DualGaze: Addressing the Midas Touch Problem in Gaze Mediated VR Interaction

Pallavi Mohan[1*] Wooi Boon Goh[1†] Chi-Wing Fu[2‡] Sai-Kit Yeung[3§]

[1]Nanyang Technological University, Singapore
[2]The Chinese University of Hong Kong, Hong Kong, China
[3]The Hong Kong University of Science and Technology, Hong Kong, China

Abstract

With the increasing acceptance of eye tracking as a viable interaction method for Virtual Reality (VR) headsets, thoughtful gaze interaction methods need to be carefully designed to meet common challenges such as the Midas Touch problem, where users unintentionally select onscreen objects by gazing upon them. This paper presents DualGaze, a novel interaction method in which users perform a distinctive two-step gaze gesture for object selection. Once users gaze upon an object that they wish to select, a confirmation flag pops up next to the object at a location where the users' gaze just passed through. This trajectory-adaptive flag placement strategy reduces the chance of unintentional confirmation by requiring a returning gaze back to the flag. We conducted a user study to compare the accuracy and selection speed of DualGaze and the popular gaze fixation method on a simple gaze-typing task. Our results show that DualGaze is significantly more accurate while maintaining a comparable selection speed that was observed to improve with familiarity of use.

Keywords: virtual reality, interaction methods, gaze interaction, Midas touch, eye tracking

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Interaction styles

1 Introduction

In their confined setting, conventional VR headsets provide only limited input modalities. Commonly-supported interaction methods include tracking the user's head orientation, external controllers and buttons, and more recently, gaze interaction by eye tracking. While gaze interaction is mostly used as an auxiliary input modality in regular desktop environments, with the current dearth in input modalities for immersive VR setups, it seems like a natural candidate for primary user input control. Hence, new and thoughtful mechanisms are needed to address common problems that the users faced in gaze-based interaction. One such problem is the Midas Touch [1], where users may unintentionally select an user interface (UI) element, e.g., a button, by accidentally looking at it. Common existing solutions to the problem include the gaze fixation, eye blinking, and action-selection [2]. While gaze fixation and eye blinking often incur errors due to unintentional selections, action-selection typically relies on a separate external trigger, which are mostly headset-specific, despite the fact that VR headsets have limited interaction modalities.

This paper presents *DualGaze*, a novel interaction method that makes sole use of gaze to address the Midas Touch problem for gaze mediated VR interaction. It can be readily employed in standard VR menu without the need of having any external triggers.

*pallavi003@e.ntu.edu.sg
†aswbgoh@e.ntu.edu.sg
‡cwfu@cse.cuhk.edu.hk
§saikit@ust.hk

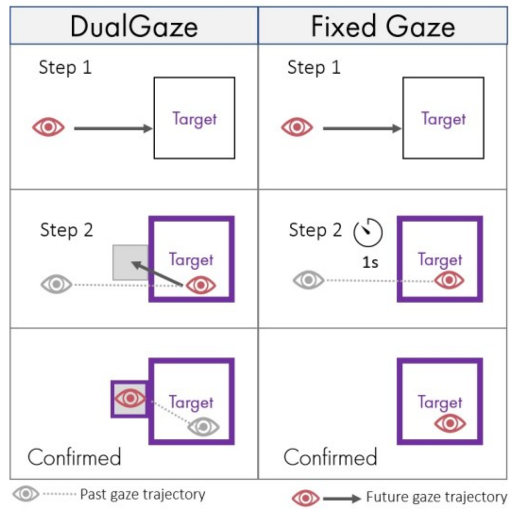

Figure 1: Two comparative interaction designs for selecting a target with one's gaze in a gaze-mediated VR user interface. Our DualGaze method (left) and the Fixed Gaze method (right).

DualGaze is designed in such a way that when the user gazes at a selectable UI element, a confirmation flag would pop up right next to the UI element for the user to willfully look at to confirm the selection; see the left column in Figure 1 for an illustration of the DualGaze interaction procedure. Instead of arbitrarily positioning the confirmation flag, we strategically adapt its location to the user's gaze trajectory, i.e., we put it at a location that user's gaze just passed through right before entering the UI element's boundary. Hence, to confirm a selection, the user needs to consciously avert their eyes back to the flag. In this way, the chance of unintended or accidental gaze at the confirmation flag is reduced, especially when one gazes at a UI element but has no intention of selecting it.

