# What Goes Bump in the Night: Learning Tactile Control for Vision-Occluded Crowd Navigation

Kyle Morgenstein[1,C] Junfeng Jiao[2] Luis Sentis[1]

Broadening the scope of social environments in which robots can be reliably deployed requires understanding how to safely navigate in contact-prone environments. While collision-free social navigation is well studied, navigation planners that incorporate safe contacts remain largely unexplored. Traditional social navigation schemes require the robot to stop short if a collision is imminent [1]. "Freezing" the robot while navigating in a crowd may cause people to trip and fall over the robot, causing more harm than the collision itself. In very dense social environments where contacts are common (e.g. public transit, narrow corridors, doorways, etc.), this control scheme would render the robot stationary and unable to traverse the environment, which would in turn prevent robots from ever being deployed in densely populated locales. Thus, if we wish to deploy robots in crowded human spaces, planning safe contacts is imperative. A further challenge is that people in the most crowded environments will occlude traditional exteroceptive sensors that are closer than the sensor's minimum resolving distance. To overcome these limitations, we propose a learning-based motion planner and control scheme to navigate ultra-dense social environments using safe contacts for an omnidirectional mobile robot. Detailed information and additional materials can be found on our project page: https://sites.google.com/view/icra-blind-social-nav.

The use of contacts in navigation planning for social spaces remains broadly under-studied. One approach to overcome the freezing-robot problem in crowds formulates a dynamical system-based controller to slide around the point of contact [2]. This results in an effective reactionary controller with maximum contact force guarantees much below human injury-causing thresholds [3]. However, this control scheme requires free space for the robot to slide into and permits only a single point of contact with the robot - both assumptions that may not be satisfied in the wild. This reactive control scheme was deployed on a standing mobility vehicle in social environments with densities of less than one person per square meter (pp$/m^2$) [4]. Another study [5] mounts a robot arm on a mobile base and explores the use of intentional touches and contacts to encourage people to take a step, clearing a path for the robot. However, this study only investigates environments with a single human and ample free space to step into. These initial solutions to model-based contact-aware navigation are rigid and impose a number of assumptions that may make them unsuitable for deployment in subways, entertainment venues, busy corridors, or other locations with high density traffic.

Alternatively, learning-based and data-driven methods have found success in collision avoidance throughout the social navigation literature [6]. While some methods learn to avoid collisions from a model-based or mechanistic description of human behavior such as the Social Force Model [7] [8] [9] [10] [11], others estimate human behavior without the use of explicit models [12] [13] [14]. However, learning-based approaches for contact-based navigation are largely unexplored. One study that applies learning methods to contact-based navigation in dense social environments was able to plan safe trajectories in crowds [15], but struggled in environments with crowd densities greater than 1 pp$/m^2$. Specifically, failure occurred when the visual sensors of the robot were occluded due to the density of the crowd or limitations in the minimum sensing distance of the sensors. In the highest density crowds, visual sensors are often occluded and are therefore unreliable for planning safe paths.

To safely navigate in crowds, we devise a learning-based motion planner and control scheme that effectively navigates dense social environments without the use of visual sensors. Critically, our planner makes no assumptions about the environment, nor does our planner prescribe an "optimal" contact. Instead, the local planner learns an implicit representation of the contact dynamics, and uses this model to estimate desired velocities and headings. Given the difficulty of explicitly modeling inter-crowd interactions and the limitations posed by occlusions to the sensors, this generality allows our approach to successfully navigate crowds of higher density than has been previously reported.

We formulate contact-based social navigation as a multi-task reinforcement learning problem. We consider the task as 1) follow a coarse set of waypoints from a global planner, and 2) minimize discomfort to humans. We define discomfort as the ratio between measured contact forces and the average pain threshold for blunt impacts between a human and a robot, as established by the ISO 15066:2016 standard. The average pain threshold for the lower legs is 130 Newtons [3]. By expressing contact forces as a ratio, the planner learns to implicitly reason about what sorts of contacts should be considered tolerable. Formulating the task in this way avoids prescribing collision avoidance behaviors and collision resolution strategies a priori. We train the policy via proximal policy optimization (PPO) [16] to output desired velocities and headings.

The local planner is modeled as a neural network policy. In particular, an observation of the robot's state and observed

[1]Department of Aerospace Engineering and Engineering Mechanics, University of Texas at Austin. [2]School of Architecture, University of Texas at Austin. [C]Corresponding Author: kylem@utexas.edu.

