Human-Supervised Autonomous Navigation of Planetary Rovers in Rough Terrain

Rui Miguel Sousa Abrantes

***Abstract*—In this dissertation, we embark on a journey through the teleoperation systems, delving into the creation and evaluation of a rover prototype named Kayjay. This research focuses on the development of a Model Predictive Controller (MPC) to provide semi-autonomy without prior knowledge of the environment. We used the Avatar interface proposed in previous studies, adapting it to function seamlessly in a realistic environment without latency. The objectives were multifaceted: to construct the Kayjay rover, develop a robust semi-autonomous controller, and integrate this system into a teleoperation framework to test the efficacy of the Avatar interface. Kayjay was meticulously designed for future upgrades, utilizing Docker containers for seamless hardware integration. The MPC, although exhibiting some overshoot, demonstrated the potential for significant advancements in teleoperation tasks. Through a preliminary user study, we explored the user interactions, comparing the performance of the Avatar and Direct modes. The findings revealed that the Avatar mode has potential in specific scenarios despite generally underperforming compared to the Direct mode. Notably, the absence of latency highlighted the preference for the Direct mode due to immediate response times. The study concludes that while the Avatar interface has room for improvement, its application in Hybrid mode indicates promising avenues for future research. The intricate interplay of the MPC and teleoperation interfaces beckons further exploration, suggesting that enhancements in the controller and user interface could certainly elevate the performance of semi-autonomous rovers in planetary exploration.**

***Index Terms*—Semi-Autonomous, Rover, Model Predictive Controller, Teleoperation, Teleoperation Interface**

I. INTRODUCTION

WITH the increasing interest and investment in space exploration, as seen with European contributions to both human and robotic exploration[1] or NASA investments[2], research in this field is expected to make significant progress in the upcoming years. As a result, the interaction between astronauts and robots, known as Human-Robot Interaction, will become increasingly important in space exploration discussions. In environments too dangerous for people to explore, as those with extremely high or low temperatures, high radiation levels, or poisonous atmospheres, teleoperated robots may provide a safer, which consequently will be safer for venturing into unexplored territories, examining geologically regions and gathering samples for scientific investigations. Furthermore, teleoperated rovers may prove to be a more economical choice than sending humans on missions or keeping complex autonomous systems up and running.

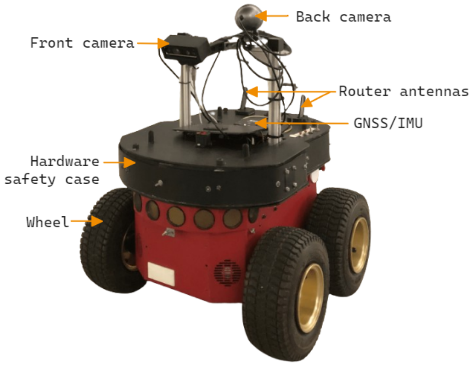

Fig. 1. Kayjay, the prototype of a Rover, this image was taken from [2]

Several organizations are already preparing for the upcoming expeditions as they get ready for future explorations of the Moon or Mars. The ÖEWF[3] is one of those organization. Together with scientists, researchers and other organizations, they have started the AMADEE analog missions. This missions replicate interactions between humans and robots in Mars-like conditions on Earth. The goal of the AMADEE mission is to assess the challenges, constraints, and benefits of future planetary missions. It also examines priorities, the experiences of the first humans on these planets, and the necessary infrastructure.

The goal of this dissertation is to develop a prototype of a rover that can be utilized in the context of planetary exploration, where the rover can be placed in planetary surface (e.g. Mars or Moon) and the operator can be in an habitat or orbiting station (e..g. deep space gateway). This rover will integrate its own semi-automation mechanisms in terms of guidance and path planning and have a teleoperation interface that was designed during the development of this study so that the operator does not get confused and is provided the greatest experience possible. This system took part in the AMADEE24 that happened in Armenia from March 7th to April 5th, 2024, as explained in [1].

II. BACKGROUND AND STATE OF THE ART

A. Kayjay, an UGV prototype

To research the area of exploratory teleoperation rovers, it is necessary to have one, so that everything that is developed can

---

[1] ”In 2023 there will be an increase of 17% of investment” - https://ptspace.pt/pt/financiamento-de-portugal-a-esa-atinge-os-e115-milhoes

[2] ”While the overall LTV Services contract has a maximum value of $4.6 billion over 15 years” - https://spacenews.com/nasa-selects-three-companies-to-advance-artemis-lunar-rover-designs/

[3] https://oewf.org/en/

be tested with a real robot and is not only simulated, therefore the Kayjay, Fig. 1, was developed with the in order to test the implemented controller and to study the usability of the avatar. This prototype is composed by 2 cameras, a router to provide its own network, a GNSS/IMU, and the Pioneer 3-AT[4].

B. Teleoperation

The term teleoperation describes the remote control that is typically a remote-control or other remote interface of a device or system. Teleoperation in the context of a rover often entails using a controller (e.g. gamepad) and some visual interface to remotely control the rover's movement, manipulators, receive information about sensors, and cameras. Space exploration, military operations and search and rescue missions are just a few of the uses for teleoperation, as can be seen in [3]–[5]. It is especially helpful when it is unsafe or impractical for a human operator to be on-site or when the operation requires specialist tools or knowledge that can be controlled remotely.

To better understand teleoperation, it is essential to study Human-Robot Interaction (HRI) and Situation Awareness (SA). HRI is a branch of research that focuses on the interactions between people and robots as well as the best ways for people and robots to cooperate. It entails both the creation of robots that can interact with people naturally and intuitively as well as the creation of interfaces and other technologies that let people efficiently command and communicate with robots. SA was explicitly described in [6] as "a person's perception of the elements of the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future". According to this concept, there are three stages of SA depending on how well one understands the characteristics and situation of the environment: Perception is the initial and most basic level of SA, where a human is capable of perceiving the pertinent data offered by the system; Comprehension is the intermediate state of SA, in which a person is able to comprehend the significance of the information observed; and Projection is the highest degree of SA, since it allows a person to forecast future occurrences.