Compared to the popularly-employed gaze fixation method (see the right column in Figure 1), DualGaze has higher accuracy in terms of avoiding the Midas touch problem. In addition, the increase in accuracy is obtained without sacrificing the interaction responsiveness in terms of selection speed. These results were confirmed by quantitative data collected from a user study, which shows that users of DualGaze can perform simple gaze-typing tasks with significantly better accuracy than gaze fixation. Further analysis shows that towards the end of the tasks, users' increasing familiarity with DualGaze resulted in DualGaze outperforming the comparative method in terms of selection speed. This also suggests that the DualGaze technique is not difficult to master. A concluding survey found most users perceive DualGaze to be their preferred gaze interaction method in scenarios where the selection task at hand is deemed critical and irreversible, such as exiting a game.

978-1-5386-7592-2/18/$31.00 ©2018 IEEE
DOI 10.1109/ISMAR-Adjunct.2018.00039

2 RELATED WORK

The rapid rise in popularity of low-end VR headsets as Google Cardboard has seen VR getting accessible to a wide range of users and applications, in fields as diverse as medicine and gaming. However, virtual reality poses several environmental constraints on users in terms of the headset build, by closing users off from the outside world and hampering their direct access to the surrounding physical objects. Such a gap reveals potential for novel interaction techniques and designs to be developed for applications in virtual and mixed reality scenarios. While all VR headsets support head tracking, other interaction modalities heavily rely on the external controllers, for which users have to operate through haptic feedback, and tap gestures on the periphery of the headset [3].

2.1 Eye tracking and Gaze-Based Interaction

The concept of eye tracking is not a recent idea. Primitive attempts to implement it can be traced as far back as 2002 [4], and research on eye typing goes even further back [5]. Of late, however, several HMDs in the market enable external IR-based eye trackers to be embedded inside, allowing eye tracking to be a whole new input modality in VR [6][7]. Studies combining VR eye tracking with BCI have also been ongoing[8][9]. With the introduction of eye tracking to VR, however, there is still a dearth of thoughtful gaze-based interaction methods, especially designed to work with the constraints of VR headsets. While there has been a considerable amount of research devoted to eye tracking in VR [10][11][12], research in gaze-based interaction in VR is yet to catch up.

An early work by Jacob and Stellmach [13] explores the issues faced by gaze-based interaction methods comprehensively. The work discusses the Midas Touch problem in detail, and presents studies on two interaction approaches to address the problem for specific cases. Later, Huckauf and Urbina [14] presented a follow-up work that takes an alternative approach using antisaccades, which are an explicit eye movement studied in cognitive psychology, for object selection. Users performing tasks involving antisaccades fixate their gaze on an object in the centre of a screen, and an eccentrically located stimulus is presented to them. They make a saccade in the direction opposite to the stimulus, with an amplitude equal to the distance between the stimulus and fixation, to confirm the selection. The work by Istance et al. [15] provides a solution to the Midas Gaze problem that uses eye dwells to select objects as in [13], called Snap Clutch, but employs different modes of operation, such as Dwell Click mode, and Eye Control Off mode.

The recent work by Kasprowski et. al. [16] sheds light on problems faced by the gaze contingent interfaces today, including the Midas Gaze problem. The work also discusses an experiment, in which mouse, touchpad and gaze contingent interfaces are employed to play a simple shooting game. However, the selection method used in the gaze contingent interface is an action-selection method based on an external touchpad trigger. Moreover, there was a significant number of users who stated that mouse input was more convenient than the proposed gaze-based selection method.

Smooth pursuit eye movements have also been suggested solutions to the Midas Gaze problem, in works such as [17] where gaze-based interaction in smart watches is facilitated by having the user's gaze follow a circularly moving target to confirm selection. This requires a more complicated eye movement sequence than that of other discussed methods since it involves a constantly moving target that the gaze has to follow. Our proposed method, on the other hand, employs a target that is fixed and requires only a simple glance-back to activate. Another work by Bednarik et al. employs machine learning to predict the intended gaze location [18].

Likewise, there have been several other works that explore gaze-based action-selection methods, e.g., the work by Velichkovsky et. al. [2], which collects all possible solutions to the problem. However, almost all solutions involve the use of an external trigger or input, such as the work by Kumar et. al. [19], which uses keyboard presses to confirm the gazed-at selections. However, gaze interaction in VR needs a selection method that is self-sufficient and independent of the external peripherals.

