Head-Mounted Augmented Reality for Explainable Robotic Wheelchair Assistance

Mark Zolotas, Joshua Elsdon and Yiannis Demiris

Abstract—Robotic wheelchairs with built-in assistive features, such as shared control, are an emerging means of providing independent mobility to severely disabled individuals. However, patients often struggle to build a mental model of their wheelchair’s behaviour under different environmental conditions. Motivated by the desire to help users bridge this gap in perception, we propose a novel augmented reality system using a Microsoft Hololens as a head-mounted aid for wheelchair navigation. The system displays visual feedback to the wearer as a way of explaining the underlying dynamics of the wheelchair’s shared controller and its predicted future states. To investigate the influence of different interface design options, a pilot study was also conducted. We evaluated the acceptance rate and learning curve of an immersive wheelchair training regime, revealing preliminary insights into the potential beneficial and adverse nature of different augmented reality cues for assistive navigation. In particular, we demonstrate that care should be taken in the presentation of information, with effort-reducing cues for augmented information acquisition (for example, a rear-view display) being the most appreciated.

I. INTRODUCTION

Independent mobility plays a significant role in our everyday activities and quality of life. However, many severely disabled individuals are incapable of exercising this fundamental ability and rely on the provision of assistive mobility platforms. By augmenting the autonomy of the disabled community, assistive robots present one of the most promising avenues for helping reduce the burden on global healthcare services [1, 2]. Within this domain of research, robotic wheelchairs have made impressive scientific advances through a wide range of unconventional input methods, such as brain-machine interfaces [3] and head motion [4], as opposed to traditional joystick control. Improvements in navigational assistance algorithms have also contributed to bettering user-technology integration for powered mobility [5, 7].

In spite of the noteworthy engineering efforts in robotic wheelchair design, the migration from controlled lab environments to widespread commercial use remains an ongoing endeavour for modern healthcare [8]. A predominant reason for this absence lies in the underlying complexity of providing assistance to people with cognitive or motor impairments, whom lack the sensorimotor capacity to steadily navigate an environment using a standard joystick-controlled wheelchair [4, 9]. One methodology of adjusting for these noisy and unpredictable signals is to engage in shared control, i.e. a continuous blending of the motor commands generated by a human operator and an intelligent controller [10].

Although typical features of a control-sharing paradigm, such as reactive obstacle avoidance, are added to ensure safety, they may instead disorient or frustrate the user whenever an input command does not elicit the intended system response [11]. In turn, this outcome could hinder a patient’s ability to learn how to navigate a wheelchair and thus lead to their failure in fulfilling the strict eligibility criteria for acquiring ownership of these mobility platforms [9].

Immersive technologies involving head-mounted displays (HMDs) are an emerging solution to help users overcome the steep learning curve associated with the adoption of navigational assistance on powered wheelchairs. Virtual reality HMDs have recently garnered attention as apt training simulators for off-line learning of wheelchair control [12, 14]. However, augmented reality (AR) HMDs could potentially serve as a more transparent mode of communicating assis-
tance, but have yet to be integrated into physical wheelchairs for on-line operation.

In this paper, we propose a novel AR system on a robotic wheelchair with built-in shared control (shown in Figure 1). A Microsoft Hololens is incorporated into our real-world setup for the purpose of highlighting to a user the inner workings of the control-sharing methodology. Using this system to explain the rationale for assistive intervention, we explore how different AR aids affect the user’s experience and learning of our robot’s internal model.

Therefore, the two main contributions presented in this paper are: 1) an AR system that renders the internal state of a shared controller for powered mobility onto the driver’s view of the world; 2) a pilot study that evaluates the acceptance rate and learning curve of an immersive training regime for wheelchair control with a variety of tested visualisations.

II. RELATED WORK

Devising policies for shared control on robotic wheelchairs is a complex and challenging process that requires careful consideration for the user’s needs and demands. Safety mechanisms inherent in shared control, such as collision avoidance, offer severely disabled individuals with the ability to manoeuvre independently and hone their navigational skill without the risk of accidents [15, 16]. In practice though, these safety mechanisms can foster a distorted interpretation of the expected system behaviour due to the lack of communication between wheelchair and user [11]. A mode of relaying back information to the user about the underlying autonomy of the wheelchair is thus a vital component of any commercial prototypes [7, 17].

