Augmented Reality Control of Smart Wheelchair Using Eye-Gaze–Enabled Selection of Affordances

Rodrigo Chacón-Quesada and Yiannis Demiris[1]

***Abstract*—In this paper we present a novel augmented reality head mounted display user interface for controlling a robotic wheelchair for people with limited mobility. To lower the cognitive requirements needed to control the wheelchair, we propose integration of a smart wheelchair with an eye-tracking enabled head-mounted display. We propose a novel platform that integrates multiple user interface interaction methods for aiming at and selecting affordances derived by on-board perception capabilities such as laser-scanner readings and cameras. We demonstrate the effectiveness of the approach by evaluating our platform in two realistic scenarios: 1) Door detection, where the affordance corresponds to a *Door* object and the *Go-Through* action and 2) People detection, where the affordance corresponds to a *Person* and the *Approach* action. To the best of our knowledge, this is the first demonstration of a augmented reality head-mounted display user interface for controlling a smart wheelchair.**

***Index terms*—** Affordances, augmented reality, head mounted display, smart wheelchairs, user interface.

I. INTRODUCTION

According to the World Health Organisation [1], 15.3 per cent of the world's population have a moderate to severe disability, 93 million of whom are children. Early utilization of electric wheelchairs for people with mobility limitations promotes integration and psycho-social development, reduces passive dependency and enhances participation, function, and independence [2].

However, controlling a wheelchair can be a cognitively challenging task for some users [3]. To enable easier control of wheelchairs, a variety of control interfaces have been proposed; the most traditional method used is a joystick, but others including electromyography and/or electroencephalogram signals have also been researched [4]. Yet, most of these control interfaces use a set of learned and unnatural commands, such as in the case of a joystick, to interact with the wheelchair. In contrast, augmented reality (AR) user interfaces (UIs) exploit the user's existing abilities, such as speech, hand gestures and head and eye movements, as a way of interacting with computers [5]. We seek to reduce the cognitive barriers associated with smart wheelchair control via UIs that introduce a more natural interaction. For these reasons, we deem AR UIs are worth exploring to control smart wheelchairs.

Figure 1: **Virtual Reality User Interface**. Composite image of the visualizations rendered on the user's view through the AR headset (1). The affordance *Go Through* associated with a door was selected by the user and is shown in green (2). The affordance approach associated with a person was not selected and therefore is shown in purple (3). The cursor and timing bar (4) give feedback to the user during the aiming and selection stages respectively. Two different aiming methods can be selected by the user using an on-board menu (5). Some affordances are saved for future usage (6).

In this paper, we propose to use an AR head-mounted display (HMD) UI as a new method to tackle the task of controlling a smart wheelchair. For this, we integrated our smart wheelchair named ARTA[1] and an AR HMD UI developed for the Microsoft Hololens. To exploit more of the user's existing abilities during the interaction with the smart wheelchair, we implemented two different methods for aiming at the UI (head-gaze and eye-gaze) and three different methods for triggering actions (hand gestures, voice commands and dwell time). In addition, we integrated an affordance inference algorithm on our platform that jointly detects the objects and people in the smart wheelchairs' surroundings and informs the user about his/her affordances, i.e. how the detected objects can be used to accomplish higher level tasks.

We demonstrate the feasibility of the proposed platform with experiments where the user can control the smart wheelchair using the AR HMD UI to allow them to go through doors and approach to other people. A composite image of the visualizations rendered on the user's view through the HMD is shown in Figure 1.

---

[1] The authors are with the Personal Robotics Laboratory, Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, United Kingdom `{r.chacon-quesada17, y.demiris}@imperial.ac.uk`. More information about the Personal Robotics Laboratory at: `http://www.imperial.ac.uk/PersonalRobotics`.
RCQ was supported by a PhD studentship jointly funded by the University of Costa Rica and Ministry of Science Technology and Telecommunications, Costa Rica.

