



Khamis, M., Oechsner, C., Alt, F. and Bulling, A. (2018) VRpursuits: Interaction in Virtual Reality Using Smooth Pursuit Eye Movements. In: 2018 International Conference on Advanced Visual Interfaces, Grosseto, Italy, 29 May - 01 Jun 2018, ISBN 9781450356169.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

© The Authors 2018. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in the Proceedings of the 2018 International Conference on Advanced Visual Interfaces, Grosseto, Italy, 29 May - 01 Jun 2018, ISBN 9781450356169 <https://doi.org/10.1145/3206505.3206522>.

<http://eprints.gla.ac.uk/170203/>

Deposited on: 5 October 2018

VRPursuits: Interaction in Virtual Reality using Smooth Pursuit Eye Movements

Mohamed Khamis
LMU Munich
Munich, Germany
mohamed.khamis@ifi.lmu.de

Florian Alt
LMU Munich
Munich, Germany
florian.alt@ifi.lmu.de

Carl Oechsner
LMU Munich
Munich, Germany
oechsner.carl@gmail.com

Andreas Bulling
Max Planck Institute for Informatics
Saarland Informatics Campus, Germany
bulling@mpi-inf.mpg.de

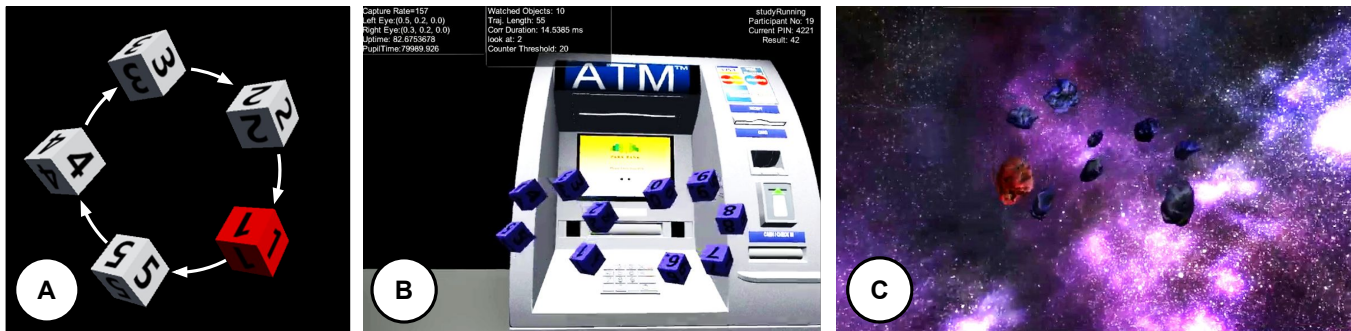


Figure 1: We investigate the selection of moving 3D targets in virtual environments (A) using smooth pursuit eye movements (arrows are for illustration only and were not shown to users). We study how parameters specific to VR settings influence the performance. We then develop and evaluate two sample applications: (B) a virtual ATM where users authenticate by following the digits with their eyes, and (C) a space shooting game where users blast asteroids by following them.

ABSTRACT

Gaze-based interaction using smooth pursuit eye movements (Pursuits) is attractive given that it is intuitive and overcomes the Midas touch problem. At the same time, eye tracking is becoming increasingly popular for VR applications. While Pursuits was shown to be effective in several interaction contexts, it was never explored in-depth for VR before. In a user study ($N=26$), we investigated how parameters that are specific to VR settings influence the performance of Pursuits. For example, we found that Pursuits is robust against different sizes of virtual 3D targets. However performance improves when the trajectory size (e.g., radius) is larger, particularly if the user is walking while interacting. While walking, selecting moving targets via Pursuits is generally feasible albeit less accurate than when stationary. Finally, we discuss the implications of

these findings and the potential of smooth pursuits for interaction in VR by demonstrating two sample use cases: 1) gaze-based authentication in VR, and 2) a space meteors shooting game.

CCS CONCEPTS

• Human-centered computing → Virtual reality;

KEYWORDS

Eye Tracking, Virtual Reality, Gaze Interaction, Pursuits

1 INTRODUCTION

Human gaze has significant potential for virtual reality (VR) applications. Not only can gaze be leveraged to learn about the user's visual attention, it also offers a natural and intuitive means for interaction. In the recent years, gaze interaction using smooth pursuit eye movements (Pursuits) has been continuously becoming popular due to its intuitiveness, and robustness against the Midas touch problem [37], i.e., the problem of distinguishing deliberate gaze input from the basic function of eye, namely to look around and perceive visual information. Pursuits is particularly relevant to VR; the dynamic nature of VR applications often requires selecting moving targets. For example, when communicating with avatars of virtual human agents in VR, a common approach is to point at them while they move [19, 22, 36]. Selection of moving targets is common

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

AVI '18, May 29–June 1, 2018, Castiglione della Pescaia, Italy

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

ACM ISBN 978-1-4503-5616-9/18/05...\$15.00

<https://doi.org/10.1145/3206505.3206522>

in VR games, and scientific simulations where users can, for example, select an object to track its development [10, 11, 34]. Although selection of moving targets in VR is important, it is challenging to point, touch or tap a target while it moves [10, 13, 21, 23]. On the other hand, Pursuits is intended for selecting moving targets.

Pursuits was extensively studied in different interaction settings, including on public displays [17], wearables [5], and smart environments [35]. Virtual reality differs from these in a number of ways. For one, unlike in desktop settings, VR users can move while interacting using Pursuits. While Khamis et al. employed Pursuits for interaction with large displays while walking [16], they studied walking parallel to the display while in VR users can move freely in different directions. Second, in contrast to 2D targets on normal displays, virtual 3D targets can be much closer and larger, span across a large movement trajectory, and move at different distances from the user's perspective. Third, previous work showed that gaze behavior when fixating on a 3D stimulus can differ from fixating on a 2D one [7, 26]. Despite these differences and its potential, Pursuits was never explored in-depth for virtual reality applications.

