1109/IECON48115.2021.9589508.

[20] K. Mehrizi. Quadrupedal Robotic Guide Dog with Vocal Human-Robot Interaction. pages 2–4, 2021. URL http://arxiv.org/abs/2111.03718.

[21] Y. Chen, Z. Xu, Z. Jian, G. Tang, Y. Yangli, A. Xiao, X. Wang, and B. Liang. Quadruped Guidance Robot for the Visually Impaired: A Comfort-Based Approach. 2022. URL http://arxiv.org/abs/2203.03927.

[22] 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*, 5(1):56–63, 1 2020. ISSN 23773766. doi:10.1109/LRA.2019.2939719.

[23] B. Shen, F. Xia, C. Li, R. Martín-Martín, L. Fan, G. Wang, C. Pérez-D'Arpino, S. Buch, S. Srivastava, L. Tchapmi, et al. igibson 1.0: A simulation environment for interactive tasks in large realistic scenes. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7520–7527. IEEE, 2021.

[24] M. Savva, A. Kadian, O. Maksymets, Y. Zhao, E. Wijmans, B. Jain, J. Straub, J. Liu, V. Koltun, J. Malik, D. Parikh, and D. Batra. Habitat: A platform for embodied AI research. *Proceedings of the IEEE International Conference on Computer Vision*, 2019-Octob:9338–9346, 2019. ISSN 15505499. doi:10.1109/ICCV.2019.00943.

[25] P. Anderson, A. Chang, D. S. Chaplot, A. Dosovitskiy, S. Gupta, V. Koltun, J. Kosecka, J. Malik, R. Mottaghi, M. Savva, and A. R. Zamir. On Evaluation of Embodied Navigation Agents. 7 2018. URL http://arxiv.org/abs/1807.06757.

[26] Guide dogs for the blind puppy raising manual-training phase descriptions. https://www.guidedogs.com/uploads/files/Puppy-Raising-Manual/Training-Phase-Descriptions.pdf, 2022.

[27] Gdb class lecture: Guidework. https://s3-us-west-1.amazonaws.com/gdb-assets/Guidework.txt?mtime=20160412125546, 2022.

[28] Gdb class lecture: Why does your guide dog work? https://www.guidedogs.com/explore-resources/alumni-resources/class-lecture-materials2/why-does-your-guide-dog-work-lecture, 2022.

[29] M. Alshiekh, R. Bloem, R. Ehlers, B. Könighofer, S. Niekum, and U. Topcu. Safe reinforcement learning via shielding. *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*, pages 2669–2678, 2018.

[30] Aliengo by unitree robotics. https://www.unitree.com/products/aliengo, 2020.

[31] A. Meliones, C. Filios, and J. Llorente. Reliable ultrasonic obstacle recognition for outdoor blind navigation. *Technologies*, 10(3):54, 2022.

[32] P. Theodorou, K. Tsiligkos, A. Meliones, and C. Filios. An extended usability and ux evaluation of a mobile application for the navigation of individuals with blindness and visual impairments outdoors—an evaluation framework based on training. *Sensors*, 22(12):4538, 2022.

[33] S. Tachi, K. Tanie, K. Komoriya, and M. Abe. Electrocutaneous communication in a guide dog robot (meldog). *IEEE transactions on biomedical engineering*, (7):461–469, 1985.

[34] A. Nanavati, X. Z. Tan, J. Connolly, and A. Steinfeld. Follow the robot: Modeling coupled human-robot dyads during navigation. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3836–3843, 2019. doi:10.1109/IROS40897.2019.8967656.

[35] A. Kulkarni, A. Wang, L. Urbina, A. Steinfeld, and B. Dias. Robotic assistance in indoor navigation for people who are blind. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 461–462. IEEE, 2016.

[36] M. Kuribayashi, T. Ishihara, D. Sato, J. Vongkulbhisal, K. Ram, S. Kayukawa, H. Takagi, S. Morishima, and C. Asakawa. Pathfinder: Designing a map-less navigation system for blind people in unfamiliar buildings. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2023.

[37] S. Kayukawa, D. Sato, M. Murata, T. Ishihara, A. Kosugi, H. Takagi, S. Morishima, and C. Asakawa. How users, facility managers, and bystanders perceive and accept a navigation robot for visually impaired people in public buildings. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 546–553. IEEE, 2022.