2.2 Gaze-based Interaction in VR

A recent work by Piumsonboon et. al. [20] similarly explores such new gaze-based selection methods. While one of them uses smooth pursuit eye movement in VR such as in [21], another of them is Duo-Reticles, a spatiotemporal-based selection method in which a moving average is made to align with the eye's current gaze position. However, since the moving average always remains at a minimum distance behind the user's current gaze, the second reticle is constantly moving, so it is distracting for the user to gaze back at; see their user study's subjective evaluation for the details. Our proposed method attempts to solve this by employing a non-moving, yet trajectory-adaptive confirmation flag, at which users can intuitively glance back to confirm the selection without being distracted, as opposed to consciously matching the current gaze reticle back to the moving reticle.

3 THE DUALGAZE INTERACTION METHOD

Gaze-based interfaces in immersive VR systems typically involve a menu selection, where users have to scan through a number of options on the screen and then select one of them afterwards. This gives rise to some very particular interaction design challenges as the travelling gaze during the visual scanning cannot be reliably distinguished from a wilful gaze on an identified choice.

There are several existing gaze interaction methods that seek to mitigate this problem by introducing a distinguishing action for confirming a choice. Such solutions in VR mainly involve gaze fixation, blinking, or an external trigger such as a remote switch [2], to confirm a selection once a desired option has been visually targeted. Each of these solutions has their respective limitations, as discussed earlier in the paper.

In this paper, we present a novel gaze-mediated interaction method called DualGaze that addresses some of these limitations. Firstly, unlike methods that employ external triggers to confirm selections, DualGaze purely employs user's gaze, thus allowing hands-free immersive VR interaction design.

Secondly, DualGaze employs a single and consistent eye gaze interaction modality to perform both the "choosing" action and "confirmation" action in the selection process. Hence, the approach avoids the many side effects of using blinking to distinguish between the visual search and selection. Frequent switching between gaze and blinking could feel unnatural and repetitive blinking can lead to rapid eye fatigue. In addition, momentary occlusion of the iris and pupil during blinking can cause the eye tracking systems to lose their smooth and continuous gaze tracking capability.

Thirdly, unlike the gaze fixation method, the gaze positions for the "choosing" action and "confirming" action in DualGaze are not co-located. As such, carefully planning successive gaze locations can help distinguish the two different but correlated actions. This addresses the limitation of the gaze fixation method that an accidental prolonged gaze on a UI element may cause an unwanted activation. A typical remedy to ameliorate the Midas Touch problem in gaze fixation is to increase the gaze duration before the confirmation, but this degrades the responsiveness of the

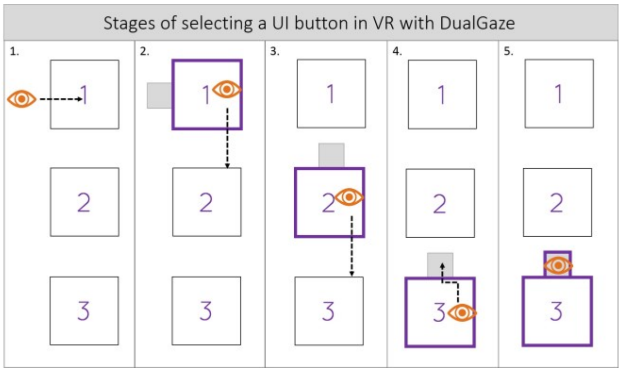

Figure 2: The stages of the DualGaze interaction method, with the red eye denoting the user's gaze point, involving a user's gaze travelling through buttons 1 and 2 to confirm the user's desired choice, button 3. The positioning of the confirmation flag does not obstruct the gaze trajectory and forces one to gaze back to confirm.

interactive experience, especially in situations where multiple selections need to be made in quick succession.

Figure 1 (left) shows the basic idea behind our DualGaze interaction design. The activation of a target onscreen choice is achieved using a quick successive two-gaze strategy. After the user's gaze enters the target region (e.g., a "clickable" button), a smaller pop-up box called the confirmation flag would appear right next to the target region. By a second gaze on the confirmation flag, the user can confirm the selection. We design the confirmation flag in the form of a simple box, roughly a quarter of the menu button in size, and coloured distinguishably from the target button. Putting the flag close to the target region reduces necessary gaze travel for the users to complete the selection action.

