2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)

Glance-Box: Multi-LOD Glanceable Interfaces for Machine Shop Guidance in Augmented Reality using Blink and Hand Interaction

Grigoris Daskalogrigorakis *
Electrical and Computer Engineering
Technical University of Crete
Chania, Greece

Ann McNamara †
Department of Visualization
Texas A&M University
United States

Angelos Marinakis ‡
Electrical and Computer Engineering
Technical University of Crete
Chania, Greece

Aristomenis Antoniadis §
Electrical and Computer Engineering
Technical University of Crete
Chania, Greece

Katerina Mania ¶
Electrical and Computer Engineering
Technical University of Crete
Chania, Greece

Figure 1: Glance-box 3-LOD interface. Simple Glanceable (SG Left). Detailed Glanceable (DG Middle). Holo-Box (HB Right).

Abstract

Glanceable User Interfaces for Augmented Reality (AR) reveal virtual content "at a glance," providing rapid information retrieval, often based on gaze interaction. They are ideal when the augmented content covers a small proportion of the view space. When the size of virtual content grows, the potential to occlude the real-world increases provoking safety concerns. Compounding this is the *Midas Touch Problem*, where users unintentionally select virtual elements by simply looking at them. Extending dwell time does not eliminate involuntary selections, impeding interaction time. In this work, we present *Glance-Box*, a novel interaction system for AR combining Glanceable interfaces and world-based 3D interfaces across three Levels-Of-Detail, including progressively more information and visuals. Glance-box combines eye-gaze and hand interactions, focusing on user safety. A 2D Glanceable interface facilitates rapid information retrieval at a glance, while extended 3D interfaces provide interaction with denser content and 3D objects. Glance-Box couples blink-based and gaze-based interactions to minimize errors arising from the Midas Touch Problem. While applicable across domains, the Glance-Box interface is designed and optimized for performing manufacturing tasks in the real world. We evaluated the Glance-Box interface using an object selection task of a manufacturing process. Participants completed tasks faster using Glance-Box, employing less dense LOD over time as they gained experience. The perceived accuracy of Glance-Box gaze-based input was high, even when the device's eye tracker accuracy was coarse.

Index Terms: Human-centered computing—Mixed / augmented reality—Augmented Reality——Human-centered computing—Interaction Desing—Gaze-based interaction Human-centered computing—Interaction Desing—User interface design—

1 Introduction

The manufacturing floor of modern machine shops are increasingly employing Augmented Reality (AR) technology as the evolution of the industry moves towards the *industry 4.0 model* [7]. A limiting factor of AR Head-Mounted Displays (HMDs) stems from the narrow Field of View (FoV) available in current hardware [18], [17]. The narrow FoV restricts the amount of virtual content that can be viewed at one time which, in turn, occludes real-world elements critical to the task at hand, jeopardizing users' safety. Workers employ AR to safely retrieve virtual information eliminating the distraction and spatial limitations of traditional 2D screens [14], [15]. AR applications require intuitive interfaces that do not impede or endanger the user aiming to be efficient at work [11]. *Glanceable User Interfaces* drive information retrieval "at a glance" based on gaze detection when using an AR HMD with integrated eye-tracking [23]. Glanceable interfaces are proven helpful in cases when the user is focused on a real-world task that directly requires keen attention, providing appropriate guidance without obstruction [1]. However, they suffer from three main drawbacks: (1) Limited Scalability: As the amount of AR content increases, Glanceable interfaces lose effectiveness due to the volume of data to be displayed, and the potential to block real-world content [23] [6]. (2) Viewing Angle: Problems can arise when viewing augmented content at an angle [20]. (3) Unintentional selection/interaction using gaze: This is the Midas Touch Problem (MTP) [12].

A Glanceable interface should be compact, enabling minimal information at a glance to avoid clutter, occlusion, and distraction [3]. When a large amount of data is necessary to guide the user through

*e-mail: gregthask.bt1@gmail.com
†e-mail: ann@viz.tamu.edu
‡e-mail: amarinakis@isc.tuc.gr
§e-mail: antoniadis@dpem.tuc.gr
¶e-mail: amania@tuc.gr

2771-1110/22/$31.00 ©2022 IEEE
DOI 10.1109/ISMAR-Adjunct57072.2022.00070

the task, AR elements occupy a large portion of the available visible space. Interaction with virtual content becomes less accurate, and in workplaces such as machine shops, obstruction of the visual field is dangerous. Voice and gesture commands have been offered as an alternative to visual instruction [16]. In a machine shop, however, AR interaction based on speech is not well-suited due to the high levels of noise, while mid-air hand gestures and wide controller motions can pose safety concerns.

