Imperial College London - HCR 2023

WHEEL-E Final Report

Authors: Maximus Wickham, Matilde Piccoli, Salman Dhaif, Bradley Stanley-Clamp, Leonardo Garofalo, Timeo Schmidt, Sam Taylor, Selin Uygun, Joshua Lim, Cristiana Gavrilescu, Bernard Benz, Tanguy Perron, and Yash Rajput

1 Abstract

Wheel-E is a smart and autonomous wheelchair aimed at children with severe disabilities. The system includes three unique features: a social navigation system that applies proxemics and 'social awareness' metrics to the autonomous path planning; a sentiment analysis model which classifies the user state and allows the wheelchair to adjust the velocity accordingly; a laser pointer that functions as a directional indicator of the wheelchair's movements. Wheel-E is expected to improve the wheelchair user's comfort navigating social environments, whilst enhancing the level of trust in the system of external people interacting with it.
This report will start by analysing the current research trends when developing smart wheelchairs, then describe the design, development and testing of our system, and, finally, evaluate the research hypotheses investigated.

2 Introduction

Young wheelchair users with severe disabilities spend the majority of their time in familiar and indoor environments, such as hospitals, their homes, or schools. Despite their familiarity with these places, the dynamism of these environments often prevents them from being able to navigate them independently. This lack of accessibility can lead to feelings of frustration, isolation, and dependence on others and negatively impact the cognitive development of the child [1]. By creating more autonomous and adaptive smart systems, wheelchair users can have greater autonomy and control over their daily lives.
Moreover, a recent study by Zhang et al. [2] has highlighted how integrating situational awareness and adaptability in smart wheelchairs can significantly help users with tasks such as social navigation and interaction. However, the primary consideration in developing such a system is finding an equilibrium between customization and user burden. Hence, the need for a design with sensory capabilities to enhance the contextual awareness of the environment and the user's emotions should be prioritised.
As a consequence of these considerations, the target use case for our wheelchair is the social navigation of pre-mapped, familiar environments, and the target customers are young users with severe disabilities, for which manoeuvring a joystick would be a challenging or impossible task. For this purpose, our wheelchair can be fully controlled by simply tapping on a screen or using vocal commands. This, in conjunction with the situational and user-state awareness of our system's movement controller and path planner, is expected to enhance the experience of children users, their carers, and even pedestrians occasionally interacting with the wheelchair.

2.1 Hypotheses

This project aims to investigate the following research hypothesis:

- H_1 Using a path planner integrated with situational awareness and proxemics measures into the planner of the wheelchair will improve:
  - the comfort of the children using the wheelchair.
  - the trust in the system of external people interacting with it (e.g. carers, pedestrians).
- H_2 Using a motion planner that takes into consideration the user's emotional state will enhance the comfort of children using the smart wheelchair.
- H_3 Using a laser-based directional indicator to anticipate the motion of the wheelchair will enhance the trust in the system of the external people interacting with it.

We use two standard metrics to evaluate our system: the level of comfort of the user, from the perspective of the children in the wheelchair, and the level of trust in the system, for external interactions. These concepts will be defined in the evaluation section.

3 Background

Assistive technology (AT) can empower individuals with disabilities to perform tasks that would otherwise be unattainable [3], expanding their opportunities for self-expression and a renewed sense of identity, as noted by [4]. In particular, according to [1], the development of critical cognitive skills in children with disabilities has been linked to increasing motor abilities that lead to greater sensory experiences. The growth in these skills is believed to result from the navigational and sensory awareness requirements associated with early child mobility. Similarly, research focused on power mobility usage by individuals with significant motor challenges has found that such use can lead to improvement in cognitive and social function, increased participation, and decreased "learned helplessness" [5]. This is believed to be a result of the navigational and perceptual experiences and learning associated with using a wheelchair [6].
However, for children with severe disabilities, the opportunities for independent mobility are often very limited due to their difficulty or impossibility to manoeuvre a joystick, constituting a further challenge when navigating a social, dynamic environment.

Another problem with smart wheelchairs, and AT in general, is that the user might be subjected to stigmatization, leading to negative social and psychological consequences. Wheelchair users often face a range of negative attitudes and stereotypes, causing social isolation and discrimination and affecting their mental health and well-being. In the long term, this may prevent individuals from embracing and utilizing the wheelchair to its fullest potential. By developing more socially-aware wheelchairs and designing a user interface that adapts to the user state, we can help reduce the stigmatization of wheelchairs and enable the full empowerment of children with disabilities, making sure the

wheelchair does not simply assist but also evaluates the appropriate time to do so.
From this, we have identified three key factors influencing the quality of the user experience and human-wheelchair interaction, which will be analysed in the following sections.

3.1 Social Navigation

A large part of this project is making the wheelchair experience more comfortable in environments involving pedestrians. Literature on this subject is split into two broad categories; predicting the interaction of humans with objects in the environment, such as walking towards a door, and predicting the interactions of humans with other mobile entities. The solutions this project investigates focus on human interaction and etiquette, mainly in order to mitigate the need for extensive object detection and overly complex models. Some papers such as [7] attempt to use machine learning techniques, such as an LSTM in this case, to learn the social etiquette rules from human motion. However, this method will not be used due to both high complexity and lack of training data available relevant to our use case.

This thesis [8] on Probabilistic Qualitative Representation in Human-Robot interaction describes the use of probabilistic distributions in cost maps for path planning around mobile people. The paper suggests using a 2D Gaussian to give the cost of moving close to a human with the Gaussian shape based on the direction and speed of a human; allowing for more natural path planning with the robot able to move into high-cost areas if unable to take any other route. The use of a Gaussian as a representation of human motion in cost maps appears to be a common approach [9] although many alternatives exist, such as Qualitative Trajectory Calculus [9]. The Gaussian approach provides a lower computational cost, however, and appears sufficient for the scenarios investigated in this project.

In addition to predicting the human location, it is also important to consider where a robot should move relative to a person. A study by Morales et al. [10] identified key elements for a more human-aware smart wheelchair, both with respect to the wheelchair passengers and surrounding pedestrians:

1. The wheelchair velocity should be limited to a human walking pace (around 1m/s in normal conditions, and acceleration should be limited.
2. The system should avoid zigzagging, which may cause pedestrians to change their course.
3. The wheelchair should be prevented from invading pedestrian personal space and minimize the need for evasive action.
4. Even when there are no pedestrians present, passengers prefer the wheelchair to stay on one side of the passage.

