PROVIDING PRIVACY FOR EYE-TRACKING DATA WITH APPLICATIONS IN XR

By

BRENDAN DAVID-JOHN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2022

To my loving wife Mikaela, our beautiful dog Kona, our family and friends, and all of the mentors who encouraged and lifted me up at different stages along my journey

ACKNOWLEDGEMENTS

This dissertation and my growth as an academic would not be possible without my advisor Eakta Jain. Thank you for your unrelenting support and motivation over the past five years. To my committee members Ben Lok, Arunava Banerjee, Kevin R.B. Butler, and Sriram Kalyanaraman: thank you for your insightful feedback and challenging me every step of the way. There is not enough room to list all of the colleagues and mentors I have gained over the years at UF, including fellow graduate students Pallavi Raiturkar, Yuzhu Dong, and Ethan Wilson, peers in my cohort, undergrads that remind me how bright the future will be, multi-disciplinary collaborators, department chair Juan Gilbert, and the former Chief Diversity Officer Antonio Farias. The knowledge I have gained from all of these interactions has both breadth and depth, introducing me to the academic landscape and teaching me how to navigate its most treacherous waters.

Outside of UF, I would like to thank the mentors and advisors that introduced me to research as an outlet for my curiosity and passion within Computer Science. I think frequently about the first day I visited a university campus that led me to attend RIT, and the conversation that I had two years later with Roger Dube that lead me to undergrad research in eye tracking. This research experience introduced me to Reynold Bailey and the RIT Graphics lab, opening my eyes to what types of research in Computer Science are possible. Special thanks to Srinivas Sridharan and Joe Geigel for working alongside me in the lab for years, and the countless other colleagues from RIT that taught me how to turn failures into accomplishments.

I would not be where I am today without the community of Native students formed by AISES National and corresponding student chapters. The distinct lack of Natives in higher education and Computer Science reminds me of my purpose to stay on this path. Attending the annual national conference is a high point of every year, refreshing my outlook on life, work, and cultural identity. Nyah:wëh to the future student leaders of Gator AISES that will work to stay resilient and keep the organization around far into the future (that's you, Hunter).

The biggest thanks go out to my family and friends. To my wife Mikaela, for sticking by my side throughout the process, and our parents, grandparents, aunts, uncles, cousins, and pets back home for pushing me every step of the way.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

Figures page

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

PROVIDING PRIVACY FOR EYE-TRACKING DATA WITH APPLICATIONS IN XR

By

Brendan David-John

August 2022

Chair: Eakta Jain
Major: Computer Science

Virtual and mixed-reality (XR) technology has advanced significantly in the last few years and will enable the future of work, education, socialization, and entertainment. Eye-tracking sensors enable the design of immersive experiences and the deployment of display hardware. Eye-tracking data is required for supporting novel modes of interaction, animating virtual avatars, and rendering or streaming optimizations. While eye-tracking enables many beneficial applications in XR, it also introduces risks related to the security and privacy of the user and their captured data. I have explored solutions that address concerns related to the leaking of biometric features that uniquely identify users at different components of the eye-tracking pipeline and evaluated the impact on XR applications.

First, I explored the risk of user identification from the iris pattern as imaged by the near-eye camera feed of the eye tracker. Image blur was applied to secure the user's iris pattern and lower identification rates from a standard iris authentication approach. Our experiments demonstrated that there is a level of image blur that achieves recognition rates lower than 5% from a standard iris authentication approach without negatively impacting the social experience of using gaze data to animate the eyes of a virtual avatar.

Second, I developed privacy mechanisms that can be applied to gaze data streamed to third-party applications. Features extracted from the time series of gaze positions enable biometric identification with high accuracy. The explored privacy mechanisms can be applied to real-time data streams to remove or reduce the risk of unique user identification depending on the

application. For applications that do not require gaze samples, relevant metrics are extracted instead of streaming raw samples. When gaze samples are required, standalone privacy mechanisms are applied that reduce the risk identification from the streamed samples. Our evaluation of mechanisms showed that identification rates could be lowered from 85% to 30% while introducing less than 40ms in Root Mean Squared Error for Area-of-Interest dwell time analysis, a Kullback-Leibler Divergence of 0.04 or lower between saliency maps, and up to 1.14° of error in gaze prediction.

Third, I applied privacy mechanisms to eye-tracking datasets that provided formal guarantees against re-identification attacks. Evaluated mechanisms adapted *k*-anonymity and plausible deniability to eye-tracking data and were compared with results from differential privacy. Our findings established a superior privacy-utility trade-off for *k*-anonymity on feature datasets used to train a classifier for document type recognition. For sample datasets, the kalεido DP mechanism performed best at retaining utility for training activity recognition models while reducing identification rates to chance; while the *k*-same-synth mechanism performed best at retaining utility for gaze prediction models.

CHAPTER 1
INTRODUCTION AND MOTIVATION

Mixed-reality devices, such as virtual reality (VR) and augmented reality (AR) displays, create immersive experiences that have an impact on education [132, 82], socialization [116], entertainment, and the future of work [6]. Mixed reality is already entering the enterprise market, led by companies such as Magic Leap and Pico targeting enterprise users [238]. Enterprise XR applications support private and government entities, evidenced by a $461 million investment by Saudi Arabia [162] and visions of a future workplace from the CEO of Qualcomm [6]. VR and AR devices, collectively referred to as XR, are rapidly becoming available to consumers. The Oculus Quest 2 consumer VR display sold between two and three million units in Q4 of 2020 [107]. The deployment of devices will grow continuously, as the Quest 2 sold 8.7 million units in 2021 [108]. XR devices are continuously engineered to make them ubiquitous and accessible, accounting for users sensitive to simulator sickness or users with disabilities. To enable this future, the XR ecosystem will collect data from users at a large scale with an array of sensors, including integrated eye trackers [32].

Eye trackers perform gaze estimation to track where a user is looking and their visual attention. Tracking gaze enables key applications for deploying mixed reality, including intelligent interfaces [92, 63], social interaction [217], activity recognition [112], and foveated streaming [157, 249, 117, 150] or rendering [192, 167, 166]. Gaze data either enhances users' interactions or comfort within mixed reality or enables the practical use of low-power and mobile devices accessible to the general population. Eye tracking has the potential to enable new applications for mixed reality; however, the captured sensor data also enables unintended inferences about users.

Sensor data and the scale at which it will be collected within a mixed-reality ecosystem introduce new concerns for privacy and security. Compared to current technology in the form of mobile devices, new challenges for mixed reality include the need to enable the high-resolution camera and depth sensors for device operation, the addition of more intrusive sensors for tracking user movement, and the ability for XR to influence perceptual behavior and user decisions [61]. Eye-tracking sensors are particularly concerning as they collect a high-resolution measure of

users' conscious and unconscious attention in where gaze was oriented most often and could be used to reveal preferences for personalized ads.

Surveys of XR users have identified that the majority are uncomfortable with having their XR data sold to advertisers [189], and most would not be comfortable sharing eye-tracking data with non-trusted entities, i.e., third-party applications [232]. Meta's next-generation Project Cambria will be the first VR display to include integrated eye tracking directly marketed to consumers [79]. The expansion of eye tracking to consumer-grade devices is dangerous, as researchers have shown current XR applications violate their privacy policies and share tracking data with third-party apps [241], linking the concerns of XR users with real-world harms. For example, researchers have recently hypothesized how tracking user data in the United States could enable tracking of women seeking abortions in states where it is illegal [163].

Eye-tracking data presents a critical risk to privacy, as it captures sensitive information about the user based on where they look and introduces the risk of re-identification from captured data. Re-identification from XR data was addressed as the first recommendation of the IEEE Global Initiative on Ethics of Extended Reality report on XR and the Erosion of Anonymity and Privacy [164]: "XR stakeholders should actively develop and/or support efforts to standardize differential privacy and/or other privacy protocols that provide for the protection of individual identities and data". Eye tracking is among the sensors that enable accurate identification and recognition of users. Eye-tracking data applied as a biometric is well studied between iris recognition [58, 118] and gaze-based biometrics [172, 93, 86, 209, 73, 222, 156]. State-of-the-art eye movement biometrics achieve an accuracy of 94% [222] and an Equal Error Rate of 2% [156], suggesting that, with high enough data quality, users are recognized as accurately as a four-digit pin with as little as five seconds of data [155]. Achieved performance from sharing gaze data with third-party applications risks leaking biometric identity. An adversarial application can collect this data and spoof the user's identity or compare it with data from known identities, linking identity to any sensitive information that can be inferred from gaze.

An example of violating privacy through eye tracking is the monitoring of employee

behavior outside of work. Gaze data captured from User X at work can be linked to their identity and used to recognize User X in other environments. A feasible scenario includes User X attending a virtual meeting anonymously to discuss forming a labor union at work and using eye-tracking data to animate the eyes of a virtual avatar as part of the social platform.[1] User X is then discovered by their employer through app collusion when the gaze data from the meeting is sold to local employers by data brokers.[2] Data brokers regularly sell tracking data on users to individuals or employers, enabling the sale of VR interactions and behavioral data in the future [213]. Employer surveillance exists today using big data to track employee behavior with respect to labor organizing [91]. Thus, the ability to identify User X is a risk when their data is shared with untrusted platforms or third parties. Current data privacy policies in most states and countries would not restrict the sharing of raw gaze data with external parties, despite being a source of biometric identification [109].

Our approach to addressing future privacy risks from eye tracking is to form plausible threat scenarios at different points in the data pipeline. Threat scenarios are inspired by previously established attacks on sensors or datasets in similar domains and define a risk for users. Threat models are a tool for developing secure systems and are derived from these scenarios to state assumptions on what data an adversary has access to and how they perform an attack [248]. A threat model defines the type of data being processed and metrics for evaluating whether an attack can be prevented by privacy solutions [65].