We present Glance-box, a novel AR interface that combines Glanceable AR and world-based 3D interfaces employing a combination of eye-gaze and hand interactions. We evaluate this in a machine shop, considering workplace safety issues, boosting productivity and safety in a manufacturing process. The contribution of this work is threefold:

Firstly, we present a novel, gaze-activated 3-tier LOD system (3-LOD) that initially launches with a compact interface and progressively expands to reveal more information in each LOD, with minimal real-world occlusion. Two initial LOD interfaces are controlled through glance (Glanceable interfaces), while the deepest level of LOD is a 3D world-based interface. The interface enables the user to view simple guidance information through the Glanceable interfaces or select to view the 3D interface to examine more detailed information. The interface accurately decodes gaze even when the accuracy of the eye tracker is low. Secondly, we present a new interaction system that utilizes a combination of eye-gaze blinking with a delay (dubbed blink-delay) for interacting with Glanceable interfaces to minimize unintended interactions. Objects remain selected and interactable even if the eye cursor exits their activation trigger unless another interactable object is selected. Glanceable interfaces are visually adjusted to minimize errors of interaction by placing buttons on opposite sides of the interface. The 3D interfaces allow hand tracking interactions with a delay (dubbed hand-delay). Users can engage interchangeably between hand-delay and blink-delay; thus, interactable objects can be placed on any layout without restrictions. Thirdly, the content shown in each LOD is adjusted, guiding users to varied levels of knowledge on its own. In the most compact Glanceable interface, we show only textual information such as the name and material of a given task. The second LOD level of the Glanceable interface offers additional text and 2D images. The subsequent 3D interface includes 3D objects and multiple interactable buttons. The Glanceable interfaces provide adequate guidance for experienced users, while the 3D interface provides additional information for inexperienced users using 3D models relevant to manufacturing tasks.

We evaluate our system by employing object selection involving milling and turning manufacturing tasks in the form of gamified missions. Users completed missions faster and used more compact interfaces as they gained more experience. Our proposed interaction system compensates for the eye tracker's inaccuracies resulting to higher perceived accuracy of blink-delay input than hand-delay.

2 Related Work

2.1 Attentive User Interfaces and Context-Aware Interaction

Attentive User Interfaces (AUIs) gather input passively to determine user's attention and then rank virtual communication based on importance according to the active task [32]. A simple signal on their periphery can prompt users. Only when they respond to the given signal the virtual information is presented. Context-aware UIs optimize interaction by providing the most relevant information to users [9]. The challenge lies in trying to model and understand users' intentions based on their actions and context and finding an optimal way to present such information. On the other hand, user-triggered adaptations could lead to increased controlability and predictability [30]. Gaze-based adaptation is useful in social environments [29]. Users find gaze-adaptive interfaces less distracting [28]. Eye-tracking can introduce spatial and temporal context to augmented information, allowing the system to show only the most relevant augmented interfaces [26]. In austere environments, context-based adaptation reduces errors in information retrieval [33]. In contrast, in noisy environments, e.g., in supermarket aisles, the context-less presentation proved to be more accurate despite the visual clutter [25] while users interacted with a VR simulation without performing a real-world task. Virtual interfaces have been adapted by repositioning or adjusting the transparency of virtual interfaces to allow the user to focus on two independent tasks simultaneously, one on the real world and one on virtual objects, as shown in a VR simulation of an AR application [4]. Participants intuitively focused primarily on the real-world task and left viewing virtual information as secondary, even when asked to do the opposite. Manual adaptation of the content of virtual interfaces during a pre-defined manufacturing maintenance task occluded multiple real-world objects, which is a safety hazard [34]. In this work, Glanceable interfaces are utilized that limit occlusions, while interaction with virtual content occurs through eye-tracking or hand-tracking only in pre-defined safe areas.

2.2 Glanceable AR and Gaze-based Interaction

Glanceable AR allows for interaction in order to access information in AR. The information is placed at the periphery to remain unobtrusive, activated by a glance when required [23]. Virtual information is either always visible, or it can be "summoned" through user interaction, often using eye-tracking, without noticeable distraction to the user. Glanceable interfaces can also be anchored in the real world for stationary tasks or follow the user when needed for mobile tasks [21]. The combination of Glanceable interfaces with context-awareness showed faster information retrieval in social contexts, such as when conversing with another person [5], being more socially acceptable than using a smartphone. Glanceable interfaces, however, are useful when the information shown on each interface is minimal. Otherwise, they take up a large area of screen space obstructing the real-world view, especially when the interfaces require precise and rapid interactions. In this work, we utilize Glanceable interfaces for fast information retrieval combined with a 3D interface showing a larger amount of information including 3D objects. The 3D interface remains at a preset point, unlike Glanceable interfaces, allowing for precise hand-based input without imposing a safety risk.

The combination of Glanceable and context-aware interfaces can lead to adaptive interfaces that adjust both their visuals and their content based on the user's cognitive load [20]. By designing interfaces in multiple Levels of Detail (LODs), less relevant interfaces are hidden or minimized when the cognitive load increases, while showing additional information when needed. In our work, multiple LOD interfaces are implemented in the form of a 3-tier LOD system (3-LOD). The LODs in our system are adapted using direct input to avoid potential distractions as occurring in past work.

The effectiveness of Glanceable gaze-adaptive interfaces varies based on the scenario [27]. The "Midas Touch" problem occurs when an eye-tracking system cannot differentiate between "viewing" and "interacting" with virtual objects [12]. In social scenarios with minimal user input, the results were favorable. In interfaces with small interactable objects, multiple interactable objects were often close to each other. The system could not correctly determine the correct focal point causing unintended interactions. When participants read superimposed text, they often interacted with the same objects after a slight delay due to the nature of the gaze-dwell interactions. Interactions followed the gaze-dwell principle, where the user looks at a specific object for a few seconds to interact with it. The gaze dwells delay time depended purely on personal preference with no globally accepted value that minimizes errors [27]. In our work, we optimize Glanceable interfaces so that interactable objects are not close to each other. We combine gaze-dwell interaction with blink-based interaction in order to differentiate between viewing and interacting with virtual objects.

