Looking for Info: Evaluation of Gaze Based Information Retrieval in Augmented Reality

Robin Piening[1,2], Ken Pfeuffer[1], Augusto Esteves[3], Tim Mittermeier[1], Sarah Prange[1,2], Philippe Schröder[2], and Florian Alt[1]

[1] Universität der Bundeswehr München, Germany, {firstname.lastname}@unibw.de
[2] LMU Munich, Munich, Germany, {firstname.lastname}@campus.lmu.de
[3] ITI / LARSyS, IST, ULisbon, Portugal, augusto.esteves@tecnico.ulisboa.pt

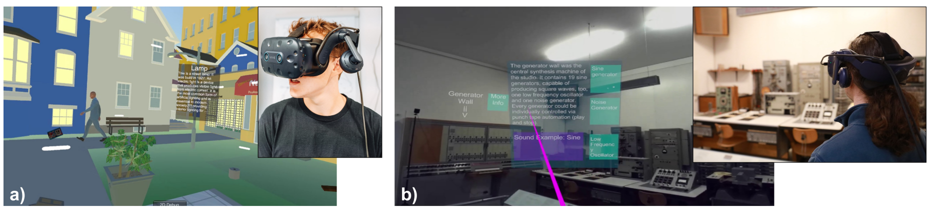

Fig. 1: Gaze-adaptive UIs in AR enable one of the most effortless and implicit access methods to context information inherent in world objects. For example, interactive objects have subtle visual indicators that, when fixated with the eyes, open information panels (a). When opened, further UI elements utilise the eyes' dwell-time as a technique to browse more pages or activate sound files on-the-fly (b).

Abstract. This paper presents the results of an empirical study and a real-world deployment of a gaze-adaptive UI for Augmented Reality (AR). AR introduces an attention dilemma between focusing on the reality vs. on AR content. Past work suggested eye gaze as a technique to open information interfaces, however there is only little empirical work. We present an empirical study comparing gaze-adaptive to an always-on interface in tasks that vary focus between reality and virtual content. Across tasks, we find most participants prefer the gaze-adaptive UI and find it less distracting. When focusing on reality, the gaze UI is faster, perceived as easier and more intuitive. When focusing on virtual content, always-on is faster but user preferences are split. We conclude with the design and deployment of an interactive application in a public museum, demonstrating the promising potential in the real world.

Keywords: Gaze Interaction · AR · Eye-tracking · Adaptive UI.

1 Introduction

Head-worn AR displays, that superimpose graphics and text over the appropriate objects in the user's view, have the potential to enable a new way of seeing the world [10]. Users can be assisted in a range of activities from everyday scenarios to obtain contextual information of the world (e.g., restaurants, streets, sights) to interactive architecture and museum tours to gain guidance and knowledge about the artifacts. This introduces a key user interface (UI) challenge of user access to contextual information through interfaces in form of panels or windows anchored to real world environments. Virtual content can occlude the visibility and distract users when attending to the real world, and environments cluttered with too many panels can feel overwhelming.

As an alternative to always-on interfaces, adaptive context-aware AR pursues the idea to present appropriate information at the right time through analysis of context information [12]. This is particularly useful in scenarios where the hands may not be available or desired to employ for the selection and deselection of casual AR information. A promising method of adaptation is to exploit the user's visual attention as context information, being highly indicative of what we are interested in [33, 45]. Eye-tracking modules are increasingly integrated with AR devices such as the Microsoft HoloLens 2 or Magic Leap One. Gaze, as the point where our eyes converge, can be harnessed to adapt information availability and density. Several prototypical gaze-adaptive UIs were proposed in the literature [1, 23, 33]. However, there is a lack of empirical evidence with regards to user experience and performance, making it difficult to assess the feasibility of the concept in the maturing AR landscape.

Our work aims to fill this gap, by presenting an empirical experiment in the lab, and a field study of the gaze-adaptive UI in a real world scenario. Figure 1 provides an overview of the concept, technical prototypes and the two studies. Users can access additional information of an object by looking at an indicator located nearby. After a time threshold, the UI element is 'activated' and opens a window of appropriate information. For example in a virtual city environment, the user can look around the scene, and fixate objects of interests such has the lamp to quickly open a definition of it (a).

Our first contributions is an empirical comparison of the gaze-adaptive to an always-on UI. the latter is representative of the default way to visualise AR contextual information in the world. This allows us to establish a ground truth for future developments and studies with gaze in AR. The comparison is two-fold, to understand both perspectives of the attention dilemma in AR. First, users focus on the reality, tasked to find "real-world" objects in this environment while AR panels can be in the way. Second, users focus on virtual content, by searching particular information in the AR panels anchored to objects in the vicinity. The results of the study provide insights on the performance, user experience, and occlusion trade-off:

- For the reality focused task, users performed faster with gaze-adaptation, found it easier, more intuitive, and preferred it.
- For the virtual UI task, as expected users were faster when seeing all information, but user preferences were split between the two techniques (11 gaze-adaptive vs. 10 always-on UI).
- Across both tasks, most users prefer using the gaze-adaptive UI, and find it less distracting.

Furthermore, to understand gaze-adaptation in practice, we present a novel gaze-adaptive AR application specifically tailored to a museum scenario. Figure 1b shows the application in action in a real museum where we deployed it. Wearing the AR headset, users view the museum artifacts and can interact with superimposed information. Complementary to physical information on a desk, users can directly read and hear about the artifacts by looking at visual indicators close to each artifact. The opened information interface has further buttons, to browse more information by dwell-time. We conducted a field study where participants report on their experience, overall showing a positive reception of this method and further ideas for improvements.

Together, these findings allow to gain a better understanding of the various trade-offs and the overall viability of gaze-activated AR information, of value for designers and researchers working to bring AR applications beyond the lab toward an efficient and natural way in everyday environments.

2 Background and Related Work

2.1 Gaze Interaction

Eye-tracking has the potential to facilitate our interactions with computers, making them more intuitive and responsive [6]. A major usability challenge with gaze input is the involuntary selection of elements when these are being gazed at by the user (”Midas Touch” problem [17]). Several selection confirmation mechanisms were investigated for objects including mouse press [17], gestures [35], touch input [34, 41], controllers [21, 42] and menus [36].

A popular approach is dwell time. It gives users a time window to look at an object before the user's gaze is interpreted as a command, e.g., to trigger a selection [4, 17, 30, 44, 40]. The limitations of dwell are response time, as the user needs to wait until the timer finishes, and false triggers, if accidentally dwelling on objects for too long. We focus on revealing information based on user interest. This is a more forgiving task, as users can simply avert their gaze if not interested, but requires a different design of timing parameters we introduce later.

Gaze has been explored for head mounted AR devices. Kyto et al. studied different selection techniques in head mounted AR displays, comparing their use of gaze or head ray as primary input with secondary refinement methods like hand gestures or external devices [21]. Ishiguro et al. used gaze as a filtering mechanism to avoid occlusion when visual information is overlaid [16]. They explored annotations at the edges of the display, with symbols that could be recognized peripherally. Upon being gazed at, these annotations would expand with detailed information. Tonnis et al. explored a mechanism, presenting contextual information at an offset from the user's gaze point. This allows the user to freely interact with an interface or the world, and only when attending to the information, it will stay in place [43]. Rivu et al. developed a prototype for a face-to-face AR interface for people, where information is indicated next to the user and gazing at them allows to reveal them [38].

