BodyPursuits: Exploring Smooth Pursuit Gaze Interaction Based on Body Motion Targets

Anja Hansen*
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
hansen@teco.edu

Sarah Makarem
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
sarah.makarem@kit.edu

Kai Kunze
Keio University
Yokohama, Japan
kai@kmd.keio.ac.jp

Yexu Zhou
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
zhou@teco.edu

Michael Thomas Knierim
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
michael.knierim@kit.edu

Christopher Clarke
University of Bath
Bath, United Kingdom
cjc234@bath.ac.uk

Hans Gellersen
Lancaster University
Lancaster, United Kingdom
h.gellersen@lancaster.ac.uk

Michael Beigl
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
michael.beigl@kit.edu

Tobias Röddiger
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
roeddiger@teco.edu

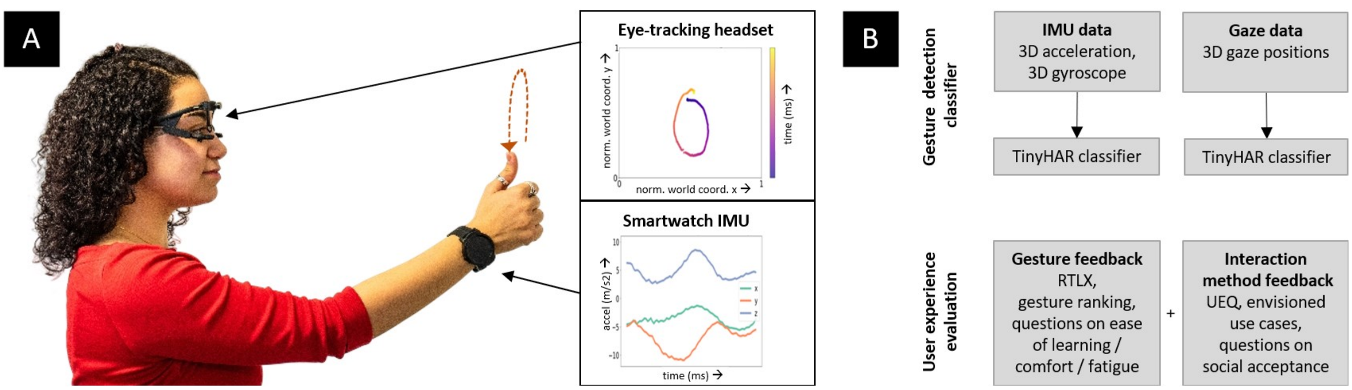


Figure 1: Feasibility evaluation of *BodyPursuits* as HCI method. (A) Setup: User study participant performing *BodyPursuits* gesture, wearing Pupil Core (mobile eye tracking hardware by Pupil Labs) [Kassner et al. 2014] and Ticwatch E2 as baseline. (B) Evaluation: Gesture classifier training and performance evaluation, as well as user experience evaluation.

Abstract

Smooth pursuits are natural eye movements that occur when we track a moving target with our gaze. While they were explored as a gaze-based input method using external screens to display moving stimuli, we propose *BodyPursuits*, a novel HCI method that eliminates additional screens. Stimuli are generated by users tracing smooth trajectories with their hand in mid-air while fixating their gaze on their thumb. We conducted a user study to collect eye-tracking and baseline IMU data from 20 participants performing 10 *BodyPursuits* gestures. Based on 1800 samples and noise data, we train a TinyHAR classifier. It achieves a macro-average F1 score of 0.772. With UEQ results indicating a positive user experience and RTLX scores showcasing low subjective workload for all 10 gestures, we successfully demonstrated *BodyPursuits*' potential as viable interaction method. We envision *BodyPursuits* could be integrated into EOG earphones to detect mid-air hand gestures without external screens or cameras.

ETRA '25, Tokyo, Japan

ACM ISBN 979-8-4007-1487-0/25/05
https://doi.org/10.1145/3715669.3723110

CCS Concepts

• **Human-centered computing** → **Gestural input**.

Keywords

smooth pursuits, gaze input, human-computer interaction, mid-air gestures, eye-tracking, inertial measurement unit, gaze-gesture combination, human activity recognition

ACM Reference Format:

Anja Hansen, Sarah Makarem, Kai Kunze, Yexu Zhou, Michael Thomas Knierim, Christopher Clarke, Hans Gellersen, Michael Beigl, and Tobias Röddiger. 2025. BodyPursuits: Exploring Smooth Pursuit Gaze Interaction Based on Body Motion Targets. In *2025 Symposium on Eye Tracking Research and Applications (ETRA '25), May 26–29, 2025, Tokyo, Japan.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3715669.3723110

1 Introduction and Related Work

Due to its intuitive and efficient nature, gaze has become an attractive modality for pervasive interaction [De Gaudenzi and Porta 2014; Vidal et al. 2015]. However, conventional gaze input methodologies are not without inherent challenges, which include eye fatigue after prolonged use, or the "Midas Touch" problem, where inadvertent selections occur [Cymek et al. 2014; Parisay et al. 2020]. A promising approach to overcome these challenges involves the use of smooth pursuits — natural, intuitive eye movements that occur when we track ("pursue") a moving target with our eyes by maintaining the line of sight [Vidal et al. 2013a]. In comparison to dwell-based methods [Parisay et al. 2020], smooth pursuits are found to be much more comfortable in their execution due to their natural occurrence in everyday eye behavior [Shimizu et al. 2016]. Also, the characteristic patterns of pursuits make them easier to distinguish from other eye movements [Esteves et al. 2015], and users are unlikely to encounter a target in their environment moving in the same way as the intended stimulus which mitigates the Midas Touch problem [Khamis et al. 2018].