Thesis statement: My work explores the following thesis statement: Privacy mechanisms are capable of de-identifying eye-tracking data while enabling XR applications in social VR, activity recognition, and gaze prediction. My thesis statement is supported by studies conducted within three different points in the eye-tracking data pipeline:

1. Securing eye-tracking sensor data by reducing the risk of leaking iris biometric in infrared eye images

[1] www.tcf.org/content/report/virtual-labor-organizing/
[2] www.pluralistic.net/2021/04/13/public-interest-pharma/#axciom

2. Protecting eye-tracking data streams by withholding gaze samples or adding noise to reduce the risk of identification from biometric features
3. Adding privacy to eye-tracking datasets with formal privacy guarantees to mitigate re-identification attacks

Capturing infrared images of the eye for gaze estimation is the first step in the typical eye-tracking pipeline. Eye images record the iris biometric, a gold standard for recognition, in high resolution. Eye trackers capture images at 60Hz or more, introducing the risk of leaking identity through eye images. Leaked eye images enable stealing identity by spoofing the iris biometric. We propose using blur to remove high-frequency iris patterns while enabling eye-tracking applications. Depending on the threat model, blur can be implemented in either software or hardware, and protect the iris biometric by reducing the risk of user recognition. Our experiments in Chapter 2 identify a trade-off between the risk of iris recognition and retaining positive perceived attributes in social VR that lowers the iris recognition rate to 5% or less. Higher blur levels further decrease the risk of iris recognition and degrade social attributes to neutral or negative responses.

Eye images are processed to extract gaze positions and generate samples which are then served to the XR platform or shared with third-party applications. Apps that have access to raw gaze samples can extract biometric features from eye movement patterns and attempt to identify the user. We propose a privacy-preserving Gatekeeper API to enable eye-tracking applications that require aggregate or event-level data. The API computes relevant metrics for each application without the risk of sharing gaze samples. For applications that require gaze samples, our privacy mechanisms add noise or downsample data to enhance privacy. Our methods in Chapter 3 enable privacy-preserving streaming by lowering the risk of biometric identification for applications such as area-of-interest analysis and event-based interaction techniques through privacy-by-design and real-time applications such as gaze prediction through standalone mechanisms.

Eye-tracking datasets aggregate gaze behavior from groups of individuals and enable training machine-learning models for activity recognition and intent or gaze prediction. Datasets are released publicly for research use or stored internally by XR companies to train proprietary

models for product deployment. Datasets are anonymized by removing personal information such as names, locations, and dates of birth; however, they are still susceptible to re-identification attacks. A prime example of re-identification is the Netflix Prize challenge [180]. Narayanan and Shmatikov took the released anonymous movie ratings and rental dates and paired them with public reviews from IMDB that were timestamped and linked to the user's real name. The risk of leaking users' identities led to a lawsuit that claimed a woman's sexual orientation could be revealed to her family as a result of the attack [227].

Re-identifying individuals contributing to an eye-tracking dataset also serves as a privacy risk. A successful re-identification attack could link identity to sensitive attributes that can be extracted from gaze data, such as medical diagnoses or sexual orientation. Current literature proposes differential privacy mechanisms to protect eye-tracking data. In Chapter 4 we introduce mechanisms that achieve privacy guarantees of k-anonymity and plausible deniability to mitigate re-identification attacks on eye-tracking datasets. Our approach for protecting eye-tracking feature data adapts existing methods from face images and location data. We then propose a novel data synthesis approach to protect datasets of eye-tracking samples while training models of gaze prediction and activity recognition. Our methods protect large-scale eye-tracking datasets with formal guarantees against re-identification while retaining utility for relevant XR applications.

CHAPTER 2
PROTECTING IRIS BIOMETRICS

2.1 Introduction

The most common eye-tracking devices utilize near-eye infrared cameras to perform gaze estimation [131]. Eye-tracking cameras have a resolution of 320×240 or more, sufficient to perform iris authentication when positioned near the eye. For example, the Microsoft Hololens v2 already enables iris authentication in this manner [4], essentially streaming the user's biometric password at 30 frames per second whenever they use the device.

Continuously recording the user's iris biometric is both a feature and a risk. If an untrusted system collects eye images for eye tracking, anonymous viewing is not guaranteed. If an adversary captures a stream of eye images containing the iris biometric, they could beat spoof detection methods and steal the user's identity [141]. Protecting biometric data is critical, as biometrics represent a physiological means for recognizing a user. Unlike text passwords, they cannot be updated in the event of a security leak.

Before this work, methods to securely collect eye-tracking images without collecting the iris biometric were not explored. Furthermore, numerous eye-tracking datasets release raw eye images, including evaluating gaze estimation approaches [137, 90, 85] and 158 datasets for iris segmentation and matching [186]. The risk of leaking the iris is unique to sensors that record infrared eye images. While alternative sensors exist for eye tracking [253, 152], video-based eye trackers that collect infrared images are still the standard in XR devices such as the HTC Vive Pro Eye, Magic Leap, and Microsoft Hololens.

We propose using blur-based mechanisms to process eye-tracking images and remove the high-frequency iris features used for authentication. Because of the networked nature of social platforms and the use of cloud-based rendering techniques for VR [175], it is expected that XR devices will follow an 'always on and connected' model. Streaming eye-tracking data makes it susceptible to attacks. Most critically, the iris pattern of the user is vulnerable. The iris image is a gold standard biometric used in high-security applications, such as border customs [2], and is recognized as such by headset manufacturers [4].

We focus on methods that can be applied in real-time to the typical eye-tracking pipeline.

The result is secure eye-tracking data flows that degrade the iris biometric pattern when streaming eye images over a network for processing or when device hardware could be compromised.

The blur-based mechanism represents a trade-off between leaking the iris biometric within the eye image and preserving the utility of gaze data generated from the eye images. Trade-offs between privacy/security and utility are vital to evaluating the impact of privacy-enhancing technologies. For example, blurring or masking out private regions of images has an impact on whether images provide helpful information, are preferred by users, or are considered visually appealing [106]. Our work explores the security-utility trade-off concerning gaze accuracy at the data-level (Study 1) and perceptual impact when animating the eyes of a virtual avatar (Study 2).

Earlier versions of this work were first published at ACM ETRA 2019 [122] and IEEE VR 2020 [120]. The initial paper at ETRA 2019 highlighted the risk of leaking the iris biometric. It spawned a sub-area of research within the eye-tracking community that has furthered methods that protect raw eye images [42, 123, 47, 77].

In this chapter, we present blur-based mechanisms for degrading high-frequency iris features that can be implemented in software within the eye-tracking pipeline or by hardware prior to image capture using optical defocus. We demonstrate that eye images can be appropriately blurred to reduce the presence of iris features while retaining social characteristics for animating social virtual avatars.

2.2 Threat Scenario

Any device that captures infrared eye images used for eye tracking could be compromised to leak raw image data. Raw sensor data is susceptible to leaking data to an application without the user being aware. In the context of mobile devices, a permissions-based model is used to protect certain data streams for apps. However, data can still be leaked through side channel attacks exposed within the permissions system or covert attacks where an app with sensitive data permissions shares data with others [201]. Thus, it is feasible that malicious applications could gain access to raw sensor data of XR devices, as current systems employ extensions of the Android operating system [241].

An adversary can take infrared images of the user's eyes and try to spoof the identity of the user through iris authentication and access sensitive information. For example, an individual who uses iris biometrics at work to access confidential information is at risk, such as a military general. Security for biometric templates in sensitive databases has been explored for fingerprint biometrics, where authentication is common for law enforcement and military agencies [251]. A user at a VR arcade with their child can have eye-tracking data collected as part of a game interface. If the arcade game is compromised then an adversary can gain access to the raw eye images collected by the system. The adversary is assumed to have read access to the stream of eye images and will then use them to try and authenticate as the victim. It is plausible that an adversary can use the unmodified stream of raw eye images to beat liveness detection and successfully spoof the user's identity if the sequence is of sufficient length (15 seconds or more) [208].

2.3 Related Work

2.3.1 Eye Tracking in Virtual Reality

Eye-tracking applications for VR include foveated rendering [192, 27], which optimizes computational resources in rendering by reducing resolution in the periphery, streaming algorithms that reduce the bandwidth of streamed 360° content [157, 115, 80], intuitive interfaces for navigation and predicting intent [190, 28], subtle gaze direction using luminance cues in the periphery to guide attention [96], redirected walking methods that take advantage of saccadic masking and blinks to orient the user within a limited physical space [236, 146], classifying neurodegenerative disease through eye movements [187], virtual experiences designed to improve joint attention of children with ASD [165], and modeling how users explore 360° content [228].

Eye-tracking hardware in VR ranges from video-based oculography [131], electro-oculography (EOG) [35], photo-sensor oculography (PS-OG) [253, 152], and magnetic sclera coils [247]. EOG, PS-OG, and sclera coil eye trackers provide gaze estimation without imaging the eye itself; however, video-based eye trackers are the most readily available solutions today. EOG and sclera coil approaches are deployed less often as they are invasive, requiring

electrodes to be attached to the user's head or a magnetic contact lens worn by the user. PS-OG trackers are limited in terms of power usage and the ability to deploy within consumer devices, as the current implementation occludes the user's field of view [253]. Companies like Facebook, HTC, and Magic Leap have opted for a non-invasive video-based eye tracker that captures images of the eye, including the iris and other identifiable features like eyebrows [67]. Thus, there is a need to investigate techniques that secure the iris during gaze estimation.