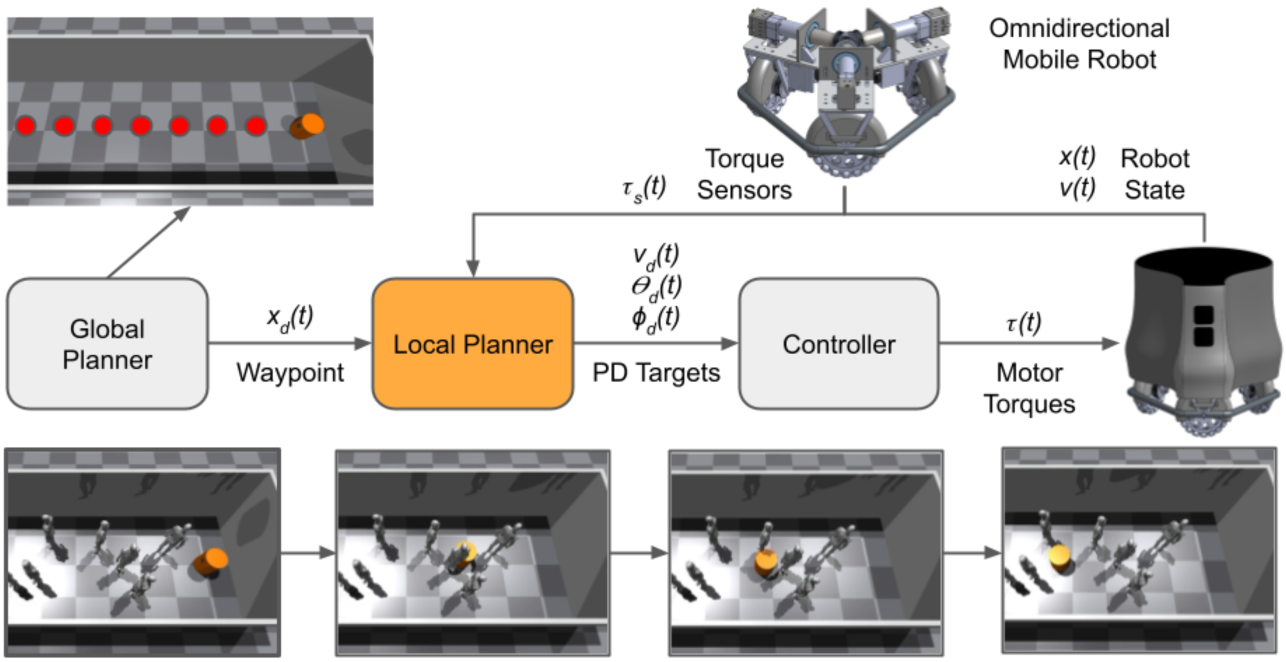


Fig. 1. Vision-occluded tactile control. A global planner (e.g. A*, RRT, etc.) sends waypoints to the local planner. The local planner also receives robot state information and contact forces. The local planner then outputs desired speed, motion heading, and orientation, which produces the desired velocity in world coordinates as $\dot{\mathbf{x}}_d(t) = v_d \left(\begin{smallmatrix}\cos\theta_d \\ \sin\theta_d\end{smallmatrix}\right)$. A low-level controller then produces motor torques to stabilize the robot's velocity and heading around the desired setpoint.

contact forces is passed through a linear projection layer before being passed to the multilayer perceptron backbone of the policy. We design the policy to be relatively shallow so that it can be evaluated in real time on the limited robot hardware. The policy has three outputs: desired speed, desired motion heading, and desired orientation. Note that the orientation of the robot and direction of motion are decoupled due to the omni wheels.

We train the local planner in simulation using NVIDIA IsaacGym [17]. Environments are procedurally generated to mimic hallways in the Anna Hiss Gymnasium robot testing facility. During training, each environment is populated by simulated humans with deformable joints at a density of 1 pp$/m^2$. Goal locations are selected five meters in front of the robot, and an A* path [18] is generated at runtime from a map of the empty environment. 2048 environments are trained for 3.28 billion total training steps with 5 mini epochs per rollout. The control frequency is 10 Hz while the physics frequency is 100 Hz with 4 substeps, resulting in 400 Hz physics updates. Training on a workstation NVIDIA 3080 12GB GPU takes 5 days. The reward function consists of three terms: one term minimizes the distance and heading to the next waypoint, one minimizes the contact force ratio, and one minimizes the slew rate, which encourages smooth policy outputs.

