SmoothMoves: Smooth Pursuits Head Movements for Augmented Reality

Augusto Esteves[1], David Verweij[1,2], Liza Suraiya[3], Rasel Islam[3], Youryang Lee[3], Ian Oakley[3]

[1]Centre for Interaction Design, Edinburgh Napier University, Edinburgh, United Kingdom
[2]Department of Industrial Design, Eindhoven University of Technology, Eindhoven, Netherlands
[3]Human and Systems Engineering, Ulsan National Institute of Science and Technology, Ulsan, Korea
{a.esteves, d.verweij}@napier.ac.uk, {liza, islam, yrlee}@unist.ac.kr, ian.r.oakley@gmail.com

ABSTRACT

SmoothMoves is an interaction technique for augmented reality (AR) based on smooth pursuits head movements. It works by computing correlations between the movements of on-screen targets and the user's head while tracking those targets. The paper presents three studies. The first suggests that head based input can act as an easier and more affordable surrogate for eye-based input in many smooth pursuits interface designs. A follow-up study grounds the technique in the domain of augmented reality, and captures the error rates and acquisition times on different types of AR devices: head-mounted (2.6%, 1965ms) and hand-held (4.9%, 2089ms). Finally, the paper presents an interactive lighting system prototype that demonstrates the benefits of using smooth pursuits head movements in interaction with AR interfaces. A final qualitative study reports on positive feedback regarding the technique's suitability for this scenario. Together, these results show SmoothMoves is viable, efficient and immediately available for a wide range of wearable devices that feature embedded motion sensing.

Author Keywords

Wearable computing; eye tracking; augmented reality; AR; input technique; smooth pursuits; motion matching; HMD.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Augmented Reality (AR) glasses are a rapidly maturing technology. The latest products, such as Microsoft HoloLens [22], include powerful computers, high resolution displays and sophisticated tracking. While these technical achievements are impressive, there is less clarity about the best ways for users to interact with AR contents and interfaces. There is an active community exploring viable
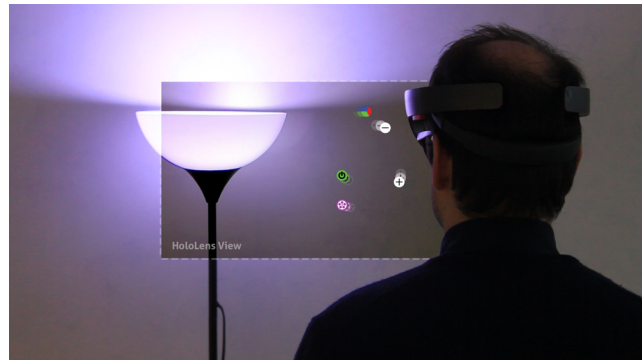
Figure 1. An interactive lighting system prototype that uses AR for displaying moving controls in space. Users make selections by tracking these movements with their heads.

modalities for head-mounted displays (HMDs) including on-headset touch [36], mid-air hand input [23] and the use of dedicated wearable peripherals such as gloves [12] or belts [8]. Within this space, we argue that input from movements of the eyes [35] and head [3] are particularly practical and appealing: in such scenarios, hands remain free and all sensing can be integrated into the headset.

Traditional approaches to head based input focus on pointing by either tracking gaze location [31] or via ray-casting techniques that infer an object of interest from the orientation of the head [25]. While the simplicity of these approaches is laudable, problems remain. Although they readily enable a user to hover over a specific icon or region, they also both require a discreet, explicit confirmation mechanism to trigger a selection. Common approaches such as dwell add a fixed time cost and decrease accuracy [13]. Alternatives such as hand gestures (as in the Microsoft HoloLens) require additional sensing equipment. Furthermore, while gaze tracking solutions exist for mobile settings, well reported challenges in accurate tracking and calibration in real world scenarios [10] makes gaze based target selection techniques practically infeasible.

To mitigate these problems, authors have proposed gaze input systems based on *smooth pursuits* [1,11,18,32] – distinctive, continuous, low latency adjustments to gaze that are naturally produced when (and only when) visually tracking a moving object. Smooth pursuits systems operate by showing a user a set of moving targets whilst tracking gaze. Statistical matching between the gaze and target

trajectories is used to infer which target a user is attending to. The technique has been shown to be useful in tasks as diverse as calibrating eye tracking systems [26] and creating novel gaze input techniques for devices large (e.g. public displays [35]) and small (e.g. smart watches [9]).

While current accounts of smooth pursuits input show its potential, we argue that key aspects of the behavior remain unstudied. In particular, we note that fundamental literature on visual tracking indicates that it involves a synergistic combination of head and eye movement [18]. Accordingly, we argue that it may be possible to reliably perform explicit smooth pursuits style tracking movements with the head instead of the eye - this extends Dhuliawala *et al.*'s [7] recent proposal that explores complementary movements of the head and eye. Using head motions accords considerable practical benefits, primarily that the Inertial Motion Units (IMUs) needed to accurately track head movements are small, cheap, low power and already integrated into the majority of AR glasses and other wearables.

In order to explore the potential of this idea, this paper contributes *SmoothMoves,* an input technique that relies on data from a head-mounted IMU to enable users to select moving targets by continuously matching the target position with the orientation of their head. To explore the viability and value of this idea, we also contribute three studies. First, a fundamental study (using a PC monitor) compares performance with IMU based head tracking against the more established baseline of gaze tracking in situations where only a single target is shown. We report strong similarities across a range of target movement conditions. Second, we compare the performance of SmoothMoves in both handheld and HMD based AR systems in situations where multiple targets are presented. Building on these results, the final sections of this paper apply SmoothMoves input to an HMD used in a smart home scenario and report on results from a qualitative user study. Together this work represents a comprehensive exploration of the potential, feasibility, reliability and experience of head motion based smooth pursuits as an input modality for augmented reality.

RELATED WORK

Gaze is the inseparable product of head movements plus eye movements. The relationship between these activities is sophisticated. At the most fundamental level, the Vestibulo-Ocular Reflex (VOR) [19] continuously stabilizes gaze by adjusting (basically inverting) eye position in response to changes in head position sensed by the vestibular system. It is key to providing a stable visual experience of objects. In contrast, during smooth pursuits tracking of rapidly moving objects [29], the head and eye move together [18] to keep an object optimally in view. Smooth pursuits movements also involve two distinct stages. Initially, the eyes and head are accelerated to align with the moving stimuli, an open-loop process that can take up to 300-500ms [27]. Subsequent closed-loop tracking closely matches the target, particularly in situations where velocities are stable.

