Hummus Team Description Paper
RoboCup@Home 2024

Chadi Salmi, Corrado Pezatto, Forough Zamani, Luzia Knoedler, Max Spahn, Saray Bakker, Yujie Tang and Martijn Wisse

November 27, 2023

Abstract. This paper provides an overview of the robot Albert and the innovative strategies devised by the Hummus team for the Open Platform League (OPL) 2024 competition held in Eindhoven, Netherlands. Albert, a robotic arm built on the Franka Emika platform and affixed to a mobile base, serves as the focal point of our efforts. Our team, consisting of PhD students from the Cognitive Department of the Technical University of Delft, embarked on its inaugural participation in the RoboCup@Home league. Addressing key challenges inherent to the RoboCup@Home domain, namely social navigation, safe manipulation, and adaptive task assignment, we present our novel approaches developed to tackle these issues. Leveraging our collective expertise, we sought to overcome obstacles and contribute to the advancement of robotic capabilities in domestic environments. Motivated by the unique opportunity provided by RoboCup to rigorously test our scientific methodologies and research findings, our involvement in the competition signifies a pivotal moment for the Hummus team. As we navigate this competitive arena for the first time, we aim to showcase the potential of our innovative solutions and contribute to the broader scientific community.
Website: `https://tud-hummus.github.io/`

1 Introduction

Hummus is an innovative and dynamic team comprised of Ph.D. scholars from the Cognitive Robotics department at TU Delft. As pioneers in the field of cognitive robotics, our team brings together a wealth of knowledge, passion, and expertise to tackle the intricate challenges of the RoboCup@Home league. Our journey is rooted in the belief that robotics, particularly in complex world scenarios, requires not only technical excellence but also a deep understanding of cognitive processes and human interaction.

At the heart of the Hummus team is the Cognitive Robotics (CoR) department at TU Delft, known for its commitment to advancing the frontiers of robotics by exploring high-level solutions to real-world problems. As Ph.D. candidates within this department, we are driven by a shared vision—to revolutionize robotics by developing intelligent systems capable of navigating the complexities of everyday environments. Our focus extends beyond mere technical functionality; we strive to imbue our robots with cognitive capabilities, enabling them to comprehend, learn, and interact seamlessly with the human world.

In a landscape marked by high-level challenges, we embrace the opportunity presented by the RoboCup@Home league. The RoboCup@Home league focuses on advancing and implementing autonomous service and assistive robot technology, which is crucial for future applications in personal domestic settings [1]. This competition serves as a testing ground for our research-driven innovations, where our robots are not only tasked with performing intricate tasks but also with understanding and responding to the diverse and unpredictable nature of human environments. As members of the Hummus team, we are excited to showcase our dedication to pushing the boundaries of cognitive robotics and providing real-world solutions to the complex scenarios posed by the competition. In the following parts of the document, we are going to elaborate more on our robot's features and our accomplishments.

2 Our team's main principles and accomplishments

The robotics team's approach to mobile manipulation is guided by a philosophy rooted in simplicity, collaboration, and the integration of human and machine capabilities. Embracing simplicity as the cornerstone of reliability, the team advocates for straightforward solutions to enhance robot robustness, ease maintenance, and streamline troubleshooting. They actively leverage human cognitive abilities in decision-making processes, avoiding sole reliance on autonomous systems and facilitating the transfer of implicit knowledge to the robot without extensive coding.

Programming excellence is a focal point for the team, emphasizing the importance of clean code for reliability and adopting widely used techniques to maintain code bases. This includes the incorporation of continuous integration and SCRUM principles into their weekly routines. Careful hardware selection is paramount, aiming for a symbiotic relationship where hardware choices intentionally simplify software design. This approach not only streamlines development but also enhances the adaptability and upgradability of robotic systems.

The team's philosophy extends to promoting cooperation over competition in robotics. They design robots to be cooperative partners with humans, emphasizing a seamless collaboration that amplifies the strengths of both parties. This vision aligns with their broader goal of technology serving as a tool to enhance human potential. By adhering to these principles, the team seeks to excel in competitions while contributing to the advancement of robotics with solutions that are innovative, reliable, and human-centric.

Team Achievements, Participations and Collaborations: In this part, we are going to highlight additional achievements and collaborations that underscore the capabilities and collaborative spirit of our robotics lab at TU Delft. These accomplishments not only showcase our commitment to excellence but also reflect our dedication to advancing the field of robotics through hosting innovative competitions.