[1] http://www.imperial.ac.uk/personal-robotics/robots/

II. RELATED WORK

AR UIs exploit the user's existing abilities as ways of interaction with computers. As a result, these interactions are easier to perform and less demanding in terms of cognitive workload in comparison with conventional computer systems [5]. For this reason, we believe that AR UIs are suitable to tackle the task of controlling a smart wheelchair. Furthermore, given the variability in user preferences, allowing users to choose between multiple interaction methods, depending on which ones are more convenient for them, might help increasing the system's acceptability and the amount of potential users.

A. Interaction Techniques for User Interfaces in Virtual Environments

Typically, two stages are needed to trigger an action while interacting with UIs in augmented environments: 1) the aiming stage, which is the process of positioning a cursor over a virtual object, and 2) the selection stage, which is the process of selecting the aimed virtual object.

For the aiming stage, external hand-held controllers have been proposed. The position and orientation of these controllers are tracked so the user can aim at virtual objects. However, since the user needs to hold the controllers with their hands, the hands are unavailable to perform other actions. A current widely used technique for the aiming stage is *head-gaze*. This technique measures the head direction of the users through a sensor in the AR system and displays a cursor in the center of the users' view. The users then need to move their heads to control this cursor. Finally, *eye-gaze* is a promising hands-free interaction technique for the aiming stage and an active field of research [5], [6], [7]. This technique tracks the users' eyes with camera sensors and in combination with the head direction infers what the eyes are looking at in the environment.

Eye-gaze has several advantages over head-gaze in augmented environments in terms of speed, task load, required head movements and user preference [7]. However, the eye-gaze based aiming technique is reported to affect the spatial memory of the users, i.e. it is more difficult for the users to remember the location of the virtual object they interacted with, therefore making it less convenient for tasks that require spatial memory [5]. The eye-gaze aiming technique is also affected by the "Midas touch" problem, i.e. it is difficult to differentiate between gaze for perception or for user action [8]. In addition, it might lead to some technical issues, such as calibration quality and robustness [6], resulting in a limited application in non-research products.

The selection stage can be done using traditional techniques such as controller- or button-based selection methods, while more recent techniques are based on hand-gestures, voice commands, eye-gaze smooth pursuit [6], and eye-gaze dwell-time. In the latter, users need to remain aiming at the virtual object during a predefined amount of time to trigger the action.

In this paper we propose a platform that integrates eye-gaze and head-gaze techniques for the aiming stage and hand-gestures, voice commands and dwell-time for the selection stage. It is important for our platform to be suitable for a wide variety of scenarios where a single interaction technique may not be enough.

B. Virtual Reality and Augmented Reality for Smart Wheelchairs

VR technologies for smart wheelchairs have been restricted to specific domains, such as simulation [9] and training [10], [11], [12]. Mobile AR applications have also been researched. However, they have only been used as a way of giving additional information to the user, for example using fiducial markers [13] and beacon technologies [14]. Our lab has recently presented a novel AR-based interaction scenario, in which a wheelchair user can become more aware of the smart wheelchair's intentions in navigation tasks trough a set of visualizations provided as feedback [15].

In this paper we enhance the VR/AR interaction to enable *control* of the wheelchair. The user is informed about entities (objects and people) in the wheelchair's surroundings, but crucially they are automatically provided with ways to interact with these objects. Furthermore, we give the users the ability to control the robots through an AR UI based around the inferred affordances, i.e. the actions that are possible to be accomplished with the detected objects and people.

III. AR HMD UI-CONTROLLED AFFORDANCES-AWARE SMART WHEELCHAIR

A. System description

To present the users of our system with a set of affordances, we developed a platform that integrates our smart wheelchair named ARTA, the Microsoft Hololens augmented with an eye tracking device from Pupil Labs[2] and an affordance inference algorithm. This algorithm performs different tasks such as object detection, currently use to obtain the 2D position of people in images, and scene mapping for depth perception. In addition, this algorithm manages the navigation information coming from the smart wheelchair as door candidates and going to the smart wheelchair as positions towards which the user wants to navigate to . The communication between all the components of this platform is made using Robot Operating System (ROS) . Our platform is depicted in Figure 2.