A number of properties make smooth pursuits movements useful as an input technique. First, they are innate. Users know how to visually track targets and can generate this kind of motion without training. Second, they are distinctive. Users are only able to generate smooth pursuits eye movements in the presence of visually moving targets. Third, they operate on movement not position. As such, they are relatively immune to changes in target size [9] and robust to tracking errors – capturing changes in gaze is much simpler than accurately determining what a user is looking at. Fourth, they are operated hands-free. And fifth, they do not require users to memorize gestures. Several systems have been recently introduced to leverage these properties. Vidal *et al.* [35] used smooth pursuits to enable quick, spontaneous interaction with public displays, while Lutz *et al.*'s [20] applied the technique to text entry on public dashboards. Cymek *et al.* [6] and Khamis *et al.* [16] explored how smooth pursuits input can create safer PIN entry systems, and Esteves *et al.* [9] and Kangas *et al.* [14] relied on the scale-independent, calibration-free nature of smooth pursuits gaze input to deliver hands-free interaction on, respectively, smart watches and glasses. Finally, Dhuliawala *et al.* [7] show that alternative eye gaze sensing modalities, such as EOG, also have the potential to support smooth pursuits input. This work demonstrates that the technique is sufficiently powerful and flexible to be deployed in a wide range of input scenarios.

However, these systems rely on smooth pursuits eye movements. We identify an opportunity to study the viability of using IMU-derived head movements to achieve the same objectives. This approach would convey a number of advantages. First and foremost is cost: wearable eye tracking remains expensive (computer vision: ~1500 USD [15]; EOG: ~1500 USD [37]) whereas head tracking can be achieved with an IMU costing no more than ten USD. The second is form-factor: eye trackers require cameras or electrodes mounted at specific locations on the user's face, with the former also requiring a clear line of sight to the eyes. In contrast, IMUs can be mounted anywhere on the head. Furthermore, IMU's are small and light enough (<10mm square, <1 gram) to be integrated into almost any wearable item: headphones, eyewear, jewelry, clothes and, indeed, existing smart glasses (e.g. Microsoft HoloLens). Optical systems are also susceptible to changing light conditions, such as those that occur outdoors, while IMU's are relatively unaffected by environmental factors.

These beneficial properties have not gone unremarked. Indeed, a range of techniques for input based on head movements has been proposed and studied. Ray-based pointing, in which users interact by projecting a ray from their head to intersect with a target of interest, is the most common [4] and has been integrated into current head-mounted displays, such as the Google Cardboard [30] and the Microsoft HoloLens. Other authors have proposed the use of head tracking in mobile contexts to provide gestural input in the form of head tilting [5] and nodding [24].

Furthermore, studies on smart TVs have explored the use of off-the-shelf webcams to capture head motion during smooth pursuits [3]. Finally, while rigorous studies are presently lacking, recent work has proposed achieving head-based input during pursuits tracking by monitoring VOR movements [7]. In sum, while this work highlights the appeal of head-based input, to the best of our knowledge, no prior studies have explored explicit head movements for target tracking input in AR.

SMOOTHMOVES

SmoothMoves is an interaction technique for selecting graphical targets in AR interfaces. The targets move in orbital trajectories and users make selections by matching these motions with movements of their head that are sensed by a worn IMU. SmoothMoves is heavily influenced by prior pursuits based gaze interaction techniques [35], but replaces the use of eye coordinates with *yaw* and *pitch* data from the IMU. The matching process is simple: for each displayed target, Pearson's correlations are computed for *xtarget-yaw* and *ytarget-pitch* relationships. If both exceed a certain *correlation-threshold* for a given target, and no other currently displayed targets attain the same result (either individually or via an average of both results), then the target is selected. The correlation takes place after *start-up time* and on a particular data rolling *window size.* The start-up time is the period immediately after the appearance of a set of SmoothMoves targets when the user is engaged in open-loop orienting behavior that marks the beginning of a smooth pursuit movement. Performing target matching in this period would not be meaningful. The window-size specifies the duration of data sampled for SmoothMoves correlations. In the eye gaze literature, longer window sizes ensure fewer erroneous selections at the cost of the lower comfort and higher performance time [9,35].

Visually, SmoothMoves closely mimics Orbits [9]. Each graphical control is comprised of a trajectory around a center point and a target (see Figure 4) that continuously traverses this trajectory. Each control can be used for either discrete input, where target acquisitions result in issuing a command, or continuous control by monitoring the time a target is tracked for. Target disambiguation is achieved in two ways. First, targets move in different *phases*. For example, with four targets, they would be spaced at 90° intervals. Second, targets can move in different *directions*: clockwise and counterclockwise.

STUDY 1: EYE AND HEAD-TRACKING

To explore the viability of SmoothMoves, we first conducted a lab study. It had three goals. First, to validate the idea that users can acquire targets using smooth pursuits head motions. To do so, we simultaneously captured eye- and head-tracking data of participants following a series of single moving targets with different instructions: to perform the tracking naturally; to track only with the eyes and; to track only with the head. This supports contrasting head and eye motion performance. Second, to explore performance variations in eye and head tracking with a variety of moving stimuli. The goal was to enable us to make recommendations about optimal stimuli to display. Finally, the third goal was to define optimal values for the key parameters of correlation threshold, start-up time and window size, to enable construction of a working system.

Participants

18 participants were recruited (12F), aged between 20 and 26 years ($M = 24$, $SD = 1.85$). All participants were undergraduate or graduate students at a local institution, and except for one, had minimal experience with eye-tracking. All had normal or corrected to normal vision. Nine participants wore contact lenses, one wore glasses, and the remaining eight did not require any visual aids.

Experimental Setup and Design

The experiment was conducted in a quiet and private laboratory space, with participants sitting 60cm away from a 27" display (1920×1080 resolution screen). Eye data was recorded using a Pupil Pro [15] wearable eye-tracker equipped with a single camera tracking the right-eye (reported mean gaze estimation accuracy of 0.6° of visual angle). The tracker was adjusted for focus and to ensure a clear field of view of the eye and a close match between the horizontal and vertical axes of the eye and the camera. No further calibration was performed; only normalized pupil locations were recorded. A GY-86 nine axis IMU was attached to the front camera mount of the pupil using a 3D printed fixture and wired to an Arduino. A complementary filter (Mahony *et al*. [21]) tracked head orientation and provided yaw and pitch data. The display and both sensors were all connected to the same computer. The display update and IMU data logging rate were 60Hz. Difficulties in capturing a reliably timed data stream from the eye tracker resulted in recording eye packets at a target rate of 90Hz, and an actual rate of between 75HZ and 90Hz.

All participants completed the same set of trials in three different input conditions: *natural*, *eyes*, and *head*. In all conditions a single moving target was displayed for four seconds and trials were presented in a random order. In the natural condition, participants were simply asked to follow the target. In the eyes condition, participants were asked to follow the target with their eyes. Similarly, in the head condition, participants were asked to follow the target with their head. All participants completed the natural condition first, to ensure there was no instructional bias in the way they opted to follow the moving target. The eyes and head conditions were counter-balanced to reduce possible fatigue and practice effects. The set of moving targets used in the study was selected to replicate previous studies of smooth pursuits eye movements [9,14]. Variations included:
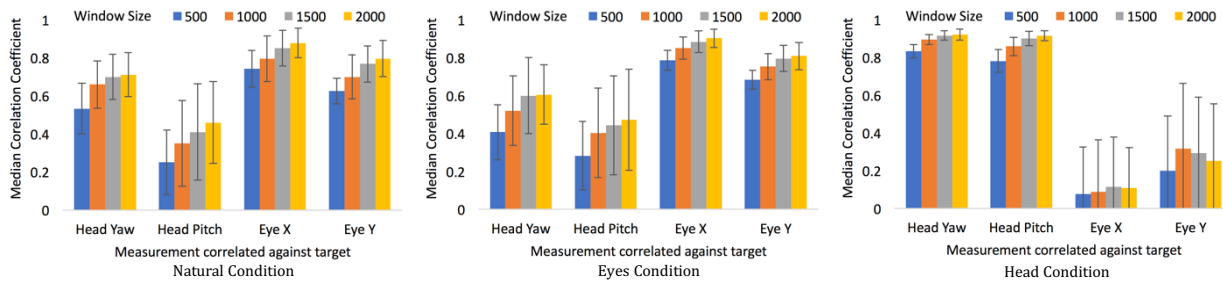
Figure 2. Absolute median correlations coefficients for head and eye movements in Natural, Eye and Head conditions. Figure omits data from a start-up time of 500ms. Bars show Median Absolute Deviation.