In this paper we fill this gap and explore pursuit interaction in VR. We study the performance of the technique for characteristic properties of virtual environments that were not explored in previous literature. This includes studying how well it performs when selecting targets of different sizes, different trajectory sizes (e.g., different radii of objects moving in circular motion), different distances from the user, and in cases where the user is stationary or walking in VR. We found that Pursuits is robust against different target sizes but that performance drops slightly when targets are too big because users do not fixate at a particular set of pixels on a 3D target's surface. We also found that larger trajectories improve performance, and that users can indeed make selections via Pursuits while on the move but performance is better when users are stationary. Finally, we implemented two use cases for Pursuits in VR (see Figures 1B and 1C) that are well perceived by users.

The contribution of this work is two-fold: (1) we report on the results of a user study (N=26) through which we investigate the performance of Pursuits in VR, and (2) we showcase and evaluate two VR applications that employ Pursuits.

2 RELATED WORK

Our work builds on two strands of prior work: (1) Eye tracking in VR and (2) interaction using smooth pursuits.

2.1 Eye Tracking in VR

The advent of affordable and high-quality VR headsets has incited the development of various VR applications. Eye tracking is a key technology for VR headsets and has therefore been integrated, for example, in the FOVE tracker¹ and the HTC Vive^{2,3}. Knowledge about the current gaze point can bring a lot of benefits to the user experience in VR. It can be used to speed up rendering of the virtual scene by limiting rendering to the user's high acuity area, so-called foveated rendering [8, 25, 28]. Eye tracking can also be used to navigate [24], enhance collaboration [1], or predict subjective

presence [38] in virtual environments. Eye gaze was also used for active interaction in VR, such as for steering [31]. Other headsets, such as the Microsoft HoloLens⁴, support head-pose tracking as an alternative to eye tracking, and was recently used to detect Pursuits-like movements using the head for AR applications [6].

In contrast to these previous works, we focus on gaze-based interaction with moving targets using smooth pursuit eye movements. Tripathi and Guenter used smooth pursuit for calibration in VR, but not for interaction [33]. Piumsomboon et al. used smooth pursuit to allow occluded objects to be selected, and hence objects to be moved on demand when fixated at [27]. While these works applied Pursuits, we contribute a deeper exploration of Pursuits in light of the unique properties of VR settings. We compared performance across different sizes of 3D virtual targets that are continuously moving, different sizes of movement trajectories, and distances between the user and the target.

2.2 Pursuits for Interaction

Until recently, the majority of work on gaze-based interaction utilized dwell time [12] or gaze gestures [4]. Smooth pursuit eye movements are increasingly becoming popular for gaze-based interaction. Initially introduced by Vidal et al. for interaction with public displays [37], the technique was subsequently studied in different contexts, including public displays [17], smartwatches [5], smart homes [35], and smart glasses [3]. It has been used for gaming [15, 37], authentication [2], voting [18] and text entry [20]. It was also successfully integrated into active eye tracking, where eye trackers follow users as they move along large interactive surfaces [16]. Using Pursuits overcomes the Midas touch problem, because it is unlikely a user would imitate a movement with their eyes without a stimulus to follow.

Piumsomboon et al. recently introduced RadialPursuit, a technique that employed smooth pursuit eye movements for interaction in VR [27]. RadialPursuit expands cluttered objects away from each other, and allows the user to select the object of interest as it moves away from the rest. In contrast to RadialPursuit, we use the Pearson's correlation coefficient rather than Euclidean distance difference to match eye movements with target positions. Pearson correlation is not influenced by poor eye tracker calibration [16, 17, 37], which makes it more robust since calibration often deteriorates when users take off VR headsets and put them on again. Furthermore, the Pearson correlation is robust against cases where the user's head bobs up and down due to movement [16]. This is particularly relevant to VR since moving in VR is an important topic that is increasingly becoming popular [32]. Finally, a core difference between our work and that of Piumsomboon et al. is that while they investigated a particular use case, we explore the basic properties of Pursuits with the aim of creating guidelines for researchers and practitioners who want to employ Pursuits in VR.

3 DESIGN SPACE OF PURSUITS IN VR

While the idea of using Pursuits for selecting moving targets was explored before, selecting 3D targets in virtual environments comes with unique properties. We identified different characteristics of Pursuits that may influence selection performance in VR:

¹<https://www.getfove.com/>

²<https://www.tobii.com/tech/products/vr/>

³<https://pupil-labs.com/blog/2016-08/htc-vive-eye-tracking-add-on/>

⁴<https://www.microsoft.com/en-us/holoLens>

1. Size of the trajectory: Previous work in smooth pursuit selection reported that performance is expected to be better when the trajectory size is bigger (e.g., bigger radius for circular trajectories) [34]. However the effect of trajectory size on Pursuits performance was never formally investigated before. In immersive VR headsets and in contrast to previously investigated interfaces, such as public displays [17, 37] and smartwatches [5], the trajectory size could vary widely; the user sees a large visual field in which targets could move, e.g., HTC Vive offers 110° of visual field.

2. Size of the target: Previous work discussed how Pursuits is expected to be independent of target size [37]. Accurate selection of very small 2D targets using Pursuits is indeed feasible as long as the target moves. However, it was also shown that gaze behavior when fixating at a 3D stimulus can be different from fixating at a 2D one [7, 26]. In VR, 3D targets are accompanied with many depth cues resulting from lighting, shadows, rotations, etc. These cues distract the user from fixating at a particular set of pixels on the target's surface when gazed at. Instead, users can freely gaze at any point on the target's surface as they follow. Selectable targets in VR can be too large than anything studied for Pursuits before.