Haptic controllers are one group of feedback tools that have been applied in smart mobility research to emulate expert human assistance. The general concept behind their utility is to circumvent the unnatural aspects of robotic autonomy and instead adjust a driver’s manual steering input onto a safer path via a remote human navigator, such as a therapist [18]. In doing so, these controllers aim to introduce a more intuitive form of assistance to the disabled community. Moreover, shared control policies can be derived from learning assistance by demonstration methodologies, hence removing the obligation for a therapist to remain in the loop [6, 19].

Despite reported successes at reproducing expert-level aid through haptic interfaces, there are a few specific assumptions underlying this approach. First and foremost, the internal plan of the remote navigator may not perfectly agree with the plan originating from the primary user. Secondly, the external navigator’s commands may not maintain consistency and effectiveness throughout the entire session. Finally, a third-person perspective on a task is a transformed frame of reference that could result in misguided assistance.

Immersive technologies aim to resolve the aforementioned challenges of remote feedback by visually motivating patients into building better mental models of the navigational assistance from an embodied perspective. Virtual reality is an increasingly prevalent means of simulating a safe powered mobility testbed for such use-cases [12, 13]. However, a user’s embodied experience when interacting in a virtual reality is largely influenced by the type of display, with HMDs yielding an increased sense of presence over monitors [14]. A physical aid also frequently outperforms the simulated counterpart in terms of learning system behaviour [20]. For example, robotic wheelchairs with a mounted humanoid robot companion [20] or a mobile AR interface [21, 22] have exhibited enhanced navigation performance and user preference.

As a result, we advocate taking the best of both worlds via a novel system that integrates an AR headset onto an actual robotic wheelchair. Under this framework, we seek to inform a user’s internal models into growing accustomed to our collaborative wheelchair assistance. In making this assistance more transparent to the driver, we hope to derive improvements in their navigational skill-set.

III. ASSISTIVE FEEDBACK VIA AUGMENTED REALITY

In this section, we present our proposed AR interface and assistive robot architecture (refer to Figure 1 for an overview). Figure 2 summarises the main components of our system, all of which are developed atop the Robot Operating System (ROS) [23] and presented in the following sections.

A. Shared Control

Based on the hybrid approach to shared control applied in prior work [6, 24], our method consists of two principal functions: trajectory generation and obstacle avoidance. Given a current odometry and an input velocity command, the trajectory generation process involves projecting the robot’s state forward in time according to differential-drive kinematic constraints. The output trajectory is then validated for safety against an obstacle map, which is represented as a polar histogram. Incoming laser scans are utilised to construct and update the state of this histogram. In line with [24], we
assume a simple binary model \( pr(m|o_{1:T}, x_{1:T}) \) to estimate the probability of an obstacle’s presence on a map \( m \). For each map location \( m_i \), an observation \( o \) at robot pose \( x \) reports a value of 1 for a detected collision in the forward simulation period \( T \), and 0 otherwise.

The shared controller subsequently determines how best to assist a user based on the binary output of this collision-checking routine. In the scenario where no collision is identified, the raw manual input of the driver is directly sent to the wheelchair’s motors. However, if the initial trajectory is at risk of collision, then an obstacle avoidance process must ensue.

To appropriately arbitrate the user’s input, we select the highest-scoring command from a range of discretely sampled command velocities in the robot’s control space. The scoring process for these prospective commands makes use of a simplified variant of the well-known dynamic window approach (DWA) for reactive collision avoidance \cite{25}. We thus compute the “optimal” velocity command \((v^*,\omega^*)\), with translational \( v \) and angular \( \omega \) velocities, according to the following objective function:

\[
(v^*,\omega^*) = \underset{(v,\omega)}{\text{argmax}} \left( \alpha \cdot \text{freezone}(v,\omega) + \beta \cdot \text{heading}(v,\omega) + \gamma \cdot \text{vel}(v,\omega) \right)
\]

Where \text{freezone} is a measure of clearance from nearby obstacles, whilst \text{heading} and \text{vel} are respectively heuristics for the angular and translational speed preferences of the user. The \( \alpha,\beta,\gamma \) parameters determine how to weight each of these objective measures.