- **Trajectory size**: there were three on-screen sizes: 4cm (~3.50° visual angle), 13cm (~11.75°) and 22cm (~20°).

- **Target speed**: targets moved in one of three angular velocities: 60°/sec, 120°/sec, or 180°/sec.

Additional novel variations were included in the study, so as to expand the design knowledge about interfaces based on smooth pursuits. These included:

- **Trajectory shape**: targets moved in either circular or rhomboidal trajectories (see Figure 4).

- **Trajectory visibility**: target trajectories were either invisible, where only the target was displayed, or visible, where the target's movement path was also shown.

- **Speed type**: targets could move with constant speeds, or increase their speed midway through the trial. Speed adjustments always involved an increase by 60°/sec.

- **Direction type**: as with speed type, targets could either move in a fixed orbital direction, or invert this halfway through the trial.

Each possible trial combination occurred once in each condition. Consequently, data from a total of 7776 trials (18 participants x 3 conditions x 3 sizes × 3 speeds × 2 trajectories × 2 visibilities × 2 speed types × 2 direction types) was recorded.

Data Pre-Processing

Prior to analysis, the separate data streams of eye, head and visual target movements were pre-processed. First, the eye-data was down-sampled to 60Hz and the three data streams were matched using timestamps. Second, eye data trials were removed in situations where there were breaks in the data of greater than 300ms, a threshold derived from typical blink durations. The goal was to include trials involving natural behavior such as blinks but exclude those trials where eye tracking was lost or degraded (as judged by the confidence statistic reported by the tracker) for reasons such as a prolonged closure of the eye, a glance away from the screen or a failure of the tracking algorithms. We opted for removing these trials as long lapses in the data would disrupt the planned rolling window correlation analysis. In total, we excluded 93 trials (1.2%). Of these, 71% were in the head condition, likely a consequence of the larger movements disrupting eye tracking. Furthermore, they were biased by participant (33% from one subject) due to

variations in the robustness of the eye tracker fit/calibration. They were evenly distributed over all other variables and are not sufficient in number, or skewed enough in distribution, to invalidate our analysis. The final stage of pre-processing involved running a rolling average filter over eye, head and target data streams (ignoring gaps in the eye data) with a window size of 64ms, or 4 samples. This smoothed out inevitable fluctuations in sampling times associated with data capture from three separate sources.

Results and Analysis

Initial analysis of the results focused on determining an appropriate configuration of SmoothMoves parameters. We adopted a 500ms startup-time, based on fundamental literature [27] indicating that initial motions in a tracking movement involve orientating actions that differ from later tracking motions. Using this figure, we ran correlations between all eye and head data in the three experimental conditions using window-sizes of 500ms, 1000ms, 1500ms and 2000ms – see Figure 2. Prior work has identified 1000ms as sufficient to achieve correlation results of 0.8 with gaze and suggested this is a viable correlation threshold for input [9]. With these baseline parameters, results from the natural condition show slightly diminished performance: a median of 0.75. We attribute this to the large range of stimulus display parameters used in the study and discussed in the next paragraph. Performance in the eyes condition matches the 0.8 recorded in prior work. In both these conditions, we note that correlations against the head data are low (0.46-0.5) and insufficient to support recognition via the algorithmic matching process proposed in this paper. It also indicates that participants more naturally followed targets with their eyes than their head, an effect which may be partly due to participants being aware of the eye-tracking equipment during the study setup. Data from the head condition, however, strongly shows that head based tracking can be readily achieved by participants; head correlation coefficients were higher than those reported for the eyes in any of the experimental conditions. Specifically, with the 1000ms window-size, participants achieved a median correlation between head and target movements of 0.88. This provides a firm basic validation of the SmoothMoves concept. Reflecting these results, we used a 1000ms window-size and 0.8 correlation threshold for all further analysis and activities in this paper.

Trajectory Size	Eyes	Head
4cm (~3.50°)	0.69 (0.13) *	0.78 (0.1) *
13cm (~11.75°)	074 (0.12)	0.84 (0.06)
22cm (~20°)	0.73 (0.11)	0.85 (0.05)
Target Speed	**Eyes**	**Head**
60°/sec	0.69 (0.13)	0.81 (0.07)
120°/sec	0.72 (0.12)	0.84 (0.06)
180°/sec	0.74 (0.10)	0.82 (0.1)
Trajectory Shape	**Eyes**	**Head**
Circle	0.70 (0.11) *	0.82 (0.06) *
Rhombus	0.74 (0.12) *	0.83 (0.07) *
Trajectory Vis.	**Eyes**	**Head**
Visible	0.70 (0.13)	0.83 (0.07)
Invisible	0.74 (0.10)	0.82 (0.07)
Speed Type	**Eyes**	**Head**
Constant	0.71 (0.12)	0.83 (0.07)
Varies	0.73 (0.11)	0.82 (0.07)
Direction Type	**Eyes**	**Head**
Constant	0.74 (0.11) *	0.84 (0.07) *
Varies	0.70 (0.12) *	0.81 (0.07) *

Table 1. Mean absolute Pearson correlations between eyes & target (eyes condition) and head & target (head condition) for the six study variables. Standard deviation in brackets. Asterisks indicate significant main effects of the trajectory variable at p<0.008.

A key goal of this paper is to characterize the performance of eye and head tracking movements with different trajectory designs. Rather than a high-dimensionality ANOVA, we opted to do this by analyzing each trajectory variable/modality pair individually with a low alpha threshold for significance. Specifically, we examined correlations from eye and head movements in, respectively, the eye and head conditions using six separate two-way repeated measures ANOVAs (either 3x2 or 2x2). For variables with three levels, the ANOVAs incorporated Greenhouse-Geisser corrections when Mauchly's test showed sphericity violations and were followed by Bonferroni-corrected *post-hoc* t-tests. In total, we ran six separate main tests using an alpha threshold of p<0.05/6, or p<0.008. Effect sizes are given as partial eta squared (η_p^2). In the interests of brevity, we report only significant results.