C. Automation

The idea of developing systems that function without immediate human involvement is known as automation. The term "automation" is frequently used in reference to autonomous vehicles such as self-driving automobiles and drones In recent years, robotics and automation have made significant strides in control systems, therefore, the refinement of control strategies plays a pivotal role in ensuring optimal performance and adaptability of robotic systems. As robots undertake increasingly intricate tasks across diverse operational contexts, the imperative for sophisticated control mechanisms becomes pronounced. Spanning from classical approaches such as Proportional-Integral-Derivative (PID) control to cutting-edge methodologies like Reinforcement Learning (RL) and Neural Network-based controllers, each paradigm offers distinctive perspectives on addressing the exigencies of robotic control.

D. Model Predictive Controller

In this study it will be developed a Model Predictive Controller which is a powerful control methodology widely used in robotics and automation. Unlike traditional control techniques, like PID's, that rely on predefined control laws or models of the system, MPC operates by iteratively solving an optimization problem over a finite time horizon to generate control actions. This predictive approach allows it to effectively handle complex dynamics, constraints, and uncertainties inherent in robotic systems.

III. METHODOLOGY

A. System Description

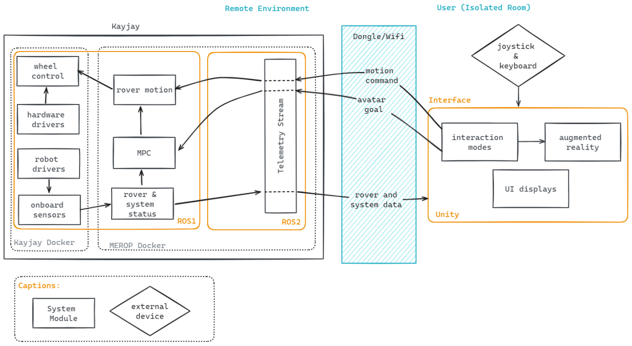

Fig. 2. Modules of the teleoperation system: (1) onboard components ensure navigation, status feedback and remote communication; (2) interface components. This image is was designed for the Thesis, but based of an image in [1].

For the whole system, presented in Fig. 2, to work correctly it is necessary to understand that there are two main components, the **Interface**, that uses Unity (C# and .NET) and the **Rover** (ROS1 and Python). To achieve a system that uses these two components, it is necessary to have a communication system that ensures that all information, either from the **Rover** or the **Interface** is feasible and in real time. In this case it was achieved by using ROS2 because of its robust communication protocols (UDP) and its compatibility with both ROS1 and Unity.

B. MPC

It was desired to develop an MPC that could be robot agnostic, but in reality there are some limitations related to the hardware that the rover possesses. Other variables, such as the robot's dynamic model, also play a crucial role of incapability of being robot agnostic. In this case, since the Kayjay rover is the main source of testing, its dynamic model was the one considered for the MPC, which is the differential drive model. The differential drive model describes the movement of a two-wheeled robot. To derive the differential equation for this model, we need to consider the kinematics of the robot's motion. Let's denote the state of the robot at time t as

---

[4] Manual of the Piooner 3-AT - http://vigir.missouri.edu/ gdesouza/Research/MobileRobotics/Software/P3OpMan5.pdf

(x_t, y_t, θ_t), where (x, y) represents the robot's position in a 2D coordinate system, and θ represents the orientation angle of the robot with respect to a reference axis. The robot's motion is governed by the following kinematic equations:

$$\dot{x} = v\cos(\theta),\ \dot{y} = v\sin(\theta),\ \dot{\theta} = \omega$$

where v is the linear velocity of the robot and ω is the angular velocity.

Now, let's discretize these continuous-time equations using a time step Δt. The position and orientation of the robot at time step k can be approximated as:

$$x_k = x_{k-1} + \Delta t \cdot v\cos(\theta_{k-1}),\ y_k = y_{k-1} + \Delta t \cdot v\sin(\theta_{k-1}),$$

$$\theta_k = \theta_{k-1} + \Delta t \cdot \omega$$

These equations represent how the robot's position and orientation change over one time step Δt in response to the linear and angular velocities. In conclusion, the Kayjay movement can be represented as shown in Eq. 1.

$$f(x_{k-1}, v, \omega) = x_{k-1} + \Delta t \cdot \begin{bmatrix} \cos(\theta) & 0 \\ \sin(\theta) & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} v \\ \omega \end{bmatrix} \quad (1)$$

where x_{k-1} represents the previous state vector, v is the linear velocity, ω is the angular velocity, Δt is the time step, and θ is the orientation angle of the robot. This equation describes how the robot's position and orientation change over time in response to linear and angular velocities. Our next step is to define a suitable cost function for creating a feasible controller. We have opted for a simple cost function that considers two primary factors: the deviation between the robot's current state and its desired goal, and the energy consumption. Hence, the controller's objective is to minimize both the positional/orientation error and energy consumption simultaneously, as outlined in Equation 2.

$$l(x, u) = |x - x_g|_Q^2 + |u|_R^2 \quad (2)$$

The x represents the robot's position, x_g denotes the desired goal position, and u encompasses the control variables, including linear and angular velocity, and in the chosen cost function, denoted as Equation 2, the variables Q and R play a significant role in emphasizing which variables have higher importance. Therefore, both Q and R have the function of indicating the prioritization within the MPC framework: The primary focus lies on achieving the goal, followed by minimizing energy consumption, and finally, ensuring correct orientation. Through an iterative tuning process, the gains for Q and R were determined as follows:

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.005 \end{bmatrix} \qquad R = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.25 \end{bmatrix} \quad (3)$$

These matrices dictate the importance assigned to each component of the cost function, ensuring the controller aligns with desired objectives. During tuning, it was found that the element Q_{33}, which corresponds to the value located in the third row and third column of matrix Q, couldn't be of the same magnitude as Q_{11} and Q_{22} because the MPC would overly minimize orientation and overshoot, never reaching the goal.