Participants could collaborate efficiently in a collaborative scenario using Magic Leap's integrated eye tracking even when the users perceived that their accuracy had significant tracking errors [13]. Overcoming accuracy errors was achieved by placing physical markers at a similar distance to the eye tracker's offset at 5cm and viewing each other's visual cursor. The effectiveness of different interaction types in Glanceable AR has been explored [22]. Five interaction techniques were evaluated on two tasks: one walking and one sitting. Interaction techniques included (1) Fixation glance (FG), (2) Head-depth (HD), (3) Hand-Overlay (HO), (4) Blink (BL), and (5) Gaze-Dwell (DW). DW and BL were among the favored interactions on both tasks. DW had the highest number of false interactions with a rate of over 70% in both tasks. BL was the preferred interaction followed by DW over all other interactions in terms of social acceptance. In our work, we use a combination of gaze-dwell and blink-based interactions, dubbed blink-delay. While dwell-based interactions were the most preferred, the act of dwelling one's gaze when reading is passive and provoked a large amount of mis-interactions [22]. To combat this, we combined it with the active action of blinking to show direct interaction intent. Another perk of this combination in our work is that dwell-based interactions can also be combined with hand tracking, dubbed hand-delay, resulting in a more streamlined combination of two different interactions using the same logic.

2.3 The Industry 4.0 model

Industry 4.0 extends the computer-oriented approach of current manufacturing machine shops and introduces new technologies for further automation of the manufacturing process [14]. Newer approaches in manufacturing education streamline manufacturing tasks as gamified missions [2]. Experts propose advanced guidance in AR by designing a gamified workspace inside the actual manufacturing machine shop [8]. AR for guidance is more direct than paper-based manuals, as the presented information can be tailored to the active task. Gamification helps keep employees focused, even on tedious, repetitive tasks, for longer [8].

Turning and Milling machines will form the context scenario of our proposed Glance-box interface. Turning machines consist of a cylindrical stock material spun rapidly while using a cutting tool that cuts the material evenly around its periphery. Milling machines consist of a rectangular stock material attached to the machine's base with a rotating cutting tool that cuts the material from above. In our work, we introduce Glanceable AR to the manufacturing machine shop for rapid information retrieval. Inexperienced users utilize the more compact interfaces of our three-tier (3-LOD) guidance system as they gain more experience.

3 Implementation

The implementation of Glance-Box, based on Unity 2020 for Magic Leap One AR HMD, comprised three key designs: (1) The interface design, (2)The design of user interactions and finally, (3) Adapting the content of the interfaces to take full advantage of the designed interface system. Our proposed design aims to achieve three primary design goals (DG):

(DG1) The system must be usable when working on a real world task while minimizing potential safety risks, improving upon information retrieval and safety practices of previous work in AR guidance in the manufacturing workspace [34]. *(DG2) The interface must show varying amounts of information at different LODs* based on whether the user's focus is on the real world or the virtual content. The user's focus is governed by direct interaction with virtual objects instead of automatically, reducing potential distractions from shifting interfaces while the user's focus is on a real-world task [20]. *(DG3) The interfaces should achieve accurate interactions and the Midas Touch problem (MTP) should be minimized when using gaze interaction*. To mitigate the impact of the MTP, a direct form of interaction, such as blinking, is preferred [22]. In dense interfaces where interactable objects are prone to overlap [27], a different interaction paradigm should be used such as hand-tracking. Speech control is not appropriate due to the noisy manufacturing workspace.

3.1 Interface Design

Glance-box employs 3-LODs (Figure 1), e.g., Simple Glanceable (SG), Detailed Glanceable (DG), and Holo-Box (HB). SG and DG are 2D display fixed Glanceable interfaces that follow the user's head orientation and are always visible. HB is a 3D interactive interface that remains fixed at a predefined anchor location in the real world, selected by the user.

Simple Glanceable (SG): (Figure 1 left) The SG interface is a compact 2D interface that shows minimal text information (a few words, for instance). SG minimizes real-world occlusion. It is the initial interface encountered notifying the user that new data is present in the virtual interface. The SG interfaces include an interactable button that transitions to the DG interface.

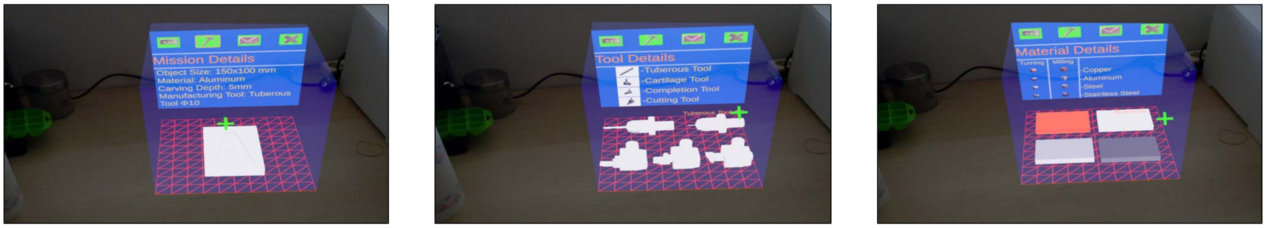
Detailed Glanceable (DG): (Figure 1 middle) The DG interface provides more detailed information to the user, including a few lines of text as well as 2D media such as images. DG is larger than the SG interface, but it is still relatively small compared to the Magic Leap's FoV aiming to not impede a large area of the available FOV. DG includes two interactable buttons, one to return to the SG and another that transitions to the HB. Placing the two buttons far from each other minimizes potential unintentional interactions.