Our work is complementary, focusing on the more general approach of providing information in the vicinity of objects through an implicit gaze activation, and then allow users to explicitly trigger more content. Such an implicit approach is closely related to Vertegaal's Attentive User Interfaces (AUIs) [45], a general concept where displays adapt to the user's attention capacities. We investigate this approach for AR, where it is a more significant problem of attention management between real world and digital augmentation.

2.2 Adaptive AR

AR display technology affords many use cases where virtual content overlays real content in everyday environments [10], and creating such mixed reality content is becoming increasingly possible for a range of users [32]. A major challenge is that the virtual content can be informative, but also obstruct the user's view and important real-world information. This is recognized in AR research as the "label placement problem" [2, 39].

Several researchers have thus investigated particular placement parameters for information close to objects [13, 24]. Bell et al. introduced view management techniques that used scaling, transparency, and repositioning to avoid occluding places of interest or other annotations, while enabling functional grouping [3]. Julier et al. used different rule-based filtering methods to avoid information overload, employing distance, visibility, object importance, and task relevance [18]. Keil et al. used camera and motion based techniques to minimize information overload [19], while Fender et al. explored automated placement optimization of visualizations based on user behavior and information quality [11].

AR systems can be improved by making them adaptive to context information on user and environment. Ajanki et al. [1] used gaze patterns, implicit speech recognition, the user's task, and physical location, to deliver highly relevant augmented information retrieved from the web. Feiner et al. described a system that took into account user's environment to determine an optimal data density [9]. Lages and Bowman investigate adaptive interfaces during walking, where information is offered through 2D windows that follow the user and adapt to the environment [22]. Lindlbauer at al. developed a context aware system that responded to cognitive load, reducing the augmentation density, positioning, and time accordingly [7]. Further potential of mobile AR has been explored by Höllerer et al. via eye-tracking and location awareness [14]. White et al. explored situated pattern visualization for urban planning and site visits, eliminating the memetic accuracy loss experienced with attention switches between, e.g., a map and the real environment [46].

Little work empirically studied gaze for context-aware AR. Kim et al.'s Shop-i concept demonstrated that users recognized the value of gaze-aware digital labels laid over the real-world in a shopping environment [20]. McNamara et al. demonstrated that having gaze dictate 'where' and 'when' to display such labels improved users' performance when searching for contextual information in VR [26, 27]. Their studies compare early prototypes of various methods such as always-on labels, labels using a clamped squared relationship between their anchor and the user's gaze point, labels remaining visible for a short amount of time after being gazed at, and labels with a trigger and exit radius [28, 29]. They found a combination of the latter two to be ideal, which we build upon.

Closely related is the research on Glanceable AR [8, 15, 23]. We share focus on temporary accessible information, and the potential integration of gaze for interaction. Lu et al. focus on a personal UI that allows notifications to be triggered and typical apps as found in a smartphone to be accessed quickly [23]. Inamov et al. studied positioning information relative to the object [15]. We extend their work by focus on 1) world-fixed interfaces which is different to a personal UI [23] (e.g., having a large number of information UIs), 2) eye based input which has been shown to be different to head direction [5] ([8, 15] do not use eye gaze), and 3) contributing insights from a controlled lab and a field study.

Lastly, our work is inspired by recent work on the ARtention design space of gaze for AR scenarios proposed in [33]. They formulate three roles for gaze: reveal information, select information, and browse information, demonstrated in three applications. We extend their work by providing 1) an empirical evaluation of using gaze to transition from reality to virtual UI focus, and 2) design a new application that integrates all three roles of the design space.

3 Looking for Info: Gaze Adaptive AR

Building on the prior work, the main idea is to utilise eye gaze input of the user to access information bits in AR. Meaning, users can look at meaningful points of the real environment, and when indicating interest by gazing at a particular point, the system provides the user with additional information. In order to better ground our design and study efforts, we formalize the properties of a gaze adaptive AR system across three dimensions:

Temporal Activation. This is defined as how long a user needs to look at a real-world object to trigger any embedded AR panel. In some tasks (e.g., information retrieval) the user might prefer immediate AR panel activations, while in others a delay can make for a more pleasant and seamless experience.

Spatial Activation. This is defined as the area a user needs to gaze at to trigger an AR panel, and its size can vary in different ways. For simple, self-contained application where the user is not required to move, each activation area can have the same size for UI consistency. For an in-the-wild application these activation areas can match the scale of the objects as they are perceived by the user, facilitating the activation of AR panels in their immediate vicinity.

User Feedback. Finally, we need to consider the various states an adaptive AR panel can be in (Figure 2). Before being gazed at, an object can already display an icon or symbol that highlights it holds a gaze-adaptive panel to explore. During dwell this icon or symbol can be used as progress bar not only so the user receives feedback the system understands its intent, but also to give users a chance to advert their eyes if they do not wish to trigger the AR panel.

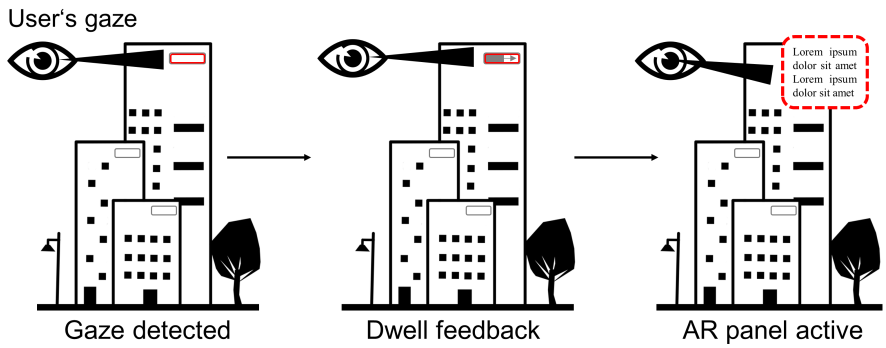

Fig. 2: The concept of gaze-adaptive AR UI from the user's perspective. The user views a city and can gaze at one of the panel's indicators. Then, the feedback indicates progress of the dwell time, which leads to opening the content.

4 User Study: AR vs World Focus

Our first study investigates the effect on user performance and experience when using a gaze-adaptive UI compared to a baseline condition. We use two tasks to vary whether the user focuses on interaction with real objects, or on interaction with the virtual content provided by AR. In the task, the user either searches for a specific AR information, or for an object in the environment.

Baseline Rationale. We use a baseline of showing information all the time, as this is a common depiction of AR use in everyday environments such as Feiner's illustration of a AR-mediated city [10]. We considered several baselines. A button-triggered UI baseline is popular, but with our focus on lightweight information access [8, 23, 33] it may not be ideal. It is possible that the hands may be occupied or otherwise unavailable. It is also overhead for simple information access, e.g., the user would point & click on each information panel in each single trial. Similarly, in peripheral vision approaches [16], the user would have to divert their eyes off to the periphery for each panel. We also considered a context adaptation baseline. But we consider these approaches as different and complementary. It uses knowledge from context data to suggest a potential selection. If context data is insufficient to make a selection, the user can utilise a method where they are in more control by gaze.