2.3.2 Iris Authentication

Infrared images of the eye with sufficient resolution capture iris patterns unique to the individual.[1] Iris recognition ranks in the top tier of biometrics as it is universal, distinct, permanent, and robust against spoofing attacks [118]. It is important to keep the iris pattern secure, as recognition methods are robust to poor lighting [127], off-axis imaging [56], occlusion [57], and distance [12] making the biometric accessible at times when the user may not consent. Iris authentication has long been established through the work of John Daugman [58] and many others as a statistically valid method for recognizing an individual.[2] As a result, iris patterns have been trusted for identification at voting booths [5], border customs [2], schools [1], and in hospitals [3]. These applications highlight the sensitivity of the information accessed if a hacker can steal an identity through a biometric. Thus, the presence of a user's iris within a dataset or application places the user's identity at risk.

2.3.3 Privacy and Security in Eye Tracking

Mobile eye trackers rely on videos from an eye camera that captures the user's eye and a front-facing scene camera that records what they see. The scene camera is akin to wearable devices that are always on and recording video data. Public perception of these devices is overwhelmingly negative, as seen with the initial release of the Google Glass, as they infringe on the privacy of both the user and bystanders [66, 179, 199]. Daily users of eye-tracking technology trade-off the privacy of their everyday actions for the benefit of activity logging, gaze-based

[1] In our eye tracker configuration the iris pattern typically spans around 150 pixels in diameter, falling between the 140 pixels recommended by Daugman et al. [57] and the 200 pixels recommended by the ISO standard for iris biometrics [100].

[2] Please see [87] for a review.

interfaces, and assistive applications [111, 11, 244, 165]. Steil et al. have developed a privacy approach specifically for the scene camera, using a controlled shutter to disable the video feed in private situations [233]. The eye camera is unique in that it captures raw eye movements and personally-identifying information without any layer of security. Previous findings for wearable-based privacy and security do not directly apply to eye-tracking images. This chapter focuses on a solution to protect against unauthorized iris-based identification from eye images.

2.3.4 Defocus Based Identity Preservation

Early work from Neustaedter et al. [182] explored adding blur to increase the privacy of a teleconference video feed. They found that no general-purpose blur level preserves utility across all scenarios in this context. For example, the participants specified a much higher amount of blur in the video that captured embarrassing activities such as picking their nose or changing clothes, compared to daily computer work. Participants were asked to identify the activities being performed, with the blur level decreasing until they could confidently classify the activity. The computed blur thresholds and classification rates determined that blur effectively increases privacy while retaining utility. Hasan et al. [106] investigated various image filters such as masking, blurring, and pixelation for their effectiveness in obscuring specific content features and retaining the utility and aesthetics of the photograph. They reported that blur was effective at obscuring the gender of the photographed person, though not the ethnicity or expression. Ultimately, they determined that there was no 'one size fits all' solution for every scenario, and object size or security context can influence the optimal method.

Pittaluga and Koppal [195] have implemented a similar blur-based privacy approach within the context of micro-scale image sensors. A hardware-based approach is used to add blur instead of a software-based Gaussian blur. The use of optics to scatter light before the image is captured creates blur on the camera sensor. Applications like head tracking, person tracking, and facial recognition are explored with several types of camera sensors (thermal, IR, RGB) imaging the user. Each camera configuration and application must be optimized and designed to balance the trade-off between security and utility. Our work investigates adding blur to eye images

pre-capture; however, the goal is to do so without modifying the stock hardware or optics. This allows consumers to control their privacy, as current consumer technology lacks any specialized privacy hardware.

2.4 Methodology

Iris authentication is performed by comparing an input eye image with a stored reference biometric from a known identity. Infrared images capture high-frequency features of the iris that are unique to each individual. Iris features are encoded as a binary pattern that is then compared to the stored template to determine if a match exists. The proposed blur mechanism is applied directly to the eye image. Blur acts as a low-pass filter for the image and obfuscates the high-frequency patterns needed for iris authentication. We propose applying blur in software when the device platform is trusted and using optical defocus to blur the image pre-capture when hardware is not trusted.

In this section, we describe the iris authentication process, define how blur degrades iris authentication, provide each threat model, and demonstrate studies that explore the impact of blur on data utility for gaze accuracy and avatar eye animation.

2.4.1 Iris Authentication

The process of performing iris-based authentication is well established, representing features extracted from the high frequency component of the iris pattern as a binary code [58]. A robust iris segmentation method [88] and standard encoding procedure [161] are used to create iris codes for each eye image. Authentication is performed by computing the Hamming distance between source and target iris codes [57]. Hamming distance is defined as the proportion of bits that disagree between source and target binary codes computed using AND ($\cap$) and XOR ($\otimes$) operations,

$$\mathrm{HD} = \frac{\|(\mathrm{S_{code}} \otimes \mathrm{T_{code}}) \cap (\mathrm{S_{mask}} \cap \mathrm{T_{mask}})\|}{\|\mathrm{S_{mask}} \cap \mathrm{T_{mask}}\|}, \tag{2-1}$$

where S_{code} and T_{code} are the input binary codes with their respective masks. The binary masks indicate which pixels contain the iris pattern, with zeros indicating eye lids, eye lashes, or any other detected noise [57]. In the subsequent data analysis, iris codes are excluded if at least 75%

of the bits are considered noise. A positive match is returned if the Hamming distance is less than a fixed threshold. A threshold of $HD_{auth} = 0.37$ was used to authenticate a match between source and target.[3]

2.4.2 Degrading Iris Authentication Using Blur

Let $I = I_C + I_R$ denote an image received from the eye-tracking camera. Figure 2-1 visualizes grayscale intensity values along the iris, pupil, and glint regions of an eye image. We define I_C as the image component that contains an eye-tracking signal, i.e., the corneal glint or pupil, and is modeled as a Gaussian distribution $I_C \sim N(\mu = 0, \sigma_C)$. The I_R component represents the iris pattern within the image, with the highest frequency in the signal limited by the camera's spatial resolution. Let us denote the highest frequency as B. While I_C contains primarily low-frequency content, I_R contains both low and high-frequency content, with the higher frequencies being the identifying features, up to a maximum frequency B.

Consider a low-pass filter F of the form $F(x) = N(\mu = 0, \sigma)$, i.e., a Gaussian blur or optical defocus. When I is convolved with $F(x)$, the result is

$$I_D(x) = I(x) * F(x) = I_C(x) * F(x) + I_R * F(x) = I'_C(x) + I'_R(x). \tag{2-2}$$

The objective is to determine an optimal parameter σ that defines $F(x)$ such that eye-tracking features are still detectable in $I'_C(x)$, while I'_R no longer contains the higher frequencies that enable iris-based authentication. A security-utility trade-off for σ balances the ability to apply iris authentication to the features I'_R and the ability to estimate accurate gaze position from $I'_C(x)$.

2.4.3 Threat Models

Iris patterns present in eye-tracking sensor data streams can serve as a password and are continuously streamed from the eye tracker. This data stream is subject to a man-in-the-middle attack if images are sent over a network [52]. In configurations where images are not streamed over a network, they are still subject to attacks when data is transferred at the hardware level [53].

[3] A threshold of 0.37 corresponds to a 0.0012% false positive rate [120].

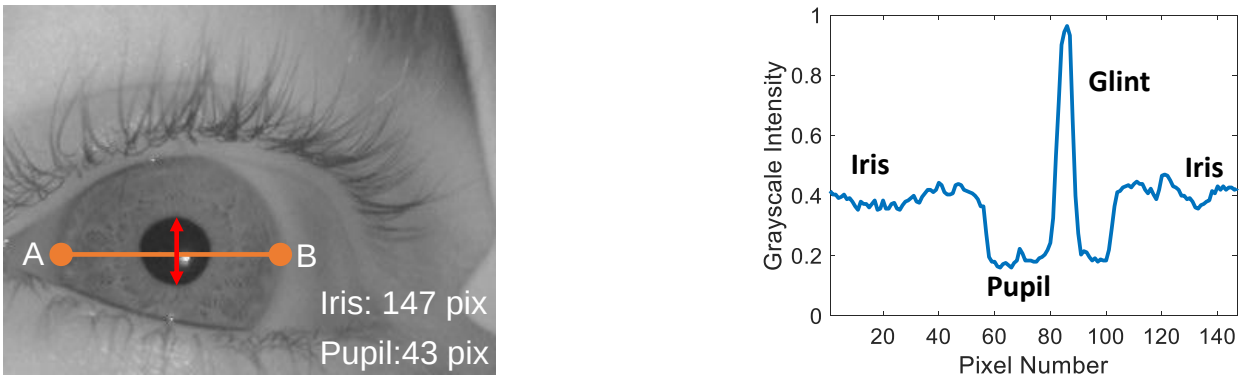

Figure 2-1. Illustration of iris signature in an eye-tracking image. Left: The iris spans ≈ 150 pixels along $\overline{AB}$. Right: Grayscale values along line $\overline{AB}$ demonstrate high-frequency features in the iris region, a step pattern representing the dark edges of the pupil, and a Gaussian shaped bump for the corneal glint. This suggests that a low pass filter $F(x)$ can degrade the high frequency iris patterns while retaining low frequency features needed for gaze estimation.

Our threat models model an adversary who obtains unauthorized eye images of an individual. The adversary then uses the eye images to spoof the user's identity. The authentication process compares the leaked eye image with that of a known eye image of the individual. The specifics of the authentication process are discussed in Section 2, producing a HD value that is compared to a threshold to determine if the identity is a match or not. The ability to spoof the authentication of a user is measured using the Correct Recognition Rate (CRR) [173]. CRR is the percentage of leaked images where $HD < HD_{auth}$ when compared to unblurred reference eye images from the individual. The threat model assumes that the platform leaks eye images before being processed for gaze estimation. Thus, any eye image modification will affect iris security and gaze estimation utility. We explored scenarios where the eye-tracking software platform that records and transmits the image is either trusted or untrusted by the user.