The raw data for each variable in the eye and head conditions are shown in Table 1. The head data (from the head condition) led to significantly higher correlation values than the eye data (from the eye condition) in all tests: (F (1, 17) = 15.7, p <0.001, η_p^2 = 0.481). This supports the idea that head condition led to improved tracking accuracy compared to the eye condition. Beyond this, as the raw figures show, the results were relatively uniform. Results varied in terms of *direction type* (F (1, 17)

= 35.747, p <0.001, η_p^2 = 0.678). This suggests that changes in target direction disrupted participant's ability to track accurately. Similarly, the data differed significant with *trajectory shape* (F (1, 17) = 20.321, p <0.001, η_p^2 = 0.544), indicating that participants tracked targets moving in rhomboidal trajectories more accurately. Finally, significant differences emerged with variations in trajectory size (F (1.241, 21.091) = 14.259, p <0.001, η_p^2 = 0.456). *Post-hoc* t-tests indicated tracking the smallest targets was more challenging that tracking those in medium (p=0.002) or large (0.004) conditions. Interactions were also observed in *trajectory visibility* (F (1, 17) = 15.052, p =0.001, η_p^2 = 0.47) and *speed type* (F (1, 17) = 11.476, p =0.004, η_p^2 = 0.403). These results suggest that tracking with the eyes modestly improves when targets move more unpredictably, an effect that is not present with head movements. This is possibly due to the eyes' faster response time.

Discussion
The study strongly confirms the idea that head motions can accurately track moving targets. In the head condition, the fidelity of the behavior, as expressed by the median correlation coefficients, exceeded that of the eyes in both the natural and eyes conditions of the current study as well as that reported in prior work [9]. This suggests that head based input can act as a surrogate for eye-based input in many smooth pursuits input scenarios; it may even be preferred in terms of performance. However, data from the natural condition also clearly indicates that participant's predilection was to track with the eyes; only when specifically instructed did they use clear, accurate and distinctive head movements.

A second goal of the study was to expand knowledge about what stimulus parameters are effective in tracking based input systems. Although a number of significant differences emerged, serving to isolate more and less effective designs, the primary message from this data is one of the robustness of the technique to variations in target movements. This is a positive outcome as it suggests that both eye and head versions of the technique can be deployed with targets moving in a broad range of patterns and thus support a large variety of graphical forms and interface designs. Specific recommendations from the study are to avoid direction changes and small target trajectories. Rhomboidal trajectories may provide some benefits. While these recommendations are sensible, we note the small absolute differences and moderate effect sizes – they may ultimately have limited impact on performance.

Beyond these analyses and recommendations, it is also worth describing the movements captured in the study. For this, we focus on data in the head condition, as this involves explicit bodily motion and represents the core idea proposed in this paper. The scale of these movements will impact a range of factors such as the obtrusiveness [39], social acceptability [28] and, possibly, long term comfort of the technique. While a full exploration of these issues goes

beyond the scope of this article, we can present and interpret basic data. The small (3.5°), medium (11.75°) and large (20°) target trajectories led to mean head rotations of 9.19° (SD 6.18), 14.85° (SD 9.35) and 17.65° (SD 9.69) and showed minimal variation (<1 degree) between yaw and pitch. This indicates participants exaggerated head movements for small targets and modestly reduced them for larger targets (see Figure 3 for examples). The movements could also be relatively subtle – for the smallest targets, median head rotations were just 6.7°. We believe these movements are sufficiently small to ensure the technique is discrete and not unduly fatiguing. Further studies will need to empirically examine these claims and formally establish how fatiguing SmoothMoves interaction is. We also note that stimuli in the current study were very simple and future work should investigate more complex situations where, for example, users would need to engage in a visual search for targets prior to performing selection.

SMOOTHMOVES VALIDATION STUDY

We opted to build on these results by validating SmoothMoves input for AR in a follow-up study deploying optimal cues in a more realistic AR setup.

Participants

A total of 16 participants completed the study (9F), aged between 21 and 26 ($M = 22.19$, $SD = 1.84$). All participants were students at a local institution and were compensated approximately ten USD for their time. In general, they rated their experience with smartphones as very high ($M = 5/5$) but their experience of wearables such as smart watches (1.8/5) and smart glasses (1.2/5) as low. Three participants were smart watch owners, resulting in the modestly higher rating for these devices.

Experimental Setup and Design

The study involved two device conditions, intended to simulate different AR viewing scenarios. These were *glasses* and *phone*. The glasses condition used the Epson Moverio AR glasses [38] which feature a pair of semi-transparent displays with a 23" field of view. In the phone condition, targets were displayed on a mobile phone (a Huawei Nexus 6P with a 5.7" display) held comfortably in participants' hands. This simulates a common current AR experience in which standard handheld devices are used as the main display device in a video-see through paradigm [33]. In both cases, participants wore the same head mounted IMU used in the first study.

The study also explored two further conditions: *trajectory-size* and *target-cardinality*, or the number of simultaneously presented targets. We re-examined the former variable as it was shown to impact performance in the first study. Furthermore, perceptual trajectory sizes in the two display devices differ substantially from each other and from those used in the first study. This reflects a more realistic deployment of SmoothMoves targets in which it is not possible to fully standardize trajectory sizes across different devices and platforms. We again selected three trajectory
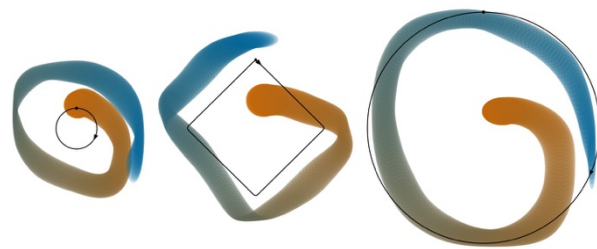

Figure 3. Normalized mean head yaw/pitch changes during three example trials showing 3.5°circular trajectory (left), 11.75° rhomboidal trajectory (center) and 20° circular trajectory (right). All examples used a target speed of 120°/second, included visible paths and did not involve speed or direction changes. Data shown to scale in colored lines with position as yaw/pitch angle and width as standard dev. Temporal progression in the trial shown from orange to blue. Black lines show target trajectory.

sizes but did so based on the available screen size of the devices (rather than visual angle). The sizes were selected so the rhomboidal target paths occupied approximately 18%, 54% and 90% of the smaller screen dimension. In the large condition this left sufficient space to display the moving target, while in the small conditions overlap of the targets remained minimal. We also examined cardinality as this is an essential practical issue for any target selection system. We displayed targets in equidistantly spaced groups of two, four, six and eight (see Figure 4) in order to determine the impact this exerts on performance.

The study was arranged so that phone and glasses were repeated and balanced: all participants completed both conditions, half in each possible order. Within each device condition, participants completed three blocks of trials. Each block contained four target-cardinality trials for each target-size. Trials in each block were randomly presented and the first block was treated as practice and discarded. As such, we retained data from 3072 trials (16 participants x 2 devices x 2 blocks x 4 target-cardinalities × 3 target-sizes x 4 repetitions). For each trial, we logged error count and successful target selection time. Errors occurred if no target selection took place within 10 seconds (a timeout) or a wrong target was selected. In these cases, trials were re-entered into the pool of remaining trials. In this way, all participants correctly completed their allotted set of trials.