We chose always-on information as baseline. This is as Feiner's vision [10] is still desirable today, i.e. that the world is augmented with informative context information. But it is an open question how to design access to those, and how adaptive methods such as eye gaze are different to such an always-on method.

Research Questions. The research questions of this study include:

- **User performance**: How is user performance affected in the object search? While the gaze-adaptive UI shows up to one AR panel at a time, they dynamically appear at gaze – it is unclear how this affects the task performance.

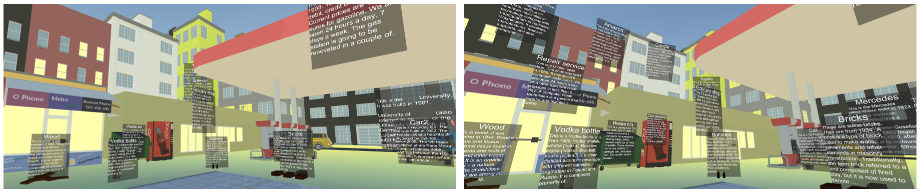

Fig. 3: The virtual city environment from the first study from a participant's perspective during the *always on* condition. A trial displaying 20 (left) and 40 AR panels (right).

– **User experience**: Even if always-on is faster in one task, performance is not necessarily linked to preference. Which method do users prefer for each task, and across both tasks as a generic UI mode?

4.1 Experimental Design

Our study followed a within-subjects design (balanced using a Latin Square) with two independent variables: AR panel behavior (all on, gaze adaptive) and search task (information, object). Participants interacted in a city-like environment that is inspired by Feiner's illustration of a AR-mediated city [10], representative of these types of search tasks while using AR. We implemented this experience using a VR setup for increased fidelity and ecological validity [25].

AR panel adaptation method. As the name implies, in the *all on* condition all AR panels were displayed concurrently. In the *gaze adaptive* condition we employed a triggering mechanism that relied on a *temporal activation* of 300 ms. We chose this as our pilot tests have shown that this provides a good compromise between fast activation and minimal accidental activation. In our pilot, 15 users tested object activation times from .3s to 3s in a simplified environment and provided preferences. The *spatial activation* was implemented by a 2×2 m which allowed to easily gaze on objects without interference from the eye-tracker's precision. *User feedback* consisted of a progress bar (see Figure 4 – top). These parameters are optimised for our study task – and can be adapted when used in other use cases (e.g., as in the following study).

Task. In the *information* condition participants were tasked with finding the year in which a building was completed. This information was present in each AR panel associated with the various buildings in the scene – the year in which each building was completed was randomized at the start of each new trial. In the *object* condition participants were tasked with finding a 2D representation of a person, animal or object (see Figure 4 – bottom) randomly placed in the scene at the start of each trial. An object was "found" when gazed at for 2 seconds.

Repetitions. Finally, to add variety to our trials in terms of potential occlusion and information overload, we varied the number of AR panels available in the scene between 20 and 40. This was randomized but equally distributed among conditions so that each user experienced each layout once (see Figure 3). This leads to 2 adaptation methods × 2 tasks × 2 repetitions = 8 trials per user.

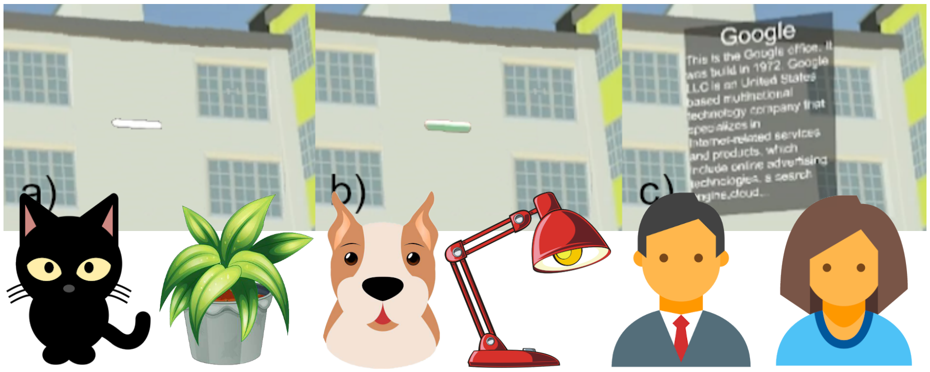

Fig. 4: Elements from the conditions in the first study. Top: three states of our gaze adaptive mechanism: (a) progress bar associated with an AR panel before being gazed at; (b) progress bar during dwell; and (c) the AR panel after being triggered. Bottom: objects a participant had to find in the *object* condition.

4.2 Experimental Setup

The city-like environment was implemented and deployed using the same setup as in the pilot study (Unity, HTC Vive Pro Eye). This included various buildings such as skyscrapers and apartments, a grocery store and a fuel station, and other pedestrians and cars that moved around. The AR panels associated with these elements had similar properties: a headline and a body of 250 characters displayed in white on a semi-transparent black background. These panels displayed appropriate content, such as opening times or the age of the building, and were displayed on average at 16.56 m ($SD = 16.6$) from participants' starting position ($\approx 7.15°$ of visual angle). This made them easy to gaze at with the eye-tracker employed in the study. In case of occasional occlusions, the AR panel closest to the participant was triggered. Finally, participants were free to move in a small area of $3{\times}3$ m at the center of the scene – the bird's eye view of this scene, together with all AR panel locations can be seen in Figure 5.

4.3 Procedure

The study took place in a quiet and empty laboratory. Hygiene measures as required in the institution were ensured for each user as of COVID-19. After providing consent and filling in the demographic questionnaire, participants calibrated the eye-tracker and started the study. Only brief instructions were given as training, as the task was straightforward for all users to perform. At the beginning of each trial, the participant's goal was displayed in VR, either as question (i.e. information search) or as the person, animal or object they had to find in the scene (i.e, object search). This was displayed for five seconds, after which we would start measuring task completion time. We also told participants that during object search trials no hints were hidden in any of the AR panels and that these were of no use to the task. After participants completed both search conditions for each of the AR panel behaviors they filled in a brief usability survey using a 5-point Likert scale.

Survey questions were specifically chosen based on what we aim to find out: how easy and intuitive the task was and how stressed and distracted participants felt. The 4 Likert questions were: "*Performing the task using this adaptation was easy/intuitive/stressful/distracting*". They were chosen to assess the trade-off between performance and experience. After completing all tasks, participants reported on their favorite conditions and provided a brief rationale.

4.4 Participants

26 people participated in the study (5F) with a mean age of 29.2 years ($SD = 10$). Using a 5-point Likert scale participants reported high affinity with technology ($M = 4.23$, $SD = 0.86$) and moderate VR / AR affinity ($M = 2.88$, $SD = 1.24$).

4.5 Results

Below we report on task completion times, usability and user feedback/preference. As two users were identified as outlier as of large times, we excluded them from the statistical analysis.

