DOI:
10.1109/TVCG.2024.3372032

Publication date:
2024

Document Version
Peer reviewed version

Privacy-Preserving Gaze Data Streaming in Immersive Interactive Virtual Reality: Robustness and User Experience

Ethan Wilson, Azim Ibragimov, Michael J. Proulx, Sai Deep Tetali, Kevin Butler, and Eakta Jain

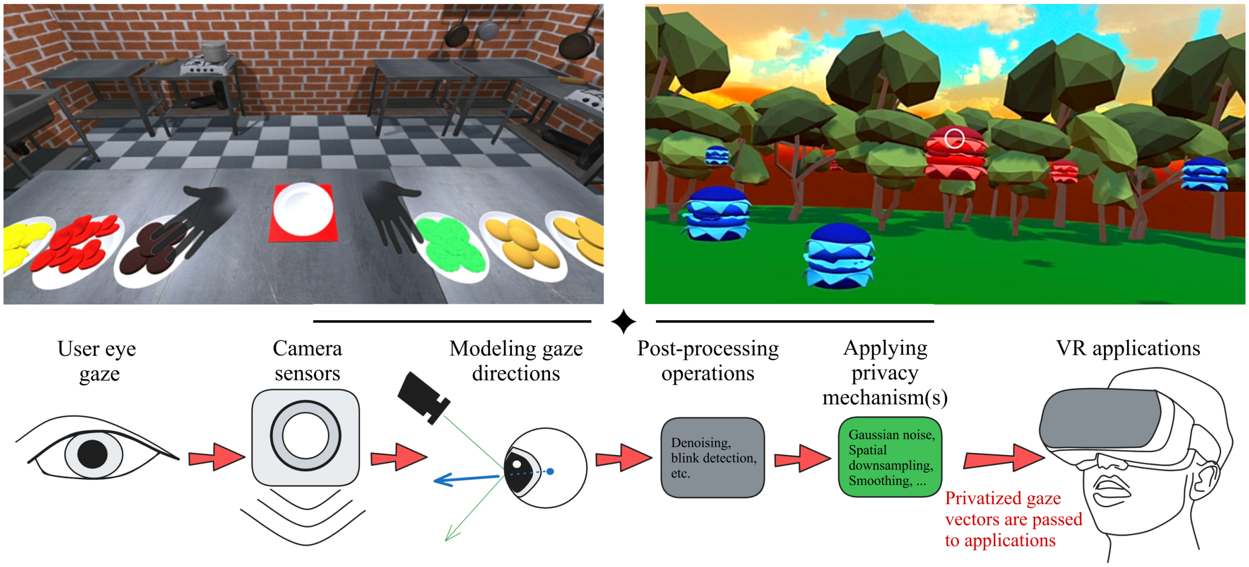
User eye gaze
Camera sensors
Modeling gaze directions
Post-processing operations
Applying privacy mechanism(s)
VR applications
Denoising, blink detection, etc.
Gaussian noise, Spatial downsampling, Smoothing, ...
Privatized gaze vectors are passed to applications

Fig. 1: **Top-left:** Participants underwent an interactive VR experience while eye movements were passively recorded; **Top-right:** Participants played an eye tracked VR game with privacy mechanisms applied to gaze streams; **Bottom:** An illustration of the eye tracking pipeline in head-mounted virtual reality systems. After eye images are processed and gaze vectors are calculated, we can apply privacy mechanisms to gaze vectors securely on the headset *before* passing values to potentially untrustworthy applications.

Abstract—Eye tracking is routinely being incorporated into virtual reality (VR) systems. Prior research has shown that eye tracking data, if exposed, can be used for re-identification attacks [14]. The state of our knowledge about currently existing privacy mechanisms is limited to privacy-utility trade-off curves based on data-centric metrics of utility, such as prediction error, and black-box threat models. We propose that for interactive VR applications, it is essential to consider user-centric notions of utility and a variety of threat models. We develop a methodology to evaluate real-time privacy mechanisms for interactive VR applications that incorporate subjective user experience and task performance metrics. We evaluate selected privacy mechanisms using this methodology and find that re-identification accuracy can be decreased to as low as 14% while maintaining a high usability score and reasonable task performance. Finally, we elucidate three threat scenarios (black-box, black-box with exemplars, and white-box) and assess how well the different privacy mechanisms hold up to these adversarial scenarios. This work advances the state of the art in VR privacy by providing a methodology for end-to-end assessment of the risk of re-identification attacks and potential mitigating solutions.

Index Terms—Virtual reality, privacy, eye tracking.

1 Introduction

Virtual reality (VR) technology has seen a rapid deployment of eye tracking-enabled headsets over the past several years. Eye tracking has many applications in VR, including as an interaction modality [9, 19, 23, 51, 54, 58], as an animation tool [53, 60, 65], for attention analysis [38, 76], for rendering optimizations [15, 26, 31, 44, 78, 79, 81], and for user authentication [28, 40, 41, 45, 57].

Recently, eye tracking movements alone have been found to function as a biometric identifier [22, 32]. Users can be uniquely identified in small directories; in VR, users have been shown to be identifiable at up to 85% accuracy [41]. This opens the risk of unwanted re-identification in online VR usage. Eye gaze is a promising input device showcasing very unique interactions and optimizations, but users should not have to choose between these interactions or their own privacy. If eye tracking data is shared with the proper safeguards, the risk of re-identification attacks or unwanted data leakage for VR users is alleviated.

Some solutions have been proposed to protect users against re-identification while still enabling eye tracking data utility. Existing analyses of these mechanisms focus on data-centric utility, including downstream processes such as area of interest analysis or gaze prediction. However, analyses focusing on interactive VR need to consider the user, who is likely to feel the effects when their eye tracking data streams are perturbed to grant privacy. It is necessary to consider the user first when designing privacy mechanisms that will be incorporated into interactive VR experiences. Users may chafe at adopting privacy

- *Ethan Wilson, Azim Ibragimov, Kevin Butler, and Eakta Jain are with the Computer & Information Science & Engineering department at the University of Florida. E-mail: {ethanwilson,a.ibragimov,butler,ejain}@ufl.edu.*
- *Michael J. Proulx is with Meta Reality Labs Research. E-mail: michaelproulx@meta.com.*
- *Sai Deep Tetali is with Meta Reality Labs. E-mail: saideept@meta.com.*

solutions that impact the interactive experience (which is the reason for being in VR in the first place). Another limitation to current eye tracking privacy methodologies is that they are framed as black-box threat scenarios. In real world settings, however, malicious 3rd parties have access to a number of strategies that aim to nullify the discussed privacy efforts.

Our contributions: We first update current knowledge on re-identification risk in interactive VR by using the state-of-the-art in eye movement identity matching and developing an interactive VR experience to serve as an evaluation testbed, re-identifying $N = 26$ users at an average accuracy of 67.3%. We measure the privacy capabilities of multiple mechanisms, identifying the following as capable of protecting privacy: Gaussian noise, spatial downsampling, and smoothing. Next, we incorporate these mechanisms into an interactive VR experience which uses eye tracking as the primary mode of control. We jointly measure re-identification accuracy, task performance, and qualitative subjective utility responses to firmly assess the privacy-utility tradeoff of these mechanisms. Our mechanisms decrease re-identification rates as low as 14.1% while retaining high subjective usability and reasonable task performance. Finally, we evaluate the mechanisms' robustness against dedicated adversaries under three plausible threat scenarios: black-box access, black-box access with exemplars, and white-box access. Our provided methodology can serve as a guide for future research of privacy mechanisms in interactive settings, measuring along multiple axes. The evaluated privacy mechanisms can be utilized as a basis for further innovation of novel privacy mechanisms.

Broader Impacts: This research contributes to the advancement of eye tracking privacy mechanisms, specifically those that must be applied to sample-level data in real-time. These mechanisms protect users against detection, especially populations that could be vulnerable if they are identified. Our work highlights a necessary shift in focus for the virtual reality research community from data-driven notions of utility to a user-centric design [48]. We additionally highlight that a calculated re-identification rate is only the first step; in situations where formal privacy guarantees can not be reached, we must consider real-world threat scenarios in order to proactively protect against adversaries. In addition, we make our collected datasets containing eye tracking data in interactive VR scenes available at https://doi.org/10.5281/zenodo.10475455.

2 Related Work

Eye tracking is becoming a prominent feature of VR experiences. Research on eye tracked VR systems began more than two decades ago [17, 72]. In recent years, many commercial VR head-mounted displays (HMDs) have released with embedded eye trackers (Magic Leap 1 [3] in 2018, Vive Pro Eye [7] in 2019, HoloLens 2 [2] in 2019, Vive Focus 3 [6] in 2021, Magic Leap 2 [4] in 2022, and Meta Quest Pro [5] in 2023. The Apple Vision Pro [1] is set to release in early 2024). These hardware advancements have created a surge in interest at the intersection between eye tracking and VR.

2.1 Applications of Eye tracking data in VR

Eye tracking enables many interactions in virtual scenes. These include using gaze to directly interact with virtual objects, improving social VR interactions, enabling foveated rendering optimizations, and gaze analysis as a research tool.

Gaze-based interaction: Eye tracking movements allow users to interact with virtual scenes, either on their own or paired with other control modes [19, 51]. Gaze direction can aim a cursor along with button presses to select objects [23, 58], or gaze fixations can be used for selection [9, 23]. For example, in a VR application interactable objects may glow when looked at, to indicate that they are dynamic. Then by fixating while pressing a button, users can select these objects[1]. The upcoming Apple Vision Pro will support gaze-controlled interfaces paired with pinching gestures [1]. Gaze-based interactions in desktop games are found to be more efficient and immersive than traditional control modes [23, 54], which may persist into VR[2].

Rendering Optimization: Foveated rendering is a critical optimization to increase resolution and frame rate of VR headsets [25, 44]. Foveated rendering sparsely renders samples outside of the fovea region, which is determined through eye tracking signals. Because peripheral vision has lower acuity than foveal vision, the image could be perceptually similar to traditional rendering but vastly less expensive to compute [31, 78, 79]. Gaze prediction algorithms [15, 26, 81] will enable proactive foveated rendering and occlusion optimizations.

Avatar Animation: Recorded gaze can be used to drive eye animations in VR. More realistic eye movements have been shown in improve the quality of interaction with virtual avatars [21] and to increase perceived presence and avatar realism [65]. Gaze can enable virtual avatars to display trust across multiple expressions and contexts [53], and multiple personality traits can be discerned solely through characters' eye motions [60]. By incorporating real gaze behaviors to embodied avatars, each avatar feels more unique and personable.

Gaze-based Analytics: Eye tracking data can be a rich tool for data analysis [67]. Examples include area of interest (AOI) calculation [14, 55], document classification and analysis [12, 62], and attention visualization [37]. Researchers from multiple fields use eye movements to analyze topics such as social behavior [59, 83], visual attention [38, 76], and simulated responses under stress [68].