The policy is evaluated in simulation over 160 trials with crowd densities varying between 0 and 1.75 pp$/m^2$. See Table I for a summary of results compared to [15][1]. We evaluate the policy with three metrics: safety factor, completion percent, and time to completion. The safety factor is defined by the percent of trials for a given crowd density in which the robot did not violate the safe contact constraint (130N). Completion percent is defined as the percent of trials for a given crowd density in which the robot successfully reached the target location. In trials with the highest densities, some crowd configurations had no possible safe paths to the target. Time to completion is the mean time of robot travel for a given policy to reach completion.

TABLE I

| Density | Method | Safety | Completion | Time ($\sigma$) |
|---|---|---|---|---|
| $<$1.0 pp$/m^2$ | [15] | 96% | 91% | 11.28 (2.32) |
| | Ours | 100% | 100% | 20.33 (1.66) |
| $\geq$1.0 pp$/m^2$ | [15] | 77.5% | 42% | 15.20 (6.52) |
| | Ours | 90% | 77% | 22.54 (5.50) |

In crowds with density less than 1 pp$/m^2$, we achieve a 100% safety factor and completion percent, beating state of the art by 4% on safety factor and 9% on completion percent. This increase comes at the cost of completion time,

[1]This work is a continuation of preliminary results that we presented at the 2nd Workshop on Human-Interactive Robot Learning (HIRL) at the 2023 ACM/IEEE International Conference on Human Robot Interaction. While the problem statements are similar, we present a distinct method to deal with the problem of sensor occlusions in very crowded environments.

which rises 80.2%. This longer mean path can be attributed to a slower maximum speed. Without visual observations, the policy must be reactive. The result is less aggressive navigation plans than the policy presented in [15]. Where [15] accelerates into free space, our policy maintains a consistent average velocity low enough that initial contact forces stay below the safety threshold. In the most crowded environments with densities greater than 1 pp$/m^2$, we achieve a 90% safety factor and a 77% completion percent, beating state of the art by 12.5% and 35%, respectively. As in the lower density case, these improvements come at the cost of a 48.3% increase in completion time. Both the increase in completion time and safety factor can largely be attributed to the slower maximum speed. However, the significant improvement in completion percent and modest improvement in safety factor are additionally benefited by the removal of the visual observation present in [15].

Ironically, the lack of information about the robot's environment produces a safer navigation plan, as the robot makes no assumption about future contacts given the current observation, and more acutely responds to sensed contacts. By comparison, a common failure mode of [15] occurs when the policy commands an aggressive acceleration because the observation indicates that the path is clear, only to run over the foot of a simulated human that is out of the depth camera's field of view. An additional failure mode of [15] occurs when the robot wedges itself between two simulated humans. The depth camera shows a clear path ahead with minimal steady state contact forces, and so the policy commands an acceleration, causing either a safety failure or for the robot to get stuck in the configuration until timeout, failing to complete the navigation task. By comparison, our policy learns a tip-toeing behavior whereby the robot attempts to slowly accelerate forwards, but, upon contact, attempts to accelerate to the left and right until a free path is found. This results in a dramatically higher completion percent, and the navigation strategy is less aggressive than the one learned in [15], leading to a higher safety factor as well. The only recorded instances of failure to meet the safety standards is caused by a secondary collision. After the robot makes a compliant contact with a simulated human, the robot accelerates aggressively in the direction opposite of the impact normal to decelerate and decrease the impact force. When another simulated human is directly adjacent to the first, this aggressive reversing behavior may cause the robot to contact the second human with enough force to violate the safety constraint. This behavior only occurs in crowds of density equal to or greater than 1 pp$/m^2$. The robot was trained in simulated crowds of 1 pp$/m^2$, and so this aggressive over-correcting acceleration may be mitigated by training in environments of higher density where such interactions are likely to occur more often. We will validate these simulation results on our research platform, Bumpybot, a mobile base with torque sensing omni wheels. The torque sensors are used to estimate the orientation and magnitude of contact forces on the body of the robot, making Bumpybot ideal for studying navigation in collision-prone environments [19]. There are many exciting opportunities to include robots in crowded human environments, and we hope this study will help explore how to best embrace contacts in these spaces.