2.4.3.1 Threat Model 1: Trusted Software Platform

A scenario in which the user trusts the eye-tracking platform depends on trusting the manufacturer of the eye-tracking hardware and the platform implementation of the core eye-tracking pipeline. In this scenario, the platform developers are trusted to implement the blurring approach immediately after the near-infrared camera captures the eye image. The

blurring approach is then applied to introduce biometric security before sending the eye image for additional processing. This threat model covers eye trackers that perform gaze estimation onboard the hardware, or stream the images to a server for further processing.

2.4.3.2 Threat Model 2: Untrusted Software Platform

We also considered a threat model in which the platform that captures the eye image is not trusted. A similar scenario was discussed by Pittaluga et al. in which privacy-preserving optics are designed to protect sensitive information prior to image capture [195]. This threat model covers the scenario where the user does not trust the manufacturer to implement iris security methods or suspects that the underlying eye-tracking hardware is compromised.

2.5 Study 1: Gaussian Blur

The first study was conducted to determine the ability of Gaussian Blur to increase iris security while retaining accurate gaze estimation. Study 1 was conducted under the assumptions of Threat Model 1, in which the Gaussian Blur was applied to eye images in software prior to gaze estimation. Gaussian Blur has a single parameter σ representing the standard deviation of the blur kernel in pixels. Defining the blur parameter in pixels depends on the resolution of the eye camera. Blur parameter results are presented in normalized units for generalization of our findings. Our evaluation produced a security-utility trade-off for different values of σ to determine an optimal value for adding security while enabling accurate gaze estimation.

2.5.1 Research Question

This study was a principal investigation into $RQ_{2.1}$:

Can we degrade identification from eye images without impacting gaze accuracy?

Specifically, this study investigated whether we can degrade the risk of iris recognition from sensor data in the form of near-infrared eye-tracking images. We considered whether we could find a value of σ in which the CRR is reduced to 0% while the introduced gaze error does not exceed 2° visual angle. We selected 2° as a target for gaze accuracy as it is an upper bound on the amount of systematic error from most commercial eye trackers and is small enough to enable most gaze-based applications [78].

2.5.2 Implementation

Gaussian blur was implemented using the *imgaussfilt* function in MATLAB to simulate an isotropic blur kernel. The input parameter σ represents pixels within the eye image. Values of σ are presented in normalized units using the diagonal resolution of the camera. The eye images collected in our experiments were 320 by 240 pixels, with a diagonal length of 400 pixels. The captured eye images were blurred prior to an offline calibration of the eye tracker and gaze estimation through the open-source Pupil Labs software [131]. An open-source version of Daugman's method that computes the *HD* between two input eye images was used to implement iris authentication [161].

2.5.3 Protocol

Five participants with normal vision wore the Pupil Labs head-mounted eye tracker (ca. 2016, 30Hz) in a lab environment and performed a 5-point calibration. The eye tracker reported a validation error of <1.5°. Users sat with a chin rest and viewed circular targets presented on a desktop monitor (Fig. 2-2, Left). They were instructed to move their eyes to look at five targets (Fig. 2-2, Right). We recorded the eye image stream, scene camera video stream, and ground truth gaze data. Target viewing took approximately 32 seconds. Three representative eye images were selected from each participant viewing targets to evaluate iris authentication performance. Data from participant S0005 had large inconsistencies in pupil detection and gaze accuracy from the unmodified eye images and was removed from the analysis.

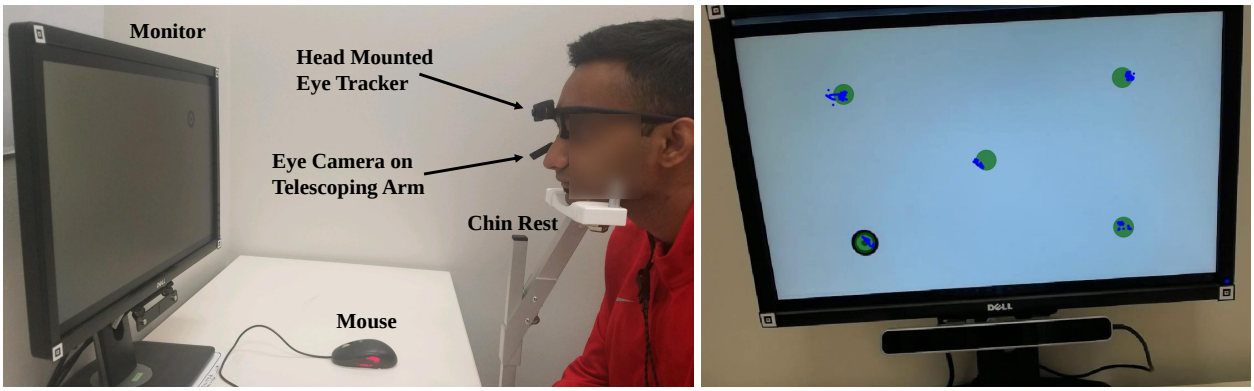

Figure 2-2. Experimental setup for eye-tracking data collection. Participants were seated with a chin rest and gaze data was collected while they viewed five on-screen circular targets.

2.5.4 Metrics

Iris security was measured as the CRR when considering each participant's three representative eye images, which act as entries in a known database for authentication. The three images for each participant are compared with the blurred eye images from all participants to determine the HD values between each pair of participant identities. CRR was computed using this set of HD values and an authentication threshold of $HD_{auth} = 0.37$.

Utility was the average angular distance between the ground truth gaze positions and the gaze positions produced by the set of blurred eye images. The angular error was computed as $\theta = cos^{-1}(x_j \cdot x'_j)$, where both x_j and x'_j were normalized 3D gaze directions produced by unblurred and blurred eye images respectively for each data point. An angular error can only be computed if the pupil was detected in both eye images. Only the frames in which the pupil detection was successful for both images were used to compute the average angular error for each participant. To validate that the pupil was being detected consistently, an additional utility measure of pupil detection rate was included. Pupil detection rate for each participant was computed as the number of blurred eye images with a detected pupil divided by the number of unblurred eye images with a detected pupil.

2.5.5 Results

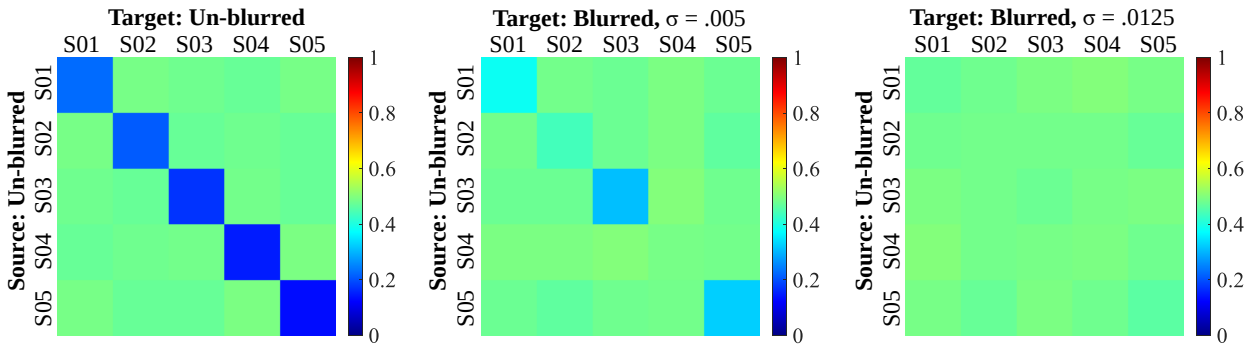

Figure 2-3. Average HD from within and between participant iris authentication. Dark blue values represent a correct recognition, while lighter values depict a failed match. The first matrix on the left shows perfect matching and the next two show that simulated defocus decreases CRR. In particular, $\sigma = .0125$ degrades iris authentication to produce a CRR of 0%.

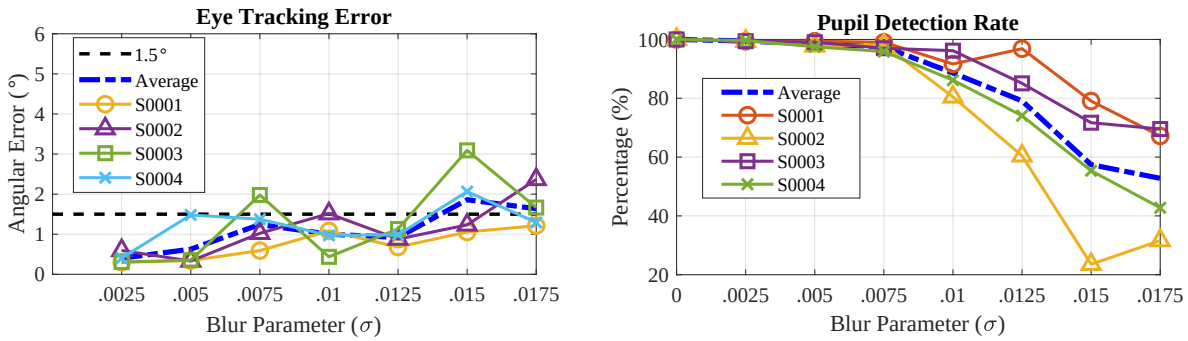

Figure 2-4. Study 1 eye-tracking data utility results. Left: Error in visual degrees of eye tracking data produced with simulated Gaussian blur up to $\sigma = .0175$. Note that at $\sigma = .0125$, where iris degradation is significant for this eye tracker, the tracking error is at or below 1.5°. Right: Detection rate for participants S0001-S0004 with simulated Gaussian blur up to $\sigma = .0175$. Note that at $\sigma = .0125$, where iris degradation is significant for this eye tracker, detection rate is still at $\approx 80\%$.

