How to Train Your Guide Dog: Wayfinding and Safe Navigation with Human-Robot Modeling

J. Taery Kim[1], Wenhao Yu[2], Jie Tan[2], Greg Turk[1], Sehoon Ha[12]

[1] Georgia Institute of Technology, [2] Robotics at Google

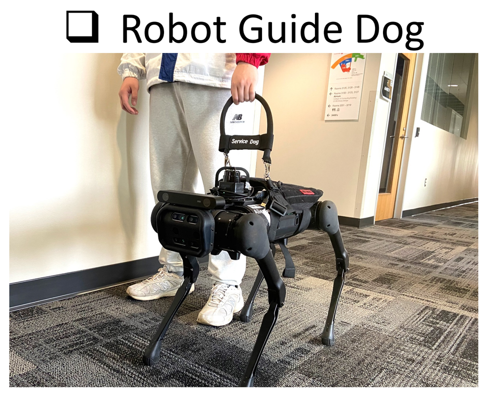
I. Motivation

- We aim to **enhance the quality of life** of blind or visually impaired individuals (BVI) by developing a **robotic guide dog**.
  - Understand **guidework** of guide dog and formally define the task
  - Present a **human-robot model** for safe navigation

❑ Guide Dog

😃 Assist BVI's safe navigation
😢 Time and resource intensive for training
😢 Not all dogs are successfully trained

❑ Robot Guide Dog



Assist BVI's safe navigation 😃
Learned skills can be transferred to other robots 😃
Can be deployed immediately 😃

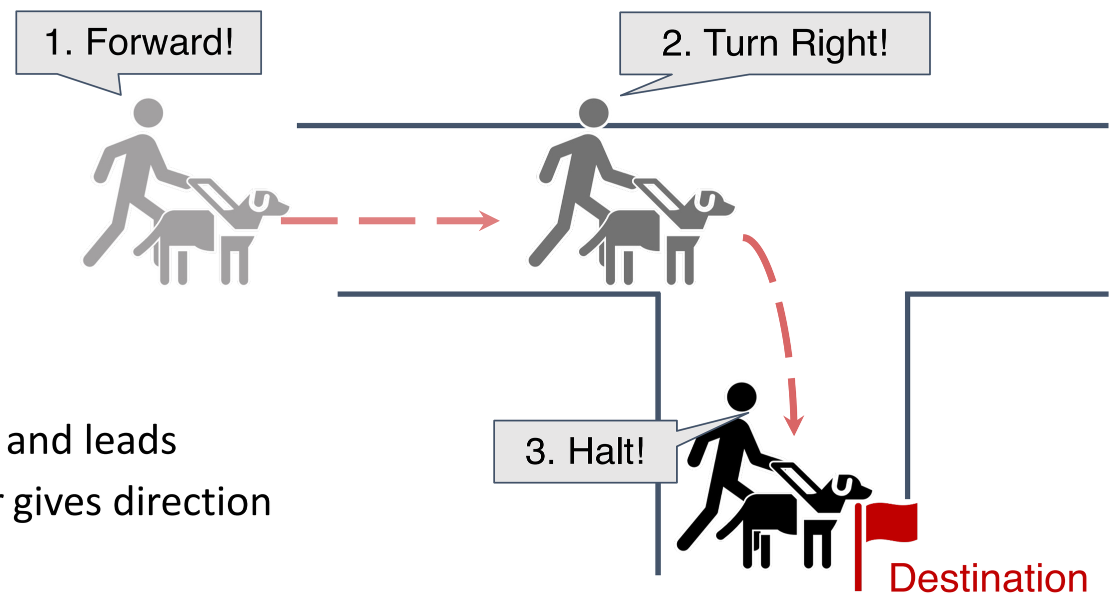
II. Task Design

Background: Guidework Mechanism

**Guide Dogs for the Blind: Meeting a New Guide Dog*
**Interview with guide dog handlers*

- Guide dog and human handler work as a **team**.
- It is the handler's responsibility to keep track of navigation and give high-level directional cues.
- The guide dog assist by leading the local path, following given directions while avoiding collisions.