The MPC problem is solved iteratively in real-time using an Interior Point Method Optimization solver, specifically the IPOPT implementation within the Casadi library[5], chosen for its robustness and efficiency in handling large-scale optimizations. To ensure robot-agnosticism, the MPC uses a short prediction horizon, with a window size (N) of 100 and a time step (Δt) of 0.1 seconds, synchronized with the system's 10Hz frequency. This results in a 10-second prediction horizon, providing robust foresight while maintaining low CPU usage. For the Kayjay rover, this prediction covers at least 7 meters, considering its maximum velocity of 0.7 m/s and 50 cm length, deemed as safe. Thresholds were set to ensure the MPC reached its target: 20 degrees for orientation and 10 centimeters for position. These thresholds, were refined through iterative adjustments, balance between minimizing overshooting and practical feasibility, accounting for factors like slippage, robot dimensions, and other relevant factors.

C. Avatar

The Avatar-based Approach to Interaction (AAI) concept, as proposed by Rute Luz and Ventura [2], which offers an innovative solution to address the limitations of traditional teleoperation methods. Unlike conventional teleoperation, where the operator directly controls the robot, the Avatar approach involves controlling an avatar model of the robot. Through this model, the operator provides high-level goals to the automation system. This approach was developed and studied in response to the common challenge of multi-second communication delays encountered during long-distance teleoperation tasks, equivalent to a future Earth-to-Moon teleoperation scenario. This study have shown that the Avatar led to a decrease in workload [7] compared to Direct Control, and resulted in higher usability when compared to Direct Control. The Avatar system is divided in two small systems, the automation and the interface.

For the **automation** it is proposed a solution leveraging the MPC to enhance the Avatar approach's adaptability. With this controller, the system achieves semi-autonomous functionality, enabling rover teleoperation in both indoor and outdoor environments without relying on preexisting maps. The MPC, will function as the semi-autonomous capability. However, to ensure its effective operation, a reliable localization system is essential. For this challenge it was implemented two solutions, one that relies on GNSS data and one that relies on Odometry data.

The **interface**, Fig. 3, incorporates various elements essential for task execution. It begins with the presentation of two camera feeds, with the primary camera (that provides the view of what's in front of the robot) occupying the whole screen and in the top left corner can be seen the image from the

---

[5] Casadi Library - https://web.casadi.org/

second camera (that provides information about what's behind of the robot). The rover's operating mode (Avatar or Direct) is prominently depicted in the top right corner, providing critical operational information. Lastly, when in Avatar mode, it is possible to see a replicate of the robot and the path that the MPC is planning to reach it.

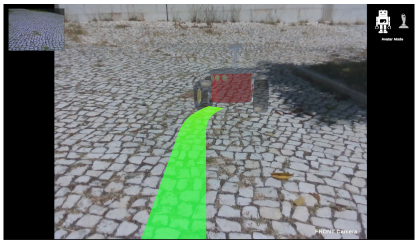

Fig. 3. Interface used for the teleoperation.

IV. EVALUATION

A. MPC

When developing the MPC, it's crucial to assess its feasibility and comprehend its strengths and weaknesses. This entails understanding the intended application of the MPC, as different settings yield varied results. Given the focus on planetary exploration, the terrain typically consists of vast spaces with minimal obstacles and rough terrain. Furthermore, the goal is to utilize the MPC as a semi-autonomous, with the controller's aims closely mirroring those of the teleoperator, who maintains real-time control over the avatar, ensuring it remains within their line of sight. Consequently, the testing scenarios do not need to cover extensive distances, as user interaction would probably not behaviour as such. Moreover, time efficiency is paramount, as prolonged durations could diminish user comfort. Therefore, the MPC will undergo testing based on the following parameters: **Trajectory**; **Behaviour of control variables** assessing their correctness and stability, with minimal oscillations; **Errors of position and orientation** assessing if the errors converge to the desired values; **Conlusion of runs**; **Time to conclude a run**; and **Distance traveled**

To acquire these metrics, several tests need to be conducted. These tests encompass various features, including the capability of the MPC to predict positions and orientations comprehensively, although some features are not necessarily metrically prioritized. Specifically, the tests aim to evaluate the MPC's ability to predict every conceivable position and orientation, though due to time constraints, exhaustive testing is impractical. Therefore, a subset of 32 permutations was examined, each tested twice. As illustrated in the Fig. 4, eight distinct positions were assessed, each at a fixed distance of 2 meters from the origin. The orientations considered were North, South, West, and East, with East corresponding to an angle of 0º in the Yaw axis, while North corresponds to 90º in the Yaw axis. Although not exhaustive, these tests endeavor to simulate a wide range of scenarios that the MPC might encounter during operation. The repetition of the test aims to assess the MPC's consistency and its ability to consistently achieve the goal across various scenarios.

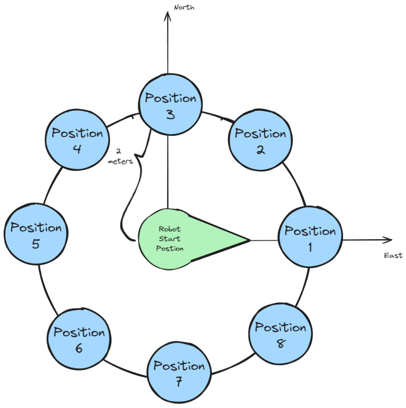

Fig. 4. Positions used as goals for testing the MPC.

It's crucial to acknowledge that all configurations and tuning of the MPC were carried out in a controlled laboratory setting, on flat ground where slippage wasn't a concern. However, for the tests to maintain practical relevance, they were conducted on traditional Portuguese paving, renowned for its significant slippage and uneven surface. Consequently, minor overshooting may arise due to the ground's nature and the localization method employed, which relies on odometry and is susceptible to errors induced by slippage and ground movement. This setup aligns with the conditions outlined in Fig. **??**, thereby ensuring the MPC is tested under realistic conditions.