## References


[1] A. J. Sathyamoorthy, U. Patel, T. Guan, and D. Manocha, "Frozone: Freezing-free, pedestrian-friendly navigation in human crowds," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4352–4359, 2020.

[2] D. Paez-Granados, V. Gupta, and A. Billard, "Unfreezing social navigation: Dynamical systems based compliance for contact control in robot navigation," *arXiv preprint arXiv:2203.01053*, 2022.

[3] I. ISO, "Ts 15066: 2016 robots and robotic devices. collaborative robots," *Geneva, Switzerland: International Organization for Standardization*, 2016.

[4] D. Paez-Granados, Y. He, D. Gonon, D. Jia, B. Leibe, K. Suzuki, and A. Billard, "Pedestrian-robot interactions on autonomous crowd navigation: Reactive control methods and evaluation metrics," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 149–156.

[5] M. C. Shrestha, Y. Nohisa, A. Schmitz, S. Hayakawa, E. Uno, Y. Yokoyama, H. Yanagawa, K. Or, and S. Sugano, "Using contact-based inducement for efficient navigation in a congested environment," *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 456–461, 2015.

[6] R. Mirsky, X. Xiao, J. Hart, and P. Stone, "Conflict avoidance in social navigation–a survey," *arXiv preprint arXiv:2106.12113*, 2021.

[7] D. Helbing and P. Molnar, "Social force model for pedestrian dynamics," *Physical review E*, vol. 51, no. 5, p. 4282, 1995.

[8] G. Ferrer, A. Garrell, and A. Sanfeliu, "Robot companion: A social-force based approach with human awareness-navigation in crowded environments," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2013, pp. 1688–1694.

[9] C. I. Mavrogiannis and R. A. Knepper, "Interpretation and communication of pedestrian intentions using braid groups," in *Workshop on Intention Recognition in Human-Robot Interaction, 11th ACM/IEEE International Conference on Human Robot Interaction (HRI)*, 2016.

[10] J. F. Fisac, E. Bronstein, E. Stefansson, D. Sadigh, S. S. Sastry, and A. D. Dragan, "Hierarchical game-theoretic planning for autonomous vehicles," in *2019 International conference on robotics and automation (ICRA)*. IEEE, 2019, pp. 9590–9596.

[11] D. Fridovich-Keil, A. Bajcsy, J. F. Fisac, S. L. Herbert, S. Wang, A. D. Dragan, and C. J. Tomlin, "Confidence-aware motion prediction for real-time collision avoidance1," *The International Journal of Robotics Research*, vol. 39, no. 2-3, pp. 250–265, 2020.

[12] V. Tolani, S. Bansal, A. Faust, and C. Tomlin, "Visual navigation among humans with optimal control as a supervisor," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2288–2295, 2021.

[13] S. Bansal, V. Tolani, S. Gupta, J. Malik, and C. Tomlin, "Combining optimal control and learning for visual navigation in novel environments," in *Conference on Robot Learning*. PMLR, 2020, pp. 420–429.

[14] M. Seo, R. Gupta, Y. Zhu, A. Skoutnev, L. Sentis, and Y. Zhu, "Learning to walk by steering: Perceptive quadrupedal locomotion in dynamic environments," in *arXiv preprint arXiv:2209.09233*, 2022.

[15] K. Morgenstein, J. Jiao, and L. Sentis, "Learning contact-based navigation in crowds," *2023 $2^{nd}$ Workshop on Human-Interactive Robot Learning (HIRL) at ACM/IEEE International Conference on Human-Robot Interaction*, 2023.

[16] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[17] V. Makoviychuk, L. Wawrzyniak, Y. Guo, M. Lu, K. Storey, M. Macklin, D. Hoeller, N. Rudin, A. Allshire, A. Handa, and G. State, "Isaac gym: High performance gpu-based physics simulation for robot learning," 2021.

[18] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.

[19] K. S. Kim, T. Llado, and L. Sentis, "Full-body collision detection and reaction with omnidirectional mobile platforms: a step towards safe human–robot interaction," *Autonomous Robots*, vol. 40, no. 2, pp. 325–341, 2016.