Task Completion Times We conducted a Shapiro-Wilk test using SPSS. It showed that the data of all conditions follow a normal distribution: information task + all on (p=.11), information task + gaze (p=.51), object task + all on (p = .39), object task + gaze (p = .2). An ANOVA showed a significant effect of task (F(1,25)=257.4, p¡.001, r=.88) and adaptation method (F(1,25)=12, p=.002, r=1) on task completion time, as well as a significant interaction effect (F(1,25)=12, p¡.001, r=.63). Post-hoc pairwise comparisons with Bonferroni corrections showed the two techniques had a significant effect on task completion time in the object (p = .035) and information search task (p = .002). In information search the particular times are *all on* (M=21.48s, SD=12.01s), *gaze adaptive* (M=37.3s, SD=17.4s); in object search *all on* (M=12.4s, SD=6.5s), *gaze adaptive* (M=9.4s, SD=3.7s). This shows that the *All_on* method was faster in information search, and the $Gaze\, t_a = 0$ method faster in the object search task.

Usability Usability results are shown in Figure 6 b) and c). We used a Friedman test with post-hoc Wilcoxon signed rank tests with Bonferroni corrections, of which we report the significant results only. For the information search task, the effect of the adaptation method was statistically significant for the category "distracting" (Z=-2.8,p=.006, r=.4). For the object search task, significant effects were revealed in categories "ease" (Z=-2, p=.047, r=.29), "intuitiveness" (Z=-2, p=.042, r=.29), and "distracting" (Z=-3.6, p¡.001, r=.52). Users found *Gaze_adaptive* (M=4.62, SD=.58) easier to use than *All_on* (M=4.17, SD=.96). Users perceived *Gaze_adaptive* (M=4.62, SD=.58) as more intuitive than *All_on*

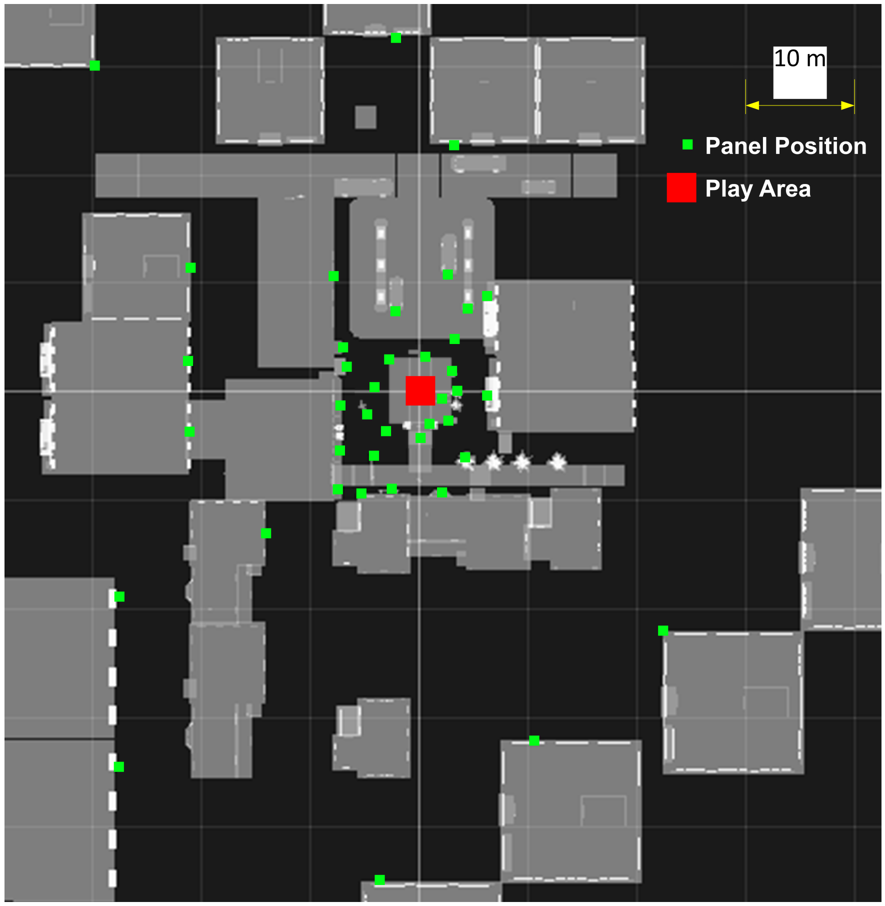

Fig. 5: The city-like environment used in the search study. All the AR panels are displayed in green, and participants' play area is displayed in red (i.e., the area in which participants were free to move in during the study).

(M=4.1, SD=1.1). Lastly, they found *All_on* (M=3.1, SD=1.35) more distracting than the *Gaze_adaptive* technique (M=1.63, SD=.86).

User Feedback The general concept of gaze-based AR activation was well received and described as "immersive" [P7, P12] and "practical" [P17]. P17 said: "this could replace looking at your phone while walking (...) to find out information about a shop or something (...)".

Information search task. 10 out of 26 participants preferred the *all on* panel behavior setting, with nine participants highlighting the speed and ease at which information could be scanned. Other reasons for this preference were the reduced visual movement necessary [P18] and the reduced likeliness of overlooking some information [P4]. The *gaze adaptive* condition was preferred more

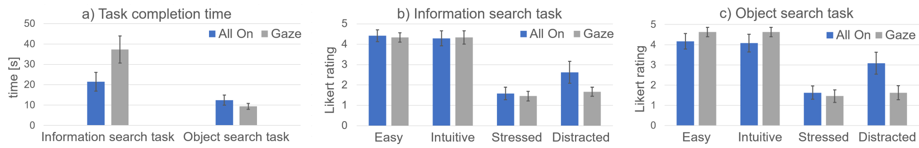

Fig. 6: Results of the search study: a) Average task completion time for the two tasks and conditions. Likert rating for information b) and object search task c). The error bars denote 95% CI.

often (11 out of 26 participants), with six participants mentioning reduced information overload. Two participants reported that this was a "more pleasant" mechanism for prolonged use, and others reported they could focus better on a single panel in this manner [P05], allowing them to build a better understanding of where they had looked already [P23]. More generally [P6, P14, P22] reported preferring the "feel of it", but [P23] suggested a more explicit gaze behavior to trigger the AR panels (e.g., a "blink or similar"). The remaining five participants did not express any preference.

Object search task. 24 out of 26 participants preferred the *gaze adaptive* AR panels, citing being less distracting (14), not getting in the way of the search task (6), and how intuitive it was [P6, P13]. On the other hand, the all on behavior was preferred by one participant who could "easily search around the panels" [P22]. Finally, one participant reported no preference as it felt the progress bars that responded to gaze could be distracting [P25].

Concurrent search tasks. Lastly, we asked participants to imagine a more representative scenario where they would often have to alternate between between information and object search tasks. In this scenario most participants still preferred the *gaze adaptive* behavior (22 out of 26), with one participant highlighting how AI could play a crucial role in this domain: "maybe the adoption of different (dwell) times for different objects (...) according to your interests." [P25]. Further, one participant referred that the dwell user feedback was quite effective at promoting exploration. Finally, two participants reported no preference in this scenario, and one reported preferring the *all on* panel behavior.