B. AMADEE24

Before delving into the metrics used, it's crucial to grasp the operational procedures of AMADEE24. Given the multitude of ongoing research activities, each day is meticulously planned in advance and time must be equitably allocated among all teams. Consequently, each run is assigned a specific time slot, with additional slots designated for testing and troubleshooting any errors or setup issues. Since teleoperation relied on a rover, the MEROP team collaborated with SAMPE [8], resulting in slightly reduced time slots due to SAMPE's own runs and the necessity for the robot to recharge batteries.

The Avatar as a whole was part of the MEROP's research during the AMADEE24, although quantitative metrics were not feasible to achieve, qualitative ones were. Therefore a questionnaire was provided at the end of each activity of the MEROP's team. It couldn't be extensive due to the team's own research agenda, and limited time slots but the Analog Astronauts were asked the following:

1) When in avatar mode, if you were the one controlling the robot (like direct mode), would you lead it the same path it went in the experiments ? Participants were asked to rate their response on a scale of 1 to 5, where 1 represents "Strongly Disagree" and 5 represents "Strongly Agree."
2) Participants were asked to rate their stress levels on a scale of 1 to 5, with 1 representing no stress and 5 indicating high levels of stress.
3) Comparing avatar and direct mode when there is network delay, did you prefer using the avatar mode ?

Participants were asked to answer with ”Yes”, ”No” and ”Did not happen”

4) Was the path generated between the robot and the avatar explanatory enough of what the robot would do in terms of movement? Participants were asked to answer with ”Yes”, ”No” and ”Did not happen”

C. User Study

*1) **Design**:* Similarly to the user study done in [2], it was employed a withing-subject design, but due to time constraints, it was only possible to have **10** participants, which all performed three experimental conditions:

- **Direct Control:** which controls directly the robot.
- **Avatar Control:** which controls an augmented avatar overlayed on the image providing information to the semi-autonomous navigation.
- **Hybrid Control:** which provides the possibilty of switching between Direct and Avatar control during the Task.

To mitigate carryover effects inherent in a within-subject design, it was essential to alter the sequence in which participants engaged with the experimental conditions. However, the Hybrid condition necessitated prior participation in both the Direct and Avatar conditions. As a result, the Hybrid condition was consistently scheduled as the final phase of the study and the Direct and Avatar conditions were permuted. Ultimately, the objective of the user study is two-fold. Firstly, to assess the influence of augmented interfaces on operator performance, addressing the following research questions:

What effect does the use of the augmented teleoperation with semi-autonomous system have on:

Q1: the task completion time?
Q2: the robot safety (number of collisions) ?
Q3: the total path length during a task ?
Q4: the workload [9] of the operator during the task ?
Q5: the ease of use of the teloperation interface ?

When using the hybrid interface:

Q6: do operator use one of the teleoperation methods more time than the other ?
Q7: in which situations do user switch between teleoperation modes ?

*2) **Experimental Apparatus**:* To achieve a realistic teleoperation, it was necessary to separate the user from the rover, in a way that the user could not, even if tried, visualize the rover and understand better its environment to accomplish the tasks.

The **teleoperation station control**, is composed by 2 different computers, one for answering the questionnaire, and one for the teleoperation. Since the teleoperation computer monitor is damaged, a monitor was necessary. To teleoperate the Kayjay, it is used the joystick.

Meanwhile, the rover would be in a **remote environment**. This space is a 28 x 24 meters that will be divided into 4 zones, of 14 x 12 meters. 3 of those zones be used for the tasks that will be described in Section IV-C3 and 1 zone for training during the study. This zones are limited by physical barriers (walls) and some imaginary barriers and the users were warned when they were about to ’collide’ with the imaginary barriers. This site is located inside the campus of Instituto Superior Técnico and the satellited view of it and the divisions of space for the study can be seen in Fig. 5.

Fig. 5. Top view, where the Kayjay while be acting during the Tasks of the User Study.

Fig. 6. Left image is the zone number 1 and right image is the zone number 2.

Fig. 7. Zone number 3.

In contrast to the scenario described in [2], where obstacles primarily comprised small objects or other components, the obstacles in this study consisted of water bottles, walls, and grass. It’s worth noting that reaching the grass entailed the rover climbing a step, which was considered a collision if the rover attempted to climb it.

As previously stated, three different configurations of the remote environment were developed, comprising: **Two configurations** for the Direct and Avatar conditions zones 1 and 2, with an open area as well as a confined space where the Kayjay can navigate Fig. 6; **One configuration** for the Hybrid conditions zone 3, characterized by open spaces but with bottle orientations necessitating cautious Kayjay maneuvering to observe the numbers Fig. 7.; and **One configuration** which is the remaining space. This area is only for practice and to ensure that users remain unaware of the other zones.

Since there is a physical structure that separates the teleoperation station control and the Kayjay, it was necessary to use USB cable extensions and a dongle[6] so that the interface can

---

[6] Wifi Dongle - https://www.tp-link.com/pt/home-networking/adapter/archer-t2u/

communicate with the Kayjay. To achieve the best possible coverage area of network, the dongle was position outside. Finally, there will be needed some waterbottles, Fig. 8, that will serve as objects to be looked for, as it was explained in Section IV-C1.

Fig. 8. Numbered water bottles, used during the user study.

*3) **Procedures**:*

Participants: Ten volunteers, aged between 18 and 50 years old, with an average age of 25, took part in the user study. The participants consisted of three females and seven males.

Instructions and Demographic Questionnaire: Prior to the commencement of the user study, all participants were provided with detailed written instructions outlining the study's objectives, experimental procedures, and the equipment used. Upon reviewing these instructions, participants were required to sign an informed consent form, granting permission for the potential publication of experimental data. Furthermore, participants completed a demographic questionnaire, which solicited information on their age and various aspects such as their frequency of traveling to new routes, use of teleoperated devices, and utilization of gamepads for playing video games or other activities.

