

MASTER'S THESIS

for

Sergej Stanovcic

Student ID 11726241, Degree Embedded Systems

Real-World Human-Robot Interaction Behavior Generation using Consistency-ModelsProblem description:

Understanding human behavior is a long-standing challenge in the AI and robotics community, involving the comprehension of complex, context-dependent actions and intentions. In the context of social interactions, the movements of individuals reflect their behavior and intentions. As humans, we predict the future movement and state of a human in the short-term future to optimize for fluent interaction. However, transferring this behavior to robots remains a challenge. This project envisions the generation of robust robot behaviors for a proper HRI.

In particular, the student will extend our previous work for deterministic robot behavior generation in HRI [1] by taking advantage of the high-end qualities observed in recent Denoising Diffusion Probabilistic Models (DDPM) in human motion generation [2]. Still, to overcome the slow inference speed of DDPM due to the recurrent denoising steps, the student will explore the use of Consistency Models (CM) for real-time inference [3]. The end goal is to generate HRI behaviors that closely resemble human-human interactions by focusing on the fidelity of the robot behavior, diversity, and robustness to real-world occlusions. Given that the end goal is to test it in real scenarios, the student will employ a synthesized occlusion strategy during training to make the model robust to wrong pose estimations that might occur in the real world.

Tasks:

- Task 1. Familiarization with DDPM and CM by adapting the code from [3] to our HRI task [1].
- Task 2. Training and evaluation of the results for HHI and HRI in the InterGen dataset [4] and, optionally, InterAct dataset [5].
- Task 3. Extension of the baseline model to account for occlusions on the source motions
- Task 4. Ablation study to improve the noise strategy in the motion domain.
- Task 5. Test in Real-world Human-Robot Interactions using the TIAGo++ robot.

Bibliography:

- [1] Esteve Valls Mascaro, Yashuai Yan, and Dongheui Lee. Robot interaction behavior generation based on social motion forecasting for human-robot interaction. *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024.
- [2] Hyemin Ahn, Esteve Valls Mascaro, and Dongheui Lee. Can we use diffusion probabilistic models for 3d motion prediction? In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9837–9843. IEEE, 2023.
- [3] Wenxun Dai, Ling-Hao Chen, Jingbo Wang, Jinpeng Liu, Bo Dai, and Yansong Tang. Motionlcm: Real-time controllable motion generation via latent consistency model. *arXiv preprint arXiv:2404.19759*, 2024.
- [4] Han Liang, Wenqian Zhang, Wenxuan Li, Jingyi Yu, and Lan Xu. InterGen: Diffusion-based multi-human motion generation under complex interactions. *arXiv preprint arXiv:2304.05684*, 2023.
- [5] Kushal Kedia, Atiksh Bhardwaj, Prithwish Dan, and Sanjiban Choudhury. Interact: Transformer models for human intent prediction conditioned on robot actions. In *IEEE International Conference on Robotics and Automation*, 2024.

Supervisor: M.Sc. Esteve Valls Mascaro
Start: June 2024
Intermediate Report: September 2024
Delivery: December 2024

(D. Lee)
Univ.-Professor