For the aiming stage, the user can choose between head-gaze and eye-gaze methods, only one method can be used at any given time and the preferred one is selected by using an on board menu (see Figure 1). The eye-gaze aiming method also works with gaze stabilization to reduce eye tracking instabilities and to avoid the loss of targets during blinking. We applied a histogram filter inspired by [7] to achieve this stabilization. For the selection stage, the user can choose between three different methods: 1) hand-gestures, 2) voice commands and 3) dwell-time. All three selection methods can be simultaneously applied and the user can select between options, even alternating between them if preferred.

Our AR UI provides feedback to the user in both the aiming and selection stages. For instance, a red dot that shows the user's gaze position is used as a cursor that changes its color

---

[2] https://pupil-labs.com/vr-ar/

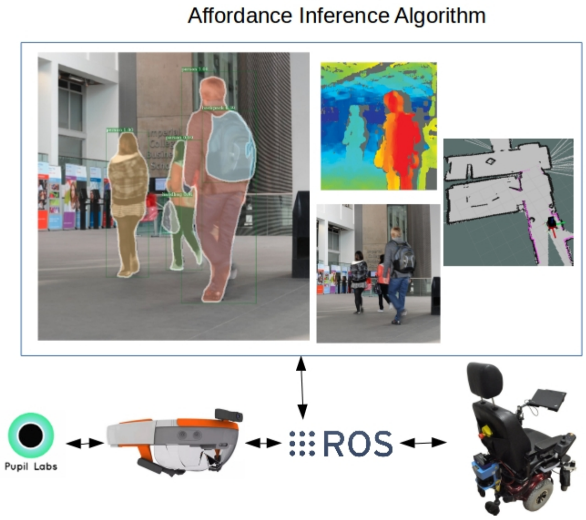


Figure 2: **System Architecture**. Schematic showing the main components of the proposed system and how they communicate between each other.

from red to green when an interactive-object (such as a UI button) is aimed. In addition, all AR UI buttons change their color when selected and a timing bar is shown for the dwell-time selection method. The detected affordances are shown to the user in the location where the detected object is and using a small hologram with a text displaying the related actions that can be performed in each case. An illustration of some of the UI interaction techniques available in our platform are shown in Figure 1.

B. Evaluation scenarios

To evaluate our platform we designed two different scenarios: 1) door detection where the affordance corresponds to a *Door* object and the *Go Through* action and 2) people detection where the affordance corresponds to a *Person* and the *Approach* action. In both scenarios, an aiming method selection menu was placed at a comfortable non-intrusive viewing angle. The world alignment, i.e. the translation and alignment between the AR HMD and the wheelchair's coordinate system, was performed as described in [15].

Based on the properties and capabilities of our robotic platform, the algorithm determines what is possible (can be afforded) in the current environmental context. For example, the action *Go Through* is only possible because the smart wheelchair can move around its surroundings (with some constraints usually handled by obstacle avoidance algorithms and shared control methods) and because an open door is detected by the affordance inference algorithm using the information coming from ARTA's navigation system. Otherwise the action is not presented to the user. In the case of the *Approach* action, it is only possible because the smart wheelchair is capable of moving around with some restrictions and because at least one person was detected in the current scene by the affordance inference algorithm using the information coming from the object detection algorithm and the depth perception methods to obtain a 3D position. If a person is not detected, the associated action is not presented to the user.

1) Door detection: In this evaluation scenario, ARTA's navigation system provides door candidate locations by means of their on-board scanning laser range-finders. These locations are then handled by our affordance inference algorithm and translated from the wheelchair's coordinate system into the AR HMD coordinate system, where our AR UI shows the option *Go Through* to the user in the door's location if an open door is detected (see Figure 3(a) for an example). The user can then select the action *Go Through* related with the door object. If this suggested action is selected, our affordance inference algorithm will send a position back to the wheelchair's navigation system (again considering the coordinate transformations needed between the smart wheelchair and the AR HMD) so the smart wheelchair can autonomously navigate through the door. As depth information is inherent to the laser scanning readings, the spatial mapping capabilities relies completely on this sensors.