- **Continuous Demo at TU Delft Campus:** Our team has successfully conducted a continuous demonstration within a small mock supermarket at the TU Delft campus over the past six months. This ongoing showcase, held multiple times per week, is a testament to our team's collaborative efforts in fine-tuning our integrated mobile robot platform for real-world applications.
- **Robothon Participation:** Participating in the Robothon competition, our team showcased a compliant robot teaching pipeline with a single 7-DOF arm, emphasizing safety and efficiency. This collaborative effort highlighted the interdisciplinary nature of our research, showcasing the expertise of our team members.
- **Incorporation of PhD Research:** The collaborative efforts of our team include the invaluable contributions of Ph.D. researchers within our lab. Their work spans motion planning, decision-making algorithms, and various other domains, enriching our mobile robot platform and positioning it as a versatile and cutting-edge solution.
- **Hosted AIRLab Stacking Challenge:** In a unique endeavor, we hosted the AIRLab Stacking Challenge—a robotics competition centered around stacking items on a shelf using a robotic manipulator arm. The aim was to simulate the process of restocking shelves. This challenge not only showcased our team's ability to organize and host events but also demonstrated our commitment to advancing the field by providing a platform for participants to engage with robotic manipulation tasks.
- **Amazon Picking Challenge:** The team is supervised by professors who led the Delft team that won the 2016 Amazon Picking Challenge.

3 Approaches used in RoboCup@Home challenges

3.1 Social Navigation

To successfully integrate and accept mobile robots in human-centered spaces, prioritizing social navigation is essential. This enhances efficiency and promotes a socially intuitive interaction, considering norms and preferences. Emphasizing the robot's ability to navigate considerately fosters a positive user experience, encouraging widespread acceptance and trust in daily use.

Our social navigation planner is based on the foundation of Model Predictive Control (MPC). This choice of utilizing MPC as the underlying framework for our social navigation system empowers the robot with the capability to dynamically plan and optimize its movements, taking into account not only the environmental constraints but also the intricacies of social interactions. By leveraging the predictive capabilities of MPC, the planner exhibits responsive motions accounting for the predicted future behavior of the surrounding people, e.g. [2]. We specifically build on the MPC formulation for navigation among dynamic obstacles presented in [3]. Furthermore, we perform free-space composition based on the lidar data to derive linear constraints for static collision avoidance. The cost function of the MPC can then be adapted to represent the desired social behavior [4].

In the context of human-centered environments, successful navigation often demands the ability to interact with specific obstacles. e.g. a laundry basket which blocks the robot's path. This concept is often termed as interactive navigation. Unlike mere collision avoidance, interactive navigation entails a more nuanced approach, allowing the robot to engage with obstacles strategically, perhaps by repositioning, to achieve its navigational objectives effectively in settings.

The interactive skill is developed based on the nonprehensile manipulation capability of the mobile base, with the onboard arm serving as the "eyes" for tracking and locating undesired obstacles. Nonprehensile manipulation, characterized by not requiring precise grasping of objects, allows the robot to manipulate objects irrespective of their shape, size, or mass. Specifically, we employ the mobile base to push the object out of its path. Through a thorough analysis of contact conditions during pushing, we have devised a stable pushing approach [5]. To enhance the flexibility of the pushing process, we leverage the physics simulator, Isaac Gym, and employ the sampling-based control method, Model Predictive Path Integral (MPPI), for motion planning during pushing maneuvers. This combination of capabilities enables the robot to navigate dynamically through its environment by intelligently interacting with obstacles and adapting its movements accordingly [6].

3.2 Safe manipulation

Our approach to trajectory generation is based on optimization fabrics [7]. This geometric approach for trajectory generation encodes different behaviors, such as collision avoidance or joint-limit avoidance into differential equations of second order. Using operator from differential geometry, namely pull-back and push-forward, it allows to combine behaviors from different task-manifold into one smooth policy that converges to the goal state.

Our recent adaptation to dynamic environments [8] allows to deploy this approach to human-shared environments. Optimization fabrics offer a versatile framework for trajectory generation in changing environments, because it is highly reactive ($\approx$ 100Hz) and safe. Despite its advantages, optimization fabrics suffer from the same problem as most other trajectory generation methods, such as sampling-based planners: it is incredibly hard to program the logic for grasps of products. Specifically, grasp must often be hand-composed of pre-grasp, grasp and post-grasp poses. We address this shortcoming, by relying on human reasoning and understanding of the scene and the product to be grasped at hand.