Training Sessions Before each experimental trial, participants underwent a learning and training session. Within this session, they received instruction on the controls of the gamepads, the behavior of the robot, and the teleoperation interface. Additionally, before each experimental condition, the specific interaction method and its impact on the robot's movement were demonstrated. Subsequently, participants practiced using the teleoperation interface until they felt adequately confident to begin the experimental trial, with a minimum practice duration of 3 minutes.

Inspection Task When determining the experimental task type, several considerations came into play. Firstly, it was essential for the task to ensure fair evaluation across all conditions. Given that the interfaces influence robot movement, whether direct or semi-autonomous, the task needed to involve navigating the robot within a remote environment. Secondly, since the interface primarily enhances the robot's image stream, it was crucial to ensure that participants focused on this image for robot control. Thus, the task involved searching for water bottles within the environment to direct participants' attention to the image. Thirdly, since it was used odometry as localization, a restart of the robot was done for each zone, so that the shift in its referential would initially be the same for everyone. Lastly, to provide participants with a realistic and engaging challenge. For these reasons, participants performed an inspection task during the experimental trials. They were required to navigate through the remote environment and locate three numbered water bottles Fig. 8. Each task had a time limit of 13 minutes before automatically ending. The number of water bottles and the time limit were determined through iterative pilot tests.

Post-Trial Questionnaire Following each experimental trial, participants completed a NASA-TLX questionnaire [10] to gauge the workload associated with the task, along with a truncated version of the USE questionnaire [11] to evaluate the interfaces' usability. The USE questionnaire employs a seven-point Likert rating scale, prompting users to express their level of agreement with various statements, ranging from strongly disagree (1) to strongly agree (7). It encompasses four dimensions: Usefulness, Satisfaction, Ease of Learning, and Ease of Use, which can be condensed to form a shorter version of the questionnaire. In this study, our aim was to assess the ease of use of the interfaces and explore potential differences in usability among them. Thus, our post-trial questionnaire featured a truncated edition of the USE questionnaire, focusing specifically on the Ease of Use dimension. This dimension comprised several pertinent statements, including: (1) It is easy to use, (2) It is user friendly, (3) It is flexible, (4) Using it is effortless, (5) I can use it without written instructions, and (6) I don't notice any inconsistencies as I use it. Upon completion of the initial two trials (Direct and Avatar), participants indicated their preferred teleoperation interface and provided reasoning for their choice.

*4) **Experimental Metrics**:* To assess the experimental conditions and address the research inquiries, we collected a range of experimental metrics, including:

1) Task completion time.
2) Collision count, serving as an indirect indicator of robot safety. A higher collision count is indicative of reduced robot safety.
3) Total path length traversed during a task, providing insights into goal overshooting tendencies. A greater path length suggests a higher likelihood of overshooting.
4) Task workload, as evaluated through the NASA-TLX questionnaire.
5) Usability of the teleoperation interfaces Direct condition and Avatar condition, assessed using the truncated USE questionnaire.
6) Interface preference.

V. RESULTS

A. MPC

1) Conclusion: For the two batches of tests, consisting of 8 different position goals and 4 different orientation (each batch with 32 tests), the Model Predictive Controller (MPC) concluded all runs, meaning that it is not expected for the MPC to not being able to anywhere in small distances (at least 2 meters, since that is what was tested). Concluding that the conclusion is of 100%.

2) Runs Analysis: Since a lot of tests were done, it would not be possible to provide images for all of the tests here, therefore it will only be presented the worst case Fig. 9 ($8.815m$ traveled).

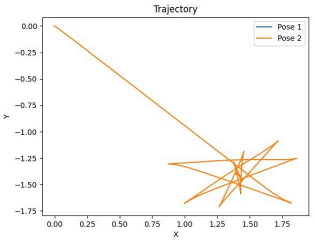

Fig. 9. Trajectory by Kayjay when the Goal was P8 with West Orientation

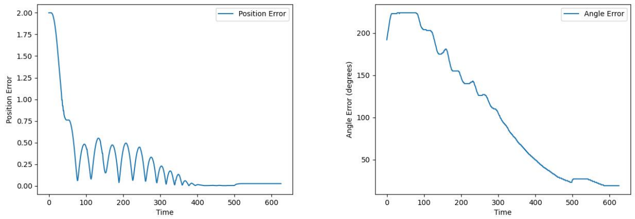

Fig. 10. Left image is Position error graph and right image is Orientation error graph.

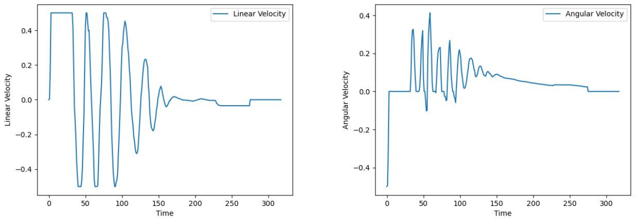

Fig. 11. Left image is linear velocity graphs and right image is Angular velocity graphs.

In Fig. 9, the trajectory to point P8 (refer to Fig. 4) with a final orientation of West is displayed. An overshoot upon reaching the goal is evident, indicating undesired behavior for the controller. In Fig.10 presents both the Position and Orientation errors, showing convergence to the values discussed in Section III-B, specifically 20 degrees and 0.1 meters. Lastly, in Fig. 11 illustrates the control variables of the Kayjay. Notably, both variables oscillate at certain points, explaining the trajectory overshoot. However, these control variables do not exhibit drastic changes within short periods, aligning with expectations for a semi-autonomous rover controller. This suggests that the cost function may require adjustments for better performance in this environment.

As it was observed in Fig. 9, the MPC exhibited significant overshoot when reaching its goal, resulting in nearly double the intended travel distance and an average time of 17 seconds, to complete a small run, which is not desired for a controller of a semi-autonomous rover that is teleoperated for planetary exploration.