Each input velocity pair \((v,\omega)\) is evaluated on the basis of the mixed-weighting operation presented in Eq. \(\text{(1)}\). To calculate the \text{freezone} value, the polar densities maintained by our obstacle map’s histogram grid are scaled to the range \( 0 \rightarrow 1 \) using a fitted exponential function. On the other hand, the \text{heading} and \text{vel} measures are generated depending on their closeness to the driver’s intended commands:

\[
\text{heading} = \exp\{-\sqrt{(\omega - \omega_s)^2}\}
\]

\[
\text{vel} = \exp\{-\sqrt{(v - v_s)^2}\}
\]

Where \((v_s,\omega_s)\) are the sampled velocities. The exponential base functions are introduced to steeply scale misaligned inputs and thus preserve similarity with the user’s original intention \cite{5}. For the purposes of this work, we set the mixing parameters to be \( \alpha = 1.2, \beta = 0.9, \gamma = 0.4 \), such that the obstacle avoidance algorithm overcompensates for collision-prone commands and prioritises user safety. Readers are referred to \cite{24} for more details on the original method.

**B. Gridmap Processing**

Whilst our shared control framework captures information relating to a driver’s navigational input, the gridmap processor instead represents the environmental context from sensor data. Given incoming rangefinder data, this module identifies dangerous obstacles in the surroundings and constructs an image view of this information to relay back to the user visually via AR (presented in Section \text{III-C}). All processing steps are entirely local and do not rely on a static map.

There are four major phases involved in the gridmap processing pipeline (shown in Figure \text{3}). Laser scan readings are first converted into a 2D occupancy grid and then translated into a binary image where occupied cells are mapped to white. This image is then morphologically dilated to enlarge these obstacle regions. The third phase of processing overlays a grey path onto the image to capture the user’s desired route, based on their current input trajectory (see Section \text{III-A}). This path represents a forward simulation of two seconds and has been dilated to match the width of the mobile base dimensions. Finally, the greyscale image is converted to RGB and green circles are drawn at coordinates where the grey path and obstacles overlap.

Two additional steps are performed to render the resulting image below the wheelchair in a user-centred way. During the third phase of the processing pipeline, the image is also rotated to align with the robot’s reference frame. Moreover, a masking and smoothing process is applied after the fourth phase to soften the harsh white boundaries of obstacles. The final image is therefore a Gaussian filtered view of solely the grey path and any potential collisions (grey circles) en route.

**C. Proposed Visualisations for Assistive Feedback**

To compensate for the potential misalignment in a user’s interpretation of their wheelchair’s behaviour, we use a Microsoft Hololens to provide visual feedback on the robot’s dynamics. We envision that an AR headset will help users
form a better mental picture of the expected system behaviour. Furthermore, this approach could potentially reduce the levels of frustration and workload experienced by users of assistive robotic wheelchairs [5].

Figure 1 provides a summary of the four visualisations implemented as AR feedback[1]. These visualisations are displayed at three different heights relative to the user: floor level, head level and floating above head level. This spatial separation was designed to help limit the likelihood of a user being overloaded with information in any given gaze direction, or to avoid overlapping visualisations.

The first visual aid is a rear-view display, which is situated directly above the user’s normal viewing direction. From the driver’s perspective, this display behaves like a large version of a rear-view mirror, such as those found in road vehicles. The camera display also renders any other graphical effects incorporated into the holistic system. This is achieved by placing a virtual camera’s view of the artificial world containing the visualisations in the same position as the real camera. The intrinsic calibration parameters of the real camera are mapped onto the captured virtual image so as to match the strong fish-eye effect applied in the real camera’s display. By applying this fish-eye effect, the user is able to view a very wide angle, which could help navigation in tight manoeuvres typical of indoor wheelchair use.

There are two kinds of visualisations rendered onto the floor. A grey path is projected either forward or backward depending on the direction of travel, which portrays the predicted future state of the wheelchair given the current input commands. If the path intersects with an obstacle then a bright green circle is rendered at that location, which is intended to help drivers identify objects that are likely to make the shared controller intervene. The construction of this image was described in Section III-B.

The last visualisation is a pair of directional arrows that float directly in front of the user. The green arrow corresponds to the user’s joystick input and the red arrow is the final command sent to wheelchair after adjustment via shared control (see Section III-A). The arrows rotate with the direction of the corresponding command velocities and lengthen to represent their magnitude.

These four visualisations fall into two categories of relative placement from a user perspective. The arrows and rear view display appear fixed to the motion of the wheelchair, behaving similarly to instruments found in an aircraft cockpit or car dashboard. On the other hand, the grey predicted path and green collision markers are perceived as fixed to the environment, not necessarily being locked to the wheelchair as it moves or rotates.