Holo-Box (HB): (Figure 1 right) The HB interface is a 3D interface. During the application setup process, the user manually places the HB on the real world, designated as a semi-transparent bounding box. Its position persists during the execution of the application or until the user manually re-positions it. Upon HB activation through the DG interface, the content appears inside the predetermined position regardless of the user's position during the activation. The content of the HB includes complex 2D interfaces, 3D objects and any number of interactable objects. HB also allows for both blink-delay and hand-delay interactions; these can be activated interchangeably. The two Glanceable interfaces allow for rapid information retrieval "at a glance" with two levels of information. SG is compact, allowing for easier focus on the real world providing guidance, while DG triggers after the user interacts with the SG interface. We can assume the user wants to focus on the virtual content when the DG interface appears. SG and DG interfaces follow the user at a set distance and viewing angle as they move around the real world. HB is anchored on a specific real-world position and does not move. HB content is presented only on the inside of its bounding box. We assume that the user has placed it at a position where they can use hand tracking without risk of injury.

3.2 User Interaction

Glance-Box allows for two types of interaction, one that uses eye-tracking and one that uses hand tracking. Both interaction types are designed around using delay-based inputs, where the user targets an object based on directed gaze using a cursor and remains on the target for a set amount of time before the interaction happens. To reduce gaze mis-interactions, the user must place the eye-tracking cursor and then blink with one eye to interact, thus, using "Blink-Delay" interactions. Hand tracking interactions, on the other hand, use different cursors which follow the users' index fingers on both hands. The user interacts with objects by positioning their fingers in a trigger zone that extends a few centimeters around the interactable object and remains inside the trigger for the same delay period for a "Hand-Delay" input. To avoid conflicting simultaneous interactions, hand tracking is prioritized over eye-tracking interaction. If both hands engage simultaneously, then the left-hand cursor is prioritized.

Interactable objects in Glance-box operate using a universal state machine. An interface can be in one of three states: Idle, High-



(a) (b) (c)

Figure 2: Holo-Box interfaces. (a) Mission Description (b) Tools Database (c) Materials Database

Figure 3: Button States

lighted or Clicked (Figure 3). Initially, all objects are in an idle state, showing their basic visuals. When a cursor enters the object's trigger zone, that object becomes highlighted enabling an additional visual effect showing the transition. After a click is initiated by blinking on a Blink-Delayed input or immediately when highlighted on Hand-Delay inputs, a progressively filling visual effect is presented showing the progress towards the interaction completion. After the delay period has passed, the click effect is triggered and the object returns to an idle state. In Glance-box, to fulfill design goal DG3, only one object can be at the highlighted and clicked states shared across all interaction types. The highlighted object is deselected only after a click, or when another object becomes highlighted and not when the user exits the trigger zone. When using Hand-Delay, the object changes from clicked to highlighted on exit. This way, the user can select an object. If their cursor moves outside the trigger zone due to tracking errors, the delayed click can still be performed by blinking if the cursor does not select a different object. To fulfill design goal DG1, Hand-Delay interactions are disabled on the SG and DG interfaces, as these interfaces move and may be placed in a dangerous location for hand motions. Based on the design, the SG and DG interfaces are designed to minimize mis-interactions. SG only comprises a single interactable button. DG includes two buttons placed on the top and bottom edges of the interface to reduce mis-interactions when using Blink-Delay inputs. HB allows for both Hand-Delay and Blink-Delay inputs. As such, there are no restrictions on the amount of interactable objects as long as they are placed so that their trigger zones are not overlapping.

3.3 Machine shop Content Adaptation

Participants were given a set of instructions, tasked with selecting the correct object corresponding to a manufacturing task, including a **cutting tool**, a **cutting material** and a **manufactured product**. Each manufacturing task is presented in the form of a gamified "mission" [2]. A mission consists of instruction required to perform a manufacturing task, e.g., tools and materials to use, a 2D blueprint of the finished product and numerical cutting parameters such as the cutting speed. The presentation of the missions is adapted to allow efficient guidance with any of the three interfaces according to the user's expertise. The HB interface contains the information required to complete any mission including mission information (Figure 2a) as well as tool (Figure 2b) and material (Figure 2c) look-up tables with added 3D objects for better visualisation. The 3D models of the tools and materials use a modified version of the

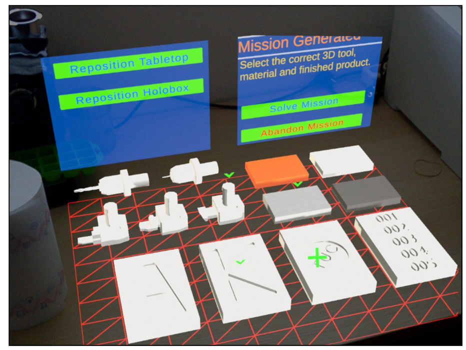


Figure 4: Evaluation Interface

interaction logic that shows the object's name above the model when highlighted. The DG interface includes mission-specific information in a 2D interface (Figure 1 middle). The SG interface only contains the name of the mission as well as the correct material (Figure 1 left). The SG interface is adequate for experienced users who have already completed a given mission and memorized it. To make the SG interface more efficient, each mission was assigned a different name, e.g., the "A plaque" corresponds to cutting the letter "A" on a square block.