2.2 Identification Risk of Recorded Eye Movements

Though users can be identified based on several cues, including head movements, body movements [50] and gestures [49], our focus is identification based on eye movements. Iris patterns are a well known biometric identifier [52], and David-John et al. examined user re-identification using iris images and presented solutions to mitigate this risk [29, 30]. Note that the Meta Quest Pro headsets used in this study do not pass on eye images or raw data to the applications. In addition to hand-crafted features derived from gaze streams [22, 32, 42, 64], there are now deep-learning methods to classify users based on short windows of eye movement data [28, 40, 41, 45, 57]. Eye Know You Too (EKYT) is currently the top performing eye movement identification model, reporting accuracies as high as 91.38% on 1000Hz data [41].

Physical and behavioral attributes such as personality [10], age [82] or gender [61] have been inferred from eye movements. Some research leverages eye movements to aid in medical diagnoses such as Autism or Alzheimer's [71, 74, 75]. While there are appropriate use cases to learn this information, users can not consciously hide the attributes embedded within eye tracking streams. The opportunity here is develop methods to block these features from being extracted without users' consent by malicious entities who acquire eye tracking data.

The threat of eye movement re-identification is larger for small sets of users [20]. This may be a particular concern for marginalized users who face disproportionate harm when privacy is compromised [63].

2.3 Gaze Data Privacy

Privacy mechanisms are mainly applied in three ways to the eye tracking pipeline. Aggregate-level mechanisms protect full datasets with operations that average across multiple users' data [11, 13, 37, 70]. Feature-level mechanisms protect users by converting raw gaze signals to features and applying privacy [11–13, 70]. Sample-level mechanisms operate on the actual data streams, perturbing gaze direction at every frame [14, 35]. In VR, we are mainly interested in sample-level mechanisms [66], which could be applied securely by the VR platform before eye tracking data is made available to third party applications [14]. With privacy mechanisms in place, users can experience novel interactions and optimizations only possible with sample-level gaze streams without risking information leakage. See Table 1 for a collection of eye tracking privacy work.

[1] https://www.uploadvr.com/polyarc-moss-psvr-2/

[2] A list of VR games which incorporate eye tracking: https://www.psfanatic.com/here-are-all-the-psvr2-games-that-use-eye-tracking-in-cool-ways/

Application	Data Domain	Data Format	Mechanism	ID Accuracy	Data-centric Utility	User-centric Utility
Gaze-based analytics	Constrained VR	Sample-level	(Kaleido) Spatial noise with adaptive sampling [13]	28% to 6%	Activity classification	-
		Sample-level	Gaussian noise [14]	85% to 30%	Dwell time RMSE	-
		Sample-level	Temporal downsampling [14]	85% to 79%	Dwell time RMSE	-
		Sample-level	Spatial downsampling [14]	85% to 48%	Dwell time RMSE	-
		Sample-level	Gaussian noise [14]	33% to 9%	KL-divergence of saliency maps	-
		Sample-level	Temporal downsampling [14]	9% to 7%	KL-divergence of saliency maps	-
		Sample-level	Spatial downsampling [14]	47% to 29%	KL-divergence of saliency maps	-
	Conventional eye tracking	Aggregate-level	Difference- and chunk-based Fourier perturbation [11]	100% to 28%	Document classification; gender classification	-
		Aggregate-level	Exponential mechanism applied to features [70]	100% to ∼10%	Document classification; gender classification	-
		Aggregate-level	k-same-synth [13]	28% to 7.5%	Activity classification	-
		Aggregate-level	Event-synth-plausible deniablity [13]	28% to 14.2%	Activity classification	-
Gaze-based interaction	Webcam eye tracking	Sample-level	(Kaleido) Spatial noise with adaptive sampling [35]	∼84% to ∼8%	scan path similarity; latency trade-off	Game enjoyment; task performance
	Interactive VR	Sample-level	**Ours**	67.3% to 14.1%	Area of interest retention	Subjective usability; task performance

Table 1: Collection of eye tracking privacy work that successfully protected against re-identification while retaining one or more measure(s) of utility.

Real-time privacy operations will be critical to ensure privacy during online eye tracking interactions [9,58] and to enable optimizations [44], but have yet to be explored in real-time VR settings. The goal of privatization is to protect sensitive attributes while keeping the data usable with respect to a given task. Existing research primarily measures data-driven utility via post-processing tasks, such as gaze-based analytics [11,12,37] or rendering optimization [15]. However, the impact of privacy mechanisms on the user's performance and subjective experience in interactive VR has not been considered prior to this work.

3 Re-identification in Interactive VR

We begin by establishing the risk of re-identification in interactive VR and identifying viable privacy mechanisms. We collect a dataset of eye tracking-enabled interactive VR tasks to serve as an evaluation testbed. We then quantify the re-identification risk on our dataset using the state of the art architecture and define and evaluate multiple privacy mechanisms on our dataset.

On conventional eye trackers at high frequencies, users can be identified at very high accuracies (91.38% identification rate, 3.66% equal error rate) [41]. In VR settings, it is less clear how reliably users can be identified, due to less precise sensors and extraneous user movements. An evaluation of 360° VR image and video datasets using a prior identification method [22] yielded identification rates ranging from 9% to 85% [14]. The same analysis also evaluated an interactive dataset where users viewed a scene of moving animals [26], yielding only a 3% identification rate. The level of interactivity and amount of user movement in VR setups could negatively correlate with the potential to be identified in VR [73]. In this paper, we present an up-to-date evaluation of re-identification risk in interactive VR tasks using the state-of-the-art identification model.

For this analysis, we distinguish eye tracking data sources into three categories. *Conventional eye tracking* utilizes high quality static sensors at 1000Hz or greater [24], and produces highly identifiable data [41]. Eye movements collected in VR can be separated into two categories: *Interactive VR* is presented in a natural way. Users directly interact with dynamic objects in the virtual scene, and different users experience the scene at their own rate. This is reflective of consumer applications, such as VR games or dynamic training scenarios. On the other hand, *constrained VR* is representative of existing experimental setups. User movements are limited, such as sitting in a chair [69] or placing the head on a chinrest [39], and the tasks are standardized such that all users experience the same stimuli at the same rates. In this section, we answer the following research questions:

- How do current state of the art eye movement re-identification algorithms perform in interactive VR?
- What real-time privacy mechanisms are effective at protecting against re-identification in interactive VR tasks?

We present an updated evaluation of the risk of identification in VR. Prior work only considered identification risk on constrained VR setups using models trained on hand-crafted features [14]. We first construct a dynamic VR game representative of interactive VR. We then present an analysis of the current state of the art architecture trained both on conventional eye tracking data and VR data, evaluating on conventional eye tracking data (GazeBase [24]), constrained VR data (GazeBaseVR [39]), and interactive VR data (our dataset). We then discuss our results, giving insights to the current risk of identification in consumer VR and the relationship between re-identification potential and the amount of data made available. Using the same dataset, we then evaluate multiple privacy mechanisms that can be applied to protect eye tracking data in VR. We evaluate these mechanisms across increasing intensities to derive well saturated privacy curves.

3.1 Data Collection Methodology

We describe the protocol for our collected dataset which serves as a testbed to evaluate identification in interactive VR.

Participants: Survey participants were recruited under IRB approved protocol via several communication channels including word of mouth and electronic mailing list advertisements ($N = 26$; 57.69% male, 42.31% female). No monetary compensation was provided, but some participants received extra credit for undergraduate courses. Eligible participants required normal or corrected-to-normal vision without the use of eye glasses. The racial-ethnic distribution is 61.54% White, 19.23% Indian Asian, 15.38% Black or African American, 7.69% Eastern Asian, and 7.69% Hispanic/Latino; 11.54% of participants report two or more races. 34.62% of participants were age 18-20, 50% 21-29, 11.54% 30-39, and 3.85% 50-59. 88.46% of participants reported some level of experience with VR, and 30.77% reported some experience with using eye tracking as a control mode.

Procedure: Participants were instructed to act as employees in a sandwich shop, and were tasked with assembling as many sandwiches as possible in a 90 second time frame. Plates of stacked ingredients were organized on the sandwich assembling counter on each side. In the center of the counter, participants would assemble their sandwiches on an empty plate (See Figure 1a). After a sandwich was completed, a small animation would play and the plate would be cleared, allowing participants to begin the next sandwich. A digital timer could be seen which displayed the amount of time remaining.

Before beginning the main task, participants were encouraged to practice picking up and assembling ingredients. In the experiment, participants did not use controllers; instead, grabbing was driven by hand tracking technology to provide a more immersive experience [36]. Participants underwent 4 trials of the same task and were allowed time to rest between trials.

As participants performed the tasks, gaze data was collected passively to evaluate the potential of re-identification attacks and to store area of interest (AOI) data. Each ingredient plate, the assembly plate, and the timer were all AOI regions with fully-covering bounding boxes being used for AOI collision detection.

Validation: Data was collected using the Meta Quest Pro [5] (1920 x 1800 pixels per eye, 72Hz refresh rate). Before the experiment, each participant underwent the headset's eye tracking calibration procedure. Participants gazed directly at spheres which appeared at random points on the screen and gradually shrank until no longer visible. To validate the accuracy of collected gaze data, participants performed an eye tracking validation task. Participants faced a 3×3 checkerboard of red targets situated 2 meters away spanning a 38.58° angle vertically and horizontally. Participants were instructed to gaze directly at the active target, which would become green. The active target cycled uniformly; each target was active for 2 seconds total, with no downtime until the next dot became active. We report a spatial accuracy error of $\mu = 2.64°$, $\sigma = 1.24$. This protocol is comparable with recent analysis of the Quest Pro [77].

3.2 Privacy Mechanisms

Collected gaze streams over a full VR session can be represented as a time series of gaze angles in spherical coordinates θ, ψ and their corresponding time stamps t: $X = \{(\theta_0, \psi_0, t_0), (\theta_1, \psi_1, t_1), \cdots, (\theta_n, \psi_n, t_n)\}$. The gaze angles are localized relative to the recorded position of the headset, i.e. head pose, thus are constrained roughly to the human field of view. As our privacy mechanisms are implemented in real-time, the operations are applied directly on the frame which the gaze is sampled.

We implement three privacy protection mechanisms proposed by David-John et al. [14]: additive Gaussian noise, temporal downsampling, and spatial downsampling. We also introduce linearly weighted average smoothing as a novel mechanism. These mechanisms are designed to be feasible in real-time settings, suitable for streamed sample-level eye tracking data in VR applications.

Gaussian Noise: Gaussian noise is sampled independently for the θ, ψ gaze angles for every frame. We use the standard deviation of the noise sample σ to control the privacy-utility trade-off provided by the method, yielding the following per-frame operation: $X'_n = (\theta_n + x \sim \mathcal{N}(0,\sigma), \psi_n + y \sim \mathcal{N}(0,\sigma), t)$.

Temporal Downsampling: Temporal downsampling effectively lowers the sampling rate of a stream of data by a factor K. Data entries with indices where K is not a factor are removed from the data stream, yielding a stream with N/K total entries after downsampling.