3. Distance to target: One parameter that was never investigated for Pursuits before is the distance to the target. Distance to the target is particularly important for VR. While the distance is often constant in desktop and public display settings, where users position themselves 60–90 cm away from the display [16, 17], the distance between the user and a virtual 3D target can vary greatly.

4. Trajectory shape: Previous work investigated multiple trajectory shapes. Most existing work established that circular trajectories perform better [5, 15, 17], and our pilot tests did not show any tendencies for different results in VR. Due to this reason, we decided to not investigate this parameter in more detail.

5. Moving user: Another unique property of selecting moving targets in VR is that users themselves can be moving. There is a growing trend towards enabling users to walk in VR environments. For example, companies are offering wireless adaptors for VR headsets to allow users to walk freely without tethering⁵. Multiple companies introduced VR Walkers to allow users to walk in VR despite space limitations⁶. There is also a large body of previous research about enabling users to move in VR by walking in place [32] or by the so-called redirected walking [29].

While tracking the eyes of moving users for diagnostic and monitoring purposes is widely adopted, gaze-based interaction while the user is moving is relatively under-investigated. To our knowledge, the only exception is EyeScout where users interacted with a large display via smooth pursuit while walking parallel to the display [16]. However, in VR users could be moving towards or away from the targets while selecting them. This motivated us to explore how the movement of the user influences selection of moving targets in VR using smooth pursuit eye movements.

4 CONCEPT AND IMPLEMENTATION

Our implementation builds on previous work on Pursuits [37]. The key idea is to show the user a set of moving targets, and to compare eye movements to movements of the targets. The target whose

⁵<https://www.tpcastvr.com/product>

⁶<http://www.virtuix.com/>

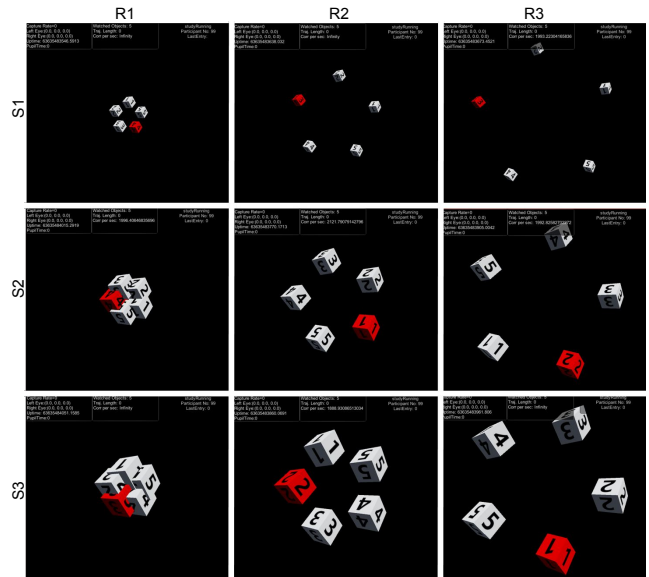


Figure 2: We experimented with 3 trajectory sizes (i.e., 3 radii lengths), 3 target sizes, and 3 distances to the target.

movement correlates most with that of the eye is deemed to be the one the user is looking at. As done in [37] and pilot tests, we opted for using the Pearson's product moment correlation for comparing the trajectory of the target with that of the gaze. We used a window size of 300 ms, and a threshold of 0.4. That is, every 300 ms the coordinates of each target in the last 300 ms are compared to the gaze coordinates in the same time window, and the correlation coefficients are stored for further analysis. Every 900 ms, we calculate the mean of the last three correlation coefficients C_{mean} for each target. The target with the highest C_{mean} is deemed to be the one the user is looking at as long as it is above the threshold (0.4). We used a sliding window, which means that instead of discarding all previous gaze points after each correlation check, we enqueue every new gaze point and dequeue the oldest one.

5 EVALUATION

We conducted a user study to derive guidelines that help designing performant and reliable Pursuits selection of moving target in virtual reality. In a first study, we narrowed down the possible properties of moving targets (e.g., target size, and trajectory size), in order to understand which of those have an impact on smooth pursuit selection in VR. Based on the results, we conducted an in-depth analysis to understand the impact of different parameters of some properties in a followup study. Afterwards, we conducted a final study to showcase two possible applications for moving target selection in VR and to collect qualitative insights from participants.

5.1 Apparatus

Participants were asked to wear an HTC Vive with a bionocular PUPIL eye tracker add-on [14] connected to a PC with an Intel Core i7 6500k processor, an NVidia GTX 1080 graphics adapter, and 16GB RAM running Windows 10. To make sure participants are not distracted by other objects in the scene, and to ensure that the recorded eye behavior is solely influenced by the manipulations

of the independent variables, we used an abstract scene with no objects other than the moving targets (see Figure 2).

5.2 Study Goal and Design

The aim of the study was to evaluate the performance of Pursuits in VR with respect to the characteristics mentioned in Section 3. Our study covered the following independent variables:

- (1) **Trajectory size**, we experimented with three radii (R1, R2, R3): 5.55°, 12.22°, and 17.14° in degrees of visual angle.
- (2) **Target size**, we experimented with cubes of three sizes (S1, S2, S3): 4.84°, 14.25°, and 22.95° in degrees of visual angle.
- (3) **Distance to targets**, we covered three distances (D1, D2, D3): 2 meters, 6 meters, and 16 meters.
- (4) **User state**: stationary and moving.

We intentionally set the target and trajectory sizes in degrees of visual angle. This was done because the perceived size of a 3D object changes depending on the distance to said object. Hence, our system adapted the absolute sizes to maintain the aforementioned sizes in degrees of visual angle across the respective conditions.