Multiple research works elicited smooth pursuit eye movements by moving targets displayed on screens [Esteves et al. 2015; Vidal et al. 2013a,b], demonstrating the feasibility of smooth pursuits as a gaze input method [Esteves et al. 2015; Shimizu et al. 2016]. Vidal et al. [2013a] initially proposed smooth pursuits as a novel gaze input method allowing users to interact with an eye-based interface spontaneously by standing in front of an external screen and displaying moving targets that are pursued by gaze. They identified smooth pursuits to be a promising and robust input technique. Khamis et al. [2015] further demonstrated that smooth pursuits input was well perceived by novice users for spontaneous interaction with public displays. "Orbits" [Esteves et al. 2015] transferred the concept to smartwatch displays, and Shimizu et al. [2016] to commercial smart glasses. Further research led to various applications of smooth pursuit input, e.g., entry of PIN codes by following moving PIN pad buttons with gaze [Cymek et al. 2014], interacting with different widgets [Špakov et al. 2016], or for hands-free spelling [Lutz et al. 2015]. Further, the potential of smooth pursuits in virtual reality (VR) was explored [Khamis et al. 2018].

Smooth pursuit eye movement inherently requires an external dynamic visual target, as it cannot be performed without an external stimulus. However, current human-computer interaction applications rely on external displays to generate these cues. This dependency limits the integration of smooth pursuits, especially in dynamic and mobile environments where screens may not always be accessible. For example, while earphones equipped with Electrooculography (EOG) sensors can capture eye movements [Lepold et al. 2024; Manabe and Fukumoto 2006; Manabe et al. 2013; Röddiger 2023], they do not provide a means to present visual cues within the user's line of sight. Furthermore, research suggests that the accuracy and engagement of smooth pursuit eye movements improve significantly when the target motion is self-generated [Kowler et al. 2019].

Therefore, we introduce *BodyPursuits* — an interaction concept that leverages user-generated body motions as stimuli for smooth pursuit input. By eliminating the need for external devices, this approach creates new opportunities for interaction in situations where external displays are not desirable or available. In this work, we report an offline classification result for *BodyPursuits* gestures, with the vision of using *BodyPursuits* as a real-time input method one day. In sum, we contribute:

- *BodyPursuits*, a novel interaction paradigm that leverages hand-generated motions as stimuli for smooth pursuit eye movements and 10 gestures specifically tailored to this approach.
- User study recordings of 20 participants performing *BodyPursuits* gestures repeatedly and engaging in gaze noise tasks, resulting in 1800 gesture samples and 600 minutes of gaze noise data.
- A lightweight classifier based on the TinyHAR architecture [Zhou et al. 2022], the best model trained on gaze data achieving a macro-average F1 score of 0.772 in recognizing our 10 *BodyPursuits* gestures and noise.
- A comprehensive usability evaluation demonstrating that performing *BodyPursuits* gestures induces low subjective workload and yields an overall positive user experience.

2 BodyPursuits

In the following we introduce the *BodyPursuits* interaction technique. We further describe the rationale behind a gesture set that we specifically conceptualized for *BodyPursuits*.

2.1 Concept

BodyPursuits uniquely combines **body** motions with smooth **pursuits**. In contrast to screen-based methods, the moving target is generated by the user's thumb tracing a trajectory mid-air which is then followed by the user's gaze. We envision that such a technique could be used in situations where external screens are not available or desirable.

2.2 Gesture Set

As our objective was not to validate the general efficacy of smooth pursuits as an input method, but rather in establishing the detectability and usability of a moving target generated through body motion, we opted for target trajectories that are well-established in literature: circular [Drewes et al. 2018; Esteves et al. 2015; Gomez and Gellersen 2017; Khamis et al. 2015, 2018, 2016; Kosch et al. 2018; Shimizu et al. 2016; Špakov et al. 2016; Vidal et al. 2013a], vertical [Cymek et al. 2014; Gomez and Gellersen 2017; Khamis et al. 2015, 2018, 2016; Lutz et al. 2015; Zhe et al. 2020] and horizontal [Collewijn and Tamminga 1984; Gomez and Gellersen 2017; Rottach