These elements will be taken into account in the determination of the movement algorithm and velocity control alongside the user's emotional state (see Section 3.5.2). Specifically, the route planning will consider the avoidance of the human Comfort Zone (CZ) by using proxemics-based metrics. Building upon a study by Ali et al. [11], we aim to predict people's CZ and trajectory by analysing the head and body positioning of the pedestrians, as detected by the ZED 2i camera. Not only will this improve the people avoidance strategy, but also it will allow us to explore novel scenarios. In particular, a possible further expansion will be to relate the pedestrians' body language with the wheelchair user's emotional data to identify situations in which the interaction is intentional, i.e. when the two parties want to engage. In such a case, the wheelchair user might benefit from an adjustment in the path planning that overrides the default assumption to avoid people's trajectory when moving towards an alternative target.

3.2 Sentiment Analysis

Previous works on the subject [12, 13] have used various biofeedback methodologies, such as electroencephalogram (EEG) or myoelectric signals-based ones, to measure the emotional and physical state of the user and control wheelchair behaviour accordingly. However, these interfaces are often invasive and can be perceived as constrictive, making them unsuitable for children. Moreover, these types of interfaces can draw unwanted attention to the user, creating a stigma around them. Hence, we believe a less invasive approach using a camera that measures the emotional state would improve user comfort. Other studies [14] have utilised facial expressions and movement to fully command smart wheelchairs (including stop, go, and turning motions). However, these also saw several drawbacks; namely, they required large head movements from the users to discern the intended motion commands. As well, testing revealed that users often intuitively looked at obstacles they were approaching; this moved the wheelchair closer towards obstacles increasing the risk of accidents. Furthermore, the above control methods required large amounts of user attention to operating the wheelchair, increasing the cognitive workload on the user and reducing engagement with their surroundings.
Following these considerations, we hypothesised that a better balance between personalisation and user burden could be achieved by restricting emotion control-based feedback to motion features, such as velocity and acceleration, rather than the trajectory planning itself.

3.3 Directional Indicators

Past studies about motion indicators include the use of a single laser pointer dot to indicate and set the close-range motion of the wheelchair [15] and the use of Hololens to display the trajectory information to the user [16]. While the projection of a single dot is an efficient means of indicating the destination of the wheelchair, it does not convey information about the path taken by the wheelchair. A path projection would allow the user and others in the environment to better anticipate the motion of the wheelchair. The use of the Hololens allowed only the user to observe the path of the wheelchair through an augmented reality headset. Along with the path information, this method displayed the possible collision points with the objects around, the users' raw input and the corrected output path of the wheelchair. Even though this implementation offers more detailed information to the user, it utilises equipment which is expensive and may draw unwanted attention to the user. The laser path indicator (proposed in H_3) offers a more inexpensive approach to display path information. As well allows for better anticipation and understanding of the motion for the user, pedestrians, and caretakers. Moreover, a paper by [17] suggests how providing explanations about a robot's behaviour affects the degree of human trust in the system, a useful result when considering our addition of a directional indicator to the wheelchair, as this can also be considered as a descriptor of the system's behaviour.

4 System Design

4.1 Control & Infrastructure

The wheelchair comprises three control subsystems; an Nvidia Jetson running object detection and SLAM, a RaspberryPi that interfaces with the CANBUS system on the wheelchair and an external laptop running both `roscore` and the central path planning and human interface programs. The RaspberryPi interfaces with the laptop through a WebSocket connection, receiving commands converted from `VelocityTwist` messages on the `/cmd_vel` topic to a data format the CANBUS can understand. These messages published to the `/cmd_vel` topic allow programs to set the forward velocity of the wheelchair in addition to the z-axis rotational velocity. Figure 4 shows the ROS nodes infrastructure of the whole system.

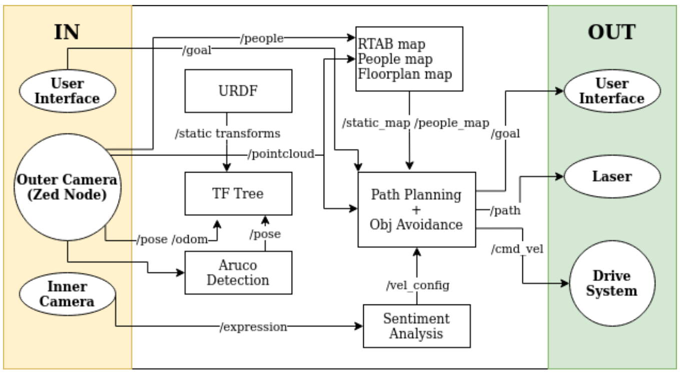

Figure 1: Overall System Architecture

4.2 Model & Simulation

To aid in design verification and functionality testing, a 3D model of the wheelchair was developed in Solidworks and ported from an STL design into a URDF model. Alternatives of this approach were to develop the model using Blender or Fusion 360, however these involved longer development times. The model was then iterated upon to tune the inertial behaviour of the model, alongside the addition of sensors and drive controllers, to reflect the physical properties and behaviour of the real wheelchair when used in a Gazebo simulation. Once the final wheelchair model was completed (Figure 2), it could be used in the Gazebo simulation. A Python program was developed to automatically generate a URDF model of the working environment, given an array of dimensions. Simulated external actors were also added to aid the implementation of algorithms interfacing the wheelchair with external users, providing a reliable platform for the development and testing of the various stacks of wheelchair.

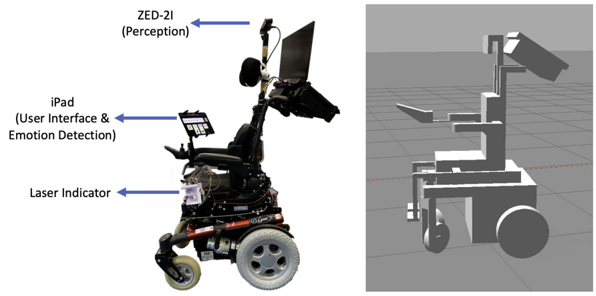

Figure 2: Wheelchair and final simulated model

4.3 Environment Perception

The purpose of the perception system is to provide a detailed map of the environment, including obstacles such as objects, and humans, while also delivering accurate odometry data for the position and orientation of the wheelchair. The perception system is crucial for the safe and efficient navigation of the robotic wheelchair in its operating environment. In this section, we discuss the choice of hardware and software components, the integration of ARUCO localisation, and the rationale behind these choices.