D. Augmented Reality System Alignment

All visualisations presented in this work require appropriate alignment with both the world and mobile platform, therefore a correspondence between the frames of reference of the Hololens and wheelchair must be determined. The Hololens maintains its own internal map for the purpose of visual odometry, however by default there is no well-defined origin for the rest of the robotic system to reference. This problem was previously solved in the context of a motion capture arena in [26], however due to the multi-room nature of indoor wheelchair use, a motion capture system is not a reasonable proposition.

To solve the registration problem, three points were manually marked using the Hololens by placing virtual objects in the environment. The Hololens has a system known as spatial anchors, which use local geometry to latch objects in place despite shifts in the global map. This enables the virtual markers to persist across multiple uses of the Hololens, whilst also allowing adaptations to be made on-the-fly given any environmental changes. These three points are compared to their equivalent coordinates on the map constructed using the localisation module’s SLAM component. Utilising singular value decomposition as outlined by [27], we obtain the transform between the Hololens world and the global frame of the mobile base. In this case there are four unknown variables accounted for, three for the position offset between coordinate systems and one representing the rotation in yaw direction. These points should not be co-linear to avoid multiple solutions and should span the experimental arena to minimise the effect of placement error.

IV. EXPERIMENTS

For the evaluation of our AR system, we used an internally developed robotic wheelchair. The underlying powered wheelchair is controlled using a joystick with a circuit board that enables an Arduino UNO to translate the user issued commands into motor signals. Two Hokuyo URG-04LX-UG1 laser scanners are situated at the front of the mobile base, and a SICK LMS200 rangefinder is equipped at the back. A Phidgets spatial 3/3/3 IMU is also equipped to improve the odometry estimate of the mobile setup. An on-board laptop was used as the main driver to control the wheelchair and all the processes included in our ROS framework. The Unity game engine was used to develop the AR application and deploy it on the Microsoft Hololens. Communication with the HMD was established over a wireless router.

A. Experimental Setup

To explore the assistive effects of AR on wheelchair control, we conducted a pilot study with 16 able-bodied participants (13 male, 3 female) aged between 20 and 31. Participants were asked to sign a consent form for the collection of data and presentation in this work. Prior and post experiment questionnaires were also handed out for completion.

Each subject was requested to complete a navigation route four times in sequence, which lasted an average total duration of 30 minutes. The trial route devised for this experiment includes a subset of tasks assessed in the Wheelchair Skills Test (WST) manual (version 4.2) [28]. This route is illustrated in Figure 4 and the evaluated criteria are shown in Table I.
The purpose of the experimental task was to investigate the effectiveness of different graphical aids and whether the AR accelerated learning of the wheelchair’s behaviour. We controlled for this by assigning individuals to one of two groups: with-visualisation and without. People in the control group also wore the Hololens but without any visualisations displayed, so that head orientation data could still be recorded and that the obtrusiveness of the HMD is kept fair for both groups.

Participants in the visualisation group were distinct from the non-visualisation counterpart in two ways. First, they were administered augmented feedback and instructed on the meaning of the visualisations, although no advice on how to interpret or make use of them was provided. Secondly, subjects in the visualisation group were requested to perform the fourth attempt at the course without any graphical aid. This was designed to observe whether a dependency on the AR formed, or if the task-learnt skills were independent of these visual cues.

### TABLE I: A summary of the assessment points for the modified WST. Each of the task-specific positions is numbered correspondingly on the map in Figure 4.

<table>
<thead>
<tr>
<th>Skill</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward motion in narrow 1m passageway</td>
<td>1</td>
</tr>
<tr>
<td>Reverse in narrow 1m passageway</td>
<td>1</td>
</tr>
<tr>
<td>Turn while rolling forwards (90°)</td>
<td>2</td>
</tr>
<tr>
<td>Turn while rolling backwards (90°)</td>
<td>2</td>
</tr>
<tr>
<td>Turn in place (180°)</td>
<td>3</td>
</tr>
<tr>
<td>Traverse through open doorway</td>
<td>4</td>
</tr>
<tr>
<td>Avoid static obstacles</td>
<td>5</td>
</tr>
<tr>
<td>Stop before walls</td>
<td>A and B</td>
</tr>
</tbody>
</table>

Fig. 5: Total time to completion for each trial. The group without visualisations performed better in every trial, even in the 4th round where both groups did not have visual aid.