A real-time application may expect a gaze vector at all frames. Thus, we preserve the original data format by simply copying the gaze directions from the prior time step on frames that would traditionally be removed. Given a downsampling factor K,

$$X'_n = \begin{cases} (\theta_n, \psi_n, t) & \text{if } n \,\%\, K = 0 \\ (\theta'_{n-1}, \psi'_{n-1}, t) & \text{otherwise} \end{cases}$$

Spatial Downsampling: Spatial downsampling lowers the spatial resolution of the gaze data, e.g., multiple nearby full-resolution points would be mapped to a single down-sampled point, lowering the spatial fidelity. To apply spatial downsampling to continuous gaze angles, we first map the angles into a set of discrete points large enough to preserve data quality, choosing 2160 points to cover a 180° field of view. We achieve spatial downsampling by remapping the gaze angles into a smaller domain equal to the reference domain divided by L. We map the θ, ψ angles into the discrete domain of $M = 2160/L$, providing a step size of $\delta = \frac{180°}{M}$. This yields the following per-frame operation: $X'_n = (\lfloor \theta/\delta \rfloor \cdot \delta, \lfloor \psi/\delta \rfloor \cdot \delta, t)$.

Smoothing: We introduce a smoothing operation as a novel privacy protection mechanism. Smoothing streamed gaze data can remove identifiable features without displacing individual samples in a jarring way as the above mechanisms can. Because of this, we hypothesize that users will be accepting of smoothing; they can consciously correct for the gaze stream's behavior by fixating at an objects for a second longer, for example.

Algorithm 1 Linearly weighted average smoothing

```
$B \leftarrow$ size of window
window $\leftarrow Queue()$        ▷ Note that this Queue pops at the $0^{th}$ index
$D \leftarrow 0$
for $i \leftarrow 1$ to $B$ do        ▷ Initializing the window
    window.$add([0,0])$
    $D \leftarrow D + i$
end for
while application is running do
    $X \leftarrow$ current gaze vector
    window.$pop()$
    window.$add(X)$
    $X' = (0,0)$
    for $i \leftarrow 1$ to $B$ do
        $X'[0] \leftarrow X'[0] +$ window$[i][0] * i$
        $X'[1] \leftarrow X'[1] +$ window$[i][1] * i$
    end for
    $X' \leftarrow X'/D$
end while
```

To operate in real-time, we define the current gaze vector as a linearly weighted average of B preceding samples. Preceding samples are stored in a sliding window that is updated every frame. The smoothed value is a weighted average of the B values in the window, with each sample weighed by its index in the buffer.

$$X'_n = \frac{X_{n-B} + 2(X_{n+1-B}) + 3(X_{n+2-B}) + \cdots + B(X_n)}{\sum_i^B (i)}$$

Larger window sizes equal a more intense smoothing operation, which is more successful at removing identifying features, but introduces a larger temporal delay between the input and output gaze vectors. Implementation details are shown in Algorithm 1.

3.3 Evaluation Methodology

We evaluate the identification potential of EKYT models trained on conventional eye tracking data (GazeBase [24]) and constrained VR data (GazeBaseVR [39]). Both models are trained at 125Hz to best match the frequency of our collected data. We follow the training and testing methodology of Lohr et al.[3] [41], and present the rank-1 identity retrieval rate averaged across all tasks in the dataset.

To our knowledge, there is no existing interactive VR dataset at the scale required to train the EKYT model effectively. Therefore, we assess the effectiveness of training on constrained VR data and evaluating on interactive VR data, versus using a model solely trained on conventional eye tracking data.

To evaluate the identification potential on our collected data, we define the following protocol. First, embeddings are generated from the raw eye movement data. Each 90-second trial of the game is separated into 5-second segments sliding over a 1-second interval. These 5-second segments are linearly interpolated to a constant 125Hz and processed by the EKYT model to create $512D$ feature embeddings Emb, which are stored along with labels L for the individuals as records.

Records are separated into folds according to the trials of the game. So, all records corresponding to the first trial exist in the first fold, and so on. We then compare pair-wise each fold as a distinct query and reference set. For each individual, all query records are compared with the records in the embedding set. So, for N individuals and M records per individuals, $Query_n = \{Emb_{n,1}, Emb_{n,2}, \ldots\ Emb_{n,m}\}$ and reference $Ref = \{E\hat{m}b_{n,m} : \forall n \in N, \forall m \in M\}$. The record from the query is

[3] https://dataverse.tdl.org/dataset.xhtml?persistentId=doi:10.18738/T8/61ZGZN

Duration	Test Data	Train Data	
		Conventional Eye Tracking	Constrained VR
5 s	Conventional Eye Tracking	66.09%	18.04%
	Constrained VR	27.04%	33.87%
	Interactive VR	22.44%	26.92%
10 s	Conventional Eye Tracking	81.42%	32.36%
	Constrained VR	36.00 %	45.14%
	Interactive VR	19.23%	22.76%
30 s	Conventional Eye Tracking	90.00%	54.58%
	Constrained VR	48.57%	55.42%
	Interactive VR	40.06%	41.35%
60 s	Conventional Eye Tracking	90.78%	58.57%
	Constrained VR	49.98%	55.09%
	Interactive VR	56.73%	54.81%
90 s	Interactive VR	68.27%	67.31%

Table 2: Average re-identification accuracy using the EKYT architecture [41] trained and evaluated on different data types. All models are trained at 125Hz.

matched to the closest reference embedding using cosine similarity as a distance metric and the associated label is stored. For an individual n, the predicted label $L_{p,n}$ is:

$$\begin{aligned} L_{p,n} &= \underset{L}{\operatorname{argmax}} \left(\sum_{m}^{M} L_z \right) \\ \text{where} \quad z &= \underset{m}{\operatorname{argmin}} \left(\sum_{n'}^{N} \sum_{m'}^{M} \left(1 - \frac{Query_{n,m} \cdot Ref_{n',m'}}{||Query_{n,m}|| \; ||Ref_{n',m'}||} \right) \right) \end{aligned} \tag{1}$$

The returned labels are aggregated between each query embedding $Query_{n,m}$ and all embeddings in the reference set to determine a final prediction $L_{p,n}$. If $L_{p,n}$ equals the true label L_n, the individual has been successfully identified.

The reported metrics are an average over all individuals and all pairwise combinations of distinct trials, readable as an average identification accuracy over the full dataset, or the likelihood of any individual being successfully identified.

When evaluating privacy mechanisms, before processing the set of trials which make up the query set, the mechanism is applied at a given strength. The privatized results are then processed and compared to the reference set. This represents the following threat scenario:

> *An adversary acquires a privatized data record without knowing the identity, then queries against a dataset of records which have known identities attached.*

To provide an initial data-centric representation of utility, we measure AOI intersections before and after privatization. We calculate the multi-class weighted precision and recall and overall F1 score which can be interpreted as the mechanism's ability to retain original AOI behavior and after privatization.

3.4 Results

We organize our results according to the initial analysis of EKYT models on across conventional and VR data domains and the performance of privacy mechanisms on our collected dataset.

3.4.1 Re-identification Capability of EKYT Models

Our findings are reported in Table 2. We report the accuracies between the 2 training and 3 test setups. We additionally report results varying the duration of each data record. These results are calculated using the explained methodology but first limiting the amount of data to the first N seconds per data record.

We find that the model trained on conventional eye tracking data performs well when evaluated in the same domain, reaching identification accuracies of 66.09% and 90.78% on 5 second and 60 second data records, respectively. However, the conventional model performs much less effectively in when applied to VR, achieving 27.04%@5s and 49.98%@60s on constrained VR and 22.44%@5s and 56.73%@60s on interactive VR data.

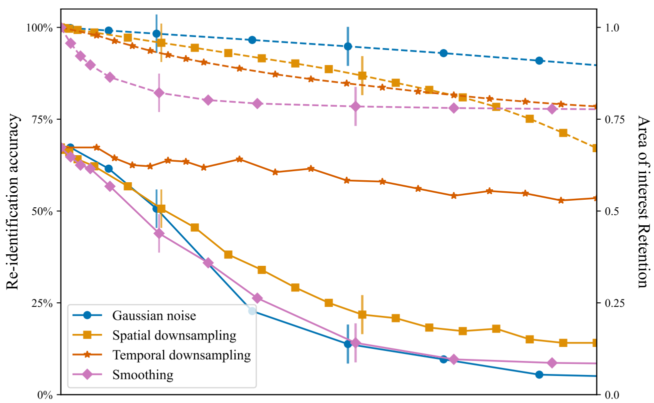

Fig. 2: Identification accuracy (solid lines) and AOI retention (dotted lines) of privacy mechanisms applied at various strengths. X axes have been scaled to provide roughly the same privacy falloff so that utility can be directly compared. Vertical lines indicate the chosen low and high strengths of each mechanism.

On the other hand, the constrained VR model favors VR data at a lower duration (18.04%@5s for conventional, 33.87%@5s for constrained VR, 26.92%@5s for interactive VR) but performs roughly equal on all domains at higher duration (58.57%@60s for conventional, 55.09%@60s for constrained VR, 54.81%@60s for interactive VR).

We find performance on interactive VR data to be roughly equal between the two models. While the conventional eye tracking model is trained on a larger dataset, the constrained VR model has closer spatial precision and more similar setup to our data. For the rest of this paper, we use the constrained VR EKYT model to compute identification accuracies.

3.4.2 Privacy Mechanism Performance

We compute identification accuracies for each mechanism at multiple strengths to define a privacy curve and measure against AOI F1 scores to estimate the anticipated utility trade-off. We measure Gaussian noise up to $\sigma = 20°$, spatial downsampling up to $L = 256$, temporal downsampling up to $K = 30$, and smoothing up to $B = 300$. The privacy/ utility trade-off can be seen in Figure 2.

Of the previously proposed mechanisms, additive Gaussian noise and spatial downsampling are effective at providing privacy. The proposed smoothing mechanism also provides privacy, albeit at a higher trade-off in AOI retention. We have not found temporal downsampling to be an effective privacy mechanism. It is likely that by retaining N/K real samples, the model can still associate real samples and distinguish users until K reaches a point far beyond usable utility.

We introduced smoothing as a novel mechanism to protect user eye movements. On our data, smoothing shows potential for privacy protection with similar levels of privacy attained as Gaussian noise and spatial downsampling. However, AOI retention is lowest for smoothing on the tested dataset, indicating that an application receiving smoothed data would have less reliable accuracy per-sample than other mechanisms. However, as we will see in Section 4, lower AOI retention is not reflective of smoothing's impact to user experience.