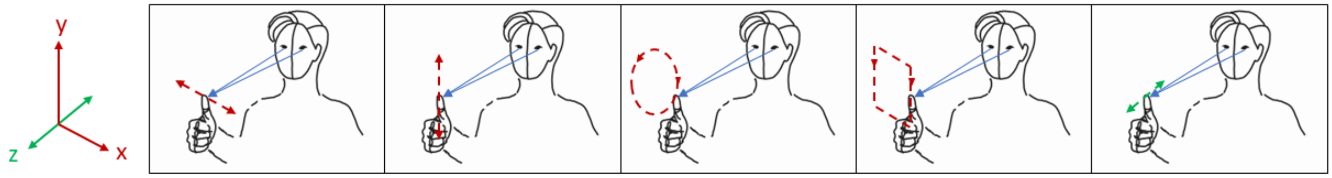


Figure 2: Gestures tailored for *BodyPursuits*, combining established trajectories from related works on smooth pursuit and mid-air gesture input, and two proof-of-concept gestures to explore depth-based 3D pursuits (near-far / far-near).

et al. 1996] trajectories. Moreover, our gesture set had to be suitable for mid-air gesture tracing. In that regard, basic geometric shapes were found to be an effective input method [Carter et al. 2016; Clarke et al. 2016; Kim et al. 2015]. More related, work on motion correlation found that screen targets followed by body motions were well feasible with circles, squares, and horizontal / vertical lines [Clarke 2020; Clarke et al. 2016].

Taking in all findings from previous work, we selected horizontal and vertical lines, as well as circular trajectories, to be included in our gesture set. Inspired from geometric trajectories that combine multiple linear strokes [Cymek et al. 2014; Drewes et al. 2018; Khamis et al. 2016; Kosch et al. 2018], we also selected a basic square shape. For each trajectory, we select both directions (clockwise and anticlockwise, to the left or right, upwards or downwards) to maintain a simple gesture vocabulary by varying already established, easier trajectories [Clarke 2020; Esteves et al. 2015]. Further, we introduce a novel near-to-far and far-to-near trajectory as a proof of concept for gestures that incorporate depth changes on the 3rd dimension. While various target speed have been explored in related work [Cymek et al. 2014; Drewes et al. 2018; Gomez and Gellersen 2017; Khamis et al. 2015; Kosch et al. 2018; Zhe et al. 2020], we chose to allow users to trace the trajectories at their own comfortable pace, to gain insights into natural interaction patterns and avoid the challenge of enforcing a fixed speed. Our final gesture set consists of 10 gestures, see Figure 2.

2.3 Apparatus

To detect various *BodyPursuits* gestures, appropriate sensing hardware is required. For eye tracking, we utilized the *Pupil Core*, a binocular, camera-based eye tracker equipped with two rotatable eye cameras and a world camera [Kassner et al. 2014]. Additionally, as baseline for mid-air gestures participants wore a smartwatch — the "Mobvoi Ticwatch E2" — which features an inertial measurement unit (IMU) to capture wrist acceleration and gyroscope data [Mobvoi 2025]. In this work, we evaluated the feasibility of *BodyPursuits* using *Pupil Core*, but the interaction method's actual applicability varies with the quality and robustness of the used devices.

2.4 Recognition Pipeline

To detect the different *BodyPursuits* gestures, we employ the TinyHAR architecture, a lightweight deep learning model designed by Zhou et al. [2022]. The name "TinyHAR" directly signifies both its lightweight nature ("tinyness") and its specific design for Human Activity Recognition (HAR). We envision that *BodyPursuits* would operate on real-time edge systems which is closely reflected by the choice of this architecture. The raw sensor data that we obtain from the eye tracker (x, y, z position for both eyes) and IMU (x, y, z for accelerometer and gyroscope) are resampled to 100 Hz with linear interpolation before they are fed into TinyHAR. In the TinyHAR architecture, we opted for four convolutional layers.

3 Evaluation

The goal of our evaluation was to understand the usefulness and comfort of *BodyPursuits*. In addition, we wanted to understand whether smooth pursuits elicited by *BodyPursuits* are distinguishable from everyday gaze. As baseline, we employ a smartwatch for mid-air gesture recognition. To facilitate the data collection, we developed a web-based tool with Flask that communicates with the eye tracker and smartwatch and logs all event data.

3.1 Procedure

The study was conducted within the guidelines of the "Declaration of Helsinki", the ethical guidelines of the university and local laws. We find no privacy or ethical risks associated with *BodyPursuits*.

Participants arrived at the lab and provided informed consent. They were then instructed to attach the eye-tracker and smartwatch, undergo the eye-tracking calibration process, and position themselves in front of a projection surface that showed the instructions given by our custom tool. The instructions included a woman performing a *BodyPursuits* gesture in mid-air, like in Figure 1. Consequently, participants typically maintained similar arm postures, slightly bending their arms rather than fully extending them. The bending angle varied among participants and throughout the study duration. We also observed that some participants preferred positioning their hands closer to their faces, while others positioned them further away.