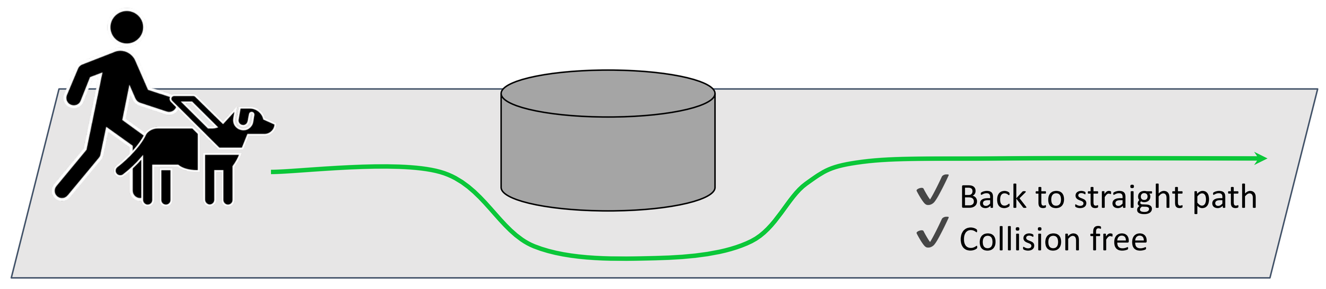
Guidework Example:



❑ Basic Principle

The *guide dog* sees and leads
The *human handler* gives direction

Wayfinding Task Definition

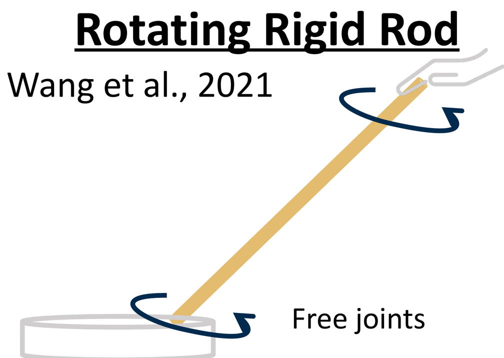


- Follow high-level human directions (cue commands)
- Assist human's navigation: Straight path and Safe path
- MDP Formulation:
  - State: sensor configurations
  - Action: commands to robot
  - Reward: $r_t = \alpha \cdot (\theta_t^{target} - \theta_{t-1}^{target}) + d_t - d_{t-1} - c_t^{collide} - \lambda$
    - Encourage to reach and maintain target rotation, travel further
    - Penalize collision and time

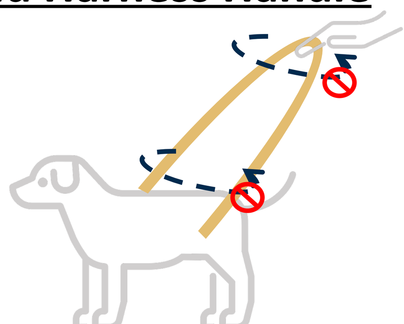
III. Human-Robot Modeling

Rotating Rigid Rod

Wang et al., 2021



Rigid Harness Handle

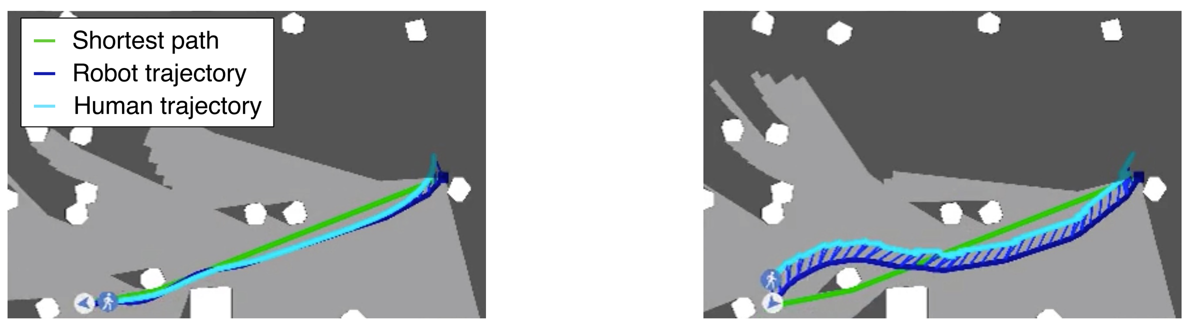


❑ Human position update

$$\begin{cases} x_t^H = \frac{r}{d_t} \cdot x_{t-1}^H + \left(1 - \frac{r}{d_t}\right) \cdot x_t^R, \\ y_t^H = \frac{r}{d_t} \cdot y_{t-1}^H + \left(1 - \frac{r}{d_t}\right) \cdot y_t^R, \\ d_t = \sqrt{(x_{t-1}^H - x_t^R)^2 + (y_{t-1}^H - y_t^R)^2} \end{cases}$$

$$\begin{cases} x_t^H = x_t^R + r\cos\theta_t, \\ y_t^H = y_t^R + r\sin\theta_t \end{cases}$$