It is important to note that the parameter tuning was conducted in a controlled laboratory environment, whereas the tests were performed in a realistic setting (the user study environment, with a slippery ground, making the measurements for odometry more difficult than necessary) without adjusting the parameters accordingly. This discrepancy indicates that parameter adjustments are necessary for the MPC to achieve optimal performance in different environments. These adjustments were not made due to time constraints for completing the thesis. This serves as a reminder that the MPC is not environment-agnostic and requires proper parameterization before use.

3) Time: In an ideal scenario, the optimal time to complete each run would be around 3 seconds, given the maximum linear velocity of the Kayjay is $0.7m/s$. However, the Kayjay, using the MPC, took an average of 16.9 seconds (ranging from 2.6 to 30.5 seconds) to complete a run, with an RMS error of 2.501. Although this performance is not ideal, it is important to note that due to the small overshoot in the best-case scenario and significant overshoot in the worst-case scenario, achieving the ideal time is unrealistic. Therefore, the observed completion times are expected to be longer than the ideal 3 seconds.

4) Path Length: The path traveled by the Kayjay averaged 3.87 meters, with a range from 2.097 to 8.815 meters, and an RMS of 0.218. This is significantly higher than ideal, since the distance between the starting position and the goal was 2 meters. Thus, an average travel distance of 3.87 meters is nearly double the intended distance.

B. AMADEE24

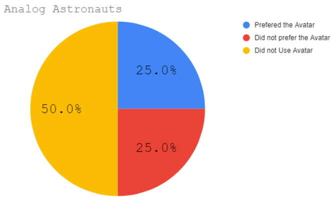

Fig. 12. Analog Astronauts preference to the Avatar Mode.

As explained in Section IV-A the number of runs, can be low, in this case, only 4 runs were possible and in Fig. 12 can be seen that half of those did not try the Avatar and in the other half, both analog astronauts had different opinions, therefore, for the question " Comparing avatar and direct mode when there is network delay, did you prefer using the avatar mode ? Participants were asked to answer with "Yes", "No" and "Did not happen" " the answer is a bit ambiguous for the generality of the study.

For the remaining questions:

1) When in avatar mode, if you were the one controlling the robot (like direct mode), would you lead it the

same path it went in the experiments ? Participants were asked to rate their response on a scale of 1 to 5, where 1 represents ”Strongly Disagree” and 5 represents ”Strongly Agree.”

2) Participants were asked to rate their stress levels on a scale of 1 to 5, with 1 representing no stress and 5 indicating high levels of stress.
3) Was the path generated between the robot and the avatar explanatory enough of what the robot would do in terms of movement? Participants were asked to answer with ”Yes”, ”No” and ”Did not happen”

The analog astronauts’ responses were as follows: **1)** One rated it as 3 and the other as 5, resulting in an average score of 4, indicating that the Avatar generally followed the expected path; **2)** Both rated their stress level as 3, suggesting that while this was not a formal workload assessment [**?**], the Avatar did not significantly increase user stress; for **3)** Both answered ”No,” highlighting a need for the interface to better clarify the path displayed to users.

After completing the run, participants were asked an open-ended question to provide additional feedback on the run and the Avatar. One participant commented, ”Avatar mode in principle was cool, but unfortunately not working right, since the avatar kept driving even after I stopped,” providing valuable insight, that the setup of the avatar might not be complete. The other participant did not provide further comments on the run or the Avatar.

C. User Study

1) Direct and Avatar Conditions:

1) Success Finishing the Inpscetion Task

The success rate of tasks was very high, (100% for the Direct mode and 90% for the Avatar mode). The majority of participants completed the tasks, and no one gave up. There was only 1 unsuccessful run (with the Avatar Interface) and was due to a time constraint, where the participant had difficulty maneuvering the robot to verify the last water bottle number.

2) Completion Time

A repeated measures one-way ANOVA with the Huynh-Feldt correction revealed a statistically significant difference in mean completion time across the different interfaces ($F(1, 9) = 40.84, p =< 0.01$). The Direct condition ($M = 217.10, SD = 60.01$) had a lower mean completion time compared to the Avatar condition ($M = 573.60, SD = 167.51$). These results address research question **Q1:** ”What effect does the use of augmented teleoperation interfaces (Direct and Avatar) have on task completion time?” and the findings indicate that, in the absence of latency, using the Direct Control mode significantly reduces task completion time compared to the Avatar mode.

3) Collisions

A repeated measures one-way ANOVA with the Huynh-Feldt correction determined that the mean of path length differed statistically significantly between interfaces ($F(1, 9) = 9.151, p = 0.014$). The Direct condition ($M = 0.40, SD = 0.70$) had a lower mean of collisions occurred compared to the Avatar condition ($M = 2.60, SD = 1.96$). These results address research question **Q2:** ”What effect does the use of the augmented teleoperation interfaces (Direct and Avatar) have on the robot safety (number of collisions)?” and the findings indicate that, in the absence of latency, using the Direct Control mode significantly reduces the occurance of collisions with the environment compared to the Avatar mode.

4) Path Length (Goal Over-Shoot)

A repeated measures one-way ANOVA with the Huynh-Feldt correction determined that the mean of collisions differed statistically significantly between interfaces ($F(1, 8) = 33.60, p =< 0.001$). The Direct condition ($M = 39.27, SD = 5.21$) had a lower mean of path lenght traveled compared to the Avatar condition ($M = 74.40, SD = 19.72$). These results address research question **Q3:** ”What effect does the use of the augmented teleoperation interfaces (Direct and Avatar) have on the total path length during a task?” and the findings indicate that, in the absence of latency, using the Direct Control mode significantly reduces the path traveled by the rover compared to the Avatar mode.