The resulting set of *HD* values with no blur produced a CRR of 100%, while CRR is reduced to 0% at $\sigma = .0125$ (Figure 2-3). A blur value of $\sigma = .0125$ introduced gaze error that was less than 1.5° on average (Figure 2-4, Left). While gaze accuracy was retained at $\sigma = .0125$, the pupil detection rate was reduced to 80% on average with a minimum of 60% (Figure 2-4, Right). Reducing blur to $\sigma = .0075$ lowers the negative impact on pupil detection with an average rate of 96%.

2.6 Study 2: Optical Defocus

Study 2 followed the assumptions of Threat Model 2, in which blur is introduced prior to image capture by the eye-tracking camera. Optical defocus is introduced in a user-controlled manner by increasing the physical distance between the eye and the eye-tracking camera. The effect of optical defocus is modeled with a Gaussian Blur kernel to measure the introduced blur in terms of σ. Our evaluation implies that by introducing enough distance between the eye and camera, the CRR can be reduced while still enabling reasonable gaze estimation. Next, we computed a security-utility trade-off for a key application of eye tracking to social VR and the impact of eye image blur on the characteristics of the resulting avatar eye animations.

Eye movements play an essential role in non-verbal communication and thus are critical in creating compelling social interactions with virtual avatars. For example, Steptoe et al. [235]

showed that eye movements caused participants to more accurately determine if an avatar was being truthful or not when compared to an avatar without eye movements. Eyes are important for conversational avatars that discuss sensitive information, such as medical diagnoses [245]. The animation of virtual eyes can be data-driven or generated by procedural algorithms that model the dynamics of the eye. Realistic eye animations may include micro-saccadic jitter, blinks, eyelid displacement, and pupil diameter [216]. Results from Garau et al. [89] suggest that a virtual avatar rendered with naturalistic eye and head movements did not improve communication over an audio-only conversation when the eye and head movements do not match the context of the conversation. The authors also showed that an avatar with eye movements based on the current conversation produced similar attentiveness and involvement to a video call with a real person. This finding implies that while models can generate natural eye movements for an avatar, they may not contain the non-verbal cues and subtleties needed to simulate a real conversation. In these cases, real eye-tracking data is critical.

2.6.1 Research Questions

This study provides further investigation into $RQ_{2.1}$:

Can we degrade identification from eye images without impacting gaze accuracy?

An optics-based approach to introducing blur was explored to secure the iris biometric during eye tracking. Adjusting the eye-tracking camera fits Threat Model 2, in which the eye-tracking platform is untrusted, and blur cannot be applied after the image is captured. The data-level utility is measured in an extension of the Study 1 protocol that features an in-focus and out-of-focus configuration. To answer the research question, we considered whether the out-of-focus camera configuration could reduce CRR to 0% while introducing less than 1.5° gaze error, as in Study 1.

The study also addresses $RQ_{2.2}$:

Can we degrade identification from eye images without impacting virtual avatars animated with gaze data?

An additional set of experiments are conducted where blurred eye images are processed to animate the eyes of a virtual avatar. Our perceptual experiments determined a detection threshold for when a viewer would notice a difference due to security blur and explored the impact on attributes of a virtual avatar. Specifically, the perceptual experiments answered the following two research sub-questions:

- $RQ_{2.2.1}$: *At what level of defocus do viewers detect a difference in the animation of a virtual avatar's eyes compared to a reference?*
- $RQ_{2}.2.2$: *What is the relationship between eye image defocus and the perception of avatar truthfulness, naturalness, attentiveness, comfort, and eye contact?*

2.6.2 Implementation

Optical defocus is introduced as a hardware configuration by adjusting the eye camera to increase the distance from the eye (Figure 2-5, Right). This solution requires that the user has access to and can manipulate the eye-tracking camera. The introduced distance between camera configurations can be mapped to a value of σ for comparison with the Gaussian Blur implementation [120]. Gaze data was collected using the same eye tracker as Study 1 and processed similarly using Pupil Labs software. A simulated conversation with a video was used to collect conversational eye movements and animate the virtual avatar.

Virtual avatars were rendered using Unity version 2017.4.24f1 and animated directly using the 3D gaze direction generated by Pupil Labs. Gaussian blur was added to the eye images from the conversational gaze data in software with blur parameter σ. Figure 2-6 shows an example of what a participant would see when viewing the virtual avatar and the difference introduced with a blur level of $\sigma = .0125$.

2.6.3 Protocol

The existing study protocol was extended with fifteen additional participants and included a target viewing session across both in-focus and out-of-focus configurations. Prior to target viewing, the participant was asked to look directly at the eye-tracking camera for five seconds, eliciting on-axis eye images that simulate a "stop-and-stare" interface to evaluate iris authentication [56, 197]. The in-focus configuration was then implemented by placing the eye

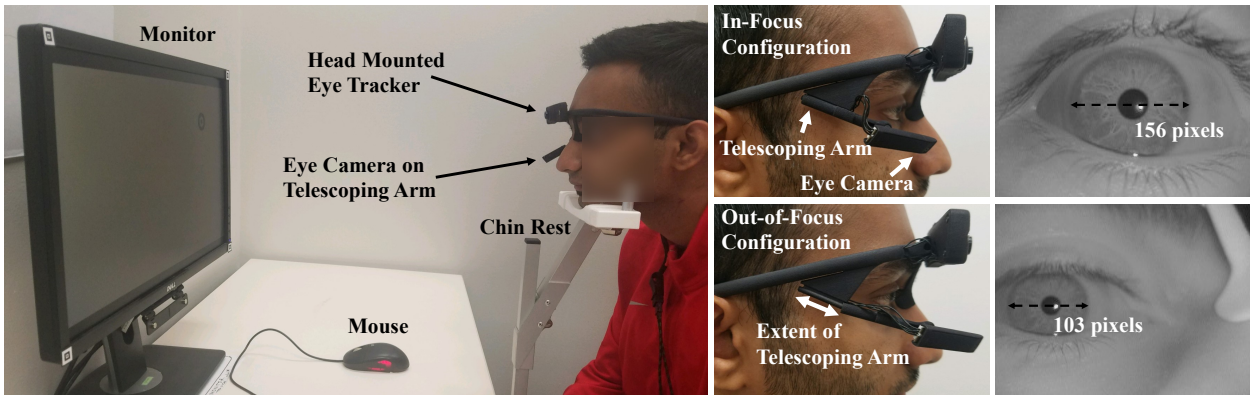

Figure 2-5. Experimental setup for optical defocus evaluation. The adjustable telescoping arm of the eye tracker is used to create an out-of-focus configuration. Eye images from an in-focus configuration (23.3mm) and out-of-focus configuration (35.4mm) are shown.

camera as close as possible to the participant's eye while keeping the eye in the center of the eye image frame. The participant then viewed the five circular targets. Next, to create the out-of-focus configuration, the experimenter adjusted the telescoping arm to the farthest point possible, again orienting the camera such that the eye stayed within the frame. The participants then viewed the circular targets for a final time. Finally, the participants repeated the "stop-and-stare" procedure for five seconds to collect a second set of on-axis eye images. Data from these fifteen participants were used to evaluate the security-utility trade-off when applying optical defocus.

Two perceptual experiments were conducted to determine the impact of applying blur to an eye image prior to gaze estimation to animate a virtual avatar's eyes. The first experiment (20 participants) determined the detection threshold using a same-different task when shown eyes animated with modified data side-by-side animations from the original data. The second experiment (20 new participants) presented eye animations one at a time and recorded five-point Likert scale responses to measure perceived eye contact, comfort in the interaction, avatar truthfulness, avatar naturalness, and avatar attentiveness. The eye images that generated the eye movements of the avatar gaze were created using Gaussian blur. Five levels of σ were considered: None ($\sigma = 0$), Low ($\sigma = .0025$), Medium ($\sigma = .0075$), High ($\sigma = .0125$) and Very High ($\sigma = .0225$).

Figure 2-6. Eye animations from blurred and unblurred images were presented side-by-side with identical avatars for the same-different evaluation. The deviation in gaze from defocus ($\sigma = .0125$) is shown in the avatar on the right.

2.6.4 Metrics

Iris security was measured as the CRR when comparing the eye images collected during our "stop-and-stare" authentication routine. Each eye image was compared with all other images to generate sets of *HD* values within and between participants. In-focus and out-of-focus eye images were compared to see at what rate an individual can be identified by eye images when using an out-of-focus configuration for an eye tracker.

Utility at the data level was measured as the average angular gaze error between the center of targets presented on-screen to the user and gaze positions recorded by the eye tracker. Unlike Study 1, gaze error was not computed using all the collected samples, only using samples collected when viewing each of the five presented targets.

The detection threshold of a stimulus determines how much can be added before the user noticed it. Captured data was used to compute a Miss Rate that measures the percentage of times a participant did not notice a difference for each blur level σ. Response data were modeled using a psychometric curve that establishes a relationship between stimulus intensity, i.e., blur, to the resulting Miss Rate. The model is characterized using the Point of Subjective Equality (PSE) and

Detection Threshold (DT), which are computed as the σ where a psychometric curve is equal to 50% and 25% respectively [234]. These thresholds capture how much blur is needed to produce a response rate of chance as well as the amount of blur where a viewer will consistently detect a difference. The second experiment of the avatar study collected Likert scale responses from a series of questions for each avatar animation. Responses were visualized using box and whisker plots to understand the summary statistics and compared using pairwise Wilcoxon signed-rank tests to determine when blur caused a significant decrease in the perception of each attribute.