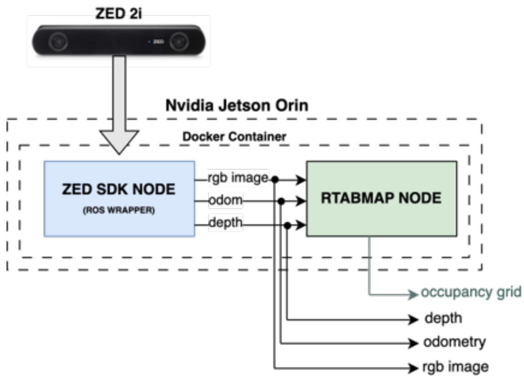

Figure 3: Single Environment Facing Camera Architecture

The robotic system utilizes the Stereolabs Zed2i stereo camera for environmental perception. This camera features an IMU, magnetometer, and barometer, providing rich sensor data for the perception system. The Zed2i camera is connected to an NVIDIA Jetson Orin platform, responsible for running the full computer vision stack and outputting the processed information to other subsystems. The perception software stack is primarily built upon the Zed SDK, wrapped in a ROS node that receives data streams from the camera and performs sensor fusion with the onboard sensors. This produces high-level ROS topics, such as odometry, depth, and point cloud data. The second component in the perception software stack is the Real-Time-Appearance-Based-Mapping (RTABMAP) [18] ROS node. This node consumes the information generated by the Zed SDK and outputs a 2D occupancy grid, representing the surrounding obstacles as a 2D map that is compatible with the path planning algorithm within the navigation stack.

The choice to use the Stereolabs Zed2i stereo camera and its accompanying Zed SDK was made after evaluating alternative options, such as using a regular camera and developing custom computer vision and localization algorithms. The Zed2i camera and Zed SDK were chosen for several reasons:

1. **Maturity and Integration:** The Zed SDK is a mature system with seamless integration with the Zed2i camera and sensor hardware, ensuring optimal performance.
2. **Modularity:** Utilizing an external, standardized system like the Zed SDK allows for easier replacement or upgrading of the computer vision and localization stack if needed in the future.
3. **Professionally Developed & Real-Time Optimised:** The Zed SDK is a professionally developed solution that offers a set of highly optimized, real-time computer vision and SLAM pipelines that are compatible with ROS.

During the development of the Perception system, different implementations were tested, to increase the overall system performance. Two possible strategies were developed and tested as part of this. The first approach was to increase the sensor count and add a second stereo camera. In this approach, the data from the two cameras would be fused to produce one single odometry data stream of higher accuracy. This was achieved by implementing an 'Extended Kalman Filter' [19] which would output a single fused odometry topic that would be consumed by the RTAB-Map Node to produce a better occupancy grid and subsequently enable better performing navigation.

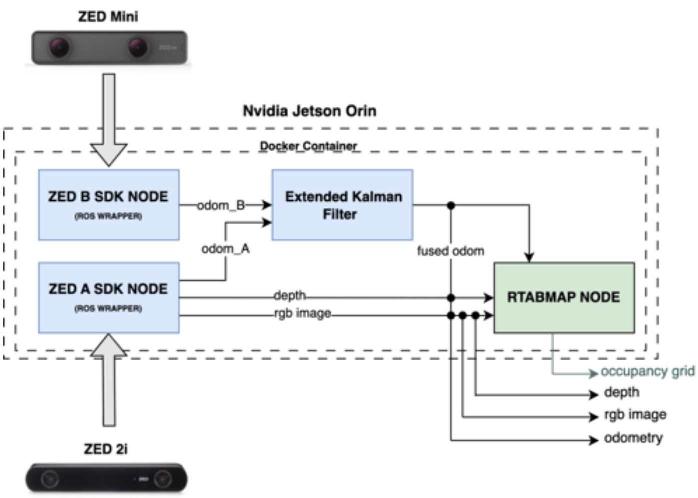

Figure 4: Dual Environment Facing Camera Architecture

The second approach to improve the performance of the odometry was an ARUCO localisation system. For this system, ARUCO markers are placed at known locations within the robot's operating environment, such as various points within a hospital. Since in environments with highly homogeneous or repetitive layouts, it may not be feasible to rely solely on visual SLAM for accurate localization, by using ARUCO markers, the robot can re-calibrate its position on the map whenever it detects a marker. This is possible because the marker's absolute position on the map is known and its relative position with respect to the wheelchair frame can be computed. Although the first dual-camera approach showed promising results in the simulated Gazebo environment, the approach was abandoned as integrating it with the full system would have involved sophisticated architecture changes. Furthermore, there were limited added benefits as it was found that ARUCO markers detected using the aruco_ros package [20] were reliable and accurate in localizing the position of the wheelchair on the map, with a precision level up to the centimetre. Furthermore, the use of ARUCO markers proved to be effective in counteracting the natural drift in visual odometry that accumulates naturally after prolonged movements of the wheelchair.

4.4 Mapping & Navigation

4.4.1 Navigation Stack

The navigation system (Figure 5) is responsible for directing the user to a goal position set via coordinates within a known map. It consists of four main components: the mapping system, localisation system, odometry and costmap generation. Based on the target use case of this wheelchair, we can assume the environment is known beforehand, therefore the wheelchair would have a floor plan of the environment and the goal positions would be expressed as known coordinates within the map. Two challenges arose as a result of using a static floor plan: the need for a mechanism to locate the robot accurately within the map and the requirement to perform additional mapping of the environment whilst moving to build a knowledge of static obstacles not given in the floor plan.

We started by considering ACML-based localisation models, which use point-cloud data received from the environment to find the most probable location of the robot within the map. However, there are several reasons why this system might not be applicable in this use case. Firstly, if the floor plan provided is particularly simple or characterised by repetitive patterns, it would be extremely hard to localise within the map accurately. This is especially true in a hospital environment where there are many objects in the surroundings obscuring the walls described by the floor plan. A solution would be for the user to input their rough location (e.g. room number) or to use other localisation tools, such as GPS data, as this would help by restricting the area within which to localise. This would, however, either require more intervention from the user or additional signal requirements to the robot, both of which were undesirable. For these reasons, it was decided to use Aruco markers placed throughout the environment to allow the robot to localise itself (see section 4.3).