5) Workload

A repeated measures one-way ANOVA with the Huynh-Feldt correction determined that the mean of the overall of workload differed statistically significantly between interfaces ($F(1, 9) = 22.860, p =< 0.001$). The Direct condition ($M = 32.80, SD = 18.58$) had a lower mean of collisions occurred compared to the Avatar condition ($M = 61.53, SD = 12.14$). These results address research question **Q4:** ”What effect does the use of the augmented teleoperation interfaces (Direct and Avatar) have on the workload of the operator during the task?” The findings indicate that, in the absence of latency, using the Direct Control mode significantly reduces the workload during the task, compared to the Avatar mode. The results for other workload metrics showed similar patterns to the overall analysis, with some metrics differing statistically significantly between interfaces and others not.

Mental: Indicated a differed statistically significantly ($F(1, 9) = 6.766$ and $p = 0.029$, between the Direct condition ($M = 30.5, SD = 25.65$) and the Avatar condition ($M = 49.00, SD = 27.57$), which clearly indicates a significantly lower mean in the Direct Interface.

Temporal: Indicated a differed statistically significantly ($F(1, 9) = 7.454$ and $p = 0.023$, between the Direct condition ($M = 26.5, SD = 22.37$) and the Avatar condition ($M = 49.00, SD = 16.12$), which clearly indicates a significantly lower mean in the Direct Interface.

Effort: Indicated a differed statistically significantly

($F(1,9) = 14.135$ and $p = 0.004$, between the Direct condition ($M = 28.00, SD = 17.03$) and the Avatar condition ($M = 63.00, SD = 25.41$), which clearly indicates a significantly lower mean in the Direct Interface.

Frustration: Indicated a differed statistically significantly ($F(1,9) = 28.210$ and $p < 0.001$, between the Direct condition ($M = 23.50, SD = 15.82$) and the Avatar condition ($M = 66.00, SD = 22.92$), which clearly indicates a significantly lower mean in the Direct Interface.

On the other hand, some metrics did not show significant differences between interfaces:

Physical: $F(1,9) = 3.939$ and $p = 0.078$ indicated no statistical significant difference, although the Direct condition ($M = 16.50, SD = 12.27$) had a lower mean than the Avatar condition ($M = 38.5, SD = 32.14$).

Performance: $F(1,9) = 1.948$ and $p = 0.196$ indicated no statistical significant difference, with the Direct condition ($M = 38.50, SD = 34.00$) compared to the Avatar condition ($M = 59.90, SD = 24.00$).

6) Ease of Use

To analyze the USE Questionnaire, the Friedman test was used, revealing a statistically significant difference in usability scores depending on the teleoperation interface, $\chi^2(1) = 10.00, p = 0.002$. A post hoc analysis was then performed using the Wilcoxon signed-rank test with a Bonferroni correction, setting the significance level at $p < 0.005$. The median reported usability for the Direct and Avatar trials was 6.00 (range: 5.17 to 6.71) and 4.17 (range: 3.29 to 4.63), respectively. These results address our fifth research question **Q5:** ”What effect does the use of the augmented teleoperation interfaces (Direct and Avatar) have on the ease of use of the teleoperation interface?” The findings indicate that, in the absence of latency, using the Direct Control significantly improves the usability of the system compared to the Avatar.

Finally, it was performance an additional analysis on each of the statements of the USE questionnaire. Participants reported:

1) the Direct is easier to use than the Avatar ($Z = -2.831, p = 0.005$)
2) the Direct is friendlier to use than the Avatar ($Z = -2.565, p = 0.010$)
3) the Direct is more flexible than the Avatar ($Z = -2.836, p = 0.005$)
4) the Direct is more effortless than the Avatar ($Z = -2.319, p = 0.020$)
5) the Direct can be better used without written instructions than the Avatar ($Z = -2.154, p = 0.031$)
6) the Avatar had more inconsistencies than the Direct ($Z = -2.214, p = 0.027$)

7) Post-Task Questionnaire

When asked about their preference regarding the tested teleoperation interfaces, 90% of the participants preferred the Direct Control Interface, while 10% preferred the Avatar Interface. Some participants who favored the Direct Control cited its ease of use, noting that commands were executed immediately without any delay. In contrast, the Avatar Interface had a noticeable delay between command input and robot response. Participant 2 stated, ”The direct control interface because was much easier to adjust and control the robot on close spaces.” Conversely, Participant 5 remarked, ”According to what I experienced it feels like that Avatar Interface is more suited for a real world scenario”, suggesting that in contexts with potential network failures and communication delays, the Avatar Interface might perform better than in an environment without communication delays.

2) Hybrid Interface Condition:

1) Usage of Teleoperation Methods

When using the Hybrid Interface, participants used, on average, the Direct mode 83.9% ($SD = 0.228$) of the task time and the Avatar 16.1%($SD = 0.228$). A t-test showed a statistically significance difference between the use of the Direct and Avatar modes during the Hybrid condition ($t(9) = 4.708, p =< 0.001$). These results address research question **Q6:** ”When using the Hybrid interface, do operators use one of the teleoperation methods during more time than the other?” The findings indicate that, in the absence of latency, the Direct Control was significantly used during longer during the Task, compared to the Avatar mode.

2) Teleoperation Method Changes

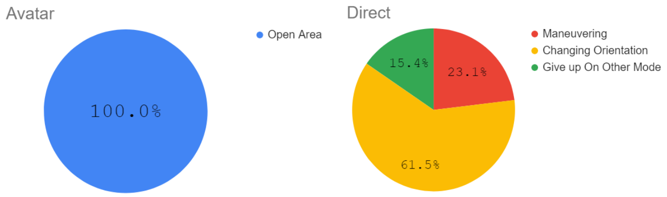

Fig. 13. Left image are the reasons to change to Avatar mode, according to participants. Right image are the Reasons to change to Direct mode, according to participants

When using the Hybrid Interface, participants switched modes an average of 2.70 times (SD = 2.901). Statistical analyses were conducted to assess whether changing the teleoperation method impacted participants' performance. A Pearson product-moment correlation was used to determine the relationship between the number of method changes and completion time. There was no statistically significant correlation found between the number of method changes and completion time ($r = 0.449, n = 10, p = 0.193$).