The study was designed as a repeated measures experiment. Each participant underwent two sessions, one per user state condition. In each of the two sessions, participants went through 27 conditions (3 trajectory sizes × 3 target sizes × 3 distances to targets) in a counter-balanced manner. Participants always started with the “stationary session” and followed it by the “walking session”.

5.3 Participants and Procedure

We recruited 26 participants aged between 21 and 37 ($M = 25.88$, 10 females). The experimenter started by explaining the study and the participant completed a consent form. Every participant was shown how to put on the headset and it was adjusted to be tight yet still comfortable. The participants were shown five numbered cubes, a random one of which was colored in red (see Figure 2). Once the experimenter hit the space button, the cubes started moving along a circular trajectory at the same speed in a clockwise direction. The participants’ task was to gaze at the red cube until all cubes disappeared. In the walking session, participants had to perform the selection while walking up to a white square that alternately appeared at two opposing edges of the available space and changed its position each time the user came within a certain range.

5.4 Results

We recorded a total of 1,512 trials (3 trajectory sizes × 3 target sizes × 3 distances to target × 26 participants). For each trial, we measured (1) the *selection time*: the time from the moment the cubes

started moving until the moment they were selected, and (2) the *percentage of detections*: whether or not the selection was correct.

5.4.1 Selection Time. The overall mean selection time was 3.07s ($SD = 1.91$). Mean selection time was 1.53 s when stationary, and 3.85 s when walking. This means that users perform Pursuits selections faster when stationary. We did not observe any consistent tendencies when comparing selection times across the different distances to targets and target sizes. As shown in Figures 3 and 4, mean selection times for the different distances to target and target sizes did not vary strongly across the conditions. However, there was a decline in mean selection times as the trajectory size increased: the radii R1, R2, R3 resulted in mean selection times of 2.8 s, 2.26 s, 1.87 s respectively. When differentiating further between stationary and walking, we found that the decline is even sharper when walking than when stationary (compare Figure 3A and Figure 4A). This means that the trajectory size has an influence on selection time in general, and particularly when the user is walking.

A generalized linear model was built to investigate the main effects of the independent variables and their pairwise interactions. A gamma distribution was assumed to match the data and selected for the model and a log-link function was applied. The results confirm a significant effect of the trajectory size (radius) in the walking scenario, while in the stationary setting, increasing the trajectory radius had no significant impact on selection time ($p_{R2} = .735$, $p_{R3} = .229$), an increase of the radius while walking significantly reduced expected selection time by factors .716 (R1 to R2) and .546 (R1 to R3, $p < .001$). In general, the change from stationary to walking setting resulted in an increase of the estimated selection time by a multiplicative factor of $exp(B)_{walking} = 2.40$ ($p < .001$). Results show no significant effects for size or distance to target.

5.4.2 Percentage of Detections. The mean values for percentage of detections did not vary widely for target size, trajectory size, and distance to target. While stationary, 79% of all entries were interpreted correctly, and 58% were correct while walking. This means that, similar to mean selection times, performance of Pursuits in VR is higher in terms of correct selection rates when stationary compared to when walking. As shown in table 1, accuracy seemed to be slightly higher when the distance to the target is shorter, but that was only the case when the user was stationary. Accuracy increased when the trajectory size was larger. This increase was sharper when walking: correct detection rate increased from 53% at R1 to 62% at the larger R3.

To find any significant main and interaction effects of the independent variables on accuracy, a binomial logistic regression was performed. Pairwise interactions between target parameters and movement type were integrated in the model. Of the four variables, “movement” had the most impact on correctness and the largest effect was detected for a change from “walking” to “not walking”: if the participant is moving, an expected odds ratio of 25.4% was estimated compared to the stationary setting. When participants are walking, a change from R1 to R2 results in an estimated odds ratio of 1.772 ($p = .069$), a change to R3 in 1.970 ($p < .05$). Target size had no significant effect on correctness.

	Trajectory Size			Target Size			Distance to Target		
	R1	R2	R3	S1	S2	S3	D1	D2	D3
Stationary	77%	78%	79%	79%	80%	75%	80%	78%	76%
Walking	53%	61%	62%	58%	58%	60%	56%	63%	57%

Table 1: Percentages of correct detections are generally higher when the user is stationary. Similar to selection times, correct detections increase when larger motion trajectories are used. This is particularly the case when walking.

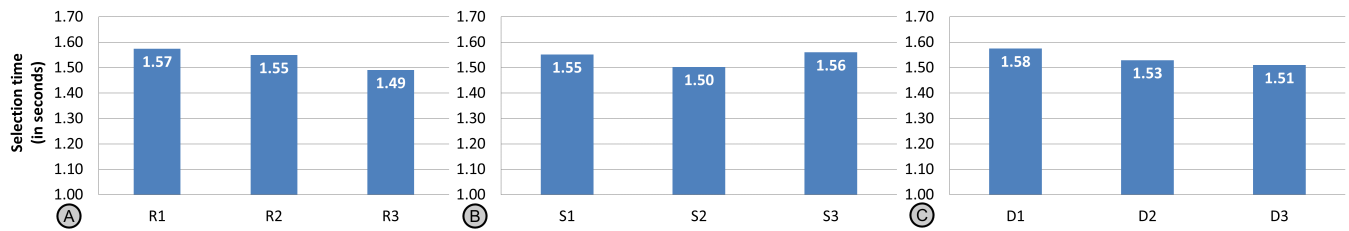


Figure 3: Selection times across the different conditions when *stationary*.