Beyond these variables, the stimuli used parameters from the first study. Targets moved at 120°/sec; their trajectories were continuously presented; there were no speed or direction changes. Three other display variables were equally distributed among each set of four cardinality/size trials. These were target direction (clockwise/anti-clockwise), trajectory shape (circle/rhombus) and target starting angle (four cardinal directions). Rather than as experimental variables, these variations increased the realism of the study – trajectories in real systems will likely vary in path (or appearance) and the study examined performance in this relatively unpredictable situation.
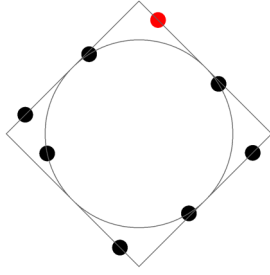
Figure 4. Screenshot from a trial in the second study, where up to eight targets were displayed in tandem. Participants were instructed to follow the target in red.

Procedure

This study implemented SmoothMoves using parameters from the first study. Participants sat at a desk holding the phone in their right hands, or wearing the AR glasses. They started each trial by tapping a key on a PC keyboard on the desk. A set of targets was then displayed, but no data was collected for 500ms of start-up time. Correlations were analyzed with a window-size of 1000ms and a selection triggered when participants reached a correlation-threshold of 0.8. In cases where the standard deviation of head movements in either axis was less than 2°, no correlations were calculated. This threshold was substantially under the mean standard deviation observed in head condition trials in the first study (6°-9° over the size conditions) and served to reduce false positives by capturing only intentional movements. Finally, if multiple targets led to correlations above the target threshold, no selection was returned.

Results and Analysis

Time and error data from the study were analyzed with a pair of three-way repeated measures ANOVAs on device, trajectory size and target cardinality. In cases where sphericity was violated, we report Greenhouse-Geisser-corrected degrees of freedom. Post-hoc pairwise comparisons include Bonferroni CI adjustments. For brevity, only results significant at $p<0.05$ are reported. We use partial eta squared (η_p^2) to express effect size.

The time data is charted in Figure 5. Only the main effect of trajectory size attained significance (F (1.07, 16.056) = 10.901, p=0.004, η_p^2=0.421), so no interactions are included in the chart. The effect size is moderate and borne out by *post-hoc* t-tests showing the smallest trajectories led to slower selections than the medium (p=0.009) and large (p=0.016) trajectories. This indicates that participants took longer to select targets moving around the smallest paths.

Error data was more diverse. Two way interactions are plotted in Figure 6 and main effects in Figure 7. The three-way interaction did lead to a significant result (F (2.162, 6.692) = 4.423, p = 0.018, η_p^2=0.228), but we opt to interpret the data in terms of the more comprehensible significant two-way interactions and main effects, as these all exhibit larger effect sizes. Specifically, the significant two way interactions were between trajectory size and
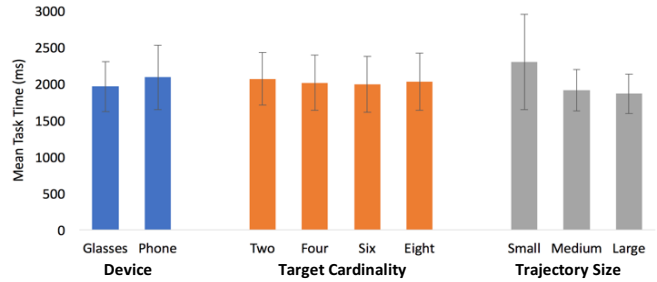

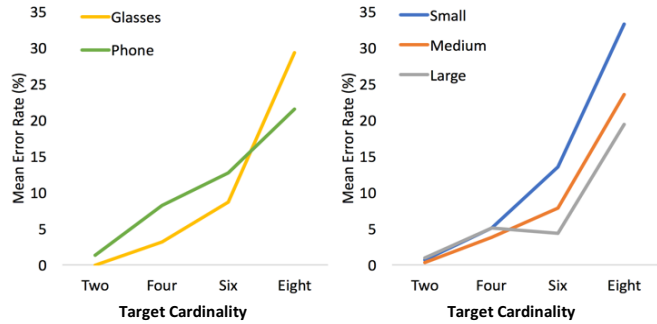**Figure 5. Main effects in validation study time data. Bars show standard deviation.**

Figure 6. Significant two-way interaction effects in error data from validation study between device and target cardinality and trajectory size and target cardinality.
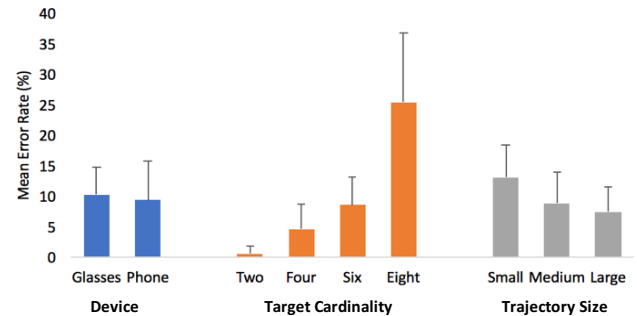

Figure 7. Main effects in validation study error data. Bars show standard deviation.

target cardinality (F (1.974, 29.607) = 6.488, p<0.05, η_p^2=0.302) and trajectory size and device (F (1.164, 17.456) = 5.082, p = 0.033, η_p^2=0.253). Looking at the charts, the first interaction suggests that while errors increase with more targets, they do so more steeply with small trajectory sizes. The second interaction indicates that performance with the glasses was superior to the phone with six or less targets, but this relationship was inverted with eight targets.

The significant main effects were trajectory size (F (1.408, 21.113) = 12.272, p = 0.001, η_p^2=0.45) and target cardinality (F = (1.204, 18.062) = 40.167, p<0.001, η_p^2=0.728). These are the largest effect sizes in the study, and relate to simple outcomes. Specifically, *post-hoc* t-tests showed that small trajectories led to more errors than medium (p=0.025) and large (p=0.001) trajectories and that all differences in cardinality were significant (at p<0.01 or less) except for a non-significant comparison between four

and six targets. Unsurprisingly, this indicates that target selection became more error-prone when more targets were displayed. These results also confirm that participants find targets moving on small trajectories more difficult to track.

DISCUSSION

The goal of this study was to explore the performance of SmoothMoves head tracking in an AR scenario in order to contrast performance with related techniques and make recommendations on how to best deploy it. Time data are simple. Mean task selection time was approximately two seconds and the only significant variation was an increase when the smallest trajectories were used: small on-display target paths should be avoided. This figure includes the 500ms start-up time and a 1000ms window-size, making it approximately half a second greater than the minimum time that the study supported. We argue this is fast enough to make the technique compelling in hands-free AR scenarios: recent studies of hand and head mediated ray based selection report task times of between 2.25 and 3.5 seconds for making a selection from a set of 16 targets [25]; and techniques based on smooth pursuits eye movements report task times ranging from 4.3 to 4.6 seconds [34]. Performance with more traditional, albeit 3D, direct selection techniques based on moving the hand to a target location within arm's reach show broadly similar results: Özacar et al. [25] examine this modality with three types of selection trigger (a physical button, dwell, and a hand gesture) and report task times of three to four seconds.