To address the Midas Touch problem, we need an appropriate placement of the confirmation flag to facilitate intentional and deliberate eye gaze trajectory, and to reduce the chance of accidental confirmation. After a brief delay of 0.3 seconds (an empirically derived value) between the gaze entry and fixation on the target region, the confirmation flag would appear, thus further suppressing the chance of unintentional confirmation. To cancel a selection, users may simply look out of the target region but not towards to the confirmation flag, as one would normally do. The process of scanning menu options and finally making a selection using DualGaze is illustrated in Figure 2.

The design goal of DualGaze is to provide a *reliable*, *accurate* and *responsive* gaze-mediated interaction method that can be used in an immersive eye-tracked VR environment for making onscreen option selections. The next section describes some preliminary studies done in deriving appropriate options in the DualGaze interaction design.

3.1 Designing the DualGaze Interaction

In designing the DualGaze interaction, we have studied and evaluated design alternatives using small-scale user studies that involve five to seven participants.

A fundamental design variable in DualGaze is the confirmation flag placement strategy for improving the reliability of gaze interaction. Two placement strategies were evaluated. The first strategy is to position the flag at a fixed place with respect to the selected button every time (in this case, the right side of the button), to allow the user to gaze at a predictable and fixed location. The second strategy is to position the flag at the entry point from which the user's gaze enters a particular selection region. In other words, we adapt the confirmation flag position to the trajectory of the user's gaze. The rationale behind the second strategy is that a conscious effort is required by the user to avert their eyes and return to the point where the user's gaze enters the selection region – such a location is salient, yet does not obstruct the user's gaze path as he/she moves forward.

We conducted a small-scale user study with seven participants to study which of the two strategies led to better user performance by using the simple number selection task shown in Figure 3 (right). The one-way ANOVA test was done, and the results showed using a trajectory-adaptive flag placement could produce significantly better accuracy (Mean(M)=9.7, $p=0.038$) when compared to the first strategy, which puts the flag at a fixed position (M=7.3). This result confirms that using the trajectory-adaptive strategy to place the confirmation flag can effectively avoid accidental confirmation as users made significantly less erroneous selections during a trial to select ten numbers out of the buttons.

Task completion time was also found to be lower for trajectory-adaptive positioning (M=23.78s) than fixed positioning (M=37.47s). However, the one-way ANOVA test (with $p=0.081$) does not show this difference to be significant. In any case, one can safely conclude that using the trajectory-adaptive placement strategy could produce a more reliable and accurate interaction design that does not compromise interaction responsiveness; see Figure 2 for an example of this flag placement strategy.

From the above user study, we complete the final DualGaze interaction design, which adopts the trajectory-adaptive confirmation flag placement strategy.

3.2 A Comparative Gaze-based Interaction Method

Gaze fixation [2] is a popular gaze interaction method, by which users stare at an onscreen target region for a pre-determined period to select and confirm his/her choice on the target. We choose this method as the baseline in our comparative study, since like DualGaze, gaze fixation is also a pure gaze-only interaction design. While gaze fixation has been in widespread use and various aspects of its design have been analyzed [22], we conducted a small-scale user study with six participants using our eye-tracking system based on the FOVE VR headset to determine the appropriate delay-to-activation time setting for the experimental setup, so that we can obtain a fair and uniform comparative evaluation against DualGaze.

Three different durations were used in the study to evaluate their relative accuracy and total time taken to input ten numbers in the same number selection task. The delay durations used were Short (0.5s), Medium (1.0s) and Long (1.5s). Ther results from the one-way ANOVA tests show that Medium (M=7.5s) has significantly better accuracy than Short (M=3.3s) with $p=0.002$, while Long (M=9.3s) has better accuracy than Medium, though the improvement was not significant ($p=0.072$). As for the task completion time, Short (M=10.4s) was significantly faster ($p=0.002$) than Medium (M=19.4s) and Medium was significantly faster ($p=0.002$) than Long (M=30.5s). As a result, a delay of 1.0 second was chosen to be the best empirical parameter to trade-off between accuracy and task completion time for the method, which we will refer to it as "Fixed Gaze." This duration was adopted in the comparative study with DualGaze.