To answer the research question **Q7:** ”When using the hybrid interface (HI), in which situations do users switch between teleoperation modes?”, participants were asked to provide a brief reason each time they changed the interaction method. These reasons, although brief, offered insight into the participants' thought processes and helped identify the strengths of both modes. The reasons provided by the participants were compiled into a list, highlighting the factors that prompted them to switch interaction methods. Fig. 13 shows that participants switched to Avatar mode primarily in open areas, while Fig. 13 indicates that the main reason for switching to Direct mode was to change the orientation of the robot in its own position.

These results are aligned with the assumptions and empirical observations made in the literature and [2]. Supervisory control has proven to be an effective and efficient method for future crew-centered teleoperation. Literature indicate that this approach can maintain operators' Situation Awareness with minimal effort and workload, while also ensuring overall mission success. However, operators report the necessity for low-level control, Direct control, to perform more harder maneuvers or to address issues that onboard autonomy presented nowadays still cannot resolve on its own.

VI. CONCLUSION

This dissertation aimed to achieve three primary objectives: create a prototype rover named Kayjay, develop a controller capable of providing semi-autonomy without prior knowledge of the environment, and integrate this semi-autonomous system into a teleoperation framework to evaluate if the Avatar interface proposed in [2] can enhance teleoperation tasks in a realistic environment without latency.

Kayjay was designed with future upgrades in mind, utilizing Docker containers for driver installations to ensure manageability and plug-and-play functionality when new hardware is added.

The developed MPC, although exhibiting some overshoot, successfully met its objectives. While it still requires some refinements, it shows promise for future teleoperation advancements.

The Avatar interface was adapted for this dissertation from the one developed by the MEROP team in [1]. Instead of using ROS's move_base, the Avatar leverages the MPC developed in this work.

With all components in place, a preliminary user study was conducted to further evaluate the Avatar interface. Unlike the study in [2], this study was performed without latency. The findings revealed that the Avatar mode had worse performance compared to the Direct mode. However, in Hybrid mode, participants still utilized the Avatar in specific areas, indicating its potential and reaffirming some conclusions from [2] even in a latency-free environment. It is also noteworthy that the Direct mode was favored due to the absence of latency, as the rover moved immediately compared to the Avatar that could not move without waiting for the Avatar to move first.

VII. FUTURE WORK

There are three main components that can be improved in the future:

Kayjay: The GNSS and IMU data can be refined for greater accuracy. Hardware improvements could enhance Situational Awareness (SA) for the teleoperator, such as adding a mechanism to control the front camera or installing a new camera, to have one for autonomous reasons and one for teleoperation purposes.

MPC: Significant improvements can be made, such as utilizing Kayjay's depth camera to create a real-time elevation map, similar to [12]. This map would facilitate obstacle detection and avoidance, and help in planning optimal travel paths considering the environment's elevation.

Avatar: Although previously discussed in [2] and [1], the setup lacked a map for teleoperation. Adding a map that integrates GNSS data and Odometry is crucial for providing excellent SA to users.

REFERENCES

[1] M. C. R. A. R. Q. M. P. I. S. J. C. J. L. S. R. V. Rute Luz, Gonçalo Coelho, "Remote operations and streamlining communication in Mars analog missions: Robot-agnostic augmented interface for live annotations during teleoperated exploration tasks," [unpublished manuscript], 2024.

[2] J. L. S. Rute Luz and R. Ventura, "Enhanced lunar exploration through earth-based teleoperation of rovers: Augmented interfaces to minimize latency impact," International Planetary Probe Workshop, 2021.

[3] L. Penin, K. Matsumoto, and S. Wakabayashi, "Force reflection for time-delayed teleoperation of space robots," in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*, vol. 4, pp. 3120-3125, 2000. doi: 10.1109/ROBOT.2000.845143.

[4] T. Kot and P. Novák, "Application of virtual reality in teleoperation of the military mobile robotic system TAROS," *International Journal of Advanced Robotic Systems*, vol. 15, no. 1, 2018. doi: 10.1177/1729881417751545.

[5] R. Luz, J. Corujeira, L. Grisoni, F. Giraud, J. L. Silva, and R. Ventura, "On the use of haptic tablets for UGV teleoperation in unstructured environments: System design and evaluation," *IEEE Access*, vol. 7, pp. 95443-95454, 2019. doi: 10.1109/ACCESS.2019.2928981.

[6] M. R. Endsley, "Measurement of situation awareness in dynamic systems," *Human Factors*, vol. 37, no. 1, pp. 65-84, 1995. doi: 10.1518/001872095779049499.

[7] D. Lovi, N. Birkbeck, A. H. Herdocia, A. Rachmielowski, M. Jägersand, and D. Cobzaş, "Predictive display for mobile manipulators in unknown environments using online vision-based monocular modeling and localization," 2010, pp. 5792-5798.

[8] R. Halatschek, K. D. Konanur Ramanna, W. Url, and G. Steinbauer, "Universal offroad robot platform for disaster response," in *2020 IEEE International Symposium on Safety, Security, and Rescue Robotics*, 2020.

[9] J. Lim, W.-C. Wu, J. Wang, J. A. Detre, D. F. Dinges, and H. Rao, "Imaging brain fatigue from sustained mental workload: An ASL perfusion study of the time-on-task effect," *NeuroImage*, vol. 49, no. 4, pp. 3426-3435, 2010.

[10] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Advances in Psychology*, vol. 52, Elsevier, Amsterdam, The Netherlands, 1988, pp. 139-183.

[11] A. M. Lund, "Measuring usability with the USE questionnaire," *Usability Interface*, vol. 8, no. 2, pp. 3-6, 2001.

[12] P. Fankhauser, M. Bloesch, and M. Hutter, "Probabilistic terrain mapping for mobile robots with uncertain localization," *IEEE Robotics and Automation Letters (RA-L)*, vol. 3, no. 4, pp. 3019-3026, 2018. doi: 10.1109/LRA.2018.2849506.