[38] H. Hwang, T. Xia, I. Keita, K. Suzuki, J. Biswas, S. I. Lee, and D. Kim. System configuration and navigation of a guide dog robot: Toward animal guide dog-level guiding work. *arXiv preprint arXiv:2210.13368*, 2022.

[39] H. Tan, C. Chen, X. Luo, J. Zhang, C. Seibold, K. Yang, and R. Stiefelhagen. Flying guide dog: Walkable path discovery for the visually impaired utilizing drones and transformer-based semantic segmentation. In *2021 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2021.

[40] J. Guerreiro, D. Sato, S. Asakawa, H. Dong, K. M. Kitani, and C. Asakawa. Cabot: Designing and evaluating an autonomous navigation robot for blind people. *ASSETS 2019 - 21st International ACM SIGACCESS Conference on Computers and Accessibility*, (October):68–82, 2019. doi:10.1145/3308561.3353771.

[41] B. Knol, C. Roozendaal, L. Van den Bogaard, and J. Bouw. The suitability of dogs as guide dogs for the blind: criteria and testing procedures. *Veterinary Quarterly*, 10(3):198–204, 1988.

[42] M. J. Tellefson. A perspective on teaching early harness travel to young blind children using children's visual companion dogs. *Journal of Visual Impairment & Blindness*, 106(5):306–312, 2012.

[43] G. E. Jan, K. Y. Chang, and I. Parberry. Optimal path planning for mobile robot navigation. *IEEE/ASME Transactions on mechatronics*, 13(4):451–460, 2008.

[44] J. Bruce and M. Veloso. Real-time randomized path planning for robot navigation. In *IEEE/RSJ international conference on intelligent robots and systems*, volume 3, pages 2383–2388. IEEE, 2002.

[45] C. Chen, Y. Liu, S. Kreiss, and A. Alahi. Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning. In *2019 international conference on robotics and automation (ICRA)*, pages 6015–6022. IEEE, 2019.

[46] E. Wijmans, A. Kadian, A. Morcos, S. Lee, I. Essa, D. Parikh, M. Savva, and D. Batra. Dd-ppo: Learning near-perfect pointgoal navigators from 2.5 billion frames. In *International Conference on Learning Representations*, 2020.

[47] M. Sorokin, J. Tan, C. K. Liu, and S. Ha. Learning to navigate sidewalks in outdoor environments. *IEEE Robotics and Automation Letters*, 7(2):3906–3913, 2022.

[48] K. Zhu and T. Zhang. Deep reinforcement learning based mobile robot navigation: A review. *Tsinghua Science and Technology*, 26(5):674–691, 2021.

[49] X. Xiao, B. Liu, G. Warnell, and P. Stone. Motion planning and control for mobile robot navigation using machine learning: a survey. *Autonomous Robots*, 46(5):569–597, 2022.

[50] N. Jansen, B. Könighofer, J. Junges, A. Serban, and R. Bloem. Safe reinforcement learning using probabilistic shields. 2020.

[51] O. Bastani and S. Li. Safe reinforcement learning via statistical model predictive shielding. 2021. doi:10.15607/rss.2021.xvii.026.

[52] I. ElSayed-Aly, S. Bharadwaj, C. Amato, R. Ehlers, U. Topcu, and L. Feng. Safe multi-agent reinforcement learning via shielding. *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, 1:483–491, 2021. ISSN 15582914.

[53] G. Thomas, Y. Luo, and T. Ma. Safe Reinforcement Learning by Imagining the Near Future. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P. S. Liang, and J. W. Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 13859–13869. Curran Associates, Inc., 2021.

[54] H.-L. Hsu, Q. Huang, and S. Ha. Improving safety in deep reinforcement learning using unsupervised action planning. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 5567–5573. IEEE, 2022.

[55] N. Yokoyama, Q. Luo, D. Batra, and S. Ha. Benchmarking augmentation methods for learning robust navigation agents: the winning entry of the 2021 igibson challenge. *arXiv preprint arXiv:2109.10493*, 2021.

[56] A. Chang, A. Dai, T. Funkhouser, M. Halber, M. Niessner, M. Savva, S. Song, A. Zeng, and Y. Zhang. Matterport3d: Learning from rgb-d data in indoor environments. *International Conference on 3D Vision (3DV)*, 2017.

[57] N. Yokoyama, S. Ha, and D. Batra. Success weighted by completion time: A dynamics-aware evaluation criteria for embodied navigation. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1562–1569. IEEE, 2021.