4.6 Discussion

We studied a gaze-adaptive UI for AR in a virtual city with a plethora of information, firstly quantifying the trade-off between performance, user experience, and occlusion issues. The main insight is that users, after experiencing both tasks, find the gaze-adaptive UI more suitable as a UI mode. Users found the always-on condition, a default AR viewing mode, more distracting in both tasks.

In the task where users focused on the VR environment, we discovered that users are faster with the gaze-adaptive UI and found it easier and more intuitive. Note that it would be misleading to assume that this is an expected result. The gaze-adaptive UI dynamically triggers information on gaze, which could as well

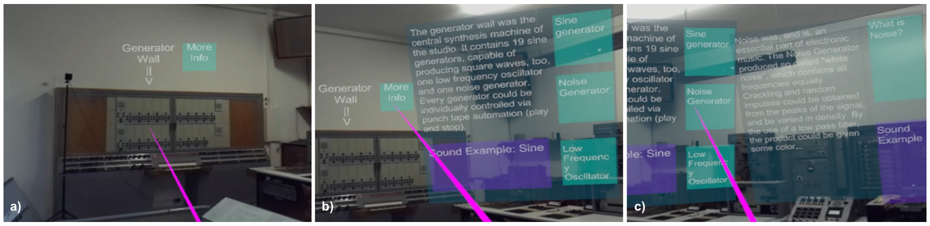

Fig. 7: Various examples of AR panels from our museum AR prototype application: (a) shows a first AR panel after it has been triggered, simply displaying the name of the object it is matched to and a green controller that can trigger further content (b and c). Gazing at content in purple plays back audio (b). Finally, the purple line is purely illustrative of the user's gaze and was not shown in the UI.

be distracting, and affect performance, compared to a static information visualisation. Yet, our study showed that this is not the case in the tested condition.

In the task where users focused on retrieving information from AR panels, interestingly we find that even when users are faster with the *All_on* condition, users preferences were split, and they also found the constant display as more distracting. Thus, in cases when usability is prioritised over performance, as it can be with many open AR environments, the gaze-adaptive UI is a promising alternative UI mode.

Our study has several limitations to consider when interpreting the findings. It was focused on an abstract virtual city task (not real AR), and a basic task of visual search. Such an abstraction allows to gather a more generalisable result, however real world situations can be different and the UI is likely to need adaptations. To better understand the utility in practice, in the following we investigate gaze-adaptation for a real use case.

5 Field Study: Museum Application

Following prior work in which deployments in museum were used as a powerful approach for in-situ studies of novel technologies and concepts and to validate lab findings [37], we implemented a prototype application for use in a museum. The application came and was a classic example of an AR application that can provide relevant information and media-rich content without getting in the way of the user's experience. This was studied during a museum's standard opening hours and lasted for three days. We recruited 50 participants on the spot as they visited one of the museum's exhibitions. We observed a balanced split between young/senior adults, no gender bias, and high diversity in nationality, but did not collect specific demographics to minimize the study duration.

5.1 Experimental Setup

Our prototype was developed for the German Museum in Munich[4], Germany and augments a music studio exhibition spanning approximately 4×2 m and featuring 15 objects. The studio featured historic sound generators/transformers and was chosen as the location due to a favorable, undisturbed position and because the available content suited our setup. Participants' movements were tracked in front of the exhibition in an area of 3.5×3 m using Vive base stations, and static hit boxes were placed over each object. The entire interaction was controlled via gaze, and unlike the previous study, each AR panel had between two to four layers of information that could contain more content and media (triggered after a dwell times of 300 ms, see Figure 7). Sound in particular was played while participants looked at the appropriate AR panel; these were displayed at 35% opacity so that the exhibition was always perceivable, and always facing the participant for improved readability.

Since the setup was not a controlled study, we used the chance to test new ideas and react to participant feedback on the fly. Adjustments and implementations were done during downtime or before visitors arrived. As such, we completely refrained from markers indicating interactable objects in initial runs, in fear they'd obscure the exhibit too much. However, they proved very necessary according to comments and behaviour of the first ¡ 10 participants (see visual feedback results). Following participant feedback from the previous study, we designed a simpler user feedback mechanism that displays a relevant character over an object when this is not being gazed at, instead of progress bar (see Figure 8). Additionally, these characters faded out once the corresponding AR panel has been interacted with. These could still be triggered, but now required a dwell time of 600ms. This way, viewed panels were less prone to accidental activation, as they are now only offering old, possibly boring information. Since users are still likely to scan these objects again in search for new, it seemed logical to raise the respective threshold of confidence for inferred interest, i.e. activation time. Finally, we used the same hardware as before with the addition of a Zed Mini Mixed-Reality[5] camera for high-resolution video see-through AR (2560×720 p, 60 fps, 110° FoV), and thus a more pleasant experience of the exhibition.

5.2 Procedure and Metrics

Participants were acquired by a sign at the exhibition, by evoked interest due to other users testing the setup, or verbal invitation by the experimenter. Instructions included a brief explanation of hardware, the UI, and gaze adaptation, as well as an instruction on how to put on and adjust the headset, which was helped with if asked for. No particular task was given to participants other than casually engaging with the augmented exhibition. This was preceded by the calibration of the eye-tracker and some simple instructions on the functionality

[4] Deutsches Museum: http://www.deutsches-museum.de/
[5] https://www.stereolabs.com/zed-mini/

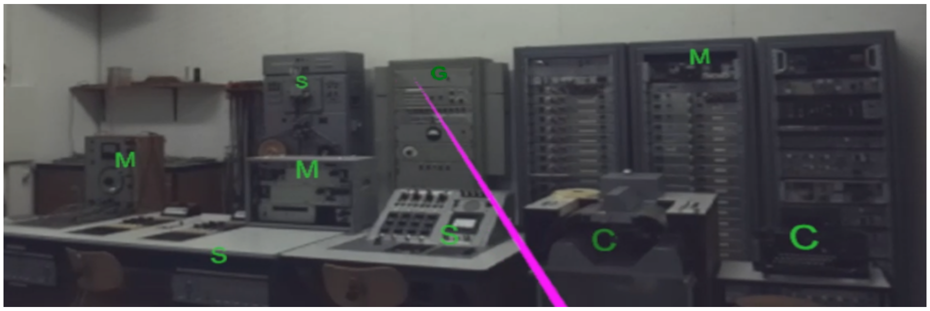

Fig. 8: The museum scene when none of the AR panels are active. A simple and relevant character is displayed over each object, indicating these can be gazed at for further content. Two of these characters appear faded out to indicate users already interacted with them (”G” and ”C”).

of our application. Afterwards, the UI was immediately active and reactive to gaze. Participants interacted with the prototype for approximately seven to 10 minutes, viewing over 80% of all AR panels developed (including its layers). We took notes of participants verbal comments and actions, and asked for optional feedback on ease of use, speed, and how natural the interaction felt (via a 5-point Likert scale). From mental reconstructions, most verbal comments were positive, with only a few negative ones. Of the 50 users, 22 left written comments not referring to technicalities (e.g. headset fit). 14 gave general positive feedback along the lines of ”interesting”, ”interactive” and ”easy to use”. Others gave constructive feedback regarding functions and behaviour, e.g. activation time. The other 28 users only left ratings.