Once the robot's location could be confirmed within the static floor plan map, it was necessary to keep track of the robot's position as it moved throughout the environment and out of sight of an Aruco marker. Many methods were considered to solve this problem, including SLAM-based approaches; in the end, we decided to use the pose tracking system available from the ZED2I camera, which provided both continuous odometry and pose tracking used to offset odometry drift, both due to its convenience and the encouraging results in initial testing. The ZED2I pose tracking algorithm makes use of key point detection in the environment. These key points are stored to maintain an accurate position of the

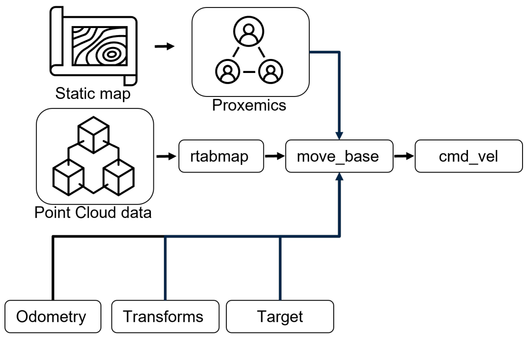

Figure 5: Overview of the navigation stack architecture

robot relative to its starting position using loop closure.
With these different components implemented, the inputs to the navigation system were a static floor plan map, point cloud data provided from the ZED2I, a known position with the static map and a 2D map built as the robot moved through the environment that was provided by the ZED2I (see section 4.3). The need for this dynamic map was to make more accurate navigation decisions based on objects that had previously been detected in the environment instead of just the simple static map.
Furthermore, we identified the MoveBase ROS node to be a well-supported and standard option to make navigation decisions based on these inputs. This tool also allowed the potential of switching to the less documented but more capable MoveBaseFlex in the future due to supposed backward compatibility. Different occupancy layers were created by splicing the static and dynamic maps using numpy to create an input for both global and local costmaps, as well as providing the point cloud data as an input for the local costmap. MoveBase could make use of these layers to generate a path from its current position to the set goal location and output "cmd_vel" messages to drive the robot.
However, during testing, we found that the behaviour shown by MoveBase was not always clearly explainable, with actions such as turning the long way around to achieve a particular orientation often performed. Unfortunately, the costmap implementations were only sufficiently free of noise late into the project, and it was assumed at first that these errors were due to the noise in the costmap. Once this noise had been removed, and some unexpected behaviours remained, there was inadequate time left to try another alternative, such as MoveBaseFlex.

4.4.2 Social Navigation

To design a wheelchair with situational awareness, we had to include proxemics considerations within the navigation planner, ensuring that the robot manoeuvres in a socially respectful manner. We explored two options: using the ROS package 'social_navigation_layers' or building the tool from scratch. The 'social_navigation_layers' package provides plugin-based layers for implementing constraints onto the path planning taking into account proxemics [21] and providing easy integration with the navigation stack; however, it does not allow for customization and improvement. Therefore, we chose to build our own tool to have greater control over its design.
The tool uses object detection from the ZED2i to obtain information about the people, such as location and velocities. It then represents these people depending on their behaviours in the form of a 'socially respectful' cost map, which is then taken into account when path planning is performed resulting in a route similar to the way a human would move through the environment.
It should be noticed that the 'social' cost map, representing the people disposition, is integrated into the global cost map of the navigation stack and not the local cost map, as the decisions the wheelchair should make about people at close range (local cost map) should be different to that of people further away (global cost map). In fact, if a person is close to the wheelchair, the latter needs to actively avoid them quickly, e.g. stopping and then moving around someone or waiting for them to move. For such a situation, we treat the nearby person as a static object, detecting it using the point cloud data from the ZED2I. Instead, for a person further away, we can warn the navigation system anticipating the possible future obstacle so that move_base can take this into account in the path planning.

Based on work discussed in section 3.1 and further research into work such as [21], [22] and [23] a useful basis for representing the proxemics of people was built. We identified the three most common scenarios to consider for our behavioural analysis (see figure 6):

1. A **stationary person** was represented using the superposition of two 2d Gaussian distributions. The first Gaussian is to represent the person in their current position and size, using a distribution to ensure that the wheelchair does not get too close to the person unless there is no other less 'expensive' path. The second is to represent their direction-specific proxemics, giving someone more room if passing in front of them to avoid obstructing their view.

2. **Stationary people interacting** (e.g. two people are facing each other) are represented by applying superposition to the distributions of the individual positions, causing the area between them to be also occupied. This is to consider that would be inappropriate to cut between people having a conversation if there is an alternative path available.

3. **People moving** are represented considering their future predicted location and whether their path will collide with the wheelchair one. Their current velocity is computed, assuming this will remain relatively constant during the next few seconds, we predict where the person will be after t seconds. Moreover, we assume that if the person's velocity is larger, their motion is more likely to be sporadic, our prediction is more uncertain and, therefore, the distribution is more spread out. The value of t was tuned by taking into account the max range that we can detect a person on the ZED2I and the average speed of the wheelchair.

4.5 Human-Robot Interaction

At the core of our design lies the will to take into account the needs of the user's physical abilities, emotional state and

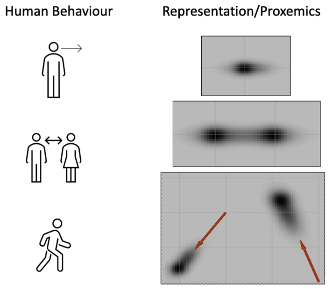

Figure 6: Distribution for human behaviours, from top to bottom; distribution of a stationary person facing to the right, distribution of two people interacting with one another, distribution of two people moving in different directions, the red arrows are the actual locations of the people at the measurement time.

the environment around them. As these change over time, the wheelchair is expected to be able to adapt through: user, social, and external agent considerations.

4.5.1 Emotion Detection

In relation to H_2, an emotion detection system utilizing automatic facial expression analysis (AFAE) was developed to constantly determine the current emotional state of the user. Various AFAE modalities currently exist in literature as discussed by [24, 25] including macro-expressions, micro-expressions, FACS, and multi-modal systems (combining facial data with physiological signals). These systems alongside corresponding various acquisition methods were evaluated against a set of criteria as show in figure 7.