4 User Study

4.1 Experimental Design and Set-up

A user study was conducted to investigate and compare the performance of DualGaze with that of Fixed Gaze. The same task was presented to each method and the task requires the user to gaze-type ten numbers that are presented as five sequential pairs of two-

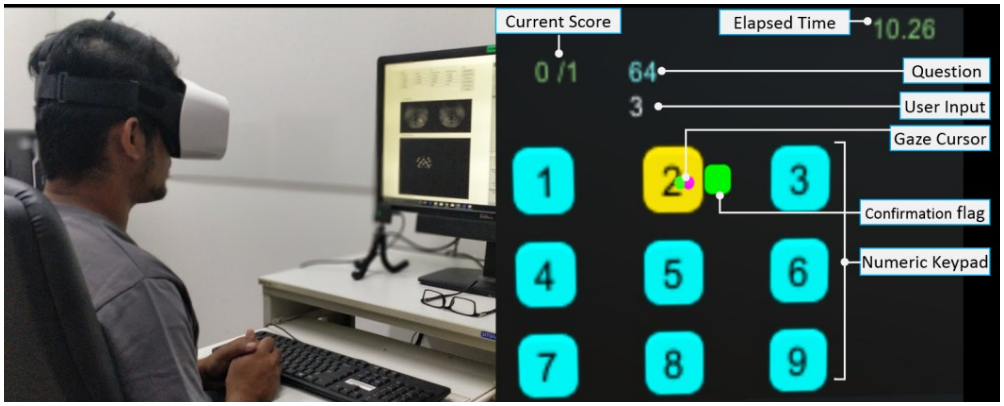

Figure 3: Left: The physical setup, with the FOVE HMD used in this work. Right: A sample layout of the UI displayed on the FOVE HMD, where users have to gaze-select ten numbers on the keypad shown.

digits target numbers. In this study, the temporal order of these ten numbers is labelled Questions 1 to 10. The sequential order of the two methods was alternated between trials to prevent biases due to asymmetric familiarity and/or exposure-induced eye fatigue.

The VR headset used for the study was FOVE VR developed by a Tokyo-based start-up. FOVE uses an infrared technology for accurate tracking of eye movements with low latency. Embedded sensors within the headset track the users' pupils, while they interact with the objects on the screen. Setting up the FOVE headset involves a calibration of an external positional tracking camera, which should always be able to see a full view of the headset, as well as staying connected to the PC through USB wires; see Figure 3 for the physical setup. The tasks given to the users were FOVE-based applications developed on Unity IDE using the FOVE's Unity-based SDK, on a stereoscopic VR view. Two separate DualGaze and Fixed Gaze applications were developed for the study, and the user's accuracy and time performance data were collected through log files at the end of each task. Figure 3 (right) shows a sample layout of the UI.

4.2 Participants and Procedure

4.2.1 Participants

The user study was conducted with 27 participants. The participants' age range from 18 to 30 years, and they were university students and staff recruited through flyers after we obtained an ethical approval for the study (IRB-2017-10-047). While 41% of the participants had used VR headsets before, 86% of them were first time users of an eye tracking system.

4.2.2 Procedure

Once a user signed a consent form, we briefly explain to him/her the order of the proceedings and strapped the VR headset comfortably around their head. Then, they were asked to perform a calibration routine, whereby their eyes have to follow a green dot moving around the screen. Once the calibration was successful, the participant was presented the first application, where he/she was required to input five 2-digit numbers from a numeric keypad (see Figure 3 (right)) using the selection method particular to the current application. In all cases, the participant used their gaze to select numbers from the numeric keypad and a beep was sounded every time the participant has successfully selected and confirmed a number. This was repeated for the other interaction method, with a minute's rest for the participant in between each task. The participant's final score (number of digits they input correctly), and other aspects of their performance such as time spent on each question, were recorded via a log file generated by the program. After completing both tasks, the participants were asked to fill a survey regarding their experience and preferences for the two different gaze selection methods.

5 Results

5.1 Accuracy

An important goal of gaze-based interaction is to allow users to perform their selections in a reliable and accurate manner. In this study, we ask the following research question:

RQ1. *Is DualGaze more accurate than Fixed Gaze?*

Accuracy was computed as the number of correct entries out of the ten digits that the user made by using each of the two interaction methods. Results from the one-way ANOVA tests showed that DualGaze is significantly more accurate than Fixed Gaze ($p=0.0004$). Participants tend to make more errors using Fixed Gaze ($M=7.7$, $SD=1.56$) than when using DualGaze ($M=9.19$, $SD=0.77$).