5.3 Results

The application was well received, with median scores of 4 for ease of use ($\sigma^2 = 0.33$), 5 for speed of use ($\sigma^2 = 0.45$), and 4 for how natural the experience was ($\sigma^2 = 0.68$). Participants described it as ”interesting” [P17, P28, P43], ”intuitive” [P28], ”effective” [P29, P36, P49], ”wonderful” [P37] and ”very cool” [P50].

Temporal Activation. As for the previous implementation, participants were positive about the short dwell times (300 ms): ”It was so fast it almost felt supernatural at times.” / ”I especially liked the fast and prompt reaction”. Only two users preferred longer dwell times [P50]: ”it was like I was being flooded by information”; ”some sort of confirmation for the activation (...) as I was having a hard time viewing the objects without AR information” [P49]. In early runs, longer dwell times were tested to allow spectating the exhibit better, but proved impractical, as users didn’t gaze long enough to trigger the UI.

Visual Feedback. At the first prototypes without visual indicators, participants reported a feeling of ”missing out”, i.e., not knowing if there was an AR panel left they had not seen yet. We changed it afterwards and added visual feedback in form of letters, as described in the description of the application – which eliminated this type of question. The experimenter also observed that with this prototype, users were revisiting the same panels in search for new and partly missing out on others. For this reason, we added another feature of whether users have already visited a panel: the corresponding indicator had lower opacity to indicate that it has already been viewed, allowing users to quickly assess their prior interaction. Also, one user suggested enhancing the letters we used. We could use descriptive symbols or short names instead of such characters. Nonetheless the response was quite positive: "it felt a little too technological at first, but turned out to be quite useful to find what you had not seen yet".

Multi-layered AR panels. One concern was the newly added multi-layered AR panels, which potentially could occlude large portions of participants FoV. This was proven not to be the case: "while you are reading the text, you do not want to view anything else anyway". Also, while participants managed to unfold AR panels up to four times, they reported that more than this could quickly become cumbersome. Finally, the main issue reported was the amount of text in each AR panel and how this demanded more attention than participants might want to provide in a casual exploration.

5.4 Discussion

Through this field study we have validated some of our findings from the prior lab experiment such as the dwell time and the overall use of gaze adaptive AR panels. We have also expanded on this mechanism by allowing participants to not only trigger these panels but navigate between layered content using solely their gaze.

The overall results are very positive, particularly taking into account these were every day museum attendees with no particular training in AR or eye-tracking technologies (and despite any novelty effects). Participants seemed motivated by our user feedback to fully engage with the exhibition, but were not shy of ideas on how to improve the visual elements of our UI. This included animations or more focused textual descriptions.

Notably, this was quite a different task from the ones proposed in the previous study, and future work should consider some of our participant feedback regarding temporal activations (i.e., it could be found to be too fast for casual, leisurely explorations). While it's possible, that results might be affected by the filter, that only people open to technology were willing to try the prototype, it's probably not significant, as most people seemed to be of average technological understanding. The exhibit, while quite technological itself, was situated in a non technological part of the museum featuring music instruments in general.

6 Overall Discussion

We discuss the main trade-offs of gaze adaptive AR interfaces in light of the study findings and our experiences across the two studies.

Search vs. Occlusion Trade-off. One of the main themes investigated in the paper is the contrast between the gaze-based adaptation with a default mode where it is always visible. Our research indicates that it depends on a search versus occlusion trade-off. The default is best for highest performance to find a specific information in one out of many panels in the environment, useful for tasks that prioritise virtual content. The gaze-adaptive UI provides a trade-off where the reality is temporally occluded, making it easier to perceive and search in reality, albeit still able to instantly access context information in the vicinity. Thus, a hybrid technique that prioritizes more real world viewing but also supports virtual content. We believe our implementation showed promising results, as users rated it significantly less distracting in both tasks. Future work could explore different gaze-adaptive UIs to improve search performance, e.g., panels that are always-on, but fade out when the user do not fixate them.

Multi-layer Information. Our museum application demonstrated a qualitative advantage to an always-on condition. If all the information panels of this application would be active, the entire field of view is occluded, whereas the gaze-adaptive UI allowed users to interact with panels on demand. Occlusion is substantially minimised, as the panels only open when fixating a small indicator dot on the object, while the remaining object area is free for viewing. Moreover, by integrating an advanced functionality where users can navigate more pages by gaze selections allowed to substantially increase the amount of information that users can experience. This points to thinking of gaze-adaptive UIs for a different class of use cases, where users often shift their focus between real objects and the information it could encapsulate.

Advanced Gaze Behaviour Adaptation. By deploying the gaze-adaptive application in the museum, we gained insights into the applicability and features of the system. Visitors who experienced the system enjoyed the subtly appearing augmentations around the museum artifacts they are looking at, and saw the potential to enrich the experience of exhibits. In addition, we believe that further mechanisms to adapt to the user's gaze behaviour can be beneficial, such as subtle indicators that communicate where AR panels reside before active, or historic gaze information to mark what has been already viewed. The order of appearance of gaze adaptive panels can be based on what the user has seen before, enabling a story-telling experience that gradually unfolds new information.

Transfer to VR interaction. As the user study shows, many of our findings regarding gaze adaptive AR interfaces are also applicable in VR. Using gaze as a modality to access context information seems to be independent from the type of mixed reality interface, as long as interactable objects are highlighted in an appropriate way, e.g. by a visual marker.

Effect of 2D targets. In the object search task of the user study, we chose to use 2D targets, that always face towards the user and stand out from the 3D background. This tends to make the task easier in comparison to 3D objects, that might be occluded or face away from the user. Assuming that the differences in the task completion time are founded on the distraction of the user by the panels, it could be stated that the gaze-adaptive method is faster for object search, even if the task gets more difficult to complete.

7 Design Recommendations

Across the two user studies, we derived the following design recommendations:

- We recommend a fast activation time (0.2-0.5 sec), because in our context users tended to miss interactive objects as they did not dwell long enough on them otherwise.
- In the context we studied, occlusion was not a major issue as people seem to want to view the occluding information anyway most of the time.
- Still, a mechanism for returning to no augmentation is necessary, which could be simply the user's action of averting gaze in a less densely augmented area, or other means of input, e.g. a gesture or gaze reactive HUD element.
- A visual marker indicating interactable objects should be employed to ensure a good applications experience where users are aware of the interactions possible. The marker can carry a low level of information, e.g. a letter, symbol or a short name, as long as it is unobstructive. Moderate colors and reduced opacity are recommended.
- To give users even more reference for navigation, markers can react to being activated and change color like a clicked link.

8 Future Work

To assess the effect of the minimal occlusion in the gaze-adaptive UI, it would be interesting to compare gaze-adaptive UIs to not using AR at all. Additionally, comparing against a manual activation of the AR panels is important to understand further trade-offs and identify the limitations of the techniques. In principle we believe manual input is more useful for highly interactive applications like games or messaging, while gaze-adaptive UIs allow light information retrieval in everyday situations. Nonetheless, our study reveals, and contributes a first empirical assessment of the contrasting interaction properties in comparison to a more basic always-on AR experience. Another challenge to consider is handling and mitigating occlusions as well as testing different methods of conveying visual information surpassing text, such as animations, 3d models, or screen space augmentations corresponding to the currently viewed object.