	Emotion Analysis Modality			
	Macro-Exp	Micro-Exp	FACS	Multi-modal
Time/Computational Complexity	Low/Low	Low/High	High/High	Varies
Interpretability	High	High	Low	High
Acquisition	RGB img	RGB/RGB-D vid	RGB vid/Wearable	Camera+Biosensor

Figure 7: Emotion detection modalities comparison

We require a system which is able to detect emotional state with accuracy and speed (to perform well in real-time), is easily interpretable to children (vital for user trust and comfort [26]), and is not too invasive in terms of acquisition. As seen in figure 7, macro-expression modality boasts a low complexity, high interpretability, and has a simple, non-invasive acquisition method. While other methods may offer better sensitivity in detecting the true user emotional state, they come with associated trade offs in other criterion [24, 27]. As discussed in [24], many macro-expression models have been developed in literature, trained and verified on datasets with varying displays of emotion, namely: posed, spontaneous, and in the wild. For the given use cases of the wheelchair, accurate performance on the later two is required. Various models discussed in [28] were compared on LFW [29] (a widely-used, in-the-wild dataset in literture). Namely, the models used (VGG[30], FaceNet[31], OpenFace[32], and DeepFace[33]) were evaluated on accuracy as seen in figure 8. VGG-Face displayed the highest accuracy alongside proving to be most robust during integration testing with the overall system setup of the wheelchair.

	Macro-expression Model			
	VGG-Face	FaceNet	OpenFace	DeepFace
Accuracy	98.78%	97.53%	93.80%	97.35%

Figure 8: Macro-expression model performance on LFW dataset

The chosen VGG-Face model was wrapped in a ROS node and received frames from the iPad also used for the user-interface. The output prediction of the model from the 7-basic emotions system would then be used to scale the velocity values of the navigation stack, as discussed in 3.2. Two possible approaches were considered to achieve this. The first is to intercept the output command velocity published by the navigation stack, scale them directly according to the last recorded emotion, and publish the adjusted values to the drive system. However, this effectively overrides the motion controller and navigation algorithms potentially introducing non-linearities and undesired behaviour into the system. The alternate approach would be to utilize the emotional state information as an input to the navigation stack via scaling the maximum and minimum velocity of the system. This can be done configuring these parameters used by the trajectory planner. This allows the navigation stack to take this change into account preventing the previously described drawbacks of the first approach. Hence, a ROS `dynamic_reconfigure` client [34] was used to achieve this as it allows for the configuration of these parameters during run-time.

4.5.2 Laser Path Indicator

The laser path indicator projects the wheelchair's intended path of travel onto the ground ahead of the wheelchair. It achieves this by rapidly sweeping the dot from a laser diode module back and forth in the x and y-axis. There were several options to consider when implementing the laser path indicator. The first one was to purchase a laser galvo to accurately control the position and the direction of the point laser to draw precise paths. However, increased precision came with increased cost as well as long shipping times making this option undesirable. The second approach was to purchase an off-the-shelf laser projector that allowed the drawing of line art images. This was more cost and time-effective but required integration with Arduino to control the projection pattern. This would have presented a more steep learning curve compared to the final option. The final approach was to create a laser galvo using DC stepper motors controlled with an Arduino, a laser diode and a 3D-printed casing to mount, allowing the setup to perform like a galvo. Although this option presented as the least precise, it was the most cost and time-effective way to project path information with enough accuracy. The path drawing was accomplished using galvos consisting of mirrors attached to stepper motors as shown in Figure 9. Two galvos are placed at right angles to each other and the laser reflects of both mirrors sequentially to cover movement in both the x and y axes.

The laser projector relies on the Persistence-of-Vision (PoV)

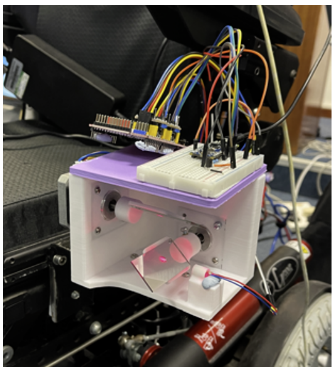

Figure 9: Laser Projector

Figure 10: User-Interface

effect to create the impression of a solid line from a single laser dot. The theoretical minimum frequency at which the dot needs to move back and forth is 24Hz, the minimum refresh rate for smooth video. However, experimentation showed that it was possible to push this value to 20Hz at the expense of some flickering. Stepper motors can be moved in small increments called steps. To remain within the timing constraints for PoV, a maximum of 400 steps can be made. This places a constraint on the maximum displacement, from the starting position, of the laser dot on each axis. The stepper motors are driven by dedicated stepper motor controllers which are controlled using an Arduino. The laser diode module can also be switched on and off by the Arduino to allow for a seamless homing of the laser dot. The motors were powered using the 12V supply from the wheelchair which was then converted to 5V using a DC step-down voltage regulator. The Arduino was plugged into the onboard computer to receive power and enable serial communication.

The short path information from Navigation was sent through a ROS topic which was then received and processed by the python code running on the computer. The coordinates received were fitted by a quadratic equation to create the line. The movement information for the x and y axes was derived from the points on the line and then sent to the Arduino through the serial port. A 3D housing was designed to mount the laser module onto the left metal bracket on the chair. The housing also included a safety cover at the top to avoid the laser shining into the user's eye in case of a loose connection in the motor wires which might make the mirrors rotate randomly. The mirror holders designed to attach the mirrors to the motors were also 3D printed. The laser path drawn is visible under any lighting condition and on any flooring. It is most suitable for polished surfaces such as wood or vinyl as it allows the laser path to be reflected away from the user's eye and towards the bystanders. A low-powered laser diode module was used to ensure the safety of both the wheelchair user and bystanders. Furthermore, as the laser is always oriented towards the ground, the risk of direct exposure to the laser beam is minimal.

4.6 User Interface

A graphical user interface (figure 10) was developed in the form of a web application coded in HTML, JavaScript and CSS. The web app communicates with the ROS system via a python intermediate which hosts the website locally and allows users to access the web page when connected to the same network. The UI was designed such that it could be easily configured and used on an iPad attached to the wheelchair. The UI has been designed with accessibility as its focus and targeted at disabled children.