5.2 Selection Speed

Another important goal that affects user experience is the responsiveness of an interaction method. This is directly related to the speed in which the required selection can be made by the user. The next question we considered was:

RQ2. *Is DualGaze faster for users to make a selection than Fixed Gaze?*

To answer this question, we computed the average time that users took to input a correct number. The time taken to input a correct number begins when the gaze moves from a previously-selected digit to the current selected digit and ends when a correct confirmation is registered. DualGaze ($M=2.13s$, $SD=0.39s$) was found to be slightly faster than Fixed Gaze ($M=2.15s$, $SD=0.42s$) in terms of the average time taken for making a correct answer, but such time difference was not statistically significant ($p=0.868$). In fact, the selection time taken is affected by factors such as the travel time as well as the selection time, and the travel time in turn may depend on successive locations of the digits, so the visual search delay and the repeated incomplete attempts. If external factors such as travel time were eliminated and only the average selection time was considered, then DualGaze ($M=0.65s$, $SD=0.36s$) is found to be significantly faster than Fixed Gaze ($M=1.01s$, $SD=0.006s$) with a ($p=0.00001$). This result implies that the time users spent doing the first gaze on a digit and then the second gaze on its

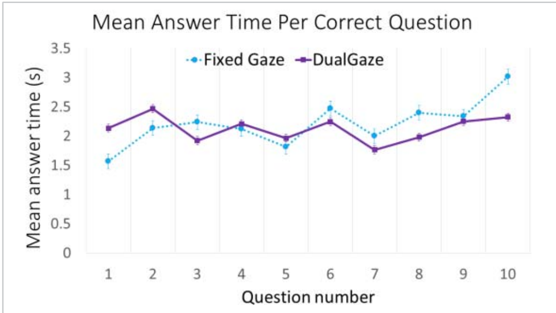

Figure 5: Temporal performance changes in selection speed, with Standard Error bars.

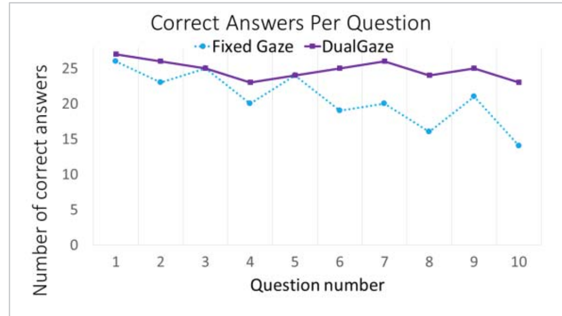

Figure 4: Temporal performance changes in accuracy.

corresponding confirmation flag in DualGaze is still faster than a single fixated gaze in Fixed Gaze.

5.3 Temporal Performance Changes

We observed in answering RQ2, users were not significantly faster in making their selections using DualGaze compared to Fixed Gaze when averaging over the entire experiment. However, DualGaze is a novel and conceptually more complicated gazed-mediated interaction method than Fixed Gaze, and users were not familiar with its operation at the start of the experiment. As such, we were interested in finding out if users were able to improve their performance in selecting numbers after more practice, as they continued to use DualGaze in the experiment. We then posed the following question:

RQ3. *Is DualGaze faster in making a selection than Fixed Gaze in the latter half of the experimental run?*

To answer this question, the average time that the users took to input a correct number for the last five questions were computed. Results from the one-way ANOVA test showed that users using DualGaze (M=2.1s, SD=0.93s) were observed to take significantly less time ($p=0.0476$) to make correct selections compared to Fixed Gaze (M=2.4s, SD=1.25s). This result testifies to DualGaze's fast learning curve. It can also be seen in Figure 4 that the selection speed improvement effect with the users' familiarity can be observed in DualGaze, but not in Fixed Gaze. While DualGaze started off on a significantly slower than Fixed Gaze, it caught up quickly and consistently surpassed Fixed Gaze with a lower task completion time towards the end.