4 Evaluating User-centric Utility

The majority of privacy literature surrounding eye tracking movements have considered privacy as a post-process operation. However, in the context of interactive VR, privacy must be provided *before* being passed to the application; the user subsequently sees the impacts of privatization. It is important to consider the user and the trade-off in *user-centric utility*, rather than solely relying on data-centric metrics. In this section, we answer the following research question:

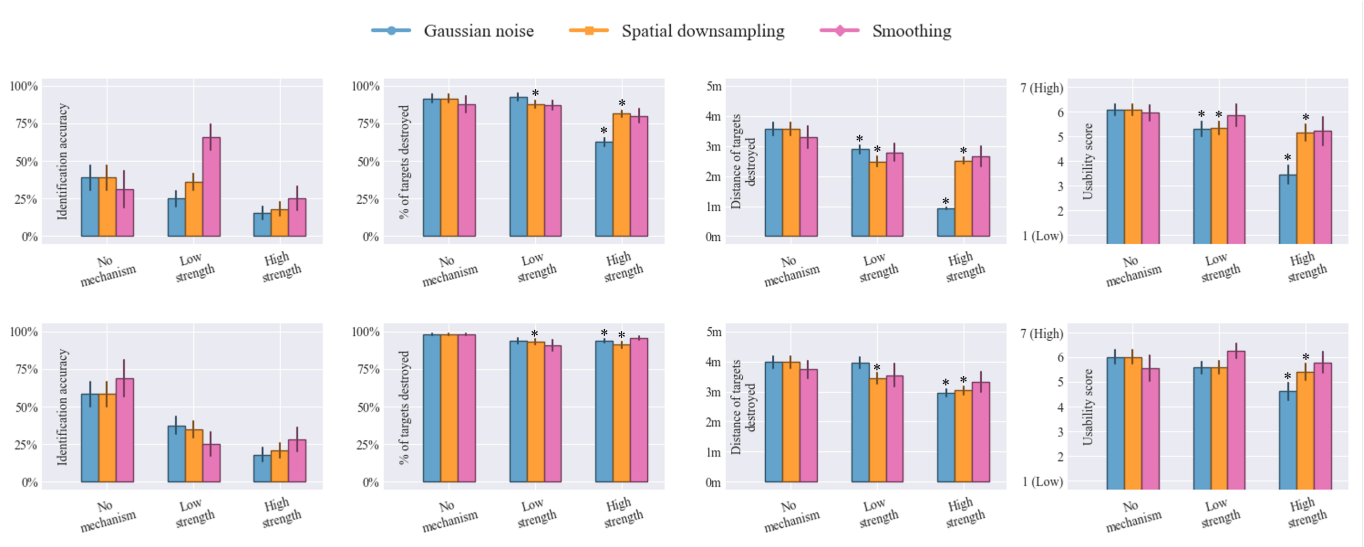

Fig. 3: Collected metrics of the immersive VR game with gaze controls. For identification accuracy (first column), lower values indicate more privacy. For all utility metrics, a higher score indicates higher utility. Each subplot illustrates no privacy mechanism compared against the low and high strengths of each mechanism. Wilcoxon signed-rank test significance for utility metrics ($p < 0.05$) are denoted with color-coded asterisks. Vertical lines indicate Standard Error of the Mean (SEM) $= \sigma/\sqrt{N}$.

- What is the impact to user-centric utility metrics when eye tracking privacy mechanisms are applied live in interactive VR?

We investigate user-centric utility through an interactive VR experience which uses eye gaze as the primary control mode. Before being processed by the application, the eye movement data is processed by the 3 viable privacy mechanisms: Gaussian noise, spatial downsampling, and smoothing.

4.1 Data Collection Methodology

We describe the protocol for our collected dataset in which privacy mechanisms were applied in real-time, allowing us to measure user-centric utility after privatization in interactive VR.

Participants: Survey participants were recruited under IRB approved protocol via several communication channels including word of mouth and electronic mailing list advertisements ($N = 18$; 77.78% male, 22.22% female). Eligible participants required normal or corrected-to-normal vision without the use of eye glasses. The racial-ethnic distribution is 33.33% White, 22.22% Indian Asian, 11.11% Black or African American, 16.67% Hispanic/Latino, 5.56% Eastern Asian, 5.56% Central Asian, and 5.56% Middle Eastern. 27.78% of participants were age 18-20 and 66.67% 21-29. 88.89% of participants reported some level of experience with VR, and 16.67% reported some experience with using eye tracking as a control mode. The smoothing mechanism was incorporated into the study after 10 participants had undergone a version with only Gaussian noise and spatial downsampling, so for all measures of smoothing, there are $N = 8$ participants represented. This detail is addressed in our results.

Procedure: Participants played a first person shooter game where they remained in a static position and defeated enemies that periodically spawned and travelled towards the player. Two types of entities would alternatively spawn in random positions and move towards the player. Friendly entities served as visual distractors [33] and enemy entities served as targets; participants were instructed to destroy targets before they could reach the player. Entities would spawn randomly from 7 uniformly spaced points on a 90° arc spaced 10 meters from the player position. Participants could see a translucent gaze cursor indicating their current gaze direction. In trials with privacy mechanisms applied, the cursor illustrated the effective gaze direction after privatization. See Figure 1b for a participant's typical view.

Each trial of the game lasted for $\sim$30 seconds. 30 entities (15 targets and 15 distractors) spawned, increasing in speed from 1 m/s to 5 m/s over the duration of the trial. By increasing difficulty over the course of the trial we can derive useful performance metrics from participants regardless of skill level. For each experimental condition, participants would undergo two trials of the game sequentially, followed by the Post-Study System Usability System Usefulness subscale (PSSUQ SYSUSE) [34].

We adopted a within-subjects design; all participants underwent every condition. Privacy mechanism conditions were either no privacy mechanism applied or {Gaussian noise, spatial downsampling, smoothing} at {low, high} strength, presented in random order. The privacy mechanism strengths chosen were derived from the re-identification accuracies found in initial analysis the data presented in Section 3.4.2[4]. For Gaussian noise, low $\sigma = 1°$ and high $\sigma = 3°$. For spatial, low $L = 48$ and high $L = 144$. For smoothing, low $B = 50$ and high $B = 150$.

We implement two control modes for facilitating the selection of targets. The first control mode is *gaze-only*; if the gaze vector from the participant's left eye was consistently within a target's bounds for 500 ms, the target was destroyed. The other control mode is *gaze-plus-gesture*; participants' gaze vectors indicated the selection of targets. If, for a given frame, the player gaze vector intersected with a target's AOI and the participant performed a pinch gesture with their left hand, the selected target was destroyed.

Participants underwent all privacy mechanism conditions in random order with one control mode, then all privacy conditions again in random order with the remaining control mode. To mitigate an order-effects bias, the order of control mode was counterbalanced among participants.

Validation: Data was collected using the Meta Quest Pro [5]. Before undergoing the experiment, each participant underwent the same validation protocol as in Section 3.1, but this time underwent 3 trials rather than 1 (spatial accuracy error $\mu = 2.78$, $\sigma = 1.45$). Participants also performed gesture validation. Again facing a 3×3 board of targets, participants were instructed to place a cursor at the center of the field of view over the target and make a pinching gesture to confirm their placement (error $\mu = 2.94$, $\sigma = 1.68$). While gestures were not explicitly calibrated per participant, this served as a primer for participants so that they were familiar with the headset's gesture recognition before undergoing the main tasks.

[4] Using preliminary results from Section 3.4.2, the parameters which initially decreased re-identification accuracy below 40% and 20% were chosen for low and high strengths.

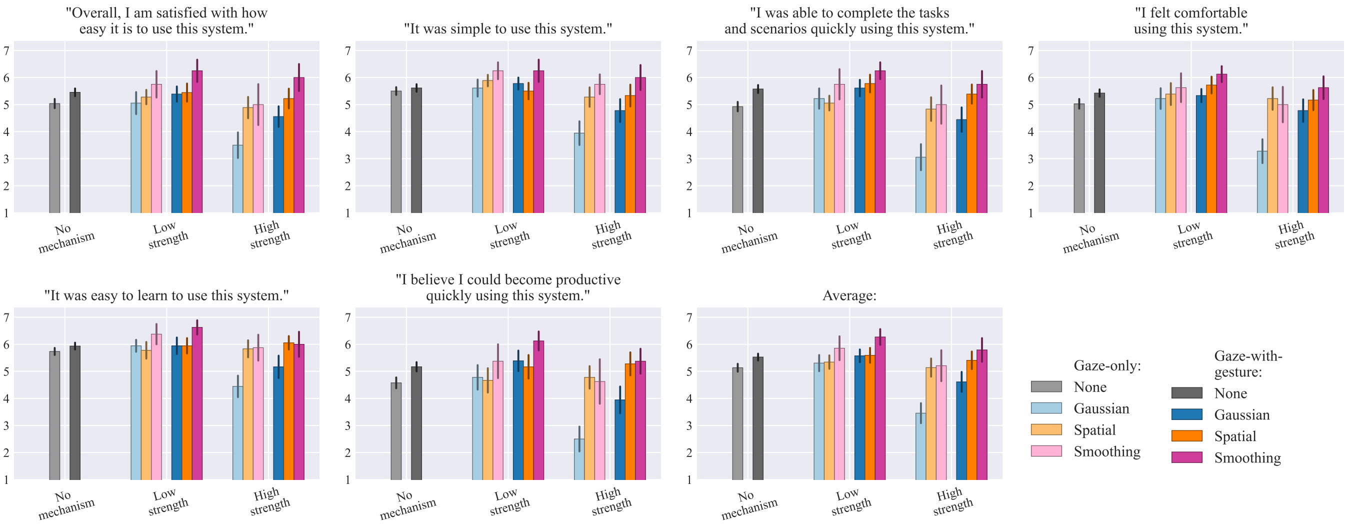

Fig. 4: Breakdown of PSSUQ SYSUSE scores [34] across conditions by individual question. Vertical lines indicate SEM.

4.2 Evaluation Methodology

We analyze identification accuracies alongside multiple notions of user-centric utility to clearly measure whether the addition of privacy mechanisms have any negative impact to users. These metrics are:

- **% of targets destroyed:** This is a simple measure of task performance, conveying on average how many targets were successfully destroyed before reaching the participant.
- **Average distance of targets destroyed:** Also measuring task performance, this metric gives a better estimate of how easy/challenging the task was for users.
- **Usability score:** The average PSSUQ SYSUSE response, representing users' perceived satisfaction with the given control mode and privacy mechanism combination.

4.3 Results

We present our results across identification accuracy and the task-specific notions of utility defined above. Note that when computing identification accuracy, we pair-wise compare the first and second trial of a given mechanism against those of no privacy mechanism to form our query and reference setups. We test for statistical significance using the Wilcoxon signed-rank test after collapsing the results of all game trials to a single data point per participant and condition, reporting values of $p < 0.05$ as significant. We compare against the non-privatized trials when testing significance. Results are visualized in Figure 3.

We first see a lower identification accuracy ($\mu = 48.61\%$) on this dataset than the dataset of Section 3.1. This is to be expected, as session lengths are ∼30 seconds rather than 90. In this new experimental setup, the high strength privacy mechanisms all provide a noticeable level of privacy.