Participants were given three common primary tasks that should elicit noise gaze data: (1) reading text, (2) looking at an image with hidden objects, (3) watching a video. While they performed these tasks, we played an auditory cue to trigger performing a *BodyPursuits* gesture in equidistant time intervals. When hearing the auditory cue, participants stopped the primary task, performed the gesture once and then resumed their primary task again. Hence, gaze noise was recorded during the primary tasks whenever users were not performing a *BodyPursuit* gesture. The order of the ten gestures was counterbalanced across participants [Zhe et al. 2020]. Per primary task, participants performed each gesture three times, and

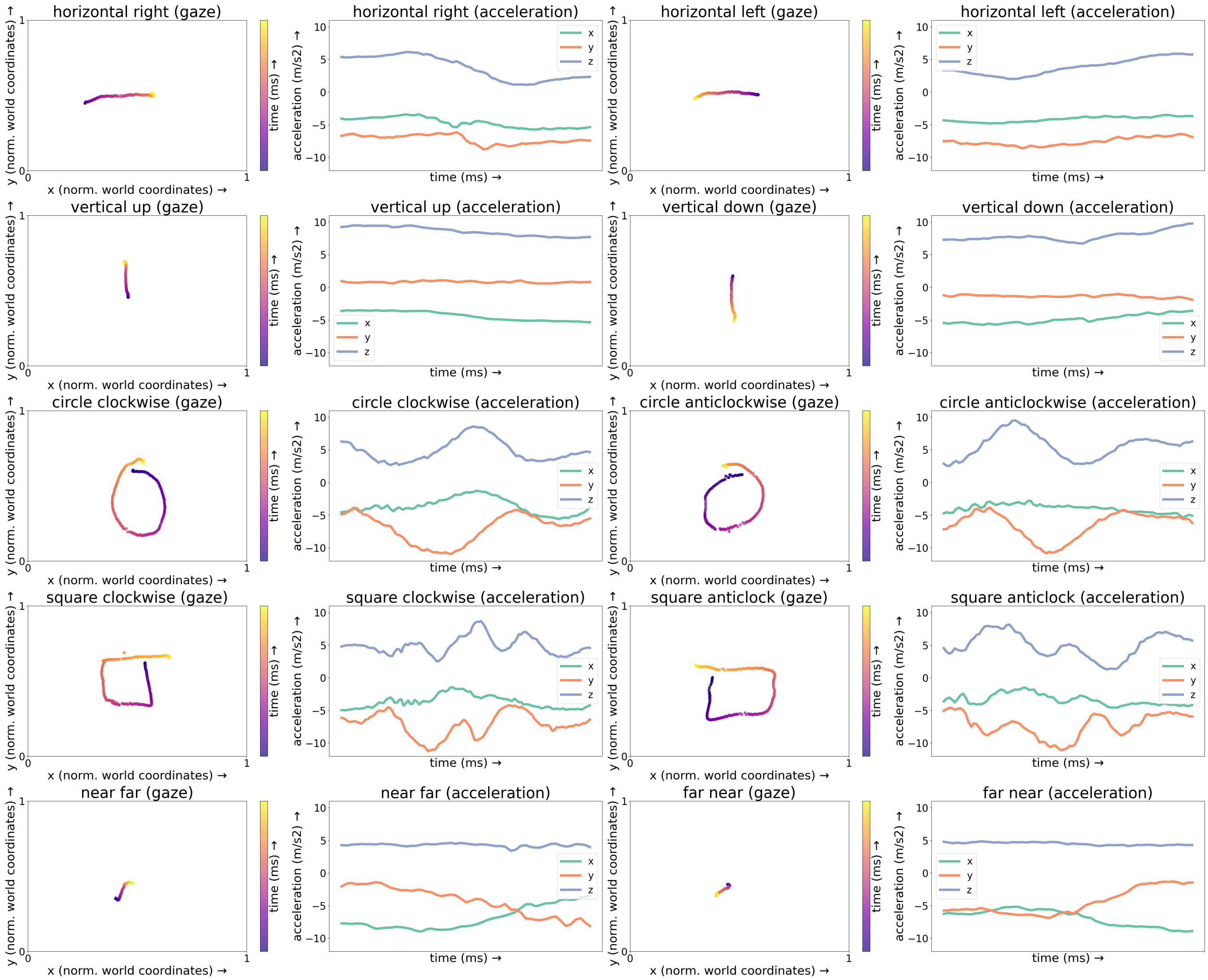


Figure 3: Selection of raw sensor data as collected during the user study for each of the ten gestures for *BodyPursuits*. Depicted: gaze data (only 2D, but we collected 3D gaze in world coordinates) and IMU acceleration data (we collected both 3D gyroscope and 3D acceleration data).

they performed one gesture for all primary tasks before performing the next gesture. The gestures were shown to the participants prior to the data collection, and they could practice until they felt comfortable. After each gesture, participants filled out the Raw NASA Task Load Index (RTLX) questionnaire [Byers et al. 1989; Hart 2006]. Given the high number of gestures, RTLX was chosen over traditional NASA-TLX as findings by Byers et al. [1989] suggest their equivalence in results. 7-point Likert scale questions were used to measure ease of learning, execution comfort and eye fatigue. After performing all gestures, participants completed the User Experience Questionnaire (UEQ) [Laugwitz et al. 2008]. We further asked questions about social acceptability. Additionally, the gestures were ranked based on personal preference.