6 Discussion

6.1 Accuracy

Our user study results show that DualGaze has a marked advantage in accuracy over Fixed Gaze. DualGaze was designed with the primary goal of minimizing involuntary selection. The design provides an additional step of confirmation by requiring the users to avert their eyes to the confirmation flag. In addition, the adaptive trajectory placement of confirmation flags also reduces the likelihood that the user's current gaze direction will accidentally trigger this flag. This resulted in a significantly better accuracy, since users had to consciously confirm a choice by reversing the current direction of their gaze to the confirmation flag. Fixed Gaze on the other hand, uses the same action (i.e. gaze fixation) to select and confirm an onscreen selectable region. This means users will make a selection whenever their eyes rest on one of these selectable regions. The only ways to avoid making a selection are to either look away to non-active regions of the screen or keep one's gaze roving all the time. The former may be challenging if the visual interface has many large active selectable regions and the latter can result in discomfort and rapid eye fatigue. Interestingly, our quantitative results of the temporal changes in accuracy plotted in Figure 5 show that the accuracy for both interaction methods starts to drop as the experimental run progresses. This could be partially due to fatigue or users' inability to sustain their attention level in the given tasks. However, what is noticeable is that the drop in accuracy for DualGaze is only marginal compared to that of Fixed Gaze. Users were asked using a survey questionnaire to rate the accuracy of DualGaze and Fixed Gaze. Their responses showed that only a slightly higher number of users perceive the accuracy of DualGaze to be better than Fixed Gaze despite the undeniable quantitative performance measures indicating DualGaze's significantly better accuracy. Those who rated DualGaze as more accurate felt that DualGaze gave them more assurance of the option they were selecting. According to one user, *"There was more accuracy in the first case* (DualGaze) *as two stage confirmation is involved."* Another user remarked that *"The selection of an option with its acknowledgment is preferable"*. These remarks seem to concur with users' preferences when they were given various use-case scenarios. For example, Scenario 1 involved having to select and confirm a single option from an onscreen multiple-choice quiz having three equally-spaced options. The major concern to address in this use-case scenario was the issue of reliability and accuracy as this was a quiz with multiple options and choosing a correct option would be deemed as a critical decision. It was observed that a majority of the users (59%) preferred to use DualGaze for their selection. Reasons given by users included: *"This allows me to confirm my answer."* and *"Getting my answers right in the quiz is very important. It feels like fixed gaze can have a lot of accidental clicks"*, which highlighted their sense that DualGaze is a more reliable and accurate interaction method.

6.2 Selection Speed

An important feature of the proposed DualGaze interaction method is that it managed to achieve high accuracy without compromising on the selection speed. Unlike the Fixed Gaze method, improved accuracy can only be obtained by increasing the delay-to-activation time (see Section 3.2) and thus the corresponding time required to complete the selection task. The temporal changes in selection speed plotted in Figure 4 also highlight the potential for improvement in selection speed as users become familiar with the DualGaze interaction design and can quickly anticipate where the confirmation flag is likely to appear based on their current gaze trajectory. Responses from the survey questionnaire suggest that some users seem to perceive that Fixed Gaze has a slightly speedier selection time than DualGaze. Some of the users who had a false sense of Fixed Gaze having better selection speed were also users with more selection errors due to the higher accidental selection with the Fixed Gaze method. As a result, they felt they completed the given task of selecting 10 numbers much faster with Fixed

Gaze, and they apparently did not take into account their high error rates in the tasks (which in turn led to them fast-forwarding through the set of 10 questions quickly). This observation highlighted to us the limited ability of a subjective survey questionnaire in evaluating the true performance of an interaction design during a user study. Quantitative measures are still the more objective assessment tool to assess non-experiential performance metrics like accuracy and selection speed.

In short, the general consensus among the users was that DualGaze would be preferred for the assurance and trust that it gave users while using it, given it felt like an interaction method akin to a mouse click or a touchscreen tap in traditional interaction modalities.

7 Conclusion

DualGaze has been shown to be more accurate than the popular Fixed Gaze interaction method, as users made significantly less errors in making selections over onscreen buttons. As a result, users also perceived DualGaze to be more reliable as the majority of users picked this method over Fixed Gaze in use-case scenarios where activating a button is deemed critical and irreversible. They were also able to perform this selection a little faster using DualGaze. Analysis of the average selection time for DualGaze showed that it is significantly faster than Fixed Gaze. However, this significant speed superiority was not evident in the total task completion time, which took into account other external factors like the time expanded in making repeated incomplete flag confirmation. We already presented preliminary evidence that this task completion time can be reduced as users became more familiar with DualGaze. Readers are also encouraged to watch an accompanying video of the DualGaze interaction method for a visual account of our work.

8 Acknowledgement

This research project is partially funded by an internal grant from HKUST (R9429).

References

[1] R. Jacob and S. Stellmach, "What you look at is what you get," *Interactions*, vol. 23, no. 5, pp. 62–65, 2016.

[2] B. Velichkovsky, A. Sprenger, and P. Unema, "Towards gaze-mediated interaction: Collecting solutions of the 'Midas touch problem,'" in *Human-Computer Interaction INTERACT '97*, Boston, MA: Springer US, 1997, pp. 509–516.