Furthermore, an examination of additional types of content will be necessary to exploit the full potential of gaze adaptive interfaces. While gaze adaptation as technique for activation is principally independent of the content's type, parameters, like dwell or fade out time may be varied.

Our system and subsequent evaluations were based on pass-through AR with a large headset. The field of view, lower resolution (than real world) and weight can have an effect on users. We envision our concept to be applied with lightweight AR devices to fulfill its full potential in everyday settings.

Prior research on interaction in public space shows, people are often in groups and that user interfaces can be designed to account for this [31]. Future work could look into how the concept could be applied to multi-user scenarios. This can point to interesting new ideas, for example, in the museum context, in a group of people it would be easy to observe if other group members are inspecting an exhibit or are reading relevant information. If both are wearing AR glasses, this might be less obvious. Future work could look at mitigating such cases, for example by providing indicators what fellow visitors are currently engaging with, or which parts of an exhibit have received more attention by the people.

9 Conclusion

We explored a gaze adaptive UI that allows users lightweight access of context information about real objects in the environment. Such a method provides a unique trade-off between reality perception and virtual UI access. We firstly present an empirical comparison that reveals how user performance and experience is affected in tasks that focus on either reality or virtual UI. We demonstrate the utility of the method by a museum application of the concept, that allows users to consume multi-level information at a glance, and report on user feedback from its field deployment. Our research offers insights into trade-offs and UI design of gaze in more open AR scenarios, and the positive user reception is a promising result to employ gaze adaptation as an always-on UI mode when AR becomes more lightweight and populated in the consumer market. In future, we aim to extend our investigation to more optical see through AR devices and to broader use cases such as tourism and shopping, as well as effects and preferences of long term use.

Acknowledgments. We thank the German Museum in Munich and in particular Claus Henkensiefken for their collaboration in the context of this project as well as all visitors who participated in our research. The presented work was funded by the German Research Foundation (DFG) under project no. 425869382 and by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr [MuQuaNet].

References

1. Ajanki, A., Billinghurst, M., Gamper, H., Järvenpää, T., Kandemir, M., Kaski, S., Koskela, M., Kurimo, M., Laaksonen, J., Puolamäki, K., et al.: An augmented reality interface to contextual information. Virtual reality **15**(2-3), 161–173 (2011)

2. Azuma, R., Furmanski, C.: Evaluating label placement for augmented reality view management. In: Proceedings of the 2nd IEEE/ACM international Symposium on Mixed and Augmented Reality. p. 66. IEEE Computer Society (2003)
3. Bell, B., Feiner, S., Höllerer, T.: View management for virtual and augmented reality. In: Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology. pp. 101–110. UIST '01, ACM, New York, NY, USA (2001). https://doi.org/10.1145/502348.502363, http://doi.acm.org/10.1145/502348.502363
4. Bernardos, A.M., Gómez, D., Casar, J.R.: A comparison of head pose and deictic pointing interaction methods for smart environments. International Journal of Human-Computer Interaction **32**(4), 325–351 (2016)
5. Blattgerste, J., Renner, P., Pfeiffer, T.: 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. pp. 1:1–1:9. COGAIN '18, ACM, New York, NY, USA (2018). https://doi.org/10.1145/3206343.3206349, http://doi.acm.org/10.1145/3206343.3206349
6. Bolt, R.A.: Gaze-orchestrated dynamic windows. In: ACM SIGGRAPH Computer Graphics. vol. 15, pp. 109–119. ACM (1981)
7. D. Lindlbauer, A. Feit, O.H.: Context-aware online adaptation of mixed reality interfaces. In: Proceedings of the 32th Annual ACM Symposium on User Interface Software and Technology. pp. 213–222. UIST '19, ACM, New York, NY, USA (2019)
8. Davari, S., Lu, F., Bowman, D.A.: Occlusion management techniques for everyday glanceable ar interfaces. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). pp. 324–330. IEEE (2020)
9. Feiner, S., Macintyre, B., Seligmann, D.: Knowledge-based augmented reality. Commun. ACM **36**(7), 53–62 (Jul 1993). https://doi.org/10.1145/159544.159587, http://doi.acm.org/10.1145/159544.159587
10. Feiner, S.K.: Augmented reality: A new way of seeing. Scientific American **286**(4), 48–55 (2002)
11. Fender, A., Herholz, P., Alexa, M., Müller, J.: Optispace: Automated placement of interactive 3d projection mapping content. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. p. 269. ACM (2018)
12. Grubert, J., Langlotz, T., Zollmann, S., Regenbrecht, H.: Towards pervasive augmented reality: Context-awareness in augmented reality. IEEE Transactions on Visualization and Computer Graphics **23**(6), 1706–1724 (2017). https://doi.org/10.1109/TVCG.2016.2543720
13. Hartmann, K., Götzelmann, T., Ali, K., Strothotte, T.: Metrics for functional and aesthetic label layouts. In: International Symposium on Smart Graphics. pp. 115–126. Springer (2005)
14. Höllerer, T., Feiner, S.: Mobile augmented reality. telegeoinformatics: location-based computing and services. Taylor and Francis Books Ltd., London, UK **21**, 2004 (2004)
15. Imamov, S., Monzel, D., Lages, W.S.: 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). https://doi.org/10.1109/VR46266.2020.00110
16. Ishiguro, Y., Rekimoto, J.: Peripheral vision annotation: Noninterference information presentation method for mobile augmented reality. In: Proceedings

of the 2Nd Augmented Human International Conference. pp. 8:1–8:5. AH '11, ACM, New York, NY, USA (2011). https://doi.org/10.1145/1959826.1959834, http://doi.acm.org/10.1145/1959826.1959834
17. Jacob, R.J.K.: What you look at is what you get: Eye movement-based interaction techniques. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 11–18. CHI '90, ACM, New York, NY, USA (1990). https://doi.org/10.1145/97243.97246, http://doi.acm.org/10.1145/97243.97246
18. Julier, S., Baillot, Y., Brown, D., Lanzagorta, M.: Information filtering for mobile augmented reality. IEEE Comput. Graph. Appl. **22**(5), 12–15 (Sep 2002). https://doi.org/10.1109/MCG.2002.1028721, http://dx.doi.org/10.1109/MCG.2002.1028721
19. Keil, J., Zoellner, M., Engelke, T., Wientapper, F., Schmitt, M.: Controlling and filtering information density with spatial interaction techniques via handheld augmented reality. In: International Conference on Virtual, Augmented and Mixed Reality. pp. 49–57. Springer (2013)
20. Kim, M., Lee, M.K., Dabbish, L.: Shop-i: Gaze based Interaction in the Physical World for In-Store Social Shopping Experience. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. pp. 1253–1258. CHI EA '15, Association for Computing Machinery, Seoul, Republic of Korea (Apr 2015). https://doi.org/10.1145/2702613.2732797, https://doi.org/10.1145/2702613.2732797
21. Kytö, M., Ens, B., Piumsomboon, T., Lee, G.A., Billinghurst, M.: Pinpointing: Precise head- and eye-based target selection for augmented reality. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. pp. 81:1–81:14. CHI '18, ACM, New York, NY, USA (2018). https://doi.org/10.1145/3173574.3173655, http://doi.acm.org/10.1145/3173574.3173655
22. Lages, W., Bowman, D.: Adjustable adaptation for spatial augmented reality workspaces. In: Symposium on Spatial User Interaction. SUI '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3357251.3358755
23. Lu, F., Davari, S., Lisle, L., Li, Y., Bowman, D.A.: Glanceable ar: Evaluating information access methods for head-worn augmented reality. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). pp. 930–939. IEEE (2020)
24. Madsen, J.B., Tatzqern, M., Madsen, C.B., Schmalstieg, D., Kalkofen, D.: Temporal coherence strategies for augmented reality labeling. IEEE transactions on visualization and computer graphics **22**(4), 1415–1423 (2016)
25. Mäkelä, V., Radiah, R., Alsherif, S., Khamis, M., Xiao, C., Borchert, L., Schmidt, A., Alt, F.: Virtual field studies: Conducting studies on public displays in virtual reality. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. p. 1–15. CHI '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3313831.3376796
26. McNamara, A., Boyd, K., George, J., Suther, A., Jones, W., Oh, S.: Information Placement in Virtual Reality. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). pp. 1078–1079 (Mar 2019). https://doi.org/10.1109/VR.2019.8797910, iSSN: 2642-5246
27. McNamara, A., Boyd, K., Oh, D., Sharpe, R., Suther, A.: Using Eye Tracking to Improve Information Retrieval in Virtual Reality. In: 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). pp. 242–243 (Oct 2018). https://doi.org/10.1109/ISMAR-Adjunct.2018.00076