3.2 Study Participants

In total 20 participants were recruited. We purposefully did not recruit participants wearing glasses to avoid interference with the accuracy of the eye tracker [Lutz et al. 2015]. 12 of the participants identify as male, eight as female. All participants were affiliated with our university as students, alumni or PhD candidates. The age distribution ranged from 22 to 34 years old, with a mean age of 26.53 years. All participants were right-handed, 80% had uncorrected vision, and 20% wore contact lenses. Most participants had no experience with eye-tracking technology (70%), gaze input techniques (90%), or mid-air gesture input techniques (55%) at all. Some had encountered eye-tracking technology or gaze input techniques at most a few times, primarily through similar user studies or research. There was notably more expertise in mid-air gesture input, with 30% having used it at least a couple of times and 15% at least once, mostly during gaming.

4 Results

In the following we break down the classifier and user experience results of *BodyPursuits*.

4.1 Classifier

We describe the training procedure and report the performance of TinyHAR to distinguish *BodyPursuits* gestures. Figure 3 shows a selection of collected unfiltered eye tracker and IMU sensor data.

4.1.1 Training. A total of 1,800 gesture samples were collected for analysis. Due to data inconsistencies, 36 samples were excluded from the dataset. Additionally, approximately 600 minutes of background noise data were gathered, from which we randomly extracted 3,000 window-size long samples evenly spread across participants.

Users controlled the gesture execution speed, leading to variability across users, gestures, and iterations. Table 1 shows the minimum, maximum, and 90th percentile (indicating when 90% of participants completed each gesture) of execution times measured during the user study.

Table 1: Gesture speed during user study: Minimum, maximum, and 90th percentile of gesture execution times in seconds.

Gesture	Min (s)	Max (s)	90th Percentile (s)
Horizontal right	0.9	5.2	2.7
Horizontal left	0.7	4.4	3.1
Vertical up	0.7	4.7	2.7
Vertical down	0.7	5.9	3.2
Circle clockwise	1.3	9.6	4.7
Circle anticlockwise	1.3	7.9	4.5
Square clockwise	1.7	7.4	4.8
Square anticlockwise	1.3	7.8	5.2
Near-far	0.9	4.4	2.7
Far-near	0.6	3.9	2.9

The longest gestures, such as circles and squares, were completed by the majority (90%) of participants within approximately 5 seconds. Given that some gestures (e.g., horizontal and square) partially overlap in their paths, we initially explored window sizes of 4, 5, and 6 seconds to ensure full gesture coverage.

Finally, we trained six TinyHAR [Zhou et al. 2022] models: three using eye-tracking data (three different window sizes), and three using IMU data as a baseline for mid-air gesture detection.

4.1.2 Performance. Since our user study investigates the feasibility of a novel interaction method, we applied leave-one-out cross-validation to evaluate how well the models generalize on new, unseen users, and computed the macro-average F1 scores for different window sizes based on the gaze and IMU data as shown in Table 2.

Table 2: Macro-averaged F1 scores for the six models trained with different window sizes, and on gaze or IMU data.

	4 seconds	5 seconds	6 seconds
Gaze data	0.764	0.772	0.760
IMU data	0.922	0.929	0.929

The confusion matrix of the eye tracker model for the 4 seconds window size is shown in the confusion matrix in Figure 5. We purposefully selected the smallest window size as we believe it is most appropriate for interaction given the trade-off between selection time and classification performance.