Large touch input space it is expected that the user of the wheelchair will have some level of motor impairment, including those related to dexterity and hand mobility. It is therefore important to have icons much larger than other interfaces on the iPad to allow ease of use.

Graphic icons the wheelchair may be used by children with cognitive disabilities as well. Hence, graphical icons are used to ease the use of the GUI.

High contrast mode taking into account possible vision impairment and colour blindness, a high contrast mode is designed with a colour pallet used as a standard for these use cases [35]; this mode is easily triggered by a toggle switch.

Voice control and talk-back some users may suffer from extremely limited to no mobility, for those individuals we include a voice control mode as well as a text-to-speech feature. This has been implemented via the web speech API toolkit which provides compatibility with a variety of browsers and sufficient efficiency to recognise the keywords, as per the use case.

5 Experiment Setup & Methodology

Testing After testing the individual subsystems and the base integrated system, we started running experiments with the system, isolating and evaluating each research hypothesis against the baseline (without the additional features). We targeted four scenarios to test the effect of the different features independently:

- *Baseline:* Autonomous wheelchair without social navigation and human interaction considerations
- H_1 : Baseline with social navigation feature
- H_2 : Baseline with emotion recognition feature
- H_3 : Baseline with directional indicator

As our target use case was a pre-mapped environment, the experiment involved setting up some pre-determined locations ('bathroom', 'kitchen', 'door' and 'table'). Each case was tested with moving and static pedestrians surrounding the wheelchair to recreate a social environment. In particular, we identified three common social scenarios:

- One or more individuals separately standing in the wheelchair trajectory.
- One or more people crossing the wheelchair trajectory.
- A group of people interacting with each other and standing in the wheelchair trajectory.

In terms of objective performance, since for each of the scenarios, the wheelchair behaved as intended, accordingly to the different features tested, we decided to follow up on the testing by evaluating the system from a research perspective, investigating the level of user comfort and trust in the system.

Hypothesis Evaluation The main limitation of the research side of this project has been the inability to run system experiments with actual users in it, making us unable to use objective measures such as task performance and robot utilisation to evaluate our system. This has been due to the impossibility of receiving ethics approval within the project time span to test it with children whilst also finding it challenging to test the system with people of greater size and weight due to the calibration and physical setup of the hardware. However, we found that using standardised questionnaires on wheelchair video demonstrations, a widely used methodology for research hypotheses of this kind, constituted a valuable starting point for the research evaluation of our system.
The survey comprised 4 sections, respectively presenting the baseline and the three research hypotheses; for each of them, a video was shown displaying the functioning of the system from the first and third-person perspective, followed by industry-standard questionnaires, together with some open-ended questions.

User Comfort Many well-established questionnaires have tackled the challenge of quantifying the user experience and comfort in particular, both with generalised metrics such as the User Experience Questionnaire (UEQ) or linking it to the cognitive workload of the user (NASA TLX index). However, in our case, the experiment is based on the perceived experience rather than a hands-on one, we identified the System Usability Scale (SUS) to be the most appropriate to evaluate user comfort.

Trust in the Wheelchair As theorised by Castelfranchi et al. [36], trust can be measured using the quantitative dimensions of its cognitive constituents, i.e. assuming that the greater the human's belief in the machine's competence and performance, the greater the human trust in machines. Following this reasoning, we use the Technology Acceptance Model (TAM) and Trust in Automation (TiA) questionnaire to quantify the level of trust of external people interacting with the wheelchair, such as carers. These two metrics considered many aspects, including the perceived usefulness, perceived ease of use, intention of use, and belief in the robot's capability and independence, which can then be combined to obtain an estimate of the general level of trust in the wheelchair of the respondent.

6 Results

From a user comfort perspective, the SUS questionnaire gave promising results, indicating how both H_1 and H_2 scored higher in terms of user experience compared to the baseline system (Figure 12). Notably, as can be seen in Figure 11, confidence in using the system rose significantly when using the social navigation modifications and was even higher when adding in the emotion detection system. This result is interesting as the emotion detection system gave a positive increase on average over simply just using the social navigation system in terms of comfort but not in terms of trust (Figure 13) and in the textual responses received. This could be due to the fact that it is easier to imagine the utility and understand the functioning of a system with social navigation, instead of a system with emotion-based motion, as the latter is less human-like. Following these considerations, it would be useful to test the system with people, letting them experience the system hands-on and observing the learning curve; this would be especially useful if we want to test the level of trust of wheelchair users and their carers, and the long-term effects on the user comfort.

When checking the eigenvalues of the covariance matrix, it found that only three eigenvalues were greater than 1, indicating that three principal components are necessary to describe the data. This suggests that some of the questions used in the surveys were overly similar and produced overly correlated information. The number of data points present in the data set, however, is not high enough statistically to apply PCA reduction (or other methods).

For the SUS questionnaire, in order for the 95% confidence interval to be above the mean score with basic navigation the average score for social navigation and social navigation with emotion detection would have to be above 3.1, which as can be seen in Figure 12 has been achieved by a significant margin. This confidence interval uses the mean and standard deviation of the data and, given the 39 samples collected, we can still assume normal distribution within the data. In the case of the TIA questionnaire, we needed above 3.2 which was again achieved by some margin as seen in Figure 13.

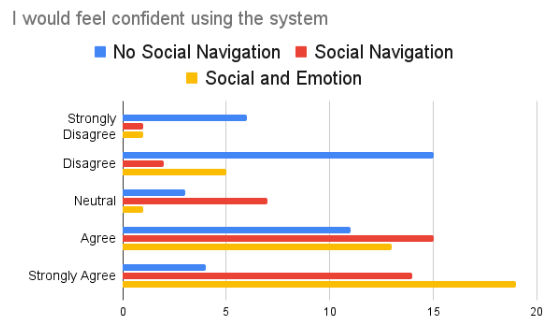

Figure 11: Highlight of the result from the SUS questionnaire: level of confidence for different system topologies