#### B. Empirical Findings

We assessed total time to completion for each trial as a performance indicator of the overall AR feedback. The results presented in Figure 5 indicate that the group without visualisations performed better across all trials and improved consistently in the first three rounds, having plateaued in skill by the third. On the other hand, the group with visualisations demonstrated more variable performance, with a greater decrease in time relative to their first trial, despite taking longer to plateau. When visualisations were removed on the fourth trial, there was a slight dip in performance but no strong claim can be made for any dependency forming on the AR assistance.

A possible explanation for this offset in absolute performance between the two groups is suboptimal placement of some of the virtual objects. This is especially true given the narrow field of view of the Hololens (estimated at 17.5° vertically and 30° horizontally). We therefore speculate that subjects could not make proper use of the AR assistive features outside of their natural field of view.

To further elaborate on how often participants made use of the different visualisations under these restrictive viewing conditions, dwell time was recorded by extending a ray directly forwards from the user’s head and registering intersections with virtual objects. We found that participants in the visualisation group spent a median proportion of 48.4% across the first three trials directed towards the rear-view display and floating arrows. The green obstacle cues were instead oriented towards for a median value of 32.6%. It is worth noting that the viewing direction of the subjects in the non-visualisation group would have also aligned with these obstacle cues for a median of 77.6% had they been rendered. This implies that the floor-based objects adopted a natural orientation angle for wheelchair navigation. Assuming that participants maintained a central eye-in-head position, we suspect that floor-plane features occupied a less salient region...
Fig. 6: Depicts the rate of head rotation in the yaw direction during the reverse passageway section in Figure 4. Participants with visual feedback rotate their head significantly less than the control group for the first three trials. When the visualisations are removed in the 4th round, both groups display similar mean rates of yaw rotational movement.

TABLE II: A summary of the user responses to the question: “Rate the following visualisation from 1-5 (1 = very poor, 5 = very good)”.  

<table>
<thead>
<tr>
<th>Visualisation</th>
<th>Mean User Rating</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rear View Display</td>
<td>4.125</td>
<td>0.64</td>
</tr>
<tr>
<td>User/Assistance Arrows</td>
<td>3.125</td>
<td>1.46</td>
</tr>
<tr>
<td>Projected Path</td>
<td>2.375</td>
<td>0.92</td>
</tr>
<tr>
<td>Highlighted Obstacles</td>
<td>2.375</td>
<td>0.92</td>
</tr>
</tbody>
</table>

The poor ratings associated with the grey path and obstacle cues are informative on floor-based renderings. Although information overlaid on an environment is a fundamental quality of AR, practical considerations should be made for the HMD’s field of view limitations. Some participants provided comments reinforcing this observation by stating that they could rarely notice these floor visualisations, supporting our quantitative analysis on dwell time. Furthermore, the embedding of this environment-based information mandates a user to perform a search of their surroundings, which itself could frustrate them.

Another noteworthy remark is on how intuitive different visualisations appear from an actual user’s perspective. Many subjects commented on how they misunderstood the purpose of the floating arrows, querying whether they should have aimed to match the corrective red arrow or simply taken both arrows into account as supplementary information. This leads us to believe that low-level cues, such as command indicators, are not necessarily an effective user-centred form of augmented assistance and would require auxiliary instruction to be provided. In contrast, highlighted obstacles are higher-level and hence potentially provide more intuitive feedback.

V. CONCLUSIONS

To the best of our knowledge, this is the first instance of an AR headset being incorporated into a powered wheelchair system. Our findings lead us to believe that there is potential benefit to be gained from the integration of AR headsets with robotic wheelchairs, as long as certain design choices are taken into account. Namely, that virtual objects are placed in easily visible locations that are not within proximity of the mobile base, and preferably do not clutter the natural viewing required for navigation. Secondly, that graphical cues are high-level and contextual enough for a typical user to perform a search of their surroundings, which itself could frustrate them.

In the post-experiment questionnaire, subjects were asked to rate the benefit of each of the provided visualisations on a 5-point scale. A strikingly positive result from this survey was the popularity of the rear-view display. Almost all subjects rated it as either “good” or “very good” (4.125 ± 0.64). Conversely, there was nearly universal disapproval of the grey path and the green obstacle markers (2.375 ± 0.92 for both). The overall average scores from 1-5 (5 being most positive) are listed in Table II.
REFERENCES