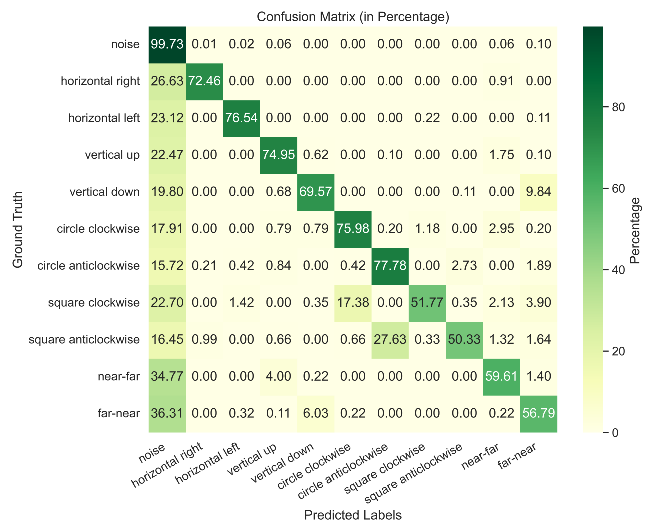
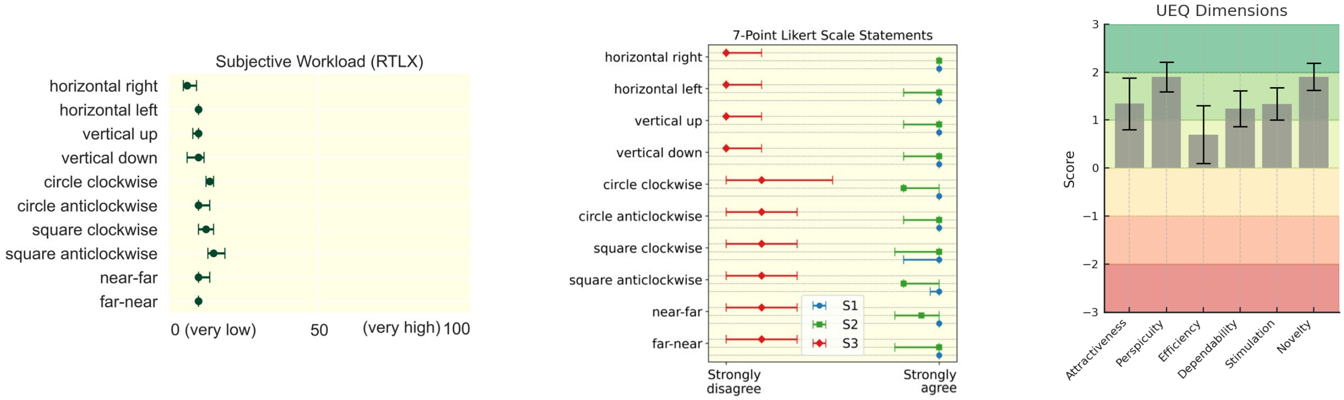


Figure 5: Confusion matrix for TinyHAR model (window size 4 seconds, trained on gaze data).

4.1.3 Discussion. Our results show that eye tracking measurement has the potential to classify hand-based gestures from the eyes following the hand, however the classification accuracies are currently lower than would be required for deployable applications. The main issue with the gaze-based models is that our gesture set leads to gestures being misclassified gestures as noise. In particular, one-dimensional line gestures are frequently misclassified as noise, with notable confusions between vertical down and far-near, as well as vertical up and near-far gestures. This is likely because saccadic eye movement will have similar trajectories and additional filtering should be used. Further, far-near / near-far vergence eye movements are also prone to be confused as noise, likely because the eye movements are so small. Two-dimensional gestures, such as circles and squares, are less likely to be misclassified as noise, but instead there is the issue that they are sometimes confused with each other when moving in the same direction (e.g., circle clockwise



(a) Gesture feedback: Raw Task Load Index (RTLX) ratings [Byers et al. 1989]. Median across 6 subdimensions (N=20).

(b) Gesture feedback: 7-point Likert scale rating on statements S1-S3. Median (N=20).

(c) *BodyPursuits* rated on the UEQ [Laugwitz et al. 2008]. Mean and confidence interval (N=20).

Figure 4: Usability evaluation results for gesture set and interaction method.

and square clockwise). This suggests that classifiers trained only on two-dimensional gestures may be required.

As expected, we find that classifiers trained solely on gaze data perform worse than those trained on the IMU data. On average, the macro-averaged F1 score for the model trained on IMU data is approximately 15% higher than that of the model trained exclusively on gaze data. However, it is important to note that no noise data was collected for the hand motions, resulting in a much noisier signal from the gaze data.

Revising the gesture set for *BodyPursuits* by selecting from the most distinct gestures could avoid some of the confusions mentioned above.

4.2 User Experience

The user experience was evaluated on a per gesture level and also based on the overall *BodyPursuits* method.

4.2.1 Gesture Feedback. We conducted a Shapiro-Wilk test for normality on the cumulative RTLX results, which suggested the collected data is not normally distributed [Shapiro and Wilk 1965]. Hence, we chose median for further evaluation [Sainani 2012]. We calculated the cumulative subjective workload score across the 20 participants and generally found low subjective workload, as depicted in Figure 4(a). This suggests that *BodyPursuits* is a low-effort and appealing input technique across all tested gestures. A subtle visible trend appears to suggest that 2D gestures (such as circles and squares) may be associated with a higher subjective workload than horizontal / vertical gestures, with far-near / near-far in-between those categories. To explore this trend, we conducted a non-parametric Friedman test [Pereira et al. 2015] which yielded a test statistic of $Q = 30.06$ with a highly significant p-value of 1.44×10^{-5}, indicating statistically significant differences in subjective workload between the 10 gestures. However, post-hoc pairwise Wilcoxon Signed-Rank tests with Bonferroni correction did not yield statistically significant results. Without correction, we find statistically significant differences between horizontal / vertical gestures and 2D gestures in terms of subjective workload. From free-text user comments, we further understood that one-dimensional gestures were preferred because the execution time is faster and that near-far / far-near gestures could make users go cross-eyed.