Following our philosophy, the human operator can actively teach the manipulator to grasp a certain product (or a class of products) by dragging the robot through the workspace. This approach, often referred to as learning-from-demonstration, is the key for successful grasping in our approach and can seamlessly be integrated with optimization fabrics.

To provide even more safety in human-shared environments, we use a compliant low-level controller for tracking the desired velocity produced by optimization fabrics. Our low-level controller is a simple PID controller in velocity space that can be adapted online if a weight is attached to the end-effector. This choice is well in line with our philosophy of favoring simple solutions of evolved methods if possible.

We have opted for suction gripper as it has shown to be sufficient for all the objects to be grasped so far. Our approach is, in principle, agnostic to the gripper, so we might change the gripper used if needed.

3.3 Adaptive task assignment

The high-level decision making in our robot is also based on novel PhD research [9]. The goal was to create flexible behavior without having to hard-code all of the contingency plans for failed atomic actions. For example, when moving to grasp a supermarket product, the action could fail because the camera might lose sight of the product, because someone moves the product, or because someone manually stops the compliant arm from moving forward.

A regular approach to program robots to handle these contingencies is to create a rich Behavior Tree (BT) [10] containing all fallback behaviors. Our approach is also based on BTs, but we introduce a novel type of leaf node to specify the desired *state* to be achieved rather than an *action* to execute. For example, the BT describes that the robot should be "holding an object" but does not specify the actions to achieve this state, because these change at runtime. These actions are determined at runtime, as explained next.

The resulting BT from our approach is simple to program and it relies on online planning through the (also novel) application of Active Inference [11,12]. Based on neuroscience, Active Inference is a Bayesian inference approach that we use to essentially continuously calculate which of the viable atomic actions has the highest probability of bringing the robot closer to the desired states. This results in continual online planning and hierarchical deliberation. By doing so, an agent can follow a predefined offline plan while still keeping the ability to locally adapt and take autonomous decisions at runtime, respecting safety constraints.

We have used our OPL robot to validate the hybrid Active Inference / Behavior Tree approach [9]. The results showed improved runtime adaptability with a fraction of the hand-coded nodes compared to classical BTs.

3.4 Intelligent perception

Detection Central to our success is an advanced computer vision pipeline that employs deep learning models for product and person detection. The product detection camera, strategically located at the end effector, enables the robot to efficiently identify and interact with grocery items on store shelves. Additionally our attachable perception tower at the rear of the robot, with it's 5 Realsense depth cameras, enables us to detect the full poses of people surrounding the robot, using Yolo based keypoint detection. Notably, our research in few-shot learning allows us to seamlessly integrate new product classes with as few as five images, ensuring adaptability and scalability.

Multi Object Tracking To ensure the precise association and tracking of both products and individuals, we employ the cutting-edge multi-object tracker, ByteTrack. In instances where products and individuals may not be visible for a given number of frames, our custom combination of the Hungarian algorithm for detection association and KalmanFilters becomes invaluable. This dynamic approach allows us to consistently keep track of the states of products and persons, contributing to the robustness of our overall system.

Visual Servoing Our capacity to swiftly detect and track the 6 DoF pose of products in the vicinity of the end effector, utilizing the relatively lightweight Yolo, sets the stage for an innovative approach known as visual servoing. Unlike the traditional sense-plan-act paradigm, this method involves performing detection in a loop, allowing for the continuous updating of tracked states. This real-time adaptation proves invaluable in mitigating noise and handling perception disturbances when the robot is engaged in the intricate task of picking items off the shelf. The implementation of visual servoing underscores our commitment to dynamic, responsive, and efficient robotic interactions in contrast to more conventional planning-based approaches.

3.5 Understanding the human instructions with the help of ChatGPT

To enhance the transparency of our robot's actions and decisions, we've seamlessly integrated ChatGPT alongside a text-to-speech model. Within our decision-making framework, we leverage the Robot Operating System (ROS) to dispatch voice commands to a dedicated ChatGPT node. Employing a tailored prompt, ChatGPT becomes a pivotal component in elucidating the robot's actions through natural language. The final layer of this integration involves a dedicated speaker on the robot, ensuring that the output from the text-to-speech model is not only generated but also audibly communicated, creating a more comprehensive and user-friendly interaction.