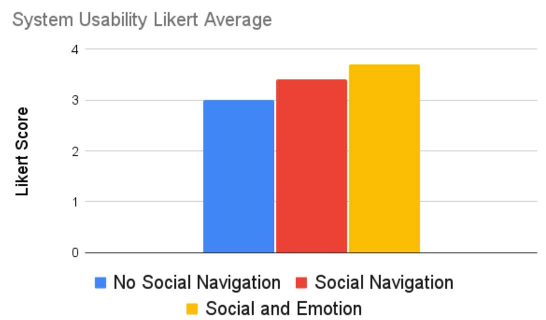

Figure 12: Grand average result from the SUS questionnaire

7 Discussion

One of the most significant challenges to the system, remains to be the odometry and pose drift. This led to difficulties in accurately localising the wheelchair within its environment – disrupting routing and navigation capabilities. To this avail,

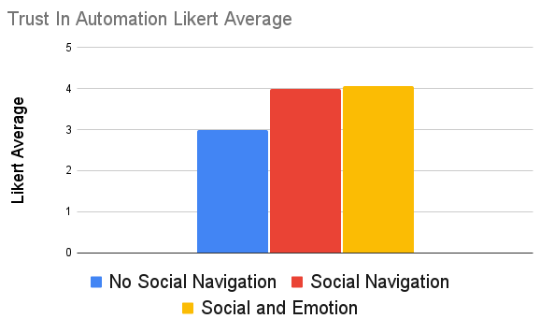

Figure 13: Grand average result from the TiA questionnaire

the integration of wheel encoders could provide a more precise measurement of the robot's displacement, mitigating the drift. Secondly, incorporating different sensor classes, such as LIDAR, may offer an alternative sensor less sensitive to environmental factors like lighting conditions. Finally, the addition of a further localisation source, such as GPS or IPS, could aid in distinguishing between homogeneous rooms or different corridors.

Initial tests found that the ZED2I could keep track of the robot with sufficient accuracy, indeed enough to maintain low enough error between Aruco detections. However, later in the project implementation, we began to notice some issues with this system, the cause of which was sometimes unclear. Occasionally, the pose tracking would jump around randomly, which was eventually diagnosed as a consequence of the camera only having a white wall in its view and so not having any key points to localise itself. During the demonstration of our robot, the pose tracking errors were noticed to be far higher than in testing. It is suspected this may be due to the extremely high number of people in the environment, which may have led to many of the key points moving around due to being located on humans in view of the camera. Using a basic SLAM system, such as the use of ACML with an environment map, would likely have resulted in the same error and mitigating this would potentially require further research into SLAM algorithms that can account for non-static objects in the environment.

One of the shortcomings of the laser path indicator is under bright lighting conditions, depending on the material and colour of the flooring, the projected path might appear less distinct. To overcome this a green laser diode can be used as it will be perceived as brighter by the eye due to its wavelength. During the implementation process of the laser path indicator, both red and green lasers were tested. It was evident that the green laser appeared brighter. However, the laser path was less visible when the green laser was used due to the dark blue carpeting in the testing room therefore, the final design includes the red laser diode. Another solution would be to use a laser diode with a higher power output. In this case, the safety of the laser would be of a larger concern as accidental exposure would be damaging to the eye.
Respondents to the questionnaire also noted a visual "drift" in the path drawn by the lasers. This can be attributed to the lack of positional feedback in the control of the stepper motors. Hence, after prolonged periods of operation, the accumulated effects of overstepping lead to the above-described phenomena. To rectify this, encoders can be used with the motors to provide a source of feedback for position control. However, the added computational demands of integrating position control for both axes exceed the capacity of the Arduino, necessitating the use of a faster and more powerful microcontroller such as an STM32.
The drift issue could also potentially be overcome by using commercially-built laser galvo assemblies. Such galvos use electromagnets, rather than stepper motors, to control the mirrors. This not only increases the precision and movement speed of the mirrors but also eliminates any overstepping and the drift associated with it. Furthermore, the enhanced precision of the mirrors in commercially built galvos will improve the resolution of the path line drawn.

8 Conclusions & Future work

This study investigated the set of hypotheses in 2.1 using a questionnaire-based approach. These preliminary findings indicate that the overall system is functioning efficiently and providing promising results in relation to the research hypotheses. However, for a more comprehensive understanding and validation of the hypotheses, it is recommended that future research involves testing the smart wheelchair with actual users. This approach allows participants to experience the system first hand leading to more representative results. Moreover, following respondents' feedback, motivation to integrate all components of the hypotheses to create a more cohesive and robust system has blossomed. This would entail refining the wheelchair's odometry to enhance its navigational capabilities, as well as developing more robust emotion and social models, ensuring that the smart wheelchair is better equipped to interact with users and their surroundings.
Despite the limitations associated with the current approach, a solid foundation for future development has been established. Further, the results obtained thus far underscore a necessity for further investigation through human experiments.

References

[1] C. Kubicek, B. Jovanovic, and G. Schwarzer, "The relation between crawling and 9-month-old infants' visual prediction abilities in spatial object processing," **Journal of Experimental Child Psychology**, vol. 158, pp. 64–76, 6 2017. pages 1

[2] T. Zhang, J. Yang, N. Liang, B. J. Pitts, K. O. Prakah-Asante, R. Curry, B. S. Duerstock, J. P. Wachs, and D. Yu, "Physiological measurements of situation awareness: A systematic review," **Human Factors**, 11 2020. [Online]. Available: https://journals.sagepub.com/doi/full/10.1177/0018720820969071 pages 1

[3] D. K. Anson, **Assistive Technology for People with Disabilities**. ABC-CLIO, 2018. pages 1

[4] H. Pedersen, S. Söderström, and P. S. Kermit, ""the fact that i can be in front of others, i am used to being a bit behind": how assistive activity technology affects participation in everyday life," **Disability and rehabilitation. Assistive technology**, vol. 16, pp. 83–91, 2021. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/31411518/ pages 1

[5] J. Damiao, M. Pizzi, C. Kiely, S. B. Are, M. Chitre, A. B. Dolan, D. Numa, and M. Yastion, "Adapting an off-the-shelf mobility device for early intervention: A case study,"

https://doi.org/10.1080/19411243.2021.2003731, vol. 15, pp. 418–428, 2021. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/19411243.2021.2003731 pages 1

[6] L. Nilsson, M. Eklund, P. Nyberg, and H. Thulesius, "Driving to learn in a powered wheelchair: the process of learning joystick use in people with profound cognitive disabilities," The American journal of occupational therapy : official publication of the American Occupational Therapy Association, vol. 65, pp. 652–660, 11 2011. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/22214109/ pages 1