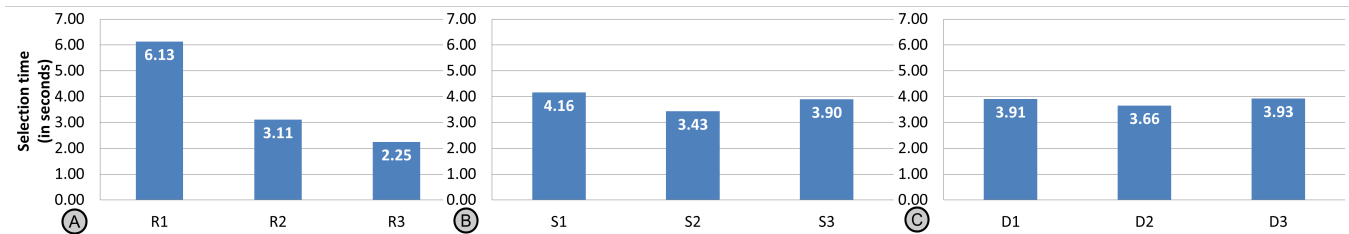


Figure 4: Selection times across the different conditions when *walking*.

6 SAMPLE APPLICATIONS

To demonstrate use cases for Pursuits in VR and gather qualitative feedback about this form of interaction, we designed two sample applications: (A) An authentication procedure at a virtual ATM that is performed while standing in front of it, and (B) a shooting game in which asteroids are destroyed while simultaneously walking. In both scenarios users have to gaze at targets moving in circular trajectories to trigger certain actions (Figures 1A and 1B). After trying each application, we collected qualitative feedback through a questionnaire, interviews, a NASA-TLX questionnaire, and a User Experience Questionnaire⁷. Both applications were tested with the same participants of the previous study.

6.1 Use Case A: Authentication

With the growing range of VR applications available on the market that require providing passwords, including those that feature in-app purchases or online shopping, authentication has recently become an important topic for VR [9]. Inspired by the work of Cymek et al. about authentication on situated displays using smooth pursuit eye movements [2], we developed an authentication scheme in which users authenticate at a virtual ATM by proving a 4-digit PIN via Pursuits (Figure 1A). Ten cubes, numbered from 0 until 9, moved in circular trajectories; half of which rotated in clockwise direction, while the other half rotated in anti-clockwise direction. The participant's task was to enter a 4-digit PIN provided by the experimenter. The cubes have a width of 0.5 meters and moved on a circular trajectory of radius 1.5 meters at an angular speed of $45^\circ/s^{-1}$. When an input was provided, visual feedback indicates to the participant that they can enter the next digit. Each participant entered 10 PINs. After each PIN, the ATM screen showed visual feedback to indicate whether or not the user was given access.

6.2 Use Case B: Gaming

Gaming is one of the most popular applications of VR. In this application, we developed a game in which the user can shoot asteroids

by pursuing them with their eyes (Figure 1B). The target asteroid was colored red while the other objects are colored blue. The centers and radii of the asteroids as well as their sizes, rotations and angular velocities are randomly determined by the software. Visual and auditory feedback is provided for each successful selection. There were ten asteroids in the scene at all times, one of which was set as the target using a different color. While the participants were "shooting" asteroids, they had to walk up to a white square that appeared alternately at two opposing edges of the available space. After the game is explained to the participant, it was started by the experimenter and ran for 180 seconds.

6.3 Results

Analyzing the NASA TLX responses revealed that participants rated the overall workload for the authentication application (47.13 ± 18.45) approximately as high as for the shooting game, in which they had to walk (47.27 ± 17.04).

6.3.1 Authentication. In the ATM use case, a total of 260 PINs were entered. Overall, participants provided 82% of all digits correctly. The mean entry time for one PIN was 21.40 s ($SD = 6.03$) and for one single digit 4.86 s ($SD = 2.11$). This is slightly faster than values reported by Cymek et al. for authentication on situated displays, where users authenticated using their system in 25 seconds [2]. When asked on a 5-point scales, participants indicated that entering PINs via Pursuits in VR was moderately easy ($Mdn = 3$, $SD = 1.93$), moderately accurate ($Mdn = 3$, $SD = 1.28$), but when asked how fast it is (1=fast; 5=slow), they indicated it was slightly slow ($Mdn = 4$, $SD = 1.4$). We followed Schrepp et al.'s approach for analyzing the results of the UEQ [30]. The data was transformed to a seven point scale ranging from -3 to 3, whereas adjectives that are generally positively connoted are projected on the highest (3) and negative adjectives on the lowest (-3) score. Values were then averaged by each of the six categories. The PIN input scenario received high mean ratings for the categories Perspicuity (1.31) and Stimulation (1.03) and the highest rating in terms of Novelty

⁷UEQ - User Experience Questionnaire <http://www.ueq-online.org>

(2.06). However, categories concerning Attractiveness (0.51), Efficiency (0.04) and Dependability (0.46) were rated neutrally. Seven participants reported that the input process takes too long, two of which believed they could forget the PIN that they are entering. Two stated that they would rather use a pattern in which they do not have to search for the respective number first, i.e., a pattern of which they could easily build a mental model of or have already built it (like the numeric pad of an ATM). Three users criticized that objects in the peripheral field of view were difficult to focus on due to blurriness. Furthermore, two participants found it frustrating when many PIN codes were recognized falsely by the software, one of which wished for an undo function. When asked how frequently they would use this scheme, 7 participants indicated they would use it daily, 13 said they would use it once a month due to the long authentication times, while 7 said they would never use it.

A medium positive correlation was found between TLX score and selection time ($r(24) = .388, p = .05$) using a Pearson correlation.