4 User Study

A study evaluated our Glance-box interface system in a real-world machine shop also guiding experienced users without using the most complex HB interface. The Evaluation Interface (EI) is a modified version of the HB with the same functionality including 2D interfaces with buttons used to operate the evaluation procedure, e.g., starting the evaluation or moving to the next mission and providing feedback to the user (Figure 4). The EI also includes 3D models of available machining tools, materials and finished products which the user can select in order to complete a given training mission.

4.1 Apparatus

The study employed the Magic Leap One AR HMD [24] and Unity 2020.1.17f1 [31]. Built-in hand, eye and blinking detection were provided by the HMD without any perceived latency. The Magic Leap Control allowed for the initial setup and the placement of the interfaces, including the rotation of the Glanceable interfaces and the position of the Holo-Box and EI. Participants sat on a rotating chair in front of an empty desk for safety. Sprites used inside the interfaces were part of Magic Leap's SDK or purchased sprite packs. The M3 CNC research lab of the Technical University of Crete provided materials for manufacturing 3D objects, tools, products, cutting blueprints and mission descriptions.

4.2 Participants

Thirteen participants (two female) between the ages of 22 and 30 years old participated in the study. Six participants had nearsightedness, one astigmatism and one hyperopia. Participants were asked not to wear any contacts or glasses as they hampered the eye tracker's accuracy. The Glanceable interfaces were positioned closer to participants with nearsightedness until they stated they were clearly visible. Six participants had no previous experience in using AR. Two participants had bright eye colors, including cyan and green.

4.3 Procedure

Participants performed visual calibration of the Magic Leap's eye tracker through the built-in calibration system. Next, participants started the demo application following a tutorial which consisted of four steps: (1) Interacting with a Glanceable button via blinking, (2) Adjusting the angle where Glanceable interfaces were visible at, (3) Placing the EI and (4) Placing the Holo-box interface. Positioning was achieved by placing the controller at the desired spot and rotation and pressing a button. Participants let go of the controller unless they wanted to re-position the interfaces at a later time.

4.4 Demos and Missions

After the initialization process, participants were asked to complete at least one demo machining training mission. The name of the mission as well as the correct tool, material and finished product were found through the 3-LOD interface. The participant selected the correct manufacturing tool, material and finished product from the EI. During demo missions, participants took as much time and missions as necessary. The main procedure consisted of an object selection task in the form of manufacturing missions. Participants were given guidance using Holo-Box's 3-LOD system while they selected the requested objects from the EI. Each mission required selecting the appropriate tool, material and finished product from the EI and pressing a button to check for correctness. The main experiment consisted of four distinct missions repeated in three epochs in randomised order, with the first and second epoch consisting of all four missions and the third epoch of two of the four missions, for a total of ten. Participants memorised each mission's content and completed each mission by using only Glanceable AR when possible.

4.5 Metrics

Quantitative data: For each participant, Magic Leap logged the following: (1) the number of tutorial missions completed (2) blink-delay interaction events (3) hand interactions events (4) time to complete each mission (5) the number of errors made for each mission and (6) the number of Glance-Box LODs used for each mission (3= all LODs, 2= only Glanceable, 1= only SG).

System Usability Scale (SUS): The SUS consists of a set of questions related to the usability of a proposed system [19]. Answers are reported on a 7-point Likert scale (1= very low, 7= very high). The questions include grading (1) Learnability, (2) Efficiency, (3) Memorability, (4) Accuracy, (5) Satisfaction, (6) Intuitiveness, (7) Naturalness and (8) Fun.

Nasa Task Load Index (Nasa TLX): The Nasa TLX covers topics regarding user experience [10]. Questions are answered on a 7-point Likert scale. The questions include (1) Mental Demand, (2) Physical Demand (3) Temporal Demand (4) Performance (5) Effort and (6) Frustration.

User Preferences: Participants selected their preferred interaction method between blink-delay and hand-delay and then graded on a 7-point Likert scale (1) the accuracy of blink-delay interactions specifically on the two Glanceable interfaces SG and DG (2) the precision of the eye-tracking cursor in regards to where they were looking and (3) the accuracy of the hand cursor in relation to their hand movements.