[7] M. Lisotto, P. Coscia, and L. Ballan, "Social and scene-aware trajectory prediction in crowded spaces." pages 2

[8] C. Dondrup, "Human-robot spatial interaction using probabilistic qualitative representations," 2016. pages 2

[9] C. Dondrup, N. Bellotto, M. Hanheide, K. Eder, and U. Leonards, "A computational model of human-robot spatial interactions based on a qualitative trajectory calculus," Robotics 2015, Vol. 4, Pages 63-102, vol. 4, pp. 63–102, 3 2015. [Online]. Available: https://www.mdpi.com/2218-6581/4/1/63/htmhttps://www.mdpi.com/2218-6581/4/1/63 pages 2

[10] Y. Morales, T. Miyashita, and N. Hagita, "Social robotic wheelchair centered on passenger and pedestrian comfort," Robotics and Autonomous Systems, vol. 87, pp. 355–362, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092188901630570X pages 2

[11] S. Ali, A. Lam, H. Fukuda, Y. Kobayashi, and Y. Kuno, "Smart wheelchair maneuvering among people," in Intelligent Computing Methodologies, D.-S. Huang, Z.-K. Huang, and A. Hussain, Eds. Cham: Springer International Publishing, 2019, pp. 32–42. pages 2

[12] A. M. Fattouh, O. Horn, and G. Bourhis, "Emotional bci control of a smart wheelchair," 2013. pages 2

[13] E. Ambarwati, A. Arifin, M. H. Fatoni, I. W. N. B. Pradivta, T. A. Sardjono, F. Arrofiqi, and A. Risciawan, "Subject intention speed control of electric wheelchair for person with disabilities using myoelectric signals," 2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM), pp. 12–17, 2020. pages 2

[14] Y. Rabhi, M. Mrabet, and F. Fnaiech, "A facial expression controlled wheelchair for people with disabilities," Computer Methods and Programs in Biomedicine, vol. 165, pp. 89–105, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260717315456 pages 2

[15] E. Rohmer, P. Pinheiro, E. Cardozo, M. Bellone, and G. Reina, "Laser based driving assistance for smart robotic wheelchairs," in 2015 IEEE 20th Conference on Emerging Technologies Factory Automation (ETFA), 2015, pp. 1–4. pages 2

[16] M. Zolotas, J. Elsdon, and Y. Demiris, "Head-mounted augmented reality for explainable robotic wheelchair assistance," in 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2018, pp. 1823–1829. pages 2

[17] M. Edmonds, F. Gao, H. Liu, X. Xie, S. Qi, B. Rothrock, Y. Zhu, Y. Wu, H. Lu, and S. Zhu, "A tale of two explanations: Enhancing human trust by explaining robot behavior," Science Robotics, vol. 4, p. eaay4663, 12 2019. pages 3

[18] "Real time appearance based mapping (rtabmap), howpublished=http://introlab.github.io/rtabmap/, year = ,." pages 3

[19] "Ros state estimation node," http://docs.ros.org/en/noetic/api/robot_localization/html/state_estimation_nodes.html#ekf-localization-node. pages 4

[20] ROS Wiki, "aruco_ros," http://wiki.ros.org/aruco_ros, accessed 2023-03-29. pages 4

[21] "social_navigation_layers - ROS Wiki." [Online]. Available: http://wiki.ros.org/social_navigation_layers pages 5

[22] V. A. L. Sosa-León and A. Schwering, "Evaluating automatic body orientation detection for indoor location from skeleton tracking data to detect socially occupied spaces using the kinect v2, azure kinect and zed 2i," Sensors, vol. 22, no. 10, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/10/3798 pages 5

[23] F. Setti, C. Russell, C. Bassetti, and M. Cristani, "F-formation detection: Individuating free-standing conversational groups in images," PloS one, vol. 10, no. 5, p. e0123783, 2015. pages 5

[24] S. Li and W. Deng, "A survey on : Facial emotion recognition," International Journal for Research in Applied Science and Engineering Technology, vol. 13, pp. 1195–1215, 2019. pages 6

[25] N. Samadiani, G. Huang, B. Cai, W. Luo, C. H. Chi, Y. Xiang, and J. He, "A review on automatic facial expression recognition systems assisted by multimodal sensor data," Sensors (Switzerland), vol. 19, 4 2019. pages 6

[26] D. Bersimas, A. Delarue, P. Jaillet, and S. Martin, "The price of interpretability a preprint," 2019. pages 6

[27] H. Guerdelli, C. Ferrari, W. Barhoumi, H. Ghazouani, and S. Berretti, "Macro- and micro-expressions facial datasets: A survey," Sensors (Basel, Switzerland), vol. 22, 2 2022. [Online]. Available: /pmc/articles/PMC8879817//pmc/articles/PMC8879817/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC8879817/ pages 6

[28] S. I. Serengil and A. Ozpinar, "Lightface: A hybrid deep face recognition framework," in 2020 Innovations in Intelligent Systems and Applications Conference (ASYU). IEEE, 2020, pp. 23–27. [Online]. Available: https://doi.org/10.1109/ASYU50717.2020.9259802 pages 6

[29] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," University of Massachusetts, Amherst, Tech. Rep. 07-49, October 2007. pages 6

[30] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in British Machine Vision Conference, 2015. pages 6

[31] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," CoRR, vol. abs/1503.03832, 2015. [Online]. Available: http://arxiv.org/abs/1503.03832 pages 6

[32] B. Amos, B. Ludwiczuk, and M. Satyanarayanan, "Openface: A general-purpose face recognition library with mobile applications," CMU-CS-16-118, CMU School of Computer Science, Tech. Rep., 2016. pages 6

[33] Y. Taigman, M. Y. Marc', A. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification." pages 6

[34] M. Ferguson, "dynamic_reconfigure - ros wiki," 7 2015. [Online]. Available: http://wiki.ros.org/dynamic_reconfigure pages 6

[35] P. Tol, "Designing for color blindness," 2021. [Online]. Available: https://personal.sron.nl/~pault/data/colourschemes.pdf pages 7

[36] C. Castelfranchi and Y.-H. Tan, "The role of trust and deception in virtual societies," Hawaii International Conference on System Sciences, vol. 7, p. 7011, 01 2001. pages 8