The error rate data paint a more complex picture. These range considerably, and the more extreme conditions studied are sufficiently challenging to render them inappropriate for use in a real system. If small trajectories are avoided, we argue the data supports the display of up to six targets simultaneously: this led to a mean error rate of 2.6% with the glasses and 4.9% with the phone. The difference between these devices is possibly due to the larger perceptual sizes for the trajectories shown on the HMD, suggesting the technique is better suited to large field of view glasses-based AR than to the perceptually smaller displays of handhelds. It is also worth noting that the experience of interacting with SmoothMoves between the two types of device is also very different. With the HMD, the screen moves with the head; with the phone, it is likely static. We note that the study results indicate that the technique is robust to this difference. It is also worth contrasting the error rates for our recommended SmoothMoves configuration with comparable selection techniques. In terms of ray pointing, Özacar et al. [25] report error rates of 6%-10%; while Esteves et al. [9] report errors in an optimally configured gaze based pursuits input system to be an average of 19% for eight targets displayed in tandem. Özacar et al.'s [25] error data from direct selection tasks ranges from 4% to ~8%. The error rates from this study suggest SmoothMoves performs well enough to act as a viable companion or alterative to these approaches.
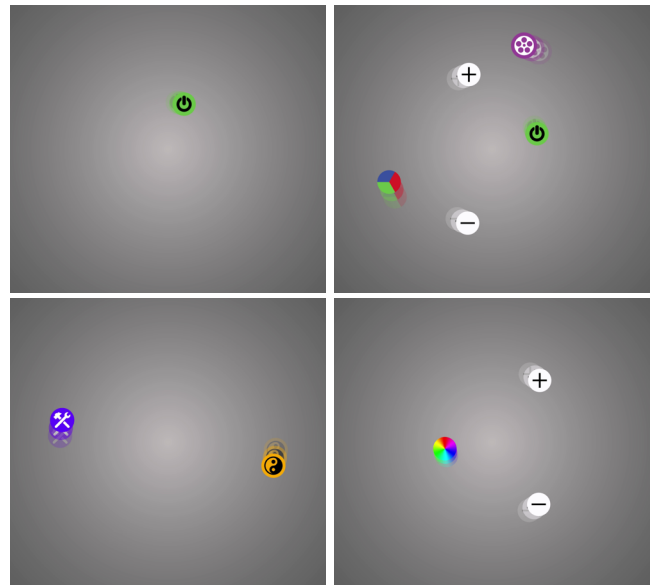

Figure 8. An AR interface built with SmoothMoves for an interactive lighting system. The moving controls are displayed in proximity to the light bulb they control, and users interact with these by tracking their movement with their heads.

In summary, the results of this study confirm that SmoothMoves targeting works well in two different AR scenarios and, in fact, may be particularly suitable for HMDs. This is useful as such systems already incorporate the required sensors to support the technique. On HMDs, and with target sets of between two and six in size, users can reliably (error rate of 2.6%) make selections in under two seconds, a level of performance that we believe is sufficient to support a rich range of possible interactions. The next section of this paper showcases these possibilities.

INTERACTIVE LIGHTS USING AR AND SMOOTHMOVES

This paper concludes with the design and evaluation of a prototype interactive lighting system that uses augmented reality for displaying moving controls, and SmoothMoves for input (see Figure 1). The system was implemented using Philips Hue smart lights [39], which were wirelessly controlled by a video see-through AR application that runs on an unmodified Microsoft HoloLens. This is a head-mounted device that combines multiple optical sensors to both sense where users are looking and map their physical surroundings. The prototype was developed using the Unity game engine [40] and the Vuforia AR platform [41]. Input was captured using the HoloLens' standard API.

The idea of the prototype is simple. 2D moving controls are displayed in space, in proximity to the lights they control. These positions are set once, using pre-defined images or real-world objects. The controls enable the user to turn the lights on or off (Figure 8, top); to control the lights' intensity (Figure 8, top-right); and to access two menus. The first is the *themes menu*, that features two pre-set light schemes: work (cool blue) and relaxing (warm yellow) (Figure 8, bottom-left). The second is the *color menu*, that

enables the user to scroll through different hue colors in the HSV/HSB model using continuous head movements, and to also adjust the color's saturation (Figure 8, bottom-right). Brightness and saturation controls have two targets moving in opposite directions. Following the clockwise target increases the value of the control (e.g., makes it brighter), while following the counterclockwise target decreases it. All selections are confirmed through audio output (a click).

The motivations underlying the prototype are threefold. First, to support immediate control of smart environments with minimal action – a requirement highlighted by Koskela *et al.* [17] in their research on smart homes. Second, to provide uniform and hands-free control over different smart devices. And third, to support direct input in physical spaces: users simply look at the system they want to control in order to start interacting.

Evaluation

We evaluated the interactive lights prototype using 10 participants (4F), aged between 21 and 47 ($M = 34.3$, $SD = 8.88$). All participants were staff or students at a local institution. Based on a 7-point scale (low to high), participants rated their experience with AR at 2.5 ($SD = 1.51$); with HMDs at 2.8 ($SD = 1.55$); with smart lights at 2 ($SD = 1.70$); and with smart rooms at 1.8 ($SD = 0.79$). Participants interacted with the prototype in a spacious and quiet environment, where they were free to move around. Each experiment took on average 30 minutes, and was based on a participatory design technique to elicit in-depth user feedback [2]. This technique includes a *sensitization* and *elaboration* phase. In the former, participants were asked about relevant past experiences; in the latter, participants commented on the demo prototype. Each experiment started with an explanation of the prototype's functionality and a small trial where participants were asked to turn the lights on and off until they felt comfortable with the SmoothMoves input technique. We recorded and transcribed audio of all sessions and performed a lightweight clustering of comments, reported below.

Overall Opinions: In general participants responded positively to the technique, describing it as a "clever" (P7), "useful" (P4), "comfortable" (P2), and a great idea overall (P1, P5, P6, P7, P10). Participants also described the interface movement as "interesting" (P1, P5, P6), "fun" (P1, P6, P7, P10), and "minimalist" (P9); and did not consider it to be invasive (P9), or much of a distraction (P1, P4, P5). P2 described the movement as "futuristic" – a way to "attract people's attention" and "impress (house) guests". P4 appreciated the technique's ability to display "different options (so) close to each other". Finally, P6 described the experience as "quite magical" – "it is almost like you are doing it psychically". This sentiment was shared by P9: "I almost feel like it is my mind; if feels that subtle, that you (…) just will it to happen.".

Target Selection with the HoloLens: Despite these positives, there were concerns about how long it took to select a target (P2), that it initially required some concentration (P6, P10) and that it was an unusual way to interact (P8). Five participants reported unintentional selection of a target at some point during the session (P3, P6, P7, P9, P10). One explanation for this is the HoloLens' limited field-of-view. This issue is exacerbated as participants move their heads to acquire different targets – especially if the headset is not properly adjusted. P6 and P10 reported constraining their head movements because the targets "tend to appear and disappear"; and P7 did the same because the HoloLens kept "slipping down". P10 also described the HoloLens as quite "heavy". To minimize field-of-view issues, participants started the interaction at roughly two meters from the targets. This caused several participants to report the target trajectories as quite small (P3, P6, P10), and "sensitive" (P10) to input. Towards the end of the (short) session, these concerns began to lessen. P10 stated that "the more I did the easier it was"; and P9 ultimately "started to find [the movement] quite calming".