The percentage of targets destroyed and distance targets are destroyed at function as measurements of task performance. For gaze-only controls, we see a decrease in performance across all privacy mechanisms. For the % of targets destroyed, Gaussian ($z = -3.73$) and spatial ($z = -3.01$) are significantly lower than the baseline at high strengths. For the average distance, Gaussian ($z = -2.94$) and spatial ($z = -3.72$) are significant at low strength, and Gaussian ($z = -3.72$) and spatial ($z = -3.1$) are significant at high strength. It is more difficult for the application to measure continuous fixations with the perturbed data, so performance decreases. Interestingly, we see less of an effect when participants used gaze-plus-gesture controls. For % of targets destroyed, spatial ($z = -2.57$) is significant at low strength, and Gaussian ($z = -2.98$) and spatial ($z = -3.06$) are significant at high strength. For average distance, spatial ($z = -2.63$) is significant at low strength and Gaussian ($z = -3.03$) and spatial ($z = -2.98$) are significant at the high strength. Generally, the average percentage of targets destroyed remains high, though the average distance decreases. This indicates that the task became slightly more difficult but remained trivial to complete. Across both task performance metrics, we see Gaussian noise utility decrease at a high rate compared to the other mechanisms.

Focusing on usability, we see noticeable differences in average performance (illustrated in Figure 4). For gaze-only controls, Gaussian ($z = -2.82$) and spatial ($z = -2.74$) are significant at low strength and Gaussian ($z = -3.64$) and spatial ($z = -2.59$) are significant at high strength. For gaze-plus-gesture controls, Gaussian ($z = -3.59$) and spatial ($z = -2.1$) are significant only at high strength. Gaussian noise again impacts utility at a higher rate than other mechanisms. Interestingly, smoothing seems to slightly increase usability when applied at a low strength. There could be a low amount of noise present in the eye tracker's raw data stream which smoothing corrects. Yet, there is not sufficient statistical evidence to prove this claim. For both non-Gaussian mechanisms, overall usability remains high even after applying the high strength variant of the mechanisms.

We conducted a three-way ANOVA on the participants which underwent all privacy mechanisms ($N = 8$) to examine the effect of privacy mechanism, strength, and type of control mode on the average PSSUQ response. We find the main effects of mechanism and strength to be significant ($p < 0.01$). There are significant interactions between mechanism and strength ($F(4,28) = 7.3979, p = 0.0003$) and between mechanism and control mode ($F(2,14) = 4.2953, p = 0.0351$). These findings evidence that the different control modes could be affected by the application of privacy mechanisms disproportionately.

Note that because smoothing was implemented midway through data collection, statistical tests of smoothing consider $N = 8$ participants, thus there is less statistical power than other conditions. However, across utility metrics, smoothing appears comparable to spatial downsampling across all categories with comparable variance.

5 Anticipated Adversarial Threats

So far, we have presented an analysis of the identification potential in interactive VR and presented privacy solutions which can mitigate the risk of identification while retaining user-centric utility. However, this is only the first step. We must also consider the robustness of any privacy solution, as a dedicated adversary will make efforts to counteract and nullify any privacy-preserving operation.

The eye tracking community has explored formal privacy guarantees such as differential privacy (DP) [11, 37, 70], k-anonymity or plausible deniability [12, 13]. However, as the privacy guarantee is tied to high level features and applied to full collections of users, these methods are not suitable for sample level eye tracking data being privatized in real time. Kalεido has developed a sample-level method with DP guarantees [35], but has a high overhead (8ms), and 15-20ms overall latency in VR can introduce sickness and nausea [8]. Thus, privacy mechanisms proposed for live gaze-based interactions must be proactively evaluated against adversarial threats. In this section, we answer the following research question:

- Are the proposed privacy mechanisms robust against malicious adversaries?

We address the robustness of our proposed mechanisms under three realistic threat models in which adversaries have varying levels of domain information. We then define and evaluate an example attack under each threat model. For this analysis, we use the larger dataset defined in Section 3.1 and implement privacy mechanisms at their high strength.

5.1 Threat Models

We explore multiple threats to eye tracking re-identification that we expect to become plausible in the next decade as eye tracking technology and VR become more mainstream. Our threat models are organized according to the information or resources that the adversary has access to; as an adversary becomes more informed of the privacy mechanism, they become increasingly able to counteract the privacy efforts. We conceptualize an example scenario and attack for each threat model.

In all cases, we make the assumption that the adversary's goal is to obtain the identity of an acquired query gaze stream. There may be sensitive information connected to either the query gaze stream or existing dataset records, and a successful attack can link the user's identity or quasi-identifier(s) to the sensitive information. These sensitive attributes could be concrete records, such as health information or group membership, or could be implicit knowledge embedded within in the actual gaze stream (such as personality [10], age [82] or gender [61]).

Black-box Access: In this scenario, an adversary has acquired a privatized gaze stream but has no knowledge regarding the mechanism applied. The adversary can attempt to query the privatized record against non-privatized records sourced from elsewhere, such as public datasets. Before querying, it is possible for the adversary to perform a filtering operation in an attempt to render the privacy mechanism ineffective [56].

> *A malicious VR gaming application (the adversary) records eye tracking data that has been securely privatized by a user's VR hardware. By enabling the posting of high scores to social media, the adversary learns the user's identity. The adversary can then query against released datasets with associated medical diagnosis (autism, alzheimer's, depression, etc.), attempting to verify the user's membership in the dataset. If successful, the adversary has a platform to perform fraud or blackmail.*

Black-box Access with Exemplars: Similar to the above scenario, the adversary does not have knowledge about the implementation of the applied mechanism. However, the adversary has access to a large number of privatized records, possibly paired with a number of non-privatized records. From here, the adversary could attempt to approximate the mechanism, or perform regression analysis to learn an inverse function of the mechanism.

> *An adversary has a new VR headset which only releases privatized gaze data streams and an older model which releases raw gaze vectors. They recruit a number of users to perform the same tasks while wearing both headsets. When they have a sufficient amount of data, they model a function to invert the privacy mechanism, increasing the chance of re-identification. From this point on, they can apply that function to other records collected through the same hardware.*

White-box Access: In this scenario, the adversary knows the exact implementation of the mechanism that has been applied. This could be learned from data leaks of design documents or code [47], or by guessing simpler mechanisms by observing a sufficient number of samples and approximating parameters.

> *An insider of a VR hardware company posts the confidential privacy algorithm to an online forum. From there, any adversary who obtains the algorithm can attempt to leverage that knowledge against privatized data records obtained from the device.*

5.2 Evaluation Methodology

We illustrate the risk of each defined threat scenario with toy example attacks, simulating an adversary with the corresponding amount of knowledge. These attacks are not optimized or exhaustive, but illustrate the additional risks of data leakage that have not been widely considered in eye movement privacy literature.

Wavelet Denoising: In the black-box threat scenario, an adversary can perform an uninformed filtering attack to attempt to nullify the effectiveness of the privacy mechanism. Time series perturbations that are implemented on independent samples are prone to filtering attacks, which can vastly reduce uncertainty if the pattern can be filtered out [56].

We illustrate this by applying a wavelet denoising filter [16] over the privatized data stream. The implementation[5] assumes a level of noise and estimates σ automatically, requiring no knowledge of the mechanism at hand.

CNN Data Regression: In the black-box with exemplars scenario, the adversary has acquired a number of data samples with and without privatization. Data driven approaches could be implemented in an attempt to approximate the privacy mechanism or to directly approximate an inverse function.

We illustrate this concept with a simple convolutional neural network (CNN) which trains to reconstruct input privatized data streams back to the original data streams. Our implementation inputs and outputs 5 seconds of data. The model consists of 4 [1D Convolution, 1D Batch normalization, Tanh] blocks. For each condition, we train a model on 50% of data then evaluate on 50%, repeat the process with reverse train-test splits, then report the averaged accuracy.

Mechanism Applied to Reference Data: In the white-box scenario, the adversary knows the implementation details at hand. Obviously, if the mechanism is deterministic it becomes possible to reconstruct the original data by reversing the process. However, in stochastic mechanisms a level of uncertainty remains.

We investigate the white-box scenario by leveraging the adversary's inside knowledge to apply the same operation to the reference set. If the adversary does not know the query identity, but does know the identities of non-privatized reference records, the adversary could apply the privacy mechanism's algorithm across the board for a more equal comparison.

5.3 Results

We report the results of our toy examples for each threat scenario in Table 3. We find that Gaussian noise is vulnerable to attacks across all threat scenarios. Spatial downsampling and smoothing are not vulnerable to the black-box scenario's filtering attack, but all mechanisms are vulnerable to a degree to each scenario in which adversaries have additional knowledge. However, these mechanisms significantly increase the amount of knowledge and effort required to successfully re-identify users. Presumably, a white-box with exemplars threat scenario would have access to attacks that are even more successful.

In the black-box with exemplars scenario, both Gaussian and smoothing identification accuracies after CNN regression are higher than the original re-identification rate. It is possible that the CNN's inverse approximation accentuated some important features from the original data stream, potentially making the undone data streams slightly more identifiable.

Smoothing is more vulnerable than spatial in the black-box with exemplars scenario, being brought above the original identification accuracy. This can be attributed to our smoothing implementation being a fully deterministic process; as a result, the original signal can be fully reconstructed if the first real data value and exact buffer

[5] https://scikit-image.org/docs/stable/api/skimage.restoration.html#skimage.restoration.denoise_wavelet

Mechanism	ID Accuracy	Black-box	Black-box with exemplars	White-box
None	67.31%	/	/	/
Gaussian	14.1%	63.14%	68.39%	64.74%
Spatial	21.79%	21.79%	61.47%	58.33%
Smoothing	14.1%	14.1%	69.49%	55.13%

Table 3: Identification accuracies before and after performing attacks on privatized data across three threat scenarios.

size is known. However, there are a number of small optimizations which could be made to the smoothing process, such as non-uniformly initializing the buffer or adding random variance to the impact of each weight.

6 Discussion

We find that there is some risk of re-identification from eye movements collected in VR applications. On 125Hz data and using models trained and/or evaluated on VR data, we report accuracies of up to 33.87%@5s, 58.57%@60s, and 68.27%@90s. Our upper measure of 90 seconds approaches reliable identification; yet, commercial VR applications such as games or virtual training scenarios can have much longer sessions. Countermeasures should be designed with this in mind.

We find that across all user-centric metrics of utility, Gaussian noise is lower than the other evaluated mechanisms. Conversely, smoothing has the lowest data-centric AOI retention but highest user-centric utility. This finding highlights that when developing privacy mechanisms for interactive VR, it is critical to be user-centric. The findings from Section 3.4.2 and prior work [14] both would suggest Gaussian noise to be the best of evaluated mechanisms, but our analysis suggests that Gaussian noise should be rejected from a user experience standpoint. Spatial downsampling and smoothing are viable as privacy mechanisms in interactive contexts, but more work should be done to increase robustness against knowledgeable adversaries.