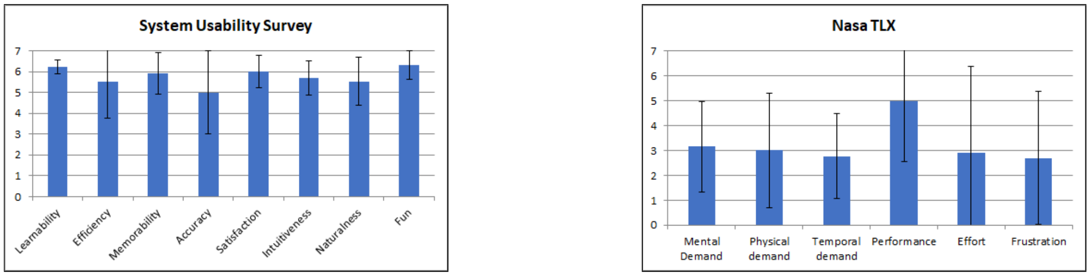
4.6 Hypotheses

(H1) Participants will complete missions faster over time, with fewer errors. (H2) Participants will be able to complete subsequent missions while using only Glanceable interfaces. (H3) Our proposed blink-delay interactions can somehow 'hide' Magic Leap's eye trackers' inaccuracies and users will not detect eye tracking failure, e.g., when the cursor freezes and does not follow the eye. (H4) Participants will prefer hand-delay interactions over blink-delay due to their higher accuracy.

4.7 Results and Discussion

Mission completion metrics: When comparing participants' completion time for one mission, this varies greatly based on individual skill. Instead, we calculated the average time improvement between epochs for each participant separately, e.g., we compared the average time on the first epoch with those of the following two. This way, we figured out whether participants improved over time regardless of their skills. Error rates and interface switches were similarly averaged. per participant. per epoch. In the second epoch, participants completed missions with an average 24,16% (var: 9,6%) time improvement, while in the third epoch, participants showed a 39,04% (var: 13,38%) improvement over the first. Two participants showed no significant improvement (negative or below 10%) on all three epochs. In contrast, three more showed no significant improvement during the second epoch but showed significant improvement (35% and above) in the third. When it comes to error rates, no significant improvement over time was shown with an average 0,46 (var: 0,45) errors per mission for the first epoch, 0,5 (var: 0,46) errors for the second and 0,46 (var: 0,82) errors for the third. Regarding interface switches, participants used an average of 2,54 (var: 0,3) interfaces during the first epoch, 1,94 (var: 0,46) during the second and 1,77 (var: 0,53) during the third. The number of participants that used only Glanceable interfaces was one during the first epoch, four during the second and nine during the third. Six participants used all three interfaces on more than two out of four missions during the first epoch, two during the second epoch, while no participant used all interfaces on both missions of the third epoch. In summary, H1 was partially supported as participants could complete missions faster over time, but the error rate showed no significant changes. H2 was supported as participants used fewer interfaces over subsequent missions and most missions during the second and third epochs were completed using only Glanceable interfaces.

Interaction preferences: Six participants preferred blink-delay interactions (over 60%), four preferred hand-delay and three used both interactions equally (both below 60%). Of the six participants who preferred blink-delay, five used blink-delay interactions (fewer than two hand-delay interactions) exclusively after the demo missions. Eight participants preferred blink-delay while five participants hand-delay. Only six participants stated they preferred the same interaction type they mostly used during the experiment. Two out of five participants that used exclusively blink-delay stated they preferred hand-delay. Concerning the question "How accurate were the blink-delay interactions in relation to the Glanceable interfaces" the answers were positive, with an average of 5.46 (var: 3,17) on the 7-point Likert scale. Only two participants had significant difficulties performing blink-delay interactions due to a combination of eye tracking errors of the built-in tracker of the device as well as user-specific eye attributes. The participants used hand-delay instead. Certain participants stated they preferred hand-delay input while they used blink-delay input during the experiment. They found eye-based input interesting for a short experiment but this did not reflect a long-term preference. Two participants could not keep only one eye closed while looking directly forward with the other to aim correctly, forcing them to keep their eyelids closed with their fingers to interact, therefore, blink-delay interactions were not possible. When rating the eye-tracking cursor's accuracy, the score was posi-



(a)

(b)

Figure 5: (a) System Usability Scale (SUS) results (b) Nasa Task Load Index (TLX) results

tive, with an average score of 4,54 (var: 1,17) and no scores below 3. Most participants stated that the blink-delay cursor was accurate as they did not notice inaccuracies in the cursor's position. Based on the supervisor's observations of the video feed and the participants' head movements, we noticed that in certain instances when the eye tracker could not detect the user's eyes and the cursor remained stationary, they instinctively rotated their neck to aim and interact correctly, resulting in higher perceived accuracy by the users. When grading the hand tracking cursor, the average score was similar at 4.39 but with a significantly higher variance of 5.09. Most scores were above six or below 2, as the hand cursor's flickering was more apparent when the users' hands were stationary. Participants noticed the flickering of the hand cursor. Four participants were not able to use hand-delay inputs. Even, though, the cursor in certain instances froze and moved away from interactable objects due to device detection failure, the perceived accuracy of blink-delay interactions was high. Participants noticed certain inaccuracies in detecting blinking. Thus, H3 was supported. We can neither confirm nor reject H4 because of varied preferences in interaction, interaction issues in relation to both interaction types for certain participants and two participants not managing to use our blink-delay input. Improvement of gaze detection offered as part of current AR devices is required.