2.6.5 Results

Table 2-1. Security and utility results for in-focus and out-of-focus eye-tracking configurations. On average, there was a difference of 8mm between in-focus and out-of-focus configurations. Defocusing the camera caused a decrease in CRR without an appreciable impact on gaze accuracy.

	Mean	Std. dev.
In-focus distance (mm)	25.1	2.8
Out-of-focus distance (mm)	33.1	2.3
In-focus CRR (%)	78.9	14.7
Out-of-focus CRR (%)	7.4	9.2
In-focus gaze error (°)	1.4	0.4
Out-of-focus gaze error (°)	1.7	0.5
In-focus gaze precision (°)	0.1	0.04
Out-of-focus gaze precision (°)	0.1	0.04

2.6.5.1 Iris Authentication

Table 2-1 reports the mean and standard deviations of computed metrics for each configuration, with an average in-focus CRR of 78.6%, while the out-of-focus images had a rate of 7.1%. Figure 2-7 (Right) demonstrates the relationship between camera distance and CRR by fitting a sigmoid function of the form $f(d) = \frac{1}{1+e^{-(a \cdot d+b)}}$, where $a = -0.43$, $b = 12.10$, and d is the input distance in mm. At 30mm CRR was 45%, and by 35mm CRR has dropped to 8%, showing only a small percentage of frames can successfully authenticate the user at increased distances.

2.6.5.2 Perceptual Studies

Our analysis determined that the defocus value of $\sigma = .0088$ is the average PSE, i.e., at this defocus level the viewer had a Miss Rate of random guessing (50%). The average DT is

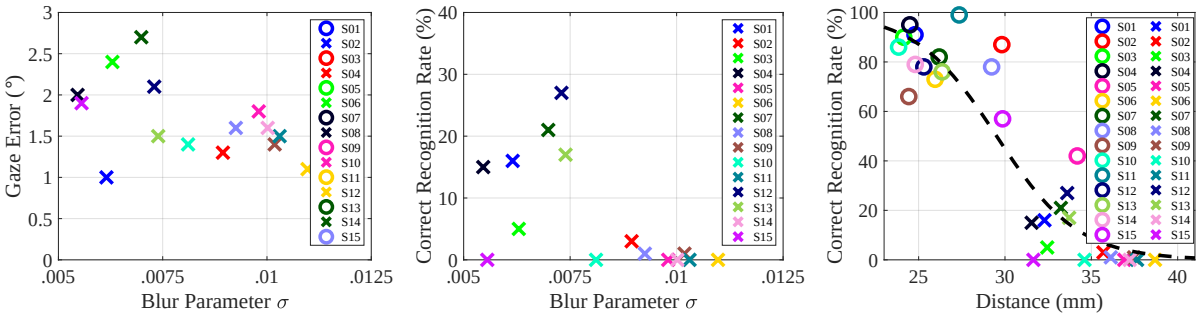

Figure 2-7. Results for security and utility show that CRR is degraded by defocus (σ) and increased camera distance. Circles indicate data from the in-focus configuration, while crosses indicate data from the out-of-focus configuration. The dashed line represents a sigmoid curve fit to CRR as a function of distance. Angular error measured between targets and gaze data for the out-of-focus configuration was at most 2.7°.

$\sigma = .0142$, which is the defocus level at which there is a 75% chance that viewers will be able to detect that the eye animation of the avatar is different compared to the original. The psychometric curves are visualized in Figure 2-8.

Likert scale responses for each dependent variable (truthfulness, naturalness, attentiveness, comfort, eye contact) represent ordinal data grouped by the defocus parameter σ. Figure 2-9 shows the average and standard error values for each attribute. The Kolmogorov-Smirnov test for normality was applied to each group and variable. Data were not normally distributed ($p < 0.001$), and therefore non-parametric statistical tests were used. A Friedman test showed a significant main effect of σ for truthfulness ($\chi^2(4)$=162.72,$p < 0.001$), naturalness ($\chi^2(4)$=290.2,$p < 0.001$), attentiveness ($\chi^2(4)$=300.41,$p < 0.001$), comfort ($\chi^2(4)$=279.15,$p < 0.001$), and eye contact ($\chi^2(4)$=199.23,$p < 0.001$). For each attribute pairwise Wilcoxon signed rank tests with Bonferroni correction showed significant differences between $\sigma = .0125$ and all other levels of σ ($p < .05$ or less); as well as between $\sigma = .02$ and all other levels of σ ($p < .001$). Additionally, for naturalness and eye contact significant differences were found between $\sigma = 0$ and $\sigma = .0075$, with ($p < .01$) and ($p < .05$) respectively.

The takeaways with respect to $RQ_{2.2}$ suggest that a defocus parameter of $\sigma = .0088$ (3.5 pixels) should be used if utility is preferred over security, and $\sigma = .0125$ (5 pixels) if security is

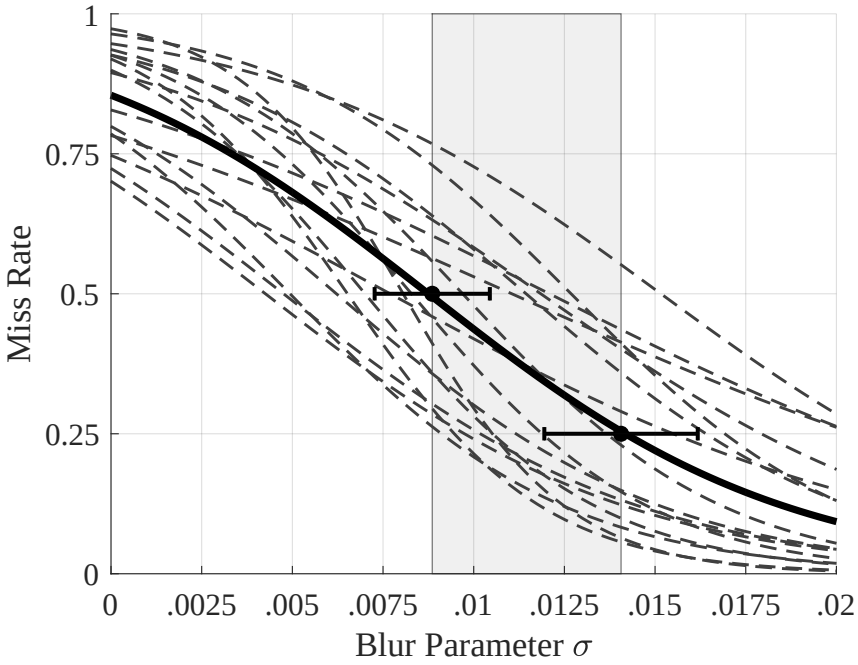

Figure 2-8. Resulting psychometric functions of the same-different task for individual and pooled responses. Gray dashed lines represent individual responses, and the solid black line represents a function fit to the average responses across individuals. Error bars represent the 95% confidence interval for PSE and DT values.

preferred based on the detection of a difference in eye animations. Results indicated that at σ values of .0125 and .02 the avatar no longer maintains eye contact, attentiveness, or naturalness. Thus, for animating an avatar's eyes in an environment similar to our experiment, a blur level less than $\sigma = .0125$ is recommended. The presented security-utility trade-off provides the opportunity to use gaze data for social VR with $\sigma = .0088$ with a lower risk ($\leq$5%) of leaking the iris biometric through captured eye images.

2.7 Discussion

To explore $RQ_{2.1}$ for Threat Model 1 (trusted software), we computed the security-utility trade-off for Gaussian Blur applied to eye-tracking images. The resulting trade-off indicates that with $\sigma = .0125$ of blur, the iris pattern is secured with a Correction Recognition Rate (CRR) of 0%, and the error introduced was less than 1.5° on average. The level of biometric security and gaze utility at $\sigma = .0125$ protects the iris pattern when considering the Daugman authentication approach while still enabling gaze applications that can tolerate up to 1.5° of additional error.

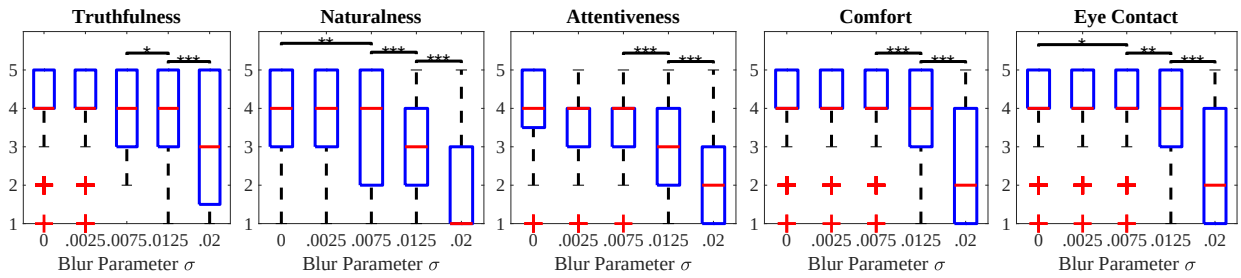

Figure 2-9. Box plots indicating the median, 25%, and 75% quartiles for Study 2 results. Significantly different groups are marked with * when $p < .05$, ** when $p < .01$, and *** when $p < .001$. For clarity ** significance brackets were established but not drawn between groups σ=[0, .0025] and σ=.0125, along with *** significance bars for groups σ=[0, .0025, .0075] and σ=.02 across all attributes.

To further explore $RQ_{2.1}$ for Threat Model 2 (untrusted software), we also computed the security-utility trade-off for hardware-based optical defocus on eye-tracking images. The secure out-of-focus configuration produced an average of $\sigma = .0083$ and CRR of 7% across individuals, compared to an average CRR of 79% before optical defocus was introduced. These findings connected well with Study 1, as software blur produced with $\sigma = .0075$ degraded iris authentication for most individuals, and $\sigma = .0125$ completely degraded iris authentication (Figure 2-3).