[58] K. A. Ingraham, C. D. Remy, and E. J. Rouse. The role of user preference in the customized control of robotic exoskeletons. *Science robotics*, 7(64):eabj3487, 2022.

[59] S. Choi, Q.-Y. Zhou, and V. Koltun. Robust reconstruction of indoor scenes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

8 Supplementary Material

8.1 Human-Robot Modeling Data Collection

Figure 7: Data collection of three blind or visually impaired (BVI) subjects.

We collect human and robot interaction data using a Vicon motion capture system with Pulsar active marker clusters (Figure 7). We recruit nine participants consisting of three blind or visually impaired (BVI) and six sighted subjects. After obtaining consent from participants, they are asked to wear a back strap to place a motion capture marker on their lower back. Then, they follow the robot, AlienGo, by using their left hand to hold onto a harness handle that is attached to the robot. For white cane users, we ask them to hold their cane in their right hand for safety purpose, but they did not actively use the cane during data collection.

Table 4: Description of Five Robot Trajectories for Interaction Data Collection

Trajectory	Description
1	2.5 m forward, 90-degree in-place left turn
2	1.2 m forward, 90-degree left turn
3	0.75 m forward, 45-degree gradual left turn, 135-degree in-place right turn
4	0.5 m forward, 90-degree right turn, 0.5 m forward, 90-degree left turn, 0.5 m forward
5	0.6 m forward, 90-degree in-place right turn, 0.6 m forward, 180-degree left u-turn, 0.6 m forward, 90-degree right turn, 0.6 m forward

Robot Trajectories. The robot is scripted to follow pre-defined trajectories as described in Table 4. The five trajectories are based on the work of Nanavati et al. [34] with modifications to fit in the motion capture space.

Experimental Protocol. The data collection process is divided into three parts. Each part is designed to examine different aspects of the interaction behavior. *Part 1:* The subjects are following the robot without vision, but purely relying on the physical interaction without any visual cues. We repeated each trajectory three times in random order. To prevent the subject from predicting the trajectory, they are told that there are 15 trajectories. *Part 2:* The subjects are following the robot without vision, but with prior information on the robot's expected movement. We provide verbal descriptions of the expected cue just before the robot executed each cue. *Part 3:* The sighted subjects are following the robot with their full vision. This is extra data collected only from sighted subjects to respond to the robot's movement with additional visual perception.

8.2 Psuedocode of Action Shielding

This section provides a more detailed algorithm of our action shielding mechanism. We compute the convex hull of the collision primitives first, compute the lidar thresholds from this hull, and then check whether the action is safe or not. Please review this material with Section 5.

Algorithm 1 Convex-hull Action Shielding.

Input current position and orientation of robot and human $\mathbf{x}_t^R$, $\mathbf{x}_t^H$, a list of possible actions $\mathbf{A}$, a vector of action probability $\mathbf{p}$, action suppression scale β, lidar reading $\mathbf{l}$.
Output A modified action probability vector $\hat{\mathbf{p}}$

procedure ACTIONSHIELDING($\mathbf{T}_{\mathrm{r}}^t$, $\mathbf{T}_{\mathrm{h}}^t$, A, s, l)
  for $i \leftarrow 1$ to len($\mathbf{A}$) **do**
    $a = \mathbf{A}[i]$
    $\hat{\mathbf{p}}[i] = \mathbf{p}[i]$
    $\mathbf{x}_{t+1}^R, \mathbf{x}_{t+1}^H \leftarrow$ EstimateNext($\mathbf{x}_t^R$, $\mathbf{x}_t^H$, a) ▷ Estimate Via Interaction Model
    $\mathcal{S} \leftarrow$ ConvexHull($\mathbf{x}_t^R$, $\mathbf{x}_t^H$, $\mathbf{x}_{t+1}^R$, $\mathbf{x}_{t+1}^H$)
    $\bar{\mathbf{l}} \leftarrow$ ComputeLidarThresholds ($\mathcal{S}$)
    if any of reading $\mathbf{l}$ is less than the threshold $\bar{\mathbf{l}}$ **then**
      $\hat{\mathbf{p}}[i] *= \beta$
    end if
  end for
  $\hat{\mathbf{p}} \leftarrow$normalize($\hat{\mathbf{p}}$)
  return the modified action probability $\hat{\mathbf{p}}$
end procedure

8.3 Real World Navigation Experiments

We include here the map information for our real world experiments (Figure 8).

Figure 8: Illustration of three routes in our experiments.