Friedman tests further revealed significant differences between gestures for users' 7-point Likert scale ratings on statements concerning ease of learning (S1: "The gesture was easy to learn.", $Q = 23.40$, p-value 5.36×10^{-3}), gesture execution comfort (S2: "It was comfortable to execute the gesture.", $Q = 30.36$, p-value 3.81×10^{-4}) and eye fatigue (S3: "The gesture execution was physically fatiguing for my eyes.", $Q = 19.95$, p-value of 1.82×10^{-2}). The median ratings are depicted in Figure 4(b).

A subjective ranking of the gestures according to the users preferences produced the following overall order: horizontal right (most favored), horizontal left, vertical up, vertical down, circle clockwise, circle anticlockwise, near-far, far-near, square clockwise, square anticlockwise (least favored).

4.2.2 Interaction Technique Feedback. At the end of the user study, participants filled out the standardized UEQ [Laugwitz et al. 2008]. The results are shown in Figure 4(c). Overall, we can observe positive trends across all dimensions, except for efficiency. This aligns with subjective feedback received in the free-text responses (e.g., "Innovative, but needs some time to get used to it.", "Creative!" and "Some gestures take too much time and coordination."). While users could imagine to perform *BodyPursuits* in a private settings ("I would be comfortable with using BodyPursuits as input technique in private.", median: 6.5/7, IQR: [5.75, 7]), they were less likely to do so in public (median: 5/7, IQR: [3.75, 5.25]). Based on qualitative feedback, users see potential for judgment or confusion by bystanders. Since the interaction method is not yet established, participants fear that bystanders might misinterpret the meaning or direction of *BodyPursuits* gestures.

4.2.3 Use Cases. Further, participants were asked to envision potential use cases for *BodyPursuits*. This included controlling music players, calls, and smart home systems, as well as applications in medical settings, food preparation, or any scenario where the user's hands are gloved or contaminated. Discussing use cases, we envision the main benefit of adding gaze-based smooth pursuit input when hand gestures are already in use in the case that eye tracking capabilities are available but lack the possibility to put a visual target in the line of sight of the user. One possible example includes earphones, which have the required eye tracking capabilities via electrooculography (EOG) [Lepold et al. 2024; Manabe and Fukumoto 2006; Manabe et al. 2013; Röddiger 2023], but no external screen to present a gaze target, or brain-computer interfaces using electroencephalography (EEG) signals like the ones used by [Knierim et al. 2025].

5 Conclusion

In this work, we introduced *BodyPursuits*, a novel interaction paradigm that leverages body-generated motions as targets for smooth pursuit interaction. We designed and evaluated 10 gestures suitable for *BodyPursuits*. We trained a classifier on sensor data from 20 participants performing 1800 *BodyPursuits* gestures, achieving a macro-averaged F1-score of 0.772 based on camera eye tracking data. All 10 gestures were rated with low subjective workloads. Further, we found that *BodyPursuits* provided a positive user experience that was perceived as attractive, reliable, and exciting. Participants particularly appreciated the intuitiveness and creativity of the approach. In addition, the interaction method was largely seen as socially acceptable, with a willingness to integrate it into everyday life, even more so in private settings. However, efficiency emerged as a potential area for improvement. Overall, we find that *BodyPursuits* proved to be a viable alternative to traditional input techniques that could be integrated EOG earphones.

Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – GRK2739/1 - Project Nr. 447089431 – Research Training Group: KD2School – Designing Adaptive Systems for Economic Decisions.

References

James C Byers, AC Bittner, and Susan G Hill. 1989. Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary. *Advances in industrial ergonomics and safety* 1 (1989), 481–485.

Marcus Carter, Eduardo Velloso, John Downs, Abigail Sellen, Kenton O'Hara, and Frank Vetere. 2016. Pathsync: Multi-user gestural interaction with touchless rhythmic path mimicry. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3415–3427.

Christopher Clarke. 2020. *Dynamic Motion Coupling of Body Movement for Input Control*. Doctoral thesis. Lancaster University.

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 *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 298–303.

Han Collewijn and Ernst P Tamminga. 1984. Human smooth and saccadic eye movements during voluntary pursuit of different target motions on different backgrounds. *The Journal of physiology* 351, 1 (1984), 217–250.

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 (2014).

Enrico De Gaudenzi and Marco Porta. 2014. Gaze Input for Ordinary Interfaces: Combining Automatic and Manual Error Correction Techniques to Improve Pointing Precision. In *Intelligent Systems for Science and Information: Extended and Selected Results from the Science and Information Conference 2013*. Springer, 197–212.

Heiko Drewes, Mohamed Khamis, and Florian Alt. 2018. Smooth pursuit target speeds and trajectories. In *Proceedings of the 17th international conference on mobile and ubiquitous multimedia*. 139–146.

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.

Argenis Ramirez Gomez and Hans-Werner Georg Gellersen. 2017. Behavioral Analysis of Smooth Pursuit Eye Movements for Interaction. (2017).

Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

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*. 1151–1160.

Mohamed Khamis, Florian Alt, and Andreas Bulling. 2015. A field study on spontaneous gaze-based interaction with a public display using pursuits. In *Adjunct proceedings of the 2015 acm international joint conference on pervasive and ubiquitous computing and proceedings of the 2015 acm international symposium on wearable computers*. 863–872.