❑ Rigid models comparison in simulation

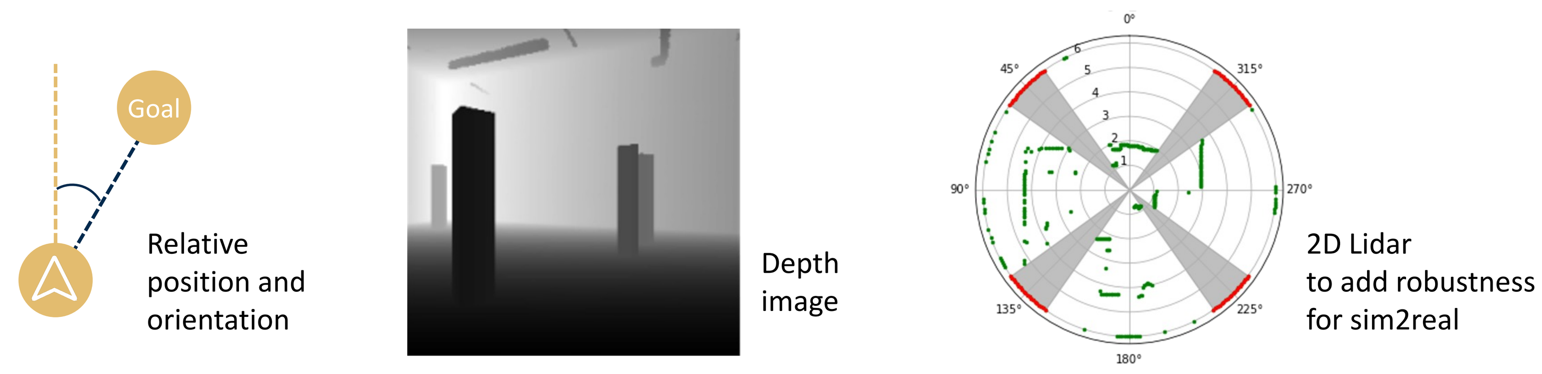


- Rotating rigid rod (left):
  - Human's motion path *follows* that of the robot
  - While robot makes stationary rotation, human *maintains* the position
- Rigid harness handle (right):
  - Human walks *side by side* and *rotates together* with robot

IV. Experiments

Training in Simulation

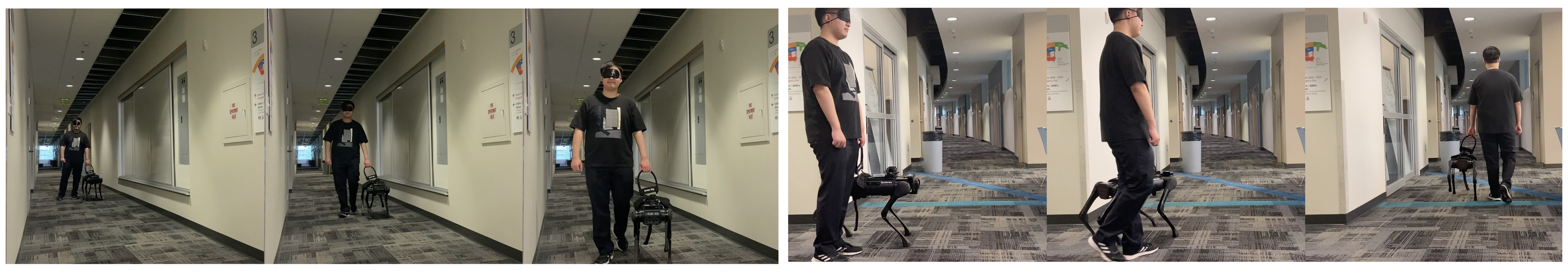
- The wayfinding policies are trained in simulation using reinforcement learning.
- RL Algorithm: Decentralized Distributed Proximal Policy Optimization (DD-PPO), Wijmans et al., 2020.



Basic Cues

- Deploy the learned policies in real world environment
- The policies can follow the basic cues and stay in a straight line
- Basic cues tested:
  - *Forward*: Maintain consistent pace and drive
  - *Turn Left/Right*: Lead in a 90 degree turn and pick up the new travel line on left/right