Our results also addressed $RQ_{2.2}$ to explore the impact of eye image blur on gaze data utility when animating the eyes of a realistic virtual avatar within a social VR application. Perceptual studies (Sec. 2.6.5.2) identified detection thresholds for the levels of image blur σ on resulting eye animations and demonstrated how increased blur negatively impacted the social characteristics of the virtual avatar. Our takeaway for the explored configuration is that at $\sigma = .0075$, viewers will not consistently notice a difference in the resulting avatar eye animations and maintain positive social characteristics. At levels of $\sigma = .0125$ or higher, viewers are likely to detect a difference and rate social characteristics with neutral or negative responses.

The presented methods provide tools to control the risk of leaking a user's iris biometric while balancing data utility for eye tracking. Early publications of our work spawned a new research sub-area in protecting users from identification based on eye images collected by an eye tracker. Our work demonstrates an approach to balancing security for raw sensor data in the

context of VR applications. Mixed-reality devices integrate multiple sensors, such as ECG or EEG sensors, that also act as biometrics for identifying the user or their sensitive attributes. We imagine that our results will inspire future research on protecting additional XR sensors while enabling cutting-edge sensor data and XR applications.

2.8 Limitations

A fundamental limitation of parameterizing blur by σ in pixels is a dependency on the resolution and field of view of the eye-tracking camera. For example, if a camera has double the resolution but an identical field of view, the blur parameter σ would have to be scaled by two to account for increased pixel resolution along each dimension. Computing σ during hardware defocus for a new camera requires a calibration procedure [120].

Using optics to introduce blur pre-capture is an efficient solution that puts the user in control of the security of their iris biometric; however, this assumes that the user can access and manipulate the focus on the camera. While this is true for most popular glasses-based eye trackers, not all head-mounted XR displays with eye-tracking sensors are accessible to the user or have adjustable focus. More robust solutions might investigate clip-on optics similar to Pittaluga and Koppal [195] that can be used to augment the eye tracker and enable blur-based iris security.

The stimuli used for our perceptual evaluation of animated virtual avatars also have limitations. Notably, our evaluation does not consider the impact of defocus on eye movement characteristics such as blinks, the dynamics of saccades with large amplitudes, or estimated pupil diameter. These characteristics play a role in complex social interactions and are more prominent the closer the user is to the avatar [68]. The stimuli did not include head or mouth movements.

CHAPTER 3
PRIVACY FOR STREAMING EYE-TRACKING DATA

3.1 Introduction

Eye-tracking biometrics extend beyond images collected from raw sensor data. Streams of gaze positions output from an eye tracker are processed to extract features that quantify behavior patterns. The resulting features are then classified to perform identification. Even if blur is applied to eye-tracking images to protect the iris, the resulting stream of gaze positions enables an adversary to recognize a known user. Furthermore, eye trackers will always output a stream of gaze positions, even if non-video-based eye trackers become standardized in the future [253]. This chapter presents a privacy-preserving approach to streaming data to gaze applications in XR.

Eye tracking is a specific case of more general behavioral tracking services in mixed reality, including head, hand, and body tracking. Mixed-reality platforms collect raw data from the native sensors, process it to perform noise removal and event detection, and pass the processed data up the software stack [64]. Current mixed-reality platforms enable third-party content through app stores [241] and web browsers [101], similar to mobile devices.

We propose a privacy-by-design approach to protecting eye-tracking sample data. The approach uses a Gatekeeper API that protects sample data by providing metrics relevant to specific application utilities. For application utilities requiring gaze samples, we propose using standalone privacy mechanisms that modify sample data to reduce the risk of identification from the data stream. Optimal privacy mechanisms reduce the risk of identification to near chance (guessing) rates while retaining utility for the respective data application.

An earlier version of this work was published at IEEE VR 2021 [62] and was nominated for a best journal track paper award. The presented framework served as a motivating template for designing privacy-preserving systems at the PrXR workshop at IEEE VR 2021 [61].

In this chapter, we present a framework for protecting privacy when streaming gaze data to XR applications. We demonstrate that aggregate and event-level eye-tracking applications are enabled by designing a Gatekeeper API. For sample-level applications, we found that an additive Gaussian noise mechanism can reduce the risk of identification from released gaze samples while introducing less than 1.2° degrees of spatial error when input to a deep gaze prediction model.

3.2 Threat Scenario

The problem of applications receiving data and passing it along to colluding apps or parent companies erodes public trust in technology and cannot be "regulated away" completely. For example, Reardon et al. highlighted data leakage attacks on Android devices from popular apps on the Google Play Store [201]. Attacks included collusion between apps that make use of SD card storage to save user information and make it available to other apps where permissions had not been restricted. Such attacks violate existing privacy laws such as GDPR and CCPA, but were still prominently featured on the platform app store. Apps that violate user privacy use the collected data for unintended purposes outside of their main application. Recently, The Weather Channel took location data it mined from users' foot traffic at different businesses, and sold it to hedge funds to inform their investments before quarterly income statements were released.[1] Even with regulation, the weather app collecting location data can collude with an advertising application that belongs to the same parent company. The user will then be served personalized ads based on location: such as car ads appearing after a visit to the car dealership for an oil change. Ads generated with information about which cars were glanced at most or that a motorcycle caught the user's eye can make the user feel that their privacy has been violated without their consent.

This problem becomes even more severe when we recognize that mixed-reality headsets are being explored as a future solution for remote meetings in work environments [125]. Employers today use social media activity outside of work when screening applicants or for firing current employees [81]. Data-oriented companies such as Amazon analyze the social demographics of their workers, such as the percentage of workers below the poverty line, to flag the Whole Foods stores most likely to unionize [102]. In response, a user could log in at work to do job-related training with their known real-world identity, but attend virtual labor union meetings as anonymous User X to avoid negative repercussions.[2,3] An adversary that connects these two

[1] www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html

[2] www.tcf.org/content/report/virtual-labor-organizing/

[3] www.foley.com/en/insights/publications/2015/09/be-careful-what-you-say-during-a-union-organizing/

identities has the power to "out" the user to their work organization. Furthermore, recent investments by the Saudi Arabian government highlight that mixed-reality data may be owned and processed by government entities [162]. Mixed-reality data, including eye tracking, can be used to infer sensitive information such as sexual orientation [206], creating the potential to harm LGBTQ+ individuals in countries where homosexuality is considered a crime [246].

3.3 Eye-Tracking Applications

3.3.1 Aggregate-Level Eye-Tracking Applications

Aggregate gaze data from many viewers are input to applications such as highlighting salient regions using heatmaps [198, 59, 228] and learning perceptual-based streaming optimizations for 360° content [157, 249]. These applications typically rely on a data collection process conducted in research lab environments for a sample of viewers. Viewer data acts as training datasets for machine-learning models, and results from the model performance are then released to inform the deployment of such methods on consumer devices. Using a research dataset provides utility to the consumer without creating privacy risks for them. However, machine-learning models still pose a risk to privacy when they leak information about individuals that contributed to the training dataset [50].

3.3.2 Event-Level Eye-Tracking Applications

Eye movement behavior captured by eye-tracking events, such as fixations, saccades, and smooth pursuit, contribute to gaze-based interfaces [190, 92], evaluating training scenarios [69, 45, 121], and identifying neurodegenerative diseases [187] or ASD [43]. Detecting eye-tracking events enables improved techniques for redirected walking [146, 134, 135], a critical application for VR that expands the usable space of a virtual environment within a confined physical environment. The most common method to quantify an individual's gaze behavior is to mark Areas of Interest (AOIs) within the content and measure how gaze interacts with this region [110]. AOIs are commonly used in UX or web design to evaluate different interfaces based on when and how long a user looks at different regions [252]. Typical metrics for these regions depend on fixation and saccade events only, recording dwell times, the number of fixations or

glances, and fixation order [148, 188].

3.3.3 Sample-Level Eye-Tracking Applications

Multiple critical mixed-reality applications depend on individual gaze samples from an eye tracker with a sampling rate of at least 60Hz. Applications include foveated rendering and streaming [192, 27, 167, 166, 51], which enable deploying immersive VR experiences on low-power and mobile devices. These applications rely on gaze samples to track where the foveal region of the user currently is and predict where it will land during an eye movement to ensure that the user does not perceive rendering artifacts [16]. Similarly, gaze prediction models are trained that predict future gaze points while viewing 360° images or video, and 3D rendered content [113, 114, 115].

Another set of applications that require sample-level data are gaze guidance techniques [214, 215]. Gaze guidance takes advantage of sensitivity to motion in the periphery to present a flicker in luminance that will attract the user's eyes, using eye tracking to remove the flicker before the user can fixate upon the region and perceive the cue [24, 97]. This technique enables manipulation of visual attention and, ultimately, user behavior. For example, gaze guidance in 2D environments improved spatial information recall [23], training of novices to identify abnormalities in mammogram images [231], and retrieval task performance in real-world environments [36]. Gaze guidance also enhances redirected walking techniques in VR by evoking involuntary eye movements and taking advantage of saccadic suppression [236]. Guiding gaze through saccades and manipulating the user allows for the use of a 6.4m×6.4m virtual space within a 3.5m×3.5m physical space, significantly expanding the usable area within VR experiences. This application requires an eye tracker sampling rate of 250Hz or more and sample-level data to know precisely when gaze moves towards the periphery cue. Providing sample-level data with high accuracy at this frequency poses a serious risk to user privacy in the form of gaze-based biometric features that can then be extracted from these gaze positions.