2) People detection: In this second scenario, a camera is used to detect people. First, our affordance inference algorithm continuously scans the environment to retrieve depth information. This step is automatically performed by the platform and does not require any intervention from the user. Then, the user instructs the platform to take a photo of the scene using hand gestures or voice commands. Next, our affordance inference algorithm will process the image with an object detection algorithm and a hologram will be instantiated through the AR HMD at the detected person's 3D position. This is done by combining the 2D location of the person provided by the object detection algorithm and the depth information obtained in the first step (where the affordance inference algorithm scans the scene). A menu displaying the option *Approach* in the detected people's locations are shown to the user (see Figure 3(b) for an example). If the user selects the *Approach* option, the position of the person will be sent to the smart wheelchair's navigation system and it will autonomously navigate towards this position.

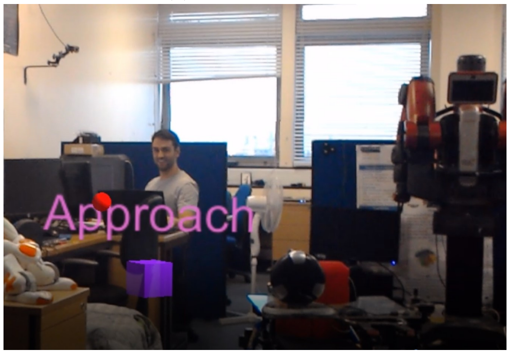
To achieve a stable display frame-rate of 60 frames per second (fps) (recommended frame-rate for mobile VR) while dealing with long-running image recognition algorithms, we followed a procedure in which the HMD manages the camera and the scene analysis requests, such as depth scanning, and sends frames to a remote computer. This remote computer processes the frames to detect people and returns a 2D position to the HMD. The HMD then transforms this 2D position from the pixel space to the application coordinate system to finally show a menu with the option *Approach* in the detected people's locations.

For this study, our affordance inference algorithm used Detectron [16] to perform the object detection task in images. These algorithms are running on a remote computer as mentioned above. For the scene mapping task in this evaluation scenario, we are currently using the Microsoft Hololens Spatial Mapping algorithms[3].

---

[3] https://docs.microsoft.com/en-us/windows/mixed-reality/spatial-mapping, last consulted on August 5, 2018

(a) Door detection



(b) People detection

Figure 3: **Test scenarios**. a) Illustration of our door detection method. b) Illustration of our people detection method.

IV. PRELIMINARY RESULTS

We ran a series of preliminary studies to evaluate the usability of our novel interface. Each of the two scenarios described in sub-section III-B have been tested with several different users in successive versions of the interface. Separate component evaluations were also performed for each affordance. For example, the door detection tests have included multiple types of doors, such as wooden-doors and wood-framed glass-doors. We conducted short interviews after each experiment to determine how the users perceived our interface, in terms of flexibility, ease of selection of interface options, ease of selection of affordances and overall impressions of the system's potential. Overall, we received very positive feedback from our users in terms of ease of use and versatility.

Based on feedback on the first iterations of the interface we enhanced it with multimodal elements including a) providing sound feedback when the user takes a photo in the people detection scenario, b) customizing the dwell-time period, among others. In ongoing work we are optimizing the aesthetics (e.g. transparency) and placement of various AR UI elements, to maximize usability while minimizing intrusion, for example through occlusion of important parts of the scene.

V. CONCLUSIONS AND FUTURE WORK

To our best knowledge, this is the first demonstration of AR HMD UI control of a smart wheelchair. The tests performed in two realistic scenarios lead us to believe that there is a substantial potential applicability of our platform to a variety of environments such as airports, hospitals, office buildings, among others. There is also potential to serve a wide variety of disability types, for example people that can not move their head but that can move their eyes, or people that does not have arms but can speak. We expect that the inclusion of natural methods of interaction (e.g. affordance-based gaze-driven action selection) as a way to control a smart wheelchair will lower the demands on the user's cognitive skills for performing useful actions, and will further advance the science of determining optimal user assistance by robots.