6.3.2 Game. In the game, a total of 349 asteroids were destroyed, of which 261 were intended. On average, every experimentee shot 13.42 asteroids, of which 73.6% ($SD = 0.16$) were correct. A selection took on average 9.91 seconds. Of 26 participants, 2 achieved a 100% correctness. Participants indicated that shooting asteroids via Pursuits in VR was moderately easy ($Mdn = 3, SD = 1.09$), moderately accurate ($Mdn = 3, SD = 0.96$), and slightly slow ($Mdn = 4, SD = 1.04$). Since the sizes of the asteroids were different, we asked participants if they found the perceived easiness is affected by how fast the object moves (1=slow; 5=fast), and how big it was (1=small; 5=big). Participants indicated that speed was not an influential factor ($Mdn = 3, SD = 1.56$), and that bigger objects were easier to select ($Mdn = 4, SD = 1.13$). Mean UEQ ratings show positive tendencies in the categories attractiveness (1.21), perspicuity (1.54), stimulation (1.44) and novelty (1.72) and neutral ratings in terms of efficiency (0.37) and dependability (0.37). Again, it was both expressed by participants and observable in the course of the game that people got more confident over time and did not have to search for the indicator on the floor. Several participants mentioned that it was rather disturbing if an object left the field of view and that it further complicated selection.

7 DISCUSSION

Overall our findings suggest that Pursuits is well suited for VR. Selection times and detection rates when stationary are comparable to previous work [15, 37]. However, selections take more time, and are less accurate when walking. We attribute this to the shaky eye images resulting from the movement; as users walk, their head bobs up and down, which causes the headset to shake, and hence the eye tracking quality deteriorates. Participants feedback also suggests that walking and gazing at moving objects is demanding.

We found that the trajectory size influences selection time and accuracy (see Figures 3 and 4). We also found that this influence is stronger when the user is walking. This can be explained as follows: although Pursuits does not require calibration, it is not completely independent from the eye tracker's accuracy. Namely, if inaccurate gaze points are collected while following a very small trajectory, the gaze points might not represent a smooth trajectory due to noise, and hence the gaze trajectory might not necessarily correlate highly with that of the target. On the other hand, if the trajectory

spans across a large area, the generated gaze trajectory might still be distorted when compared to that of the target, but they would still correlate highly since the noise limitation is less likely to result in, for example, two consecutive gaze points where the second is in the opposite direction of the trajectory's direction.

In contrast, accuracy and selection times for different target sizes did not vary significantly across the conditions. Previous work argued that correlating the eye movements to movements of the target would not be affected by the target size no matter how small it is [34, 37]. In our work, we also found that Pursuits is robust against excessively large targets (22.95° of visual angle). As objects become larger, users could follow them by gazing at random points on its surface, and hence their gaze trajectory would not necessarily match that of the target's center. However, in contrast to our expectations, we found that users indeed gaze at a subset of the pixels as they follow a moving target, which in turn results in robust performance of Pursuits with large targets. However, note that participants of our study were aware that the used interaction techniques relies on them "following" the desired object by gaze. Hence, it is not yet clear if Pursuits selections would still succeed with large objects if the user is not told how the technique works.

Similarly, accuracy and selection times did not vary greatly depending on the distance to the target. Hence, we attribute the lack of differences in performance to the same arguments above: Users in VR seem to follow a specific set of pixels when instructed to follow a moving target.

Overall, users enjoyed the shooting game, and the majority indicated they would use an authentication scheme. However participants also criticized the long selection times. Future work can build on top of this work by introducing a multimodal approach for Pursuits. For example, instead of waiting for the correlation to exceed a certain threshold, the user would click a button on a controller to indicate that she is gazing at the desired object. In case the correlation was too low to deem an object to be gazed at, the system could prompt the user to select from a smaller set of the targets using a different modality.

While in many applications it is important to optimize selection times and accuracy, for other applications, such as VR games, this might not be desired. For example, designers can make games more challenging and hence more engaging by giving more points when selecting objects while walking faster. However, such decisions must be carefully crafted to avoid annoying the user.

7.1 Design Implications

We recommend researchers and practitioners to:

- (1) Use larger trajectory sizes whenever possible, and particularly when the user is expected to use Pursuits while walking. In our study, a trajectory radius of R3 resulted in the shortest selection time (1.87 s) and highest accuracy (79%) among the tested trajectory sizes.
- (2) The distance to target and its size do not heavily influence the performance of Pursuits. However, users need to be informed that they need to "follow" the target to select it. To date, it is not clear if users intuitively gaze at a particular area of a moving target (as was the case in our study), or if the lack of instructions could result in them choosing random points to gaze at.

- (3) Long selection times frustrate users. One way to work around this is to reduce the correlation threshold, and allow users to undo their inputs. Previous work showed that users of Pursuits can tolerate fast inaccurate selections if an undo feature is available [18]. Another approach is to use an additional modality to confirm selection.

8 CONCLUSION AND FUTURE WORK

In this work we investigated the use of Pursuits in virtual environments. We studied its performance with respect to different properties that are specific for VR. We found that using a large trajectory size results in higher accuracy and faster selections. We also found that the target size and its distance to the user do not strongly influence the performance. Based on these findings, we derived several design implications for the use of Pursuits in VR. For future work, we will investigate Pursuits along with other modalities. For example, users could confirm selections using a controller (e.g., HTC Vive Controller) rather than waiting until the correlation exceeds a certain threshold. We plan to study Pursuits for other VR scenarios, such as menu selection in virtual shopping apps.

ACKNOWLEDGMENTS

This work was funded, in part, by the Cluster of Excellence on Multimodal Computing and Interaction (MMCI) at Saarland University, Germany, and by the Bavarian State Ministry of Education, Science and the Arts in the framework of the Center Digitization.Bavaria (ZD.B). This research was supported by the German Research Foundation (DFG), Grant No. AL 1899/2-1.