Use Scenarios: In response to a question on practical uses of the technique, participants P1 and P7 described how SmoothMoves would be useful for the "quick things": "I do not want to think, as you need in a smart phone application (…) I just want a button that turns on something, and then I can go back to work" (P1). P4 states that "it would definitely be useful" during hands busy activities in the home such as cleaning. Other participants saw value in terms of accessibility (P3, P5, P7, P8), or for professionals working with both hands, such as surgeons or bakers (P3). Finally, several participants envision using the technique when the hardware improves: when it is lighter (P1); when the field-of-view improves (P2); or when the device has the form factor of a normal pair of glasses (P3, P7, P9).

Gaze = Eyes + Head: Participants frequently commented on the naturalness and unobtrusiveness of the head movements and their tight coupling to gaze. P9 said it simply: "I do not feel I am moving my head". Similarly, P1 observed "I do not have to [mimics a very explicit head motion], I just have to look"; and P4 "notice[d] now that while I am just trying to do it with my eyes, my head unconsciously moves in the way [of the targets]". These quotes strongly reinforce the fundamental idea that gaze is a combination of eye and head motion – for several participants, even with instructions to move their heads, these modalities were hard to separate and distinguish.

Multi-modal input: Participants felt the technique could easily be integrated with other input modalities. Recognizing the potential problem of inadvertent activation, P3 and P6 proposed coarse mid-air gestures to trigger SmoothMoves controls. Other participants suggested integrating the technique with voice to specify more precise, important or detailed instructions (P6, P8, P10). Combining and comparing SmoothMoves with other input techniques is a compelling direction for future work.

Stimulus Parameters: Many participants were concerned about the size of both targets (P2, P3, P5, P6, P10), and trajectories (P3, P6, P8) and the speed at which targets moved (P8). Other participants were positive, feeling that that small trajectories would require only small head movements (P5). These concerns were largely alleviated when participants moved closer to the light and targets. Suggestions for dealing with this issue included various techniques for scaling targets and trajectories based on the distance to a user. Designing and refining such techniques is clearly a next step for this work.

Continuous Input Designs: Six participants specifically appreciated the flexibility of being able to set precise colors using the continuous color adjustment menus, but there were numerous reports that the implementation was confusing. For hue, a core problem was a lack of feedback as to how this parameter would vary over time (P5, P6, P8) – one solution was to control more well understood qualities such as separate RGB channels (P7). Other users reported uncertainty they were maintaining a selection during hue adjustment (P1, P7, P8, P10), likely due to the gradual rate of change in this parameter. Situations in which two controls moved in opposite directions around the same trajectory also led to trouble for P3: "it looks like they are bouncing off each other". In general, while participants also appreciated the audio feedback accompanying continuous parameter adjustment (P3, P7, P9), they also wanted more information in the form of visual or haptic (P8) cues.

Command Input Designs: Participants, in general, preferred the command input over the continuous input. Customizing lighting via choosing preset themes was reported to be more useful than continuous parameter adjustment (P1, P2, P3, P4, P5, P6), reflecting the general idea that SmoothMoves is more suited to quick and direct interaction (P1, P4, P7). Nesting menus was also viewed as appropriate as it avoided presenting too many simultaneous targets (P1, P4, P5, P6, P7, P8, P9) while still affording access to the most common commands quickly and easily (P1, P4, P5, P6, P8). The approach also kept things neat and tidy (P4, P5, P7, P8) and was reported to be consistent with traditional desktop computer interfaces (P7). Despite the proximity of the targets to the physical light, participants also explicitly suggested that feedback on selection be incorporated into the graphical interface (P1, P7, P8, P10).

In summary, SmoothMoves was well received by participants. Although there were some reports and worries regarding false activations, gripes about the headset and concerns about some of the specific control designs, the technique was viewed as convenient, relaxing, well suited to quick interactions in hands free situations and unobtrusive. This data provides evidence supporting the viability of the technique for real world input and points at key directions for improvement. Topics for future work include exploring integration with alternative input modalities (e.g. voice, ray pointing) and creating graphical

feedback to better support different selection and activation mechanisms, such as continuous parameter adjustment.

CONCLUSION
This paper introduced SmoothMoves, the first technique that supports smooth pursuits input using head movements. The paper described a pair of lab studies. The initial study generated three contributions. First, by looking at novel movement behaviors it expanded the design knowledge of smooth pursuits input systems. Second, it demonstrated that smooth pursuits input can be easily (and affordably) supported by head-tracking. And third, it generated ideal algorithm parameters for the SmoothMoves technique. The follow-up study grounded the technique in the domain of augmented reality, capturing the error rates and acquisition times on different types of AR device (head-mounted and hand-held). Finally, a prototype system was developed to demonstrate the benefits of using smooth pursuits head movements for interaction with AR applications in the context of an interactive lighting system. A final qualitative study led to positive reports of the system's suitability for this scenario. In contrast to smooth pursuits input systems based on eye-tracking, the SmoothMoves approach proposed in this paper can be immediately implemented on a wide range of devices that feature embedded motion sensing, such as AR headsets. The contributions of the paper, in terms of implementation, data and designs, represent concrete steps towards achieving this goal.

REFERENCES
1. Emilio Bizzi. 2011. Eye-Head Coordination. In *Comprehensive Physiology*, Ronald Terjung (ed.). John Wiley & Sons, Inc., Hoboken, NJ, USA. Retrieved April 5, 2016 from http://doi.wiley.com/10.1002/cphy.cp010229

2. Derya Ozcelik Buskermolen and Jacques Terken. 2012. Co-constructing Stories: A Participatory Design Technique to Elicit In-depth User Feedback and Suggestions About Design Concepts. In *Proceedings of the 12th Participatory Design Conference: Exploratory Papers, Workshop Descriptions, Industry Cases - Volume 2* (PDC '12), 33–36. https://doi.org/10.1145/2348144.2348156

3. Christopher Clarke, Alessio Bellino, Augusto Esteves, Eduardo Velloso, and Hans Gellersen. 2016. TraceMatch: a computer vision technique for user input by tracing of animated controls. In *UbiComp '16:*

Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 298–303. https://doi.org/10.1145/2971648.2971714

4. Nicholas Cooper, Aaron Keatley, Maria Dahlquist, Simon Mann, Hannah Slay, Joanne Zucco, Ross Smith, and Bruce H. Thomas. 2004. Augmented Reality Chinese Checkers. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology* (ACE '04), 117–126. https://doi.org/10.1145/1067343.1067357

5. Andrew Crossan, Mark McGill, Stephen Brewster, and Roderick Murray-Smith. 2009. Head Tilting for Interaction in Mobile Contexts. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI '09), 6:1–6:10. https://doi.org/10.1145/1613858.1613866

6. Dietlind Helene Cymek, Antje Christine Venjakob, Stefan Ruff, Otto Hans-Martin Lutz, Simon Hofmann, and Matthias Roetting. Entering PIN codes by smooth pursuit eye movements. *Journal of Eye Movement Research* 7(4), 1: 1–11.