[3] X. Yan, C.-W. Fu, P. Mohan, and W. B. Goh, "Cardboardsense: Interacting with DIY cardboard VR headset by tapping," in *DIS 2016 - Proceedings of the 2016 ACM Conference on Designing Interactive Systems: Fuse*, 2016.

[4] M. Sodhi, B. Reimer, J. L. Cohen, E. Vastenburg, R. Kaars, and S. Kirschenbaum, "On-road driver eye movement tracking using head-mounted devices," *Proc. Symp. Eye Track. Res. Appl. - ETRA '02*, p. 61, 2002.

[5] P. Majaranta and K.-J. R~ih~, "Twenty Years of Eye Typing: Systems and Design Issues."

[6] "Oculus Patented Eye Tracking Device Days After Acquiring The Eye Tribe." [Online]. Available: https://uploadvr.com/oculus-patented-new-eye-tracking-device-days-acquiring-eye-tribe/. [Accessed: 27-Feb-2018].

[7] "FOVE Eye Tracking Virtual Reality Headset." [Online]. Available: https://www.getfove.com/. [Accessed: 15-Jul-2017].

[8] Bonkon Koo, Hwan-Gon Lee, Yunjun Nam, and Seungjin Choi, "Immersive BCI with SSVEP in VR head-mounted display," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 1103–1106.

[9] X. Ma, Z. Yao, Y. Wang, W. Pei, and H. Chen, "Combining Brain-Computer Interface and Eye Tracking for High-Speed Text Entry in Virtual Reality," vol. 1, no. 212, pp. 18–22, 2018.

[10] T. Pfeiffer, M. E. Latoschik, I. Wachsmuth, and J. Herder, "Evaluation of binocular eye trackers and algorithms for 3d gaze interaction in virtual reality environments," *J. Virtual Real. Broadcast.*, vol. 5, no. 16, p. 16, 2008.

[11] V. Tanriverdi and R. J. K. Jacob, "Interacting with eye movements in virtual environments," in *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '00*, 2000, pp. 265–272.

[12] J. Orlosky, T. Toyama, K. Kiyokawa, and D. Sonntag, "ModulAR: Eye-Controlled Vision Augmentations for Head Mounted Displays," *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 11, pp. 1259–1268, Nov. 2015.

[13] R. Jacob and S. Stellmach, "Interaction technologies: What you look at is what you get: Gaze-based user interfaces," *Interactions*, vol. 23, no. 5, pp. 62–65, 2016.

[14] A. Huckauf and M. H. Urbina, "Object selection in gaze controlled systems," *ACM Trans. Appl. Percept.*, vol. 8, no. 2, pp. 1–14, 2011.

[15] H. Istance, R. Bates, A. Hyrskykari, and S. Vickers, "Snap clutch, a moded approach to solving the Midas touch problem," in *Proceedings of the 2008 symposium on Eye tracking research & applications - ETRA '08*, 2008, p. 221.

[16] P. Kasprowski, K. Harezlak, and M. Niezabitowski, "Eye movement tracking as a new promising modality for human computer interaction," *Proc. 2016 17th Int. Carpathian Control Conf. ICCC 2016*, pp. 314–318, 2016.

[17] A. Esteves, E. Velloso, A. Bulling, and H. Gellersen, "Orbits," in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*, 2015, pp. 457–466.

[18] R. Bednarik, H. Vrzakova, and M. Hradis, "What do you want to do next," in *Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA '12*, 2012, p. 83.

[19] M. Kumar, A. Paepcke, and T. Winograd, "EyePoint," *Proc. SIGCHI Conf. Hum. factors Comput. Syst. - CHI '07*, p. 421, 2007.

[20] T. Piumsomboon, G. Lee, R. W. Lindeman, and M. Billinghurst, "Exploring natural eye-gaze-based interaction for immersive virtual reality," in *2017 IEEE Symposium on 3D User Interfaces, 3DUI 2017 - Proceedings*, 2017, pp. 36–39.

[21] M. Khamis, C. Oechsner, F. Alt, and A. Bulling, "VRpursuits: Interaction in virtual reality using smooth pursuit eye movements," in *Proceedings of the 2018 International Conference on Advanced Visual Interfaces - AVI '18*, 2018, pp. 1–8.

[22] A. M. Penkar, C. Lutteroth, and G. Weber, "Designing for the eye: design parameters for dwell in gaze interaction," *Proc. 24th Aust. Comput. Interact. Conf.*, pp. 479–488, 2012.