REFERENCES

- [1] Sean Andrist, Michael Gleicher, and Bilge Mutlu. 2017. Looking Coordinated: Bidirectional Gaze Mechanisms for Collaborative Interaction with Virtual Characters. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2571–2582. <https://doi.org/10.1145/3025453.3026033>
- [2] Dietlind Helene Cymek, Antje Christine Venjakob, Stefan Ruff, Otto Hans-Martin Lutz, Simon Hofmann, and Matthias Roetting. 2014. Entering PIN codes by smooth pursuit eye movements. *Journal of Eye Movement Research* 7(4):1 (2014), 1–11.
- [3] Murtaza Dhuliawala, Juyoung Lee, Junichi Shimizu, Andreas Bulling, Kai Kunze, Thad Starner, and Woontack Woo. 2016. Smooth Eye Movement Interaction Using EOG Glasses. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction (ICMI 2016)*. ACM, New York, NY, USA, 307–311. <https://doi.org/10.1145/2993148.2993181>
- [4] Heiko Drewes, Alexander De Luca, and Albrecht Schmidt. 2007. Eye-gaze Interaction for Mobile Phones. In *Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems and the 1st International Symposium on Computer Human Interaction in Mobile Technology (Mobility '07)*. ACM, New York, NY, USA, 364–371. <https://doi.org/10.1145/1378063.1378122>
- [5] Augusto Esteves, Eduardo Velloso, Andreas Bulling, and Hans Gellersen. 2015. Orbits: Gaze Interaction for Smart Watches Using Smooth Pursuit Eye Movements. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 457–466. <https://doi.org/10.1145/2807442.2807499>
- [6] Augusto Esteves, David Verweij, Liza Suraiya, Rasel Islam, Youryang Lee, and Ian Oakley. 2017. SmoothMoves: Smooth Pursuits Head Movements for Augmented Reality. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software & Technology (UIST '17)*. ACM, New York, NY, USA, 12.
- [7] Helmut Flitter, Thies Pfeiffer, and Gert Rickheit. 2006. Psycholinguistic experiments on spatial relations using stereoscopic presentation. In *Situated Communication*, Gert Rickheit and Ipke Wachsmuth (Eds.). De Gruyter, Chapter 10, 128–153.
- [8] Florian Geiselhart, Michael Rietzler, and Enrico Rukzio. 2016. EyeVR: Low-cost VR Eye-based Interaction. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, New York, NY, USA, 277–280. <https://doi.org/10.1145/2968219.2971384>
- [9] Ceenu Goerge, Mohamed Khamis, Emanuel von Zezschwitz, Marinus Burger, Henri Schmidt, Florian Alt, and Heinrich Hussmann. 2017. Seamless and Secure VR: Adapting and Evaluating Established Authentication Systems for Virtual Reality. In *Proceedings of the Network and Distributed System Security Symposium (NDSS 2017) (USEC '17)*. NDSS. <https://doi.org/10.14722/usc.2017.23028>
- [10] Khalad Hasan, Tovi Grossman, and Pourang Irani. 2011. Comet and Target Ghost: Techniques for Selecting Moving Targets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 839–848. <https://doi.org/10.1145/1978942.1979065>
- [11] Konrad Hinsén. 2000. The molecular modeling toolkit: A new approach to molecular simulations. *Journal of Computational Chemistry* 21, 2 (2000), 79–85. [https://doi.org/10.1002/\(SICI\)1096-987X\(20000130\)21:2<79::AID-JCC1>3.0.CO;2-B](https://doi.org/10.1002/(SICI)1096-987X(20000130)21:2<79::AID-JCC1>3.0.CO;2-B)
- [12] Robert J. K. Jacob. 1991. The Use of Eye Movements in Human-computer Interaction Techniques: What You Look at is What You Get. *ACM Trans. Inf. Syst.* 9, 2 (April 1991), 152–169. <https://doi.org/10.1145/123078.128728>
- [13] Richard J. Jagacinski, Daniel W. Reppeger, Sharon L. Ward, and Martin S. Moran. 1980. A Test of Fitts' Law with Moving Targets. *Human Factors* 22, 2 (1980), 225–233. <https://doi.org/10.1177/001872088002200211> arXiv:<http://dx.doi.org/10.1177/001872088002200211> PMID: 7390506.
- [14] Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In *Adj. Proc. ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*. 1151–1160. <https://doi.org/10.1145/2638728.2641695>
- [15] Mohamed Khamis, Florian Alt, and Andreas Bulling. 2015. A Field Study on Spontaneous Gaze-based Interaction with a Public Display using Pursuits. In *Adj. Proc. of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015)* (2015-09-07), 865–874. <https://doi.org/10.1145/2800835.2804335>
- [16] Mohamed Khamis, Alexander Klimczak, Martin Reiss, Florian Alt, and Andreas Bulling. 2017. EyeScout: Active Eye Tracking for Position and Movement Independent Gaze Interaction with Large Public Displays. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software & Technology (UIST '17)*. ACM, New York, NY, USA, 12. <https://doi.org/10.1145/3126594.3126630>
- [17] Mohamed Khamis, Ozan Saltuk, Alina Hang, Katharina Stolz, Andreas Bulling, and Florian Alt. 2016. TextPursuits: Using Text for Pursuits-Based Interaction and Calibration on Public Displays. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 12. <https://doi.org/10.1145/2971648.2971679>
- [18] Mohamed Khamis, Ludwig Trotter, Markus Tessimann, Christina Dannhart, Andreas Bulling, and Florian Alt. 2016. EyeVote in the Wild: Do Users Bother Correcting System Errors on Public Displays?. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia (MUM '16)*. ACM, New York, NY, USA, 57–62. <https://doi.org/10.1145/3012709.3012743>
- [19] Fu-Shun Li and Sai-Keung Wong. 2016. Animating Agents Based on Radial View in Crowd Simulation. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (VRST '16)*. ACM, New York, NY, USA, 101–109. <https://doi.org/10.1145/2993369.2993373>
- [20] Otto Hans-Martin Lutz, Antje Christine Venjakob, and Stefan Ruff. 2015. SMOOVs: Towards calibration-free text entry by gaze using smooth pursuit movements. *Journal of Eye Movement Research* 8(1):2 (2015), 1–11.
- [21] David Mould and Carl Gutwin. 2004. The Effects of Feedback on Targeting with Multiple Moving Targets. In *Proceedings of Graphics Interface 2004 (GI '04)*. Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 25–32. <http://dl.acm.org/citation.cfm?id=1006058.1006062>
- [22] Sahil Narang, Andrew Best, Tanmay Randhavana, Ari Shapiro, and Dinesh Manocha. 2016. PedVR: Simulating Gaze-based Interactions Between a Real User and Virtual Crowds. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (VRST '16)*. ACM, New York, NY, USA, 91–100. <https://doi.org/10.1145/2993369.2993378>
- [23] Michael Ortega. 2013. Hook: Heuristics for selecting 3D moving objects in dense target environments. In *2013 IEEE Symposium on 3D User Interfaces (3DUI)*. 119–122. <https://doi.org/10.1109/3DUI.2013.6550208>
- [24] Yun Suen Pai, Benjamin I. Outram, Benjamin Tag, Megumi Isogai, Daisuke Ochi, and Kai Kunze. 2017. GazeSphere: Navigating 360-degree-video Environments in VR Using Head Rotation and Eye Gaze. In *ACM SIGGRAPH 2017 Posters (SIGGRAPH '17)*. ACM, New York, NY, USA, Article 23, 2 pages. <https://doi.org/10.1145/3102163.3102183>
- [25] Anjul Patney, Marco Salvi, Joohwan Kim, Anton Kaplanyan, Chris Wyman, Nir Benty, David Luebke, and Aaron Lefohn. 2016. Towards Foveated Rendering for Gaze-tracked Virtual Reality. *ACM Trans. Graph.* 35, 6, Article 179 (Nov. 2016), 12 pages. <https://doi.org/10.1145/2980179.2980246>
- [26] Thies Pfeiffer, Marc Erich Latoschik, and Ipke Wachsmuth. 2008. Evaluation of binocular eye trackers and algorithms for 3D gaze interaction in virtual reality environments. *JVRB - Journal of Virtual Reality and Broadcasting* 5, 16 (2008), 1660.