**Guide Dogs for the Blind: Training Phase*

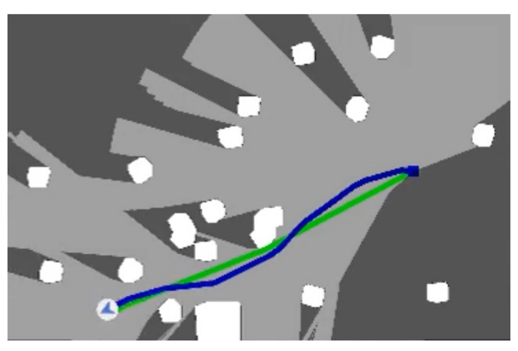


(a) Cue: Forward (b) Cue: Turn Left

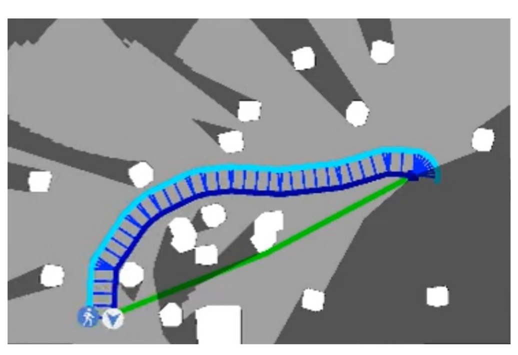
- The policy trained without a proper human-robot model avoids collisions ONLY for the robot itself.
  - Finds a short path
- On the other hand, the policy trained with the human-robot model, the team successfully arrives at the goal point.
  - Takes a detour for safety considering both robot and human

Why do we need human-robot modeling?

❑ In Simulation



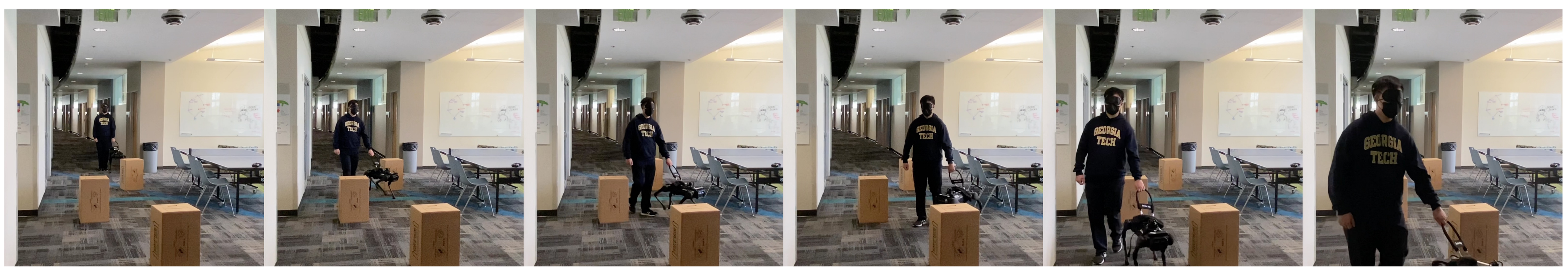
(a) w/o human model



(b) with human model

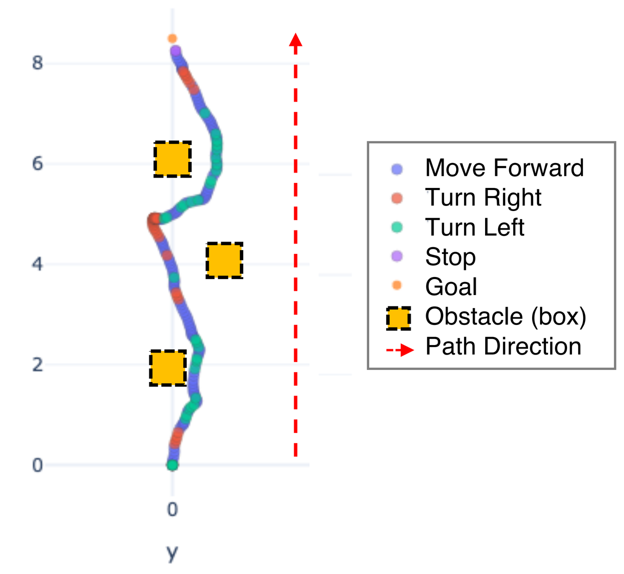
❑ Real-World Demo

(a) w/o human model



(b) with human model

Trajectory Visualization



✓ Back to straight path
✓ Collision free for both

V. Conclusion and Future Work

- We present **a learning framework** for training an effective robotic guide dog based on **guidework** of real guide dogs.
- We define the **wayfinding** task and model **human-robot interaction** for safe navigation.
- We train navigation policies using **reinforcement learning** and deploy the policy to **unseen real-world environments**.
- In the future, we plan to develop more accurate human-robot interaction models using **motion-capture data** and learn safer policies.

Visit our project page☺

Acknowledgements This work was supported by Google Research Collabs and Google Cloud. JK was supported by the NSF GRFP under Grant No. DGE-2039655.