References

1. Thomas Wisspeintner, Tijn Van Der Zant, Luca Iocchi, and Stefan Schiffer. Robocup@ home: Scientific competition and benchmarking for domestic service robots. *Interaction Studies*, 10(3):392–426, 2009.
2. Luzia Knoedler, Chadi Salmi, Hai Zhu, Bruno Brito, and Javier Alonso-Mora. Improving pedestrian prediction models with self-supervised continual learning. *IEEE Robotics and Automation Letters*, 7(2):4781–4788, 2022.
3. Bruno Brito, Michael Everett, Jonathan P How, and Javier Alonso-Mora. Where to go next: Learning a subgoal recommendation policy for navigation in dynamic environments. *IEEE Robotics and Automation Letters*, 6(3):4616–4623, 2021.
4. Luzia Knoedler, Bruno Brito, Michael Everett, Jonathan P How, and Javier Alonso-Mora. Learning a guidance policy from humans for social navigation.
5. Yujie Tang, Hai Zhu, Susan Potters, Martijn Wisse, and Wei Pan. Unwieldy object delivery with nonholonomic mobile base: A stable pushing approach. *IEEE Robotics and Automation Letters*, 2023.
6. Corrado Pezzato, Chadi Salmi, Max Spahn, Elia Trevisan, Javier Alonso-Mora, and Carlos Hernandez Corbato. Sampling-based model predictive control leveraging parallelizable physics simulations. *arXiv preprint arXiv:2307.09105*, 2023.
7. Karl Van Wyk, Mandy Xie, Anqi Li, Muhammad Asif Rana, Buck Babich, Bryan Peele, Qian Wan, Iretiayo Akinola, Balakumar Sundaralingam, Dieter Fox, Byron Boots, and Nathan D. Ratliff. Geometric fabrics: Generalizing classical mechanics to capture the physics of behavior. *IEEE Robotics and Automation Letters*, 7(2):3202–3209, 2022.
8. Max Spahn, Martijn Wisse, and Javier Alonso-Mora. Dynamic optimization fabrics for motion generation. *IEEE Transactions on Robotics*, 39(4):2684–2699, 2023.
9. Corrado Pezzato, Carlos Hernández Corbato, Stefan Bonhof, and Martijn Wisse. Active inference and behavior trees for reactive action planning and execution in robotics. *IEEE Transactions on Robotics*, 39(2):1050–1069, 2023.
10. Michele Colledanchise and Petter Ögren. *Behavior trees in robotics and AI: An introduction*. CRC Press, 2018.
11. Karl Friston, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, and Giovanni Pezzulo. Active inference: a process theory. *Neural computation*, 29(1):1–49, 2017.
12. Pablo Lanillos, Cristian Meo, Corrado Pezzato, Ajith Anil Meera, Mohamed Baioumy, Wataru Ohata, Alexander Tschantz, Beren Millidge, Martijn Wisse, Christopher L Buckley, et al. Active inference in robotics and artificial agents: Survey and challenges. *arXiv preprint arXiv:2112.01871*, 2021.

Robot Albert Hardware Description

Our custom robot is a mobile manipulator consisting of a mobile base (ClearPath Boxer: `https://clearpathrobotics.com/boxer/`) and an advanced robotic arm (Franka Emika Panda: `https://franka.de/production`) designed for diverse applications. The specifications are as follows:

- Base: Clearpath Boxer Differential Drive with two actuated wheels and 2D Lidar for SLAM
- Arms: Equipped with integrated force/torque sensors in all joints, the Panda arm allow for compliant controllers ensuring safe and precise interactions with the environment. It has a payload capacity of up to 3 kg, making it suitable for various tasks.
- Customised vacuum gripper with two suction cups and on-board vacuum pump
- RealSense D435 camera mounted on the end-effector

Robot's Software Description

For our robot we are using the following software:

- Platform: Ubuntu 20.04, ROS Noetic
- Localization/Navigation/Mapping: SLAM
- Speech generation: Using ChatGPT for the explainability and speech generation.
- Object recognition: Using YOLO V6.3
- People recognition: Using Keypoint based pose detection (YOLO V5)
- Arms control: Optimization fabrics
- Simulation environment: Gazebo

External Devices

Albert robot relies on the following external hardware:

- Laptop with a 3070 TI GPU

Cloud Services

Albert connects the following cloud services:

- Text to speech and ChatGPT api.

Fig. 1. Robot Albert

Robot software and hardware specification sheet