3.4 Related Work

3.4.1 Inferences From Eye Movements

Human eyes reflect their physical attributes. For example, algorithms can estimate the ages of users by monitoring the change in the gaze patterns as they age [177, 255], their gender based on the temporal differences in gaze patterns while viewing faces [221], and their race from the racial classification of faces they tend to look at [25].

Beyond physical attributes, gaze allows rich insights into psychological attributes, such as neurological [149] and behavioral disorders [54, 193, 176]. The eyes can also reveal whether an individual suffers from an affective disorder, as anxious individuals' gaze patterns are characterized by vigilance for threats during free viewing. In contrast, depressed individuals' gaze is characterized by reduced maintenance of gaze on positive stimuli [18]. Eye tracking has also been used to investigate gaze behavior in individuals on the autism spectrum, finding that they generally tend to fixate less on faces and facial features [37, 48]. Body mass index (BMI) status appears to influence gaze parameters that are not under conscious control, allowing BMI estimation when presenting individuals with images of foods of differing caloric content [94]. These risks involve knowledge of eye position and stimuli, whereas user identification can be applied to raw eye movements without knowledge of the stimuli.

3.4.2 State-of-the-Art in User Identification Based on Eye Movements

Gaze patterns can be used to identify individuals as they contain unique signatures that are not under a user's voluntary control [129, 130]. The Eye Movement Verification and Identification Competitions in 2012 and 2014 challenged researchers to develop algorithms that identified users based on their eye movements when they followed a jumping dot (2012) and when they looked at images of human faces (2014). The best models' accuracy ranged from 58% to 98% for the jumping dot stimuli, and nearly 40% accuracy compared to a 3% random guess probability for viewing faces.

Based on recent surveys on eye movements biometrics [86, 209] and literature search, we identified algorithms that successfully identify individual users from their eye movements in

Table 3-1. State-of-the-art gaze-based biometric methods. Key: RBF = Radial Basis Function Network, RDF = Random Decision Forests, CNN = Convolutional Neural Network, STAT = Statistical test, SVM = Support Vector Machine.

Method	Features	Classifier	Dataset	Results
Schroder et al. [222]	Fixation, saccade	RBF	BioEye 2015, MIT data set	IR: 94.1% 86.8%
Schroder et al. [222]	Fixation, saccade	RDF	BioEye 2015, MIT data set	IR: 90.9% 94.7%
George&Routray [93]	Fixation, saccade	RBF	BioEye 2015	IR: 93.5%
Lohr et al. [154]	Windows of gaze velocities	CNN	GazeBase, BioEye 2015	EER: 11.3% 21.1%
Lohr et al. [156]	Fixation, saccade	STAT	VREM-R1, SBA-ST	EER: 10.0% 2.0%
Lohr et al. [156]	Fixation, saccade	RBF	VREM-R1, SBA-ST	EER: 14.4% 5.1%
Eberz et al. [73]	Fixations, binocular pupil	SVM	[73]	EER: 1.9%
Rigas et al. [207]	Fixations, saccades, density maps	Multi-score fusion	[207]	EER: 5.8% IR: 88.6%
Monaco [172]	Gaze velocity & acceleration	STAT	EMVIC 2014	IR: 39.6%

Table 3-1. These algorithms have been evaluated on existing gaze-biometric challenge datasets and the natural viewing of image stimuli in 2D (MIT data set). The method with the best biometric performance produces an Equal Error Rate of 1.88% using pupil-based features [73]; however, the majority of consumer applications in mixed reality do not require pupil diameter. Thus, we selected to implement the RBF approach proposed by George and Routray [93], as it relies only on fixation and saccade events. This method also produces impressive results with VR eye-tracking data [156] and natural viewing of 2D images [222].

3.4.3 State-of-the-Art in Eye-Tracking Security and Privacy

In recent years, privacy concerns related to eye-tracking applications has grown significantly [142, 233, 122, 120, 42, 151]. In response, researchers have developed methods to enhance privacy of aggregate features, like saliency heatmaps [153] and event statistics [232, 40, 83]. These methods have been shown to reduce performance in the classification of gender and identity. However, the methods operate only on extracted gaze

features after processing raw data. Recent work by Li et al. has applied a differential privacy noise mechanism to raw streams of gaze to obfuscate the viewer's gaze relative to AOIs within stimuli over time [151]. The noise mechanism protects gaze positions relative to the size of objects viewed in the scene, so an adversary would not be able to determine what a user was looking at by observing the gaze stream and respective content. The ability to protect biometric identity was also evaluated empirically on the 360_em dataset [10], reducing identification to chance rate. While the chance rate of identification was achieved for one dataset, the formal differential privacy guarantee specifically targets the privacy of what was looked at, not the risk of identifying the user from their data. This chapter develops a threat model based on the streaming of gaze samples and the privacy risk related to biometric identification within an XR ecosystem.

3.5 Methodology

The typical architecture and data flow for an eye-tracking platform is to provide an API to request gaze data or access a stream of data by listening for specific types of events [131, 239]. The API provides access to gaze data and metrics for third-party applications. The eye-tracking platform performs access control for applications and the implementation of gaze estimation.

3.5.1 Eye Movement Biometrics

We define two classifiers for biometric identification using a Radial Basis Function (RBF) network [93, 222, 156], with one network to classify fixation events and one to classify saccade events. This method is analogous to a traditional neural network with an input layer representing a feature vector $\vec{x} \in \mathbb{R}^p$ containing p fixation or saccade features from a single event, one hidden layer consisting of m nodes, and an output layer containing c class scores, one for each unique individual in the dataset. The output class scores are used to measure the individual to which the input feature vector is most similar. Thus, larger scores indicate a higher probability of the fixation or saccade event being from that class or individual. The classifier is implemented identically to the prior works [93]. Classification probabilities are equally weighted between the fixation and saccade event classifiers and summed to classify the eye movements captured from each stimulus in the testing set. The protocol for splitting training and testing data is described in Section 3.6.3.

The single hidden layer is implemented using nodes defined by an activation function $\phi_i(\vec{x})$ and a set of real-valued activation weights $w_{i,c}$, where $i \in [1,2,\ldots,m]$ and $j \in [1,2,\ldots,C]$. The similarity score for a given class c in the output layer is computed as a weighted sum of all activation functions in the hidden layer,

$$Score_c(\vec{x}) = \sum_{i=1}^{m} w_{i,c} \cdot \phi_i(\vec{x}). \tag{3-1}$$

The activation function of each hidden node takes the form of a Gaussian distribution centered around a prototype vector $\vec{\mu}_i$ with spread coefficient β_i. The function is defined as

$$\phi_i(\vec{x}) = e^{-\beta_i ||\vec{x}-\vec{\mu}_i||^2}, \tag{3-2}$$

with shape coefficient β_i and prototype feature vector $\vec{\mu}_i$ defined prior to training the network. Thus, an RBF network must be constructed in two stages by first defining the prototypes and then optimizing the activation weights.

First, k-means clustering is applied to a training set of n feature vectors to determine k representative feature vectors per individual [93, 156]. Through this process β_i and $\vec{\mu}_i$ are defined for each of the $m = k \cdot c$ hidden nodes. The activation function $\phi_i(\vec{x})$ is then defined using the cluster centroid as $\vec{\mu}_i$, and β_i as $\frac{1}{2\sigma}$, where σ is the average distance between all points in the cluster and the centroid $\vec{\mu}_i$.

Second, the activation weights $w_{i,c}$ are learned from the same training data used to define the activation functions. Weights are trained using only fixation or saccade features from the training set. Training can be implemented using gradient descent [224], or by the Moore–Penrose inverse when setting up the network as a linear system [93]. The latter method is implemented in this work by defining the RBF network using an activation output matrix $A_{n\times m}$, where rows consist of the n training feature vectors input to the m previously defined activation functions, weight matrix $W_{m\times c}$ comprised of activation weights $w_{i,c}$, and an output matrix $Y_{n\times c}$ generated as a one-hot encoding of the ground truth identity labels. The RBF Network is defined as $A \cdot W = Y$ using matrix multiplication.

The weight matrix W is then learned by computing $W = A^* \cdot Y$, where A^* is the Moore-Penrose inverse of A computed using MATLAB's *pinv* implementation. Class score predictions $\hat{Y}$ are then generated for the testing data $\hat{A}$ by computing $\hat{A} \cdot W = \hat{Y}$. Every sample in the testing set is classified as the class label with the maximum score. The class scores from all events are summed together, and then the class with the maximum value is returned to classify a stream of events. Scores from the fixation RBF and saccade RBF are combined by summing the average of scores from each network for equal contribution to the final classification.

3.5.2 Threat Model

We assume that the components comprising the eye-tracking platform and API are trusted, i.e., the integrity of the hardware and software could be attested through mechanisms such as secure boot [17] and integrity measurement [218], and we assume that the operating system is protected, e.g., through SELinux mandatory access controls [174]. The adversary can examine all data transmitted to the eye-tracking applications and seeks to use this information to re-identify the user. A malicious application can collude with other applications by sharing information through overt or covert channels [159] in order to re-identify users.

Our privacy-preserving solution is focused on preventing biometric identification of users from their gaze data when collected by colluding applications. User data comes from a known set of identities linked by computing gaze features while performing the same task in both applications. The success criteria for our privacy mechanisms is to reduce the successful identification rate from eye-tracking features to chance rates. Chance rates represent the same accuracy as randomly guessing the identity, or $\frac{1}{\#Ppts}$. A mechanism that reduces identification rates to chance can defend against an adversary's attack for a certain task, the volume of data, feature set, and classification model.

3.5.3 Gatekeeper API

The simplest way to provide a gaze API would be to pass the raw gaze data to applications. At any point in time, the application would be able to request `getGazePosition()`. From this, the application would be able to compute fixations, saccades, and dwell time; in particular, an