In ongoing work we are adding more affordances (objects and related actions) to our platform. We will also take into account the user's surroundings for better placement of the AR UI elements to avoid occlusion and enhance user experience.

REFERENCES

[1] WHO (World Health Organization), "World report on disability 2011," *American journal of physical medicine rehabilitation Association of Academic Physiatrists*, vol. 91, p. 549, 2011.

[2] L. Rosen, T. Plummer, A. Sabet, M. L. Lange, and R. Livingstone, "Position on the Application of Power Mobility Devices for Pediatric Users," *Assistive Technology*, pp. 1–9, dec 2017.

[3] S. Massengale, D. Folden, P. McConnell, L. Stratton, and V. Whitehead, "Effect of visual perception, visual function, cognition, and personality on power wheelchair use in adults," *Assistive Technology*, vol. 17, no. 2, pp. 108–121, 2005.

[4] J. Leaman, H. M. La, and L. Nguyen, "Development of a smart wheelchair for people with disabilities," in *2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pp. 279–284, Sept 2016.

[5] V. Tanriverdi and R. J. Jacob, "Interacting with eye movements in virtual environments," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 265–272, ACM, 2000.

[6] S. Schenk, P. Tiefenbacher, G. Rigoll, and M. Dorr, "Spock: A smooth pursuit oculomotor control kit," in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 2681–2687, ACM, 2016.

[7] J. Blattgerste, P. Renner, and T. Pfeiffer, "Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views," in *Proceedings of the Workshop on Communication by Gaze Interaction*, p. 1, ACM, 2018.

[8] R. J. Jacob, "Eye tracking in advanced interface design," *Virtual environments and advanced interface design*, pp. 258–288, 1995.

[9] I. Ktena, W. Abbott, and A. A. Faisal, "A virtual reality platform for safe evaluation and training of natural gaze-based wheelchair driving," *IEEE Neural Eng (NER)*, vol. 7, pp. 236–239, 2015.

[10] L. Devigne, M. Babel, F. Nouviale, V. K. Narayanan, F. Pasteau, and P. Gallien, "Design of an immersive simulator for assisted power wheelchair driving," in *Rehabilitation Robotics (ICORR), 2017 International Conference on*, pp. 995–1000, IEEE, 2017.

[11] A. Alshaer, H. Regenbrecht, and D. O'Hare, "Immersion factors affecting perception and behaviour in a virtual reality power wheelchair simulator," *Applied ergonomics*, vol. 58, pp. 1–12, 2017.

[12] N. W. John, S. R. Pop, T. W. Day, P. D. Ritsos, and C. J. Headleand, "The implementation and validation of a virtual environment for training powered wheelchair manoeuvres," *IEEE transactions on visualization and computer graphics*, vol. 24, no. 5, pp. 1867–1878, 2018.

[13] L. C. d. Oliveira, A. B. Soares, A. Cardoso, A. d. O. Andrade, and E. A. Lamounier Júnior, "Mobile augmented reality enhances indoor navigation for wheelchair users," *Research on Biomedical Engineering*, vol. 32, no. 2, pp. 111–122, 2016.

[14] L. C. d. Oliveira, A. O. Andrade, E. C. de Oliveira, A. Soares, A. Cardoso, and E. Lamounier, "Indoor navigation with mobile augmented reality and beacon technology for wheelchair users," in *Biomedical & Health Informatics (BHI), 2017 IEEE EMBS International Conference on*, pp. 37–40, IEEE, 2017.

[15] M. Zolotas, J. Elsdon, and Y. Demiris, "Head-mounted augmented reality for explainable robotic wheelchair assistance," *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2018)*, p. To be published, 2018.

[16] R. Girshick, I. Radosavovic, G. Gkioxari, P. Dollár, and K. He, "Detectron," 2018.