7. 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), 307–311. https://doi.org/10.1145/2993148.2993181

8. David Dobbelstein, Philipp Hock, and Enrico Rukzio. 2015. Belt: An Unobtrusive Touch Input Device for Head-worn Displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), 2135–2138. https://doi.org/10.1145/2702123.2702450

9. 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*, 457–466. https://doi.org/10.1145/2807442.2807499

10. Karen M. Evans, Robert A. Jacobs, John A. Tarduno, and Jeff B. Pelz. 2012. Collecting and Analyzing Eye-Tracking Data in Outdoor Environments. *Journal of Eye Movement Research* 5, 2. https://doi.org/10.16910/jemr.5.2.6

11. David G. Fleming, Gehard W. Vossius, George Bowman, and Ensign L. Johnson. 1969. Adaptive Properties Of The Eye-tracking System As Revealed By Moving-head And Open-loop Studies. *Annals of the New York Academy of Sciences* 156, 2 Rein Control,: 825–850. https://doi.org/10.1111/j.1749-6632.1969.tb14017.x

12. Yi-Ta Hsieh, Antti Jylhä, Valeria Orso, Luciano Gamberini, and Giulio Jacucci. 2016. Designing a Willing-to-Use-in-Public Hand Gestural Interaction Technique for Smart Glasses. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16), 4203–4215. https://doi.org/10.1145/2858036.2858436

13. Aulikki Hyrskykari, Howell Istance, and Stephen Vickers. 2012. Gaze Gestures or Dwell-based Interaction? In *Proceedings of the Symposium on Eye Tracking Research and Applications* (ETRA '12), 229–232. https://doi.org/10.1145/2168556.2168602

14. Jari Kangas, Oleg Špakov, Poika Isokoski, Deepak Akkil, Jussi Rantala, and Roope Raisamo. 2016. Feedback for Smooth Pursuit Gaze Tracking Based Control. In *Proceedings of the 7th Augmented Human International Conference 2016* (AH '16), 6:1–6:8. https://doi.org/10.1145/2875194.2875209

15. Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (UbiComp '14 Adjunct), 1151–1160. https://doi.org/10.1145/2638728.2641695

16. 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), 274–285. https://doi.org/10.1145/2971648.2971679

17. Tiiu Koskela and Kaisa Väänänen-Vainio-Mattila. 2004. Evolution towards smart home environments: empirical evaluation of three user interfaces. *Personal and Ubiquitous Computing* 8, 3–4: 234–240. https://doi.org/10.1007/s00779-004-0283-x

18. Jeremy Lanman, Emilio Bizzi, and John Allum. 1978. The coordination of eye and head movement during smooth pursuit. *Brain Research* 153, 1: 39–53. https://doi.org/10.1016/0006-8993(78)91127-7

19. R. John Leigh and David S. Zee M.D. 2015. *The Neurology of Eye Movements*. Oxford University Press.

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: 1–11.

21. Robert Mahony, Tarek Hamel, and Jean-Michel Pflimlin. 2008. Nonlinear Complementary Filters on the Special Orthogonal Group. *IEEE Transactions on Automatic Control* 53, 5: 1203–1218. https://doi.org/10.1109/TAC.2008.923738

22. Microsoft. Microsoft HoloLens. *Microsoft HoloLens*. Retrieved September 20, 2016 from https://www.microsoft.com/microsoft-hololens/en-us

23. Pranav Mistry and Pattie Maes. 2009. SixthSense: A Wearable Gestural Interface. In *ACM SIGGRAPH ASIA 2009 Sketches* (SIGGRAPH ASIA '09), 11:1–11:1. https://doi.org/10.1145/1667146.1667160

24. Louis-Philippe Morency and Trevor Darrell. 2006. Head Gesture Recognition in Intelligent Interfaces: The Role of Context in Improving Recognition. In *Proceedings of the 11th International Conference on Intelligent User Interfaces* (IUI '06), 32–38. https://doi.org/10.1145/1111449.1111464

25. Kasım Özacar, Juan David Hincapié-Ramos, Kazuki Takashima, and Yoshifumi Kitamura. 2017. 3D Selection Techniques for Mobile Augmented Reality Head-Mounted Displays. *Interacting with Computers* 29, 4: 579–591. https://doi.org/10.1093/iwc/iww035

26. Ken Pfeuffer, Melodie Vidal, Jayson Turner, Andreas Bulling, and Hans Gellersen. 2013. Pursuit Calibration: Making Gaze Calibration Less Tedious and More Flexible. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (UIST '13), 261–270. https://doi.org/10.1145/2501988.2501998

27. Christoph Rasche and Karl R. Gegenfurtner. 2009. Precision of speed discrimination and smooth pursuit eye movements. *Vision Research* 49, 5: 514–523. https://doi.org/10.1016/j.visres.2008.12.003

28. Julie Rico and Stephen Brewster. 2010. Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '10), 887–896. https://doi.org/10.1145/1753326.1753458

29. D A Robinson. 1965. The mechanics of human smooth pursuit eye movement. *The Journal of Physiology* 180, 3: 569–591.

30. Boris Smus and Christopher Riederer. 2015. Magnetic Input for Mobile Virtual Reality. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers* (ISWC '15), 43–44. https://doi.org/10.1145/2802083.2808395

31. Sophie Stellmach and Raimund Dachselt. 2012. Look & Touch: Gaze-supported Target Acquisition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '12), 2981–2990. https://doi.org/10.1145/2207676.2208709

32. Noboru Sugie and Makoto Wakakuwa. 1970. Visual Target Tracking with Active Head Rotation. *IEEE Transactions on Systems Science and Cybernetics* 6, 2: 103–109. https://doi.org/10.1109/TSSC.1970.300283

33. Gabriel Takacs, Vijay Chandrasekhar, Natasha Gelfand, Yingen Xiong, Wei-Chao Chen, Thanos Bismpigiannis, Radek Grzeszczuk, Kari Pulli, and Bernd Girod. 2008. Outdoors Augmented Reality on Mobile Phone Using Loxel-based Visual Feature Organization. In *Proceedings of the 1st ACM International Conference on Multimedia Information Retrieval* (MIR '08), 427–434. https://doi.org/10.1145/1460096.1460165

34. 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*, 812–817. https://doi.org/10.1145/2901790.2901867

35. 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), 439–448. https://doi.org/10.1145/2493432.2493477

36. Chun Yu, Ke Sun, Mingyuan Zhong, Xincheng Li, Peijun Zhao, and Yuanchun Shi. 2016. One-Dimensional Handwriting: Inputting Letters and Words on Smart Glasses. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16), 71–82. https://doi.org/10.1145/2858036.2858542

37. ES | PRODUCTS. *JINS MEME*. Retrieved July 5, 2017 from https://jins-meme.com/en/products/es/

38. Epson Moverio Next Generation Smart Eyewear - Epson America, Inc. Retrieved September 20, 2016 from http://www.epson.com/cgi-bin/Store/jsp/Landing/moverio-augmented-reality-smart-eyewear-technology.do

39. Philips Hue. Retrieved September 21, 2016 from http://www2.meethue.com/en-gb/

40. Unity - Game Engine. *Unity*. Retrieved April 4, 2017 from https://unity3d.com

41. Vuforia | Augmented Reality. Retrieved April 4, 2017 from https://www.vuforia.com/