- [27] Thammathip Piumsomboon, Gun Lee, Robert W. Lindeman, and Mark Billinghurst. 2017. Exploring natural eye-gaze-based interaction for immersive virtual reality. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 36–39. <https://doi.org/10.1109/3DUI.2017.7893315>
- [28] Daniel Pohl, Xucong Zhang, and Andreas Bulling. 2016. Combining Eye Tracking with Optimizations for Lens Astigmatism in Modern Wide-Angle HMDs. In *Proc. of the IEEE Conference on Virtual Reality (VR)* (2016-03-19). https://perceptual.mpi-inf.mpg.de/files/2016/01/Pohl16_VR.pdf
- [29] Sharif Razzaque. 2001. Redirected Walking. In *Proceedings of Eurographics*, pp 289–294. ACM Press.
- [30] Martin Schrepp, Andreas Hinderks, and Joëlrg Thomaschewski. 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence* 4, 6 (12/2017 2017), 103–108. <https://doi.org/10.9781/ijimai.2017.09.001>
- [31] Sophie Stellmach and Raimund Dachselt. 2012. Designing Gaze-based User Interfaces for Steering in Virtual Environments. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*. ACM, New York, NY, USA, 131–138. <https://doi.org/10.1145/2168556.2168577>
- [32] Sam Tregillus and Eelke Folmer. 2016. VR-STEP: Walking-in-Place Using Inertial Sensing for Hands Free Navigation in Mobile VR Environments. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1250–1255. <https://doi.org/10.1145/2858036.2858084>
- [33] Subarna Tripathi and Brian Guenter. 2017. A Statistical Approach to Continuous Self-Calibrating Eye Gaze Tracking for Head-Mounted Virtual Reality Systems. In *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 862–870. <https://doi.org/10.1109/WACV.2017.101>
- [34] Eduardo Velloso, Marcus Carter, Joshua Newn, Augusto Esteves, Christopher Clarke, and Hans Gellersen. 2017. Motion Correlation: Selecting Objects by Matching Their Movement. *ACM Trans. Comput.-Hum. Interact.* 24, 3, Article 22 (April 2017), 35 pages. <https://doi.org/10.1145/3064937>
- [35] Eduardo Velloso, Markus Wirth, Christian Weichel, Augusto Esteves, and Hans Gellersen. 2016. AmbiGaze: Direct Control of Ambient Devices by Gaze. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, New York, NY, USA, 812–817. <https://doi.org/10.1145/2901790.2901867>
- [36] Mélodie Vidal, Remi Bismuth, Andreas Bulling, and Hans Gellersen. 2015. The Royal Corgi: Exploring Social Gaze Interaction for Immersive Gameplay. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 115–124. <https://doi.org/10.1145/2702123.2702163>
- [37] Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2013. Pursuits: Spontaneous Interaction with Displays Based on Smooth Pursuit Eye Movement and Moving Targets. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. ACM, New York, NY, USA, 439–448. <https://doi.org/10.1145/2493432.2493477>
- [38] Bartholomäus Wissmath, Daniel Stricker, David Weibel, Eva Siegenthaler, and Fred W. Mast. 2010. The Illusion of Being Located in Dynamic Virtual Environments. Can Eye Movement Parameters Predict Spatial Presence? *Journal of Eye Movement Research* 3, 5 (2010). <https://bop.unibe.ch/index.php/JEMR/article/view/2307>