SUS: SUS questionnaire (see Figure 5a): (1) Learnability avg = 6.23, var = 0.33 (2) Efficiency avg = 5.53, var = 1.78 (3) Memorability avg = 5.92, var = 0.99 (4) Accuracy avg = 5, var = 2 (5) Satisfaction avg = 6, var = 0.76 (6) Intuitiveness avg = 5.69, var = 0.82 (7) Naturalness avg = 5.53, var = 1.17 and (8) Fun avg = 6.3, var = 0.7. Glance-box has positive scores in all fields, with (4) Accuracy having a lower score and higher variance than other values, further retaining the uncertainty in relation to H4 being valid due to inaccuracies in both interaction systems.

Nasa TLX: Nasa TLX questionnaire's results (see Figure 5b): (1) Mental Demand avg = 3.15, var = 1.82 (2) Physical Demand avg = 3, var = 2.30 (3) Temporal demand avg = 2.76, var = 1.71 (4) Performance avg = 5, var 2.46 (5) Effort avg = 2.92, var = 3.45 and (6) Frustration avg = 2.69, var 2.67. Results are mostly mid-range on the Likert scale, with performance being slightly higher than mid-range. Demand, effort and frustration scores signified that participants performed well with a low task load. The high variance in effort shows that certain users struggled while using Glance-box due to the occasional eye and hand tracking device inaccuracies.

5 Conclusion

We introduced Glance-Box, a novel interface system for information retrieval through Glanceable and 3D interfaces. Glance-Box uses a 3-tier (3-LOD) interface design that reveals guidance information in progressing levels of detail as necessary. Interaction with virtual objects use blink-delay input for all interfaces, while 3D interfaces also use hand-delay input. These interaction types allow for precise interaction with virtual objects while limiting mid-air hand movements inside the bounding boxes of 3D interfaces for safety constraints. Interface content is adapted so that experienced users can complete assigned tasks using more compact interfaces, while inexperienced users can activate a denser LOD for additional guidance. Evaluation of Glance-box occurred in a safe environment; users completed a selection task of required items in a manufacturing machine shop. Results showed that our system is easy to learn and users could complete tasks faster by using more compact interfaces over time. Both hand-delay and blink-delay interactions performed comparably well, with tracking issues of the device present in both. The blink-delay cursor was perceived to be accurate. Users could adjust their position using eye and head movements and they instinctively combined both activities when selecting objects. Our results show that our Glance-box AR interface is promising. The major drawback of our interface system was eye and hand tracking errors of the device. Our choice of blink-delay interaction compensated for these errors to a degree. Future work should compare our proposed interactions with other variations, especially hand-based input, in safety critical environments. The application of our 3-LOD system under scenarios where more than one set of interfaces is needed, will provide fruitful avenues for future research.

Acknowledgments

This research is co-financed by the EU and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code: T2EDK-03649).

References

[1] W. Buxton. Integrating the periphery and context: A new taxonomy of telematics. In *Proc. of graphics interface*, vol. 95, pp. 239–246, 1995.

[2] G. Daskalogrigorakis, S. Kirakosian, A. Marinakis, V. Nikolidakis, I. Pateraki, A. Antoniadis, and K. Mania. G-code machina: A serious game for g-code and cnc machine operation training. *IEEE EDUCON 2021, Games in Engineering Education*, 2021.

[3] G. Daskalogrigorakis, A. McNamara, and K. Mania. Holo-box: Level-of-detail glanceable interfaces for augmented reality. In *ACM SIGGRAPH 2021 Posters*, SIGGRAPH '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3450618.3469175

[4] S. Davari, F. Lu, and D. A. Bowman. Occlusion management techniques for everyday glanceable ar interfaces. In *2020 IEEE Conf. on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 324–330. IEEE, 2020.

[5] S. Davari, F. Lu, and D. A. Bowman. Validating the benefits of glanceable and context-aware augmented reality for everyday information access tasks. In *2022 IEEE Conf. on Virtual Reality and 3D User Interfaces (VR)*, pp. 436–444. IEEE, 2022.

[6] C. Elmadjian and C. H. Morimoto. Gazebar: Exploiting the midas touch in gaze interaction. In *Extended Abstracts of 2021 CHI Conf. on Human Factors in Computing Systems*, pp. 1–7, 2021.

[7] J. Flynn, S. Dance, and D. Schaefer. Industry 4.0 and its potential impact on employment demographics in the uk. *Advances in Transdisciplinary Engineering*, 6:239–244, 2017.

[8] M. Funk, T. Dingler, J. Cooper, and A. Schmidt. Stop helping me-i'm bored!: why assembly assistance needs to be adaptive. In *Adjunct Proc. of 2015 ACM Intl. Joint Conf. on Pervasive and Ubiquitous Computing and Proc. of 2015 ACM Intl. Symp. on Wearable Computers*, pp. 1269–1273. ACM, 2015.

[9] J. Grubert, T. Langlotz, S. Zollmann, and H. Regenbrecht. Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE transactions on visualization and computer graphics*, 23(6):1706–1724, 2016.

[10] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, pp. 904–908. Sage publications, 2006.

[11] S. Imamov, D. Monzel, and W. S. Lages. Where to display? how interface position affects comfort and task switching time on glanceable interfaces. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 851–858, 2020. doi: 10.1109/VR46266.2020.00110

[12] R. J. Jacob. What you look at is what you get: eye movement-based interaction techniques. In *Proc. of SIGCHI Conf. on Human factors in Computing Systems*, pp. 11–18, 1990.