Broader Impacts: This work extends the discussion of privacy in eye tracking and VR by placing a focus on the real-world implications when applying privacy mechanisms to future applications. A large emphasis should be placed on user-centric notions of utility for gaze-based interaction applications, rather than only evaluating data-centric utility. If user experience is compromised, users will not be willing to engage in VR experiences in the first place. When evaluating privacy mechanisms, on top of the simplest case where re-identification accuracies are compared before and after privatization, researchers should test against more challenging threat models grounded in real world scenarios.

The evaluated privacy mechanisms could be applied to a larger set of eye tracking applications, such as augmented reality (AR) settings or to webcam eye trackers. We believe VR technology to be the most pressing use case currently; eye gaze is a promising input device showcasing unique interactions and enabling critical optimizations such as foveated rendering, but users should not have to choose between these features or their own privacy.

Limitations: Our user-centric evaluation relies on a single interactive VR dataset. This dataset includes gaze-only selection and gaze-with-gesture selection. This does not represent the full diversity of gaze-based interactions in VR, each of which may have their own nuances and thresholds for what is a reasonable level of utility traded for privacy gained. For example, consider an eye tracking-enabled competitive gaming context. Users in that context are unlikely to accept any privacy mechanism that compromises performance.

The analysis in Section 5 highlights the importance of robustness; however, our list of threat models are not exhaustive. There are a large number of adversaries and attack methods yet to be considered.

Future Work: In Section 5.3, we mentioned some improvements to smoothing that could introduce randomness. These improvements could increase smoothing's robustness while retaining a level of usability. Additionally, composition of simple operations may yield better mechanisms (passing spatial downsampled values into the smoothing buffer, for example, could be explored).

A methodology for further improved mechanisms could be to introduce temporal perturbations alongside sample-level perturbations. As the operation needs to be possible in real-time, it is difficult to introduce temporal inconsistencies without some form of delay. Potential avenues to explore would be to leverage context of objects from the scene [18] to modulate dwell times and durations between fixations, or to jointly perform privatization and gaze prediction to offset any delays.

There is a strong correlation between the ability to be identified and the amount of data available. An evaluation could be done on long data sessions (say, 30+ minutes continuously) to further quantify this risk at durations expected in VR applications. One mitigation would be to apply a random mechanism from a selection of viable mechanisms every 1 or 2 minutes, making the full session unreliable for queries.

It is currently unclear what potential impacts our mechanisms would impose on other user-experience enhancing research focused on eye tracking in VR. Future work could explore concepts such as spatial perception [80], cybersickness [27,43], and multisensory perception [46] and the level of impact that eye tracking privacy mechanisms would have on these metrics.

7 Conclusion

We analyzed the re-identification risk associated with eye tracking-enabled interactive VR applications and evaluated multiple privacy mechanisms which could serve as potential solutions for mitigating risk. In this work, a large emphasis was placed on user-centric notions of utility. While prior work has focused on data-centric utility, there is a necessary shift towards user-first design for applications where the user directly interfaces with the eye tracking functionality. We also further investigate real-world feasibility, modeling multiple threat scenarios where adversaries can attempt to counteract privacy efforts.

This work shows that there is re-identification risk associated with eye tracking in interactive VR, though the risk is less prominent than prior analyses with conventional eye tracking systems. Of the mechanisms evaluated, we found spatial down sampling and smoothing to be viable for practical applications. These mechanisms provide privacy while retaining high subjective usability and reasonable task performance, yet each mechanism is vulnerable against highly informed adversaries.

We hope that this work will aid further research regarding privacy protection in interactive VR applications. By placing a focus on user-centric utility and highlighting real world threat scenarios, we provide a methodology for the analysis of privacy mechanisms that puts the user first.

Acknowledgments

We acknowledge funding support from NSF Award CNS-2206950 and a Privacy-Enhancing Technology Award from Meta.

References

[1] Apple Vision Pro. https://www.apple.com/apple-vision-pro/. 2

[2] HoloLens 2. https://www.microsoft.com/en-us/hololens/hardware. 2

[3] Magic Leap 1 Devices. https://www.magicleap.com/ml1-devices. 2

[4] Magic Leap 2 Devices. https://www.magicleap.com/ml2-devices. 2

[5] Meta Quest Pro. https://www.meta.com/quest/quest-pro/. 2, 4, 6

[6] Vive Focus 3 Eye Tracker. https://business.vive.com/eu/product/vive-focus-3-eye-tracker/. 2

[7] Vive Pro Eye Overview. https://www.vive.com/sea/product/vive-pro-eye/overview/. 2

[8] M. Abrash. Latency – the sine qua non of AR and VR. https://web.archive.org/web/20130102023747/http://blogs.valvesoftware.com/abrash/latency-the-sine-qua-non-of-ar-and-vr/. 8

[9] R. Atienza, R. Blonna, M. I. Saludares, J. Casimiro, and V. Fuentes. Interaction techniques using head gaze for virtual reality. In *2016 IEEE Region 10 Symposium (TENSYMP)*, pp. 110–114, May 2016. doi: 10.1109/TENCONSpring.2016.7519387 1, 2, 3

[10] S. Berkovsky, R. Taib, I. Koprinska, E. Wang, Y. Zeng, J. Li, and S. Kleitman. Detecting Personality Traits Using Eye-Tracking Data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pp. 1–12. Association for Computing Machinery, New York, NY, USA, May 2019. doi: 10.1145/3290605.3300451 2, 8

[11] E. Bozkir, O. Günlü, W. Fuhl, R. F. Schaefer, and E. Kasneci. Differential privacy for eye tracking with temporal correlations. *PLOS ONE*, 16(8):e0255979, Aug. 2021. doi: 10.1371/journal.pone.0255979 2, 3, 7

[12] B. David-John, K. Butler, and E. Jain. For Your Eyes Only: Privacy-preserving eye-tracking datasets. In *2022 Symposium on Eye Tracking Research and Applications*, ETRA '22, pp. 1–6. Association for Computing Machinery, New York, NY, USA, June 2022. doi: 10.1145/3517031.3529618 2, 3, 7

[13] B. David-John, K. Butler, and E. Jain. Privacy-preserving datasets of eye-tracking samples with applications in XR. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2774–2784, May 2023. doi: 10.1109/TVCG.2023.3247048 2, 3, 7

[14] B. David-John, D. Hosfelt, K. Butler, and E. Jain. A privacy-preserving approach to streaming eye-tracking data. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2555–2565, May 2021. doi: 10.1109/TVCG.2021.3067787 1, 2, 3, 4, 9

[15] B. David-John, C. Peacock, T. Zhang, T. S. Murdison, H. Benko, and T. R. Jonker. Towards gaze-based prediction of the intent to interact in virtual reality. In *ACM Symposium on Eye Tracking Research and Applications*, ETRA '21 Short Papers, pp. 1–7. Association for Computing Machinery, New York, NY, USA, May 2021. doi: 10.1145/3448018.3458008 1, 2, 3

[16] D. L. Donoho and I. M. Johnstone. Adapting to Unknown Smoothness via Wavelet Shrinkage. *Journal of the American Statistical Association*, 90(432):1200–1224, Dec. 1995. doi: 10.1080/01621459.1995.10476626 8

[17] A. T. Duchowski, V. Shivashankaraiah, T. Rawls, A. K. Gramopadhye, B. J. Melloy, and B. Kanki. Binocular eye tracking in virtual reality for inspection training. In *Proceedings of the 2000 symposium on Eye tracking research & applications*, ETRA '00, pp. 89–96. Association for Computing Machinery, New York, NY, USA, Nov. 2000. doi: 10.1145/355017.355031 2

[18] E. Erdemir, P. L. Dragotti, and D. Gündüz. Privacy-Aware Time-Series Data Sharing With Deep Reinforcement Learning. *IEEE Transactions on Information Forensics and Security*, 16:389–401, 2021. doi: 10.1109/TIFS.2020.3013200 9

[19] A. S. Fernandes, T. S. Murdison, and M. J. Proulx. Leveling the playing field: A comparative reevaluation of unmodified eye tracking as an input and interaction modality for VR. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2269–2279, 2023. 1, 2

[20] L. Friedman, H. Stern, V. Prokopenko, S. Djanian, H. Griffith, and O. Komogortsev. Biometric Performance as a Function of Gallery Size. *Applied Sciences*, 12(21):11144, Jan. 2022. doi: 10.3390/app122111144 2

[21] M. Garau, M. Slater, V. Vinayagamoorthy, A. Brogni, A. Steed, and M. A. Sasse. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pp. 529–536. Association for Computing Machinery, New York, NY, USA, Apr. 2003. doi: 10.1145/642611.642703 2

[22] A. George and A. Routray. A score level fusion method for eye movement biometrics. *Pattern Recognition Letters*, 82:207–215, Oct. 2016. doi: 10.1016/j.patrec.2015.11.020 1, 2, 3

[23] T. Gowases, R. Bednarik, and M. Tukiainen. Gaze vs. mouse in games: The effects on user experience. In *Proceedings of the International Conference on Advanced Learning Technologies, Open Contents & Standards (ICCE)*, pp. 773–777, 2008. 1, 2

[24] H. Griffith, D. Lohr, E. Abdulin, and O. Komogortsev. GazeBase, a large-scale, multi-stimulus, longitudinal eye movement dataset. *Scientific Data*, 8(1):184, July 2021. doi: 10.1038/s41597-021-00959-y 3, 4

[25] B. Guenter, M. Finch, S. Drucker, D. Tan, and J. Snyder. Foveated 3D graphics. *ACM Transactions on Graphics*, 31(6):164:1–164:10, Nov. 2012. doi: 10.1145/2366145.2366183 2

[26] Z. Hu, S. Li, C. Zhang, K. Yi, G. Wang, and D. Manocha. DGaze: CNN-Based Gaze Prediction in Dynamic Scenes. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):1902–1911, May 2020. doi: 10.1109/TVCG.2020.2973473 1, 2, 3

[27] R. Islam, K. Desai, and J. Quarles. Cybersickness Prediction from Integrated HMD's Sensors: A Multimodal Deep Fusion Approach using Eye-tracking and Head-tracking Data. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 31–40, Oct. 2021. ISSN: 1554-7868. doi: 10.1109/ISMAR52148.2021.00017 9

[28] S. Jia, D. H. Koh, A. Seccia, P. Antonenko, R. Lamb, A. Keil, M. Schneps, and M. Pomplun. Biometric Recognition Through Eye Movements Using a Recurrent Neural Network. In *2018 IEEE International Conference on Big Knowledge (ICBK)*, pp. 57–64, Nov. 2018. doi: 10.1109/ICBK.2018.00016 1, 2

[29] B. John, S. Jörg, S. Koppal, and E. Jain. The Security-Utility Trade-off for Iris Authentication and Eye Animation for Social Virtual Avatars. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):1880–1890, May 2020. doi: 10.1109/TVCG.2020.2973052 2

[30] B. John, A. Liu, L. Xia, S. Koppal, and E. Jain. Let It Snow: Adding pixel noise to protect the user's identity. In *ACM Symposium on Eye Tracking Research and Applications*, ETRA '20 Adjunct, pp. 1–3. Association for Computing Machinery, New York, NY, USA, June 2020. doi: 10.1145/3379157.3390512 2