Mohamed Khamis, Carl Oechsner, Florian Alt, and Andreas Bulling. 2018. VRpursuits: Interaction in virtual reality using smooth pursuit eye movements. In *Proceedings of the 2018 international conference on advanced visual interfaces*. 1–8.

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*. 274–285.

Kwangtaek Kim, Joongrock Kim, Jaesung Choi, Junghyun Kim, and Sangyoun Lee. 2015. Depth camera-based 3D hand gesture controls with immersive tactile feedback for natural mid-air gesture interactions. *Sensors* 15, 1 (2015), 1022–1046.

Michael T Knierim, Christian Zimny, Gabriel Ivucic, and Tobias Röddiger. 2025. Advancing Wearable BCI: Headphone EEG for Cognitive Load Detection in Lab and Field. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 1 (2025), 1–26.

Thomas Kosch, Mariam Hassib, Paweł W Woźniak, Daniel Buschek, and Florian Alt. 2018. Your eyes tell: Leveraging smooth pursuit for assessing cognitive workload. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–13.

Eileen Kowler, Jason F Rubinstein, Elio M Santos, and Jie Wang. 2019. Predictive smooth pursuit eye movements. *Annual review of vision science* 5, 1 (2019), 223–246.

Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, November 20-21, 2008. Proceedings 4*. Springer, 63–76.

Philipp Lepold, Tobias Röddiger, Tobias King, Kai Kunze, Christoph Maurer, and Michael Beigl. 2024. OpenEarable ExG: Open-Source Hardware for Ear-Based Biopotential Sensing Applications. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 916–920.

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 (2015).

Hiroyuki Manabe and Masaaki Fukumoto. 2006. Full-time wearable headphone-type gaze detector. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems*. 1073–1078.

Hiroyuki Manabe, Masaaki Fukumoto, and Tohru Yagi. 2013. Conductive rubber electrodes for earphone-based eye gesture input interface. In *Proceedings of the 2013 International Symposium on Wearable Computers*. 33–40.

Mobvoi. 2025. *TicWatch E2*. Retrieved January 12, 2025 from https://www.mobvoi.com/tw/pages/ticwatche2

Mohsen Parisay, Charalambos Poullis, and Marta Kersten-Oertel. 2020. Felix: Fixation-based eye fatigue load index a multi-factor measure for gaze-based interactions. In *2020 13th International Conference on Human System Interaction (HSI)*. IEEE, 74–81.

Dulce G Pereira, Anabela Afonso, and Fátima Melo Medeiros. 2015. Overview of Friedman's test and post-hoc analysis. *Communications in Statistics-Simulation and Computation* 44, 10 (2015), 2636–2653.

Tobias Röddiger. 2023. *Earables: Wearable Computing on the Ears*. Ph. D. Dissertation. Dissertation, Karlsruhe, Karlsruher Institut für Technologie (KIT), 2023.

Klaus G Rottach, Ari Z Zivotofsky, Vallabh E Das, LEA Averbuch-Heller, Alfred O Discenna, Anuchit Poonyathalang, and R JOHN LEIGH. 1996. Comparison of horizontal, vertical and diagonal smooth pursuit eye movements in normal human subjects. *Vision research* 36, 14 (1996), 2189–2195.

Kristin L Sainani. 2012. Dealing with non-normal data. *Pm&r* 4, 12 (2012), 1001–1005.

Samuel Sanford Shapiro and Martin B Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52, 3-4 (1965), 591–611.

Junichi Shimizu, Juyoung Lee, Murtaza Dhuliawala, Andreas Bulling, Thad Starner, Woontack Woo, and Kai Kunze. 2016. Solar system: Smooth pursuit interactions using EOG glasses. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. 369–372.

Oleg Špakov, Poika Isokoski, Jari Kangas, Deepak Akkil, and Päivi Majaranta. 2016. PursuitAdjuster: an exploration into the design space of smooth pursuit–based widgets. In *Proceedings of the ninth biennial acm symposium on eye tracking research & applications*. 287–290.

Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2013a. 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*. 439–448.

Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2015. Pursuits: Spontaneous eye-based interaction for dynamic interfaces. *GetMobile: Mobile Computing and Communications* 18, 4 (2015), 8–10.

Mélodie Vidal, Ken Pfeuffer, Andreas Bulling, and Hans W Gellersen. 2013b. Pursuits: eye-based interaction with moving targets. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. 3147–3150.

Zeng Zhe, Felix Wilhelm Siebert, Antje Christine Venjakob, and Matthias Roetting. 2020. Calibration-free gaze interfaces based on linear smooth pursuit. *Journal of Eye Movement Research* 13, 1 (2020).

Yexu Zhou, Haibin Zhao, Yiran Huang, Till Riedel, Michael Hefenbrock, and Michael Beigl. 2022. Tinyhar: A lightweight deep learning model designed for human activity recognition. In *Proceedings of the 2022 ACM International Symposium on Wearable Computers*. 89–93.