[13] A. Jing, K. May, G. Lee, and M. Billinghurst. Eye see what you see: Exploring how bi-directional augmented reality gaze visualisation influences co-located symmetric collaboration. *Frontiers in Virtual Reality*, 2:79, 2021.

[14] S. Karadayi-Usta. An interpretive structural analysis for industry 4.0 adoption challenges. *IEEE Transactions on Engineering Management*, 67(3):973–978, 2019.

[15] M. Katsiokalis, L. Ragia, and K. Mania. Outdoors mobile augmented reality for coastal erosion visualization based on geographical data. In *XR@ ISS*, 2020.

[16] S. Kim, M. A. Nussbaum, and J. L. Gabbard. Augmented reality "smart glasses" in the workplace: industry perspectives and challenges for worker safety and health. *IIE transactions on occupational ergonomics and human factors*, 4(4):253–258, 2016.

[17] S. Kirakosian, G. Daskalogrigorakis, E. Maravelakis, and K. Mania. Near-contact person-to-3d character dance training: Comparing ar and vr for interactive entertainment. In *2021 IEEE Conference on Games (CoG)*, pp. 1–5. IEEE, 2021.

[18] G. A. Koulieris, K. Akşit, M. Stengel, R. K. Mantiuk, K. Mania, and C. Richardt. Near-eye display and tracking technologies for virtual and augmented reality. In *Computer Graphics Forum*, vol. 38, pp. 493–519. Wiley Online Library, 2019.

[19] J. R. Lewis. The system usability scale: past, present, and future. *International Journal of Human–Computer Interaction*, 34(7):577–590, 2018.

[20] D. Lindlbauer, A. M. Feit, and O. Hilliges. Context-aware online adaptation of mixed reality interfaces. In *Proc. of 32nd annual ACM Symp. on user interface software and technology*, pp. 147–160, 2019.

[21] F. Lu and D. A. Bowman. Evaluating the potential of glanceable ar interfaces for authentic everyday uses. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 768–777. IEEE, 2021.

[22] F. Lu, S. Davari, and D. Bowman. Exploration of techniques for rapid activation of glanceable information in head-worn augmented reality. In *Symp. on Spatial User Interaction*, pp. 1–11, 2021.

[23] F. Lu, S. Davari, L. Lisle, Y. Li, and D. A. Bowman. Glanceable ar: Evaluating information access methods for head-worn augmented reality. In *2020 IEEE Conf. on Virtual Reality and 3D user interfaces (VR)*, pp. 930–939. IEEE, 2020.

[24] magicleap. Magic leap: Enterprise augmenter reality (ar) platform. https://www.magicleap.com/en-us/. Accessed: 2022-06-03.

[25] A. McNamara, K. Boyd, J. George, W. Jones, S. Oh, and A. Suther. Information placement in virtual reality. In *2019 IEEE Conf. on Virtual Reality and 3D User Interfaces (VR)*, pp. 1765–1769. IEEE, 2019.

[26] A. McNamara, C. Kabeerdoss, and C. Egan. Mobile user interfaces based on user attention. In *Proceedings of the 2015 Workshop on Future Mobile User Interfaces*, pp. 1–3, 2015.

[27] K. Pfeuffer, Y. Abdrabou, A. Esteves, R. Rivu, Y. Abdelrahman, S. Meitner, A. Saadi, and F. Alt. Artention: A design space for gaze-adaptive user interfaces in augmented reality. *Computers & Graphics*, 95:1–12, 2021.

[28] R. Piening, K. Pfeuffer, A. Esteves, T. Mittermeier, S. Prange, P. Schröder, and F. Alt. Looking for info: Evaluation of gaze based information retrieval in augmented reality. In *IFIP Conf. on Human-Computer Interaction*, pp. 544–565. Springer, 2021.

[29] R. Rivu, Y. Abdrabou, K. Pfeuffer, A. Esteves, S. Meitner, and F. Alt. Stare: gaze-assisted face-to-face communication in augmented reality. In *ACM Symp. on Eye Tracking Research & Applications*, pp. 1–5, 2020.

[30] Q. Roy, F. Zhang, and D. Vogel. Automation accuracy is good, but high controllability may be better. In *Proc. of 2019 CHI Conf. on Human Factors in Computing Systems*, pp. 1–8, 2019.

[31] Unity. Unity game engine. https://unity.com/. Accessed: 2022-06-03.

[32] R. Vertegaal et al. Attentive user interfaces. *Communications of the ACM*, 46(3):30–33, 2003.

[33] J. Wolf, Q. Lohmeyer, C. Holz, and M. Meboldt. Gaze comes in handy: Predicting and preventing erroneous hand actions in ar-supported manual tasks. In *2021 IEEE Intl. Symp. on Mixed and Augmented Reality (ISMAR)*, pp. 166–175. IEEE, 2021.

[34] J. Zhu, S.-K. Ong, and A. Y. Nee. A context-aware augmented reality assisted maintenance system. *Intl. Journal of Computer Integrated Manufacturing*, 28(2):213–225, 2015.