[31] A. S. Kaplanyan, A. Sochenov, T. Leimkühler, M. Okunev, T. Goodall, and G. Rufo. DeepFovea: neural reconstruction for foveated rendering and video compression using learned statistics of natural videos. *ACM Transactions on Graphics*, 38(6):212:1–212:13, Nov. 2019. doi: 10.1145/3355089.3356557 1, 2

[32] P. Kasprowski and J. Ober. Eye Movements in Biometrics. In D. Maltoni and A. K. Jain, eds., *Biometric Authentication*, Lecture Notes in Computer Science, pp. 248–258. Springer, Berlin, Heidelberg, 2004. doi: 10.1007/978-3-540-25976-3_23 1, 2

[33] O. V. Komogortsev, Y. S. Ryu, and D. H. Koh. Fast Target Selection via Saccade-driven Methods. Technical Report TXSTATE-CS-TR-2012-6, Texas State University, June 2012. 6

[34] J. R. Lewis. Psychometric Evaluation of the Post-Study System Usability Questionnaire: The PSSUQ. *Proceedings of the Human Factors Society Annual Meeting*, 36(16):1259–1260, Oct. 1992. doi: 10.1177/154193129203601617 6, 7

[35] J. Li, A. R. Chowdhury, K. Fawaz, and Y. Kim. Kalεido: Real-Time privacy control for Eye-Tracking systems. In *30th USENIX Security Symposium (USENIX Security 21)*, pp. 1793–1810. USENIX Association, Aug. 2021. 2, 3, 8

[36] L. Lin, A. Normoyle, A. Adkins, Y. Sun, A. Robb, Y. Ye, M. Di Luca, and S. Jörg. The Effect of Hand Size and Interaction Modality on the Virtual Hand Illusion. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 510–518, Mar. 2019. doi: 10.1109/VR.2019.8797787 4

[37] A. Liu, L. Xia, A. Duchowski, R. Bailey, K. Holmqvist, and E. Jain. Differential privacy for eye-tracking data. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, ETRA '19, pp. 1–10. Association for Computing Machinery, New York, NY, USA, June 2019. doi: 10.1145/3314111.3319823 2, 3, 7

[38] R. Liu, X. Xu, H. Yang, Z. Li, and G. Huang. Impacts of Cues on Learning and Attention in Immersive 360-Degree Video: An Eye-Tracking Study. *Frontiers in Psychology*, 12, 2022. 1, 2

[39] D. Lohr, S. Aziz, L. Friedman, and O. V. Komogortsev. GazeBaseVR, a large-scale, longitudinal, binocular eye-tracking dataset collected in virtual reality. *Scientific Data*, 10(1):177, Mar. 2023. doi: 10.1038/s41597-023-02075-5 3, 4

[40] D. Lohr, H. Griffith, and O. V. Komogortsev. Eye Know You: Metric Learning for End-to-End Biometric Authentication Using Eye Movements From a Longitudinal Dataset. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 4(2):276–288, Apr. 2022. doi: 10.1109/TBIOM.2022.3167633 1, 2

[41] D. Lohr and O. V. Komogortsev. Eye Know You Too: Toward Viable End-to-End Eye Movement Biometrics for User Authentication. *IEEE Transactions on Information Forensics and Security*, 17:3151–3164, 2022. doi: 10.1109/TIFS.2022.3201369 1, 2, 3, 4, 5

[42] D. J. Lohr, S. Aziz, and O. Komogortsev. Eye Movement Biometrics Using a New Dataset Collected in Virtual Reality. In *ACM Symposium on Eye Tracking Research and Applications*, ETRA '20 Adjunct, pp. 1–3. Association for Computing Machinery, New York, NY, USA, June 2020. doi: 10.1145/3379157.3391420 2

[43] P. Lopes, N. Tian, and R. Boulic. Eye Thought You Were Sick! Exploring Eye Behaviors for Cybersickness Detection in VR. In *Proceedings of the 13th ACM SIGGRAPH Conference on Motion, Interaction and Games*, MIG '20, pp. 1–10. Association for Computing Machinery, New York, NY, USA, Nov. 2020. doi: 10.1145/3424636.3426906 9

[44] P. Lungaro, R. Sjöberg, A. J. F. Valero, A. Mittal, and K. Tollmar. Gaze-Aware Streaming Solutions for the Next Generation of Mobile VR Experiences. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1535–1544, Apr. 2018. doi: 10.1109/TVCG.2018.2794119 1, 2, 3

[45] S. Makowski, P. Prasse, D. R. Reich, D. Krakowczyk, L. A. Jäger, and T. Scheffer. DeepEyedentificationLive: Oculomotoric Biometric Identification and Presentation-Attack Detection Using Deep Neural Networks. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 3(4):506–518, Oct. 2021. doi: 10.1109/TBIOM.2021.3116875 1, 2

[46] M. Marucci, G. Di Flumeri, G. Borghini, N. Sciaraffa, M. Scandola, E. F. Pavone, F. Babiloni, V. Betti, and P. Aricò. The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload and presence. *Scientific Reports*, 11(1):4831, Mar. 2021. Number: 1 Publisher: Nature Publishing Group. doi: 10.1038/s41598-021-84196-8 9

[47] M. McCormick. Data Theft: A Prototypical Insider Threat. In S. J. Stolfo, S. M. Bellovin, A. D. Keromytis, S. Hershkop, S. W. Smith, and S. Sinclair, eds., *Insider Attack and Cyber Security: Beyond the Hacker*, Advances in Information Security, pp. 53–68. Springer US, Boston, MA, 2008. doi: 10.1007/978-0-387-77322-3_4 8

[48] M. R. Miller, F. Herrera, H. Jun, J. A. Landay, and J. N. Bailenson. Personal identifiability of user tracking data during observation of 360-degree VR video. *Scientific Reports*, 10(1):17404, Oct. 2020. doi: 10.1038/s41598-020-74486-y 2

[49] R. Miller, N. K. Banerjee, and S. Banerjee. Combining Real-World Constraints on User Behavior with Deep Neural Networks for Virtual Reality (VR) Biometrics. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 409–418, Mar. 2022. doi: 10.1109/VR51125.2022.00060 2

[50] R. Miller, N. K. Banerjee, and S. Banerjee. Temporal Effects in Motion Behavior for Virtual Reality (VR) Biometrics. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 563–572, Mar. 2022. doi: 10.1109/VR51125.2022.00076 2

[51] P. Monteiro, G. Gonçalves, H. Coelho, M. Melo, and M. Bessa. Hands-free interaction in immersive virtual reality: A systematic review. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2702–2713, May 2021. doi: 10.1109/TVCG.2021.3067687 1, 2

[52] M. Negin, T. Chmielewski, M. Salganicoff, U. von Seelen, P. Venetainer, and G. Zhang. An iris biometric system for public and personal use. *Computer*, 33(2):70–75, Feb. 2000. doi: 10.1109/2.820042 2

[53] A. Normoyle, J. B. Badler, T. Fan, N. I. Badler, V. J. Cassol, and S. R. Musse. Evaluating perceived trust from procedurally animated gaze. In *Proceedings of Motion on Games*, MIG '13, pp. 141–148. Association for Computing Machinery, New York, NY, USA, Nov. 2013. doi: 10.1145/2522628.2522630 1, 2

[54] P. A. Orlov and N. Apraksin. The Effectiveness of Gaze-Contingent Control in Computer Games. *Perception*, 44(8-9):1136–1145, 2015. doi: 10.1177/0301006615594910 1, 2

[55] J. L. Orquin, N. J. S. Ashby, and A. D. F. Clarke. Areas of Interest as a Signal Detection Problem in Behavioral Eye-Tracking Research. *Journal of Behavioral Decision Making*, 29(2-3):103–115, 2016. doi: 10.1002/bdm.1867 2

[56] S. Papadimitriou, F. Li, G. Kollios, and P. S. Yu. Time series compressibility and privacy. In *Proceedings of the 33rd international conference on Very large data bases*, VLDB '07, pp. 459–470. VLDB Endowment, Vienna, Austria, Sept. 2007. 8

[57] S. Peng and N. Al Madi. An Eye Opener on the Use of Machine Learning in Eye Movement Based Authentication. In *2022 ACM Symposium on Eye Tracking Research and Applications*, ETRA '22, pp. 1–2. Association for Computing Machinery, New York, NY, USA, June 2022. doi: 10.1145/3517031.3531631 1, 2

[58] T. Piumsomboon, G. Lee, R. W. Lindeman, and M. Billinghurst. Exploring natural eye-gaze-based interaction for immersive virtual reality. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 36–39, Mar. 2017. doi: 10.1109/3DUI.2017.7893315 1, 2, 3

[59] J. Reichenberger, M. Pfaller, and A. Mühlberger. Gaze Behavior in Social Fear Conditioning: An Eye-Tracking Study in Virtual Reality. *Frontiers in Psychology*, 11, 2020. 2

[60] K. Ruhland, K. Zibrek, and R. McDonnell. Perception of personality through eye gaze of realistic and cartoon models. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception*, SAP '15, pp. 19–23. Association for Computing Machinery, New York, NY, USA, Sept. 2015. doi: 10.1145/2804408.2804424 1, 2

[61] N. Sammaknejad, H. Pouretemad, C. Eslahchi, A. Salahirad, and A. Alinejad. Gender Classification Based on Eye Movements: A Processing Effect During Passive Face Viewing. *Advances in Cognitive Psychology*, 13(3):232–240, Sept. 2017. doi: 10.5709/acp-0223-1 2, 8

[62] C. L. Sanches, O. Augereau, and K. Kise. Using the Eye Gaze to Predict Document Reading Subjective Understanding. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 08, pp. 28–31, Nov. 2017. doi: 10.1109/ICDAR.2017.377 2

[63] S. Sannon and A. Forte. Privacy Research with Marginalized Groups: What We Know, What's Needed, and What's Next. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):455:1–455:33, Nov. 2022. doi: 10.1145/3555556 2

[64] C. Schröder, S. M. K. Al Zaidawi, M. H. Prinzler, S. Maneth, and G. Zachmann. Robustness of Eye Movement Biometrics Against Varying Stimuli and Varying Trajectory Length. In *Proceedings of the 2020 ACM CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–7. Association for Computing Machinery, New York, NY, USA, Apr. 2020. doi: 10.1145/3313831.3376534 2

[65] S. Seele, S. Misztal, H. Buhler, R. Herpers, and J. Schild. Here's Looking At You Anyway! How Important is Realistic Gaze Behavior in Co-located Social Virtual Reality Games? In *Proceedings of the ACM Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '17, pp. 531–540. Association for Computing Machinery, New York, NY, USA, Oct. 2017. doi: 10.1145/3116595.3116619 1, 2

[66] E. Selinger, E. Altman, and S. Foster. Eye-Tracking in Virtual Reality: A Visceral Notice Approach for Protecting Privacy. *Privacy Studies Journal*, 2:1–34, Mar. 2023. doi: 10.7146/psj.v2i.134656 2