28. McNamara, A., Kabeerdoss, C.: Mobile Augmented Reality: Placing Labels Based on Gaze Position. In: 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct). pp. 36–37 (Sep 2016). https://doi.org/10.1109/ISMAR-Adjunct.2016.0033
29. McNamara, A., Kabeerdoss, C., Egan, C.: Mobile User Interfaces based on User Attention. In: Proceedings of the 2015 Workshop on Future Mobile User Interfaces. pp. 1–3. FutureMobileUI '15, Association for Computing Machinery, Florence, Italy (May 2015). https://doi.org/10.1145/2754633.2754634
30. Miniotas, D., ?pakov, O., Tugoy, I., MacKenzie, I.S.: Speech-augmented Eye Gaze Interaction with Small Closely Spaced Targets. In: Proceedings of the 2006 Symposium on Eye Tracking Research & Applications. pp. 67–72. ETRA '06, ACM, New York, NY, USA (2006). https://doi.org/10.1145/1117309.1117345
31. Müller, J., Walter, R., Bailly, G., Nischt, M., Alt, F.: Looking glass: A field study on noticing interactivity of a shop window. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. p. 297–306. CHI '12, Association for Computing Machinery, New York, NY, USA (2012). https://doi.org/10.1145/2207676.2207718, https://doi.org/10.1145/2207676.2207718
32. Müller, L., Pfeuffer, K., Gugenheimer, J., Pfleging, B., Prange, S., Alt, F.: Spatialproto: Exploring real-world motion captures for rapid prototyping of interactive mixed reality. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI '21, Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3411764.3445560, https://doi.org/10.1145/3411764.3445560
33. Pfeuffer, K., Abdrabou, Y., Esteves, A., Rivu, R., Abdelrahman, Y., Meitner, S., Saadi, A., Alt, F.: Artention: A design space for gaze-adaptive user interfaces in augmented reality. Computers & Graphics (2021). https://doi.org/https://doi.org/10.1016/j.cag.2021.01.001, http://www.sciencedirect.com/science/article/pii/S0097849321000017
34. Pfeuffer, K., Alexander, J., Chong, M.K., Gellersen, H.: Gaze-touch: Combining gaze with multi-touch for interaction on the same surface. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology. pp. 509–518. UIST '14, ACM, New York, NY, USA (2014). https://doi.org/10.1145/2642918.2647397, http://doi.acm.org/10.1145/2642918.2647397
35. Pfeuffer, K., Mayer, B., Mardanbegi, D., Gellersen, H.: Gaze + pinch interaction in virtual reality. In: Proceedings of the 5th Symposium on Spatial User Interaction. pp. 99–108. ACM (2017)
36. Pfeuffer, K., Mecke, L., Delgado Rodriguez, S., Hassib, M., Maier, H., Alt, F.: Empirical evaluation of gaze-enhanced menus in virtual reality. In: 26th ACM Symposium on Virtual Reality Software and Technology. VRST '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3385956.3418962, https://doi.org/10.1145/3385956.3418962
37. Prange, S., Müller, V., Buschek, D., Alt, F.: Quakequiz: A case study on deploying a playful display application in a museum context. In: Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia. p. 49–56. MUM '17, Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3152832.3152841, https://doi.org/10.1145/3152832.3152841

38. Rivu, R., Abdrabou, Y., Pfeuffer, K., Esteves, A., Meitner, S., Alt, F.: Stare: Gaze-assisted face-to-face communication in augmented reality. In: ACM Symposium on Eye Tracking Research and Applications. ETRA '20 Adjunct, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3379157.3388930, https://doi.org/10.1145/3379157.3388930
39. Rosten, E., Reitmayr, G., Drummond, T.: Real-time video annotations for augmented reality. In: International Symposium on Visual Computing. pp. 294–302. Springer (2005)
40. Sibert, L.E., Jacob, R.J.K.: Evaluation of eye gaze interaction. In: Proc. SIGCHI Conference on Human Factors in Computing Systems. pp. 281–288. CHI '00, ACM, New York, USA (2000). https://doi.org/10.1145/332040.332445, http://doi.acm.org/10.1145/332040.332445
41. Stellmach, S., Dachselt, R.: Look & touch: Gaze-supported target acquisition. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 2981–2990. CHI '12, ACM, New York, NY, USA (2012). https://doi.org/10.1145/2207676.2208709, http://doi.acm.org/10.1145/2207676.2208709
42. Tanriverdi, V., Jacob, R.J.K.: Interacting with eye movements in virtual environments. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 265–272. CHI '00, ACM, New York, NY, USA (2000). https://doi.org/10.1145/332040.332443, http://doi.acm.org/10.1145/332040.332443
43. Tönnis, M., Klinker, G.: Boundary conditions for information visualization with respect to the user's gaze. In: Proceedings of the 5th Augmented Human International Conference. pp. 44:1–44:8. AH '14, ACM, New York, NY, USA (2014). https://doi.org/10.1145/2582051.2582095, http://doi.acm.org/10.1145/2582051.2582095
44. Velichkovsky, B., Sprenger, A., Unema, P.: Towards gaze-mediated interaction: Collecting solutions of the "midas touch problem". In: Human-Computer Interaction INTERACT'97. pp. 509–516. Springer (1997)
45. Vertegaal, R., et al.: Attentive user interfaces. Communications of the ACM **46**(3), 30–33 (2003)
46. White, S., Feiner, S.: Sitelens: Situated visualization techniques for urban site visits. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1117–1120. CHI '09, ACM, New York, NY, USA (2009). https://doi.org/10.1145/1518701.1518871, http://doi.acm.org/10.1145/1518701.1518871