[67] R. Shadiev and D. Li. A review study on eye-tracking technology usage in immersive virtual reality learning environments. *Computers & Education*, 196:104681, Apr. 2023. doi: 10.1016/j.compedu.2022.104681 2

[68] Y. Shi, Y. Zhu, R. K. Mehta, and J. Du. A neurophysiological approach to assess training outcome under stress: A virtual reality experiment of industrial shutdown maintenance using Functional Near-Infrared Spectroscopy (fNIRS). *Advanced Engineering Informatics*, 46:101153, Oct. 2020. doi: 10.1016/j.aei.2020.101153 2

[69] V. Sitzmann, A. Serrano, A. Pavel, M. Agrawala, D. Gutierrez, B. Masia, and G. Wetzstein. Saliency in VR: How Do People Explore Virtual Environments? *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1633–1642, Apr. 2018. doi: 10.1109/TVCG.2018.2793599 3

[70] J. Steil, I. Hagestedt, M. X. Huang, and A. Bulling. Privacy-aware eye tracking using differential privacy. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, ETRA '19, pp. 1–9. Association for Computing Machinery, New York, NY, USA, June 2019. doi: 10.1145/3314111.3319915 2, 3, 7

[71] J. Sun, Y. Liu, H. Wu, P. Jing, and Y. Ji. A novel deep learning approach for diagnosing Alzheimer's disease based on eye-tracking data. *Frontiers in Human Neuroscience*, 16, 2022. 2

[72] V. Tanriverdi and R. J. K. Jacob. Interacting with eye movements in virtual environments. In *Proceedings of the ACM SIGCHI conference on Human Factors in Computing Systems*, CHI '00, pp. 265–272. Association for Computing Machinery, New York, NY, USA, Apr. 2000. doi: 10.1145/332040.332443 2

[73] P. Ugwitz, O. Kvarda, Z. Juříková, Šašinka, Čeněk, and S. Tamm. Eye-Tracking in Interactive Virtual Environments: Implementation and Evaluation. *Applied Sciences*, 12(3):1027, Jan. 2022. doi: 10.3390/app12031027 3

[74] N. I. Vargas-Cuentas, A. Roman-Gonzalez, R. H. Gilman, F. Barrientos, J. Ting, D. Hidalgo, K. Jensen, and M. Zimic. Developing an eye-tracking algorithm as a potential tool for early diagnosis of autism spectrum disorder in children. *PLOS ONE*, 12(11):e0188826, Nov. 2017. doi: 10.1371/journal.pone.0188826 2

[75] G. Wan, X. Kong, B. Sun, S. Yu, Y. Tu, J. Park, C. Lang, M. Koh, Z. Wei, Z. Feng, Y. Lin, and J. Kong. Applying Eye Tracking to Identify Autism Spectrum Disorder in Children. *Journal of Autism and Developmental Disorders*, 49(1):209–215, Jan. 2019. doi: 10.1007/s10803-018-3690-y 2

[76] C.-C. Wang, J. C. Hung, and H.-C. Chen. How Prior Knowledge Affects Visual Attention of Japanese Mimicry and Onomatopoeia and Learning Outcomes: Evidence from Virtual Reality Eye Tracking. *Sustainability*, 13(19):11058, Jan. 2021. doi: 10.3390/su131911058 1, 2

[77] S. Wei, D. Bloemers, and A. Rovira. A Preliminary Study of the Eye Tracker in the Meta Quest Pro. In *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences*, IMX '23, pp. 216–221. Association for Computing Machinery, New York, NY, USA,

Aug. 2023. doi: 10.1145/3573381.3596467 4

[78] M. Weier, T. Roth, A. Hinkenjann, and P. Slusallek. Foveated Depth-of-Field Filtering in Head-Mounted Displays. *ACM Transactions on Applied Perception*, 15(4):26:1–26:14, Sept. 2018. doi: 10.1145/3238301 1, 2

[79] M. Weier, T. Roth, E. Kruijff, A. Hinkenjann, A. Pérard-Gayot, P. Slusallek, and Y. Li. Foveated Real-Time Ray Tracing for Head-Mounted Displays. *Computer Graphics Forum*, 35(7):289–298, 2016. doi: 10.1111/cgf.13026 1, 2

[80] H. Xiang, K. Kim, and J. Ryu. Work-in-Progress—Preliminary Analysis of Spatial Perception in AR application with Eye-Tracking Data. In *2022 8th International Conference of the Immersive Learning Research Network (iLRN)*, pp. 1–3, May 2022. doi: 10.23919/iLRN55037.2022.9815935 9

[81] Y. Xu, Y. Dong, J. Wu, Z. Sun, Z. Shi, J. Yu, and S. Gao. Gaze Prediction in Dynamic 360° Immersive Videos. pp. 5333–5342, 2018. 1, 2

[82] A. T. Zhang and B. O. Le Meur. How Old Do You Look? Inferring Your Age from Your Gaze. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 2660–2664, Oct. 2018. doi: 10.1109/ICIP.2018.8451219 2, 8

[83] G. A. Zito, D. Cazzoli, L. Scheffler, M. Jäger, R. M. Müri, U. P. Mosimann, T. Nyffeler, F. W. Mast, and T. Nef. Street crossing behavior in younger and older pedestrians: an eye- and head-tracking study. *BMC Geriatrics*, 15(1):176, Dec. 2015. doi: 10.1186/s12877-015-0175-0 2

Appendix

A Description of Privacy Mechanisms

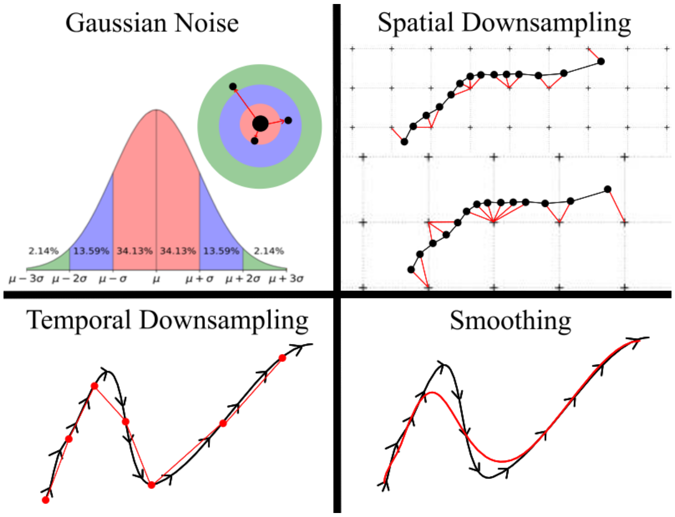

Fig. 5: Visualizations of the privacy mechanisms implemented in our experiments.

Here we provide illustrations and short descriptions of the concepts of our privacy mechanisms. See Figure 5 for illustrations. Gaussian noise offsets each gaze data point independently at each frame by drawing upon a Gaussian distribution. As the noise sampling is independent, no temporal patterns can be discerned to identify individuals; however, the mechanism is susceptible to filtering to recover the original signal. Spatial downsampling maps the continuous range of gaze values to an equirectangular grid of discrete points. The true gaze angle is mapped to the closest discrete value at every frame. Temporal downsampling effectively decreases the sampling rate, copying the true value of a frame into the next N frames. Smoothing applies a linear weighted moving average to a range of gaze values. The weighting gives recent frames a higher weight than less recent frames. The result appears to have a slight delay to the user, but by fixating on an object, the smoothed gaze value quickly reaches the intended point.

B Runtime Analysis of Privacy Mechanisms

We provide a simple performance analysis of each of our mechanisms, measuring the impact on device memory and on execution time. We measure the impact on device memory by reporting the average memory across trials within each privacy mechanism. We report the total reserved memory by the application and the system memory reported by Unity3D's Memory Profiler[6]. To measure execution time, we report the average frames per seconds of each trial for each privacy mechanism.

Mechanism	System Memory (kb)	Runtime (FPS)
No Mechanism	311825 ± 12443	70.94 ± 0.076
Gaussian Noise	318181 ± 12490	70.95 ± 0.083
Spatial Downsampling	316606 ± 13338	70.94 ± 0.076
Smoothing	322041 ± 4818	70.94 ± 0.080

Table 5: Performance analysis of privacy mechanisms on our prototype Unity environment.

We see in Table 4 that these mechanisms have little to no impact to performance. Runtime is not impacted, and system memory increases only slightly on average. We do see a more noticeable increase in memory usage with smoothing, as it is the only mechanism which stores a continuous array of past gaze samples needed to compute the current gaze sample. Overall, the mechanisms evaluated are quick to compute and provide negligible performance overhead.

In our experiments, we process gaze samples provided by the Oculus SDK and apply privacy mechanisms before utilizing the gaze samples in the application. In a real-world deployment, these mechanisms could be securely implemented on VR HMDs before passing gaze samples to applications. Figure 1 illustrates the general eye tracking pipeline for VR headsets. Along with other operations taken to model and process gaze vectors, these privacy mechanisms can be applied on the device securely, then privatized gaze vectors can be provided to potentially untrustworthy applications opened by the VR user. With current mechanisms' low overhead, these could be implemented on headset software with little performance impact. Hardware-accelerated implementations could also be explored, and may be more necessary as privacy mechanisms become more complex.

C Additional Data Collection Details

The dataset collected for this publication is available at `https://doi.org/10.5281/zenodo.10475455`. In this section we discuss the implementation of our data logging process and the data collected.

The experiments are implemented in Unity3D using the Oculus SDK to interface with the Meta Quest Pro headsets used for data collection. At every frame, we query the SDK for gaze samples, process these gaze samples using the current privacy mechanism, then pass the privatized gaze sample for use by the rest of the application.

We log data at every visual frame. Because the eye gaze provided by the Oculus SDK is only available at every frame's update, gaze data is collected at the application's frame rate. Data was logged locally to the VR headset's internal storage, then moved after each participant's session.

In `headset_data.csv` we log frames, timestamps, active trial conditions, and the position and rotations of the headset, hands, and eyes at every frame. We also log the non-privatized rotations of the eyes in Experiment 2 for possible comparison. `event_data.csv` contains experiment-relevant events, such as area of interest intersection info, the start and end of trial periods, and the completion of tasks, such as completing a sandwich in Experiment 1 and destroying an enemy in Experiment 2. `survey_data.csv` reports user responses to the PSSUQ questionnaire after each condition. The dataset also included processed files, which contain streamlined information on conditions, frame, and timestamp and the localized rotations of the eyes. This processed information is trimmed to only contain frames in which the experiment was active.

In Experiment 1, each participant underwent 4 identical trials of 90 seconds, yielding 9360 seconds of active trial data for $N = 26$ participants. In Experiment 2, each participant underwent 2 trials of 30 seconds for each privacy mechanism, strength, and control mode pairing, yielding 28 trials for participants which underwent smoothing

[6] `https://docs.unity3d.com/Manual/ProfilerMemory.html`

and 20 for those who did not, yielding 13,200 seconds of active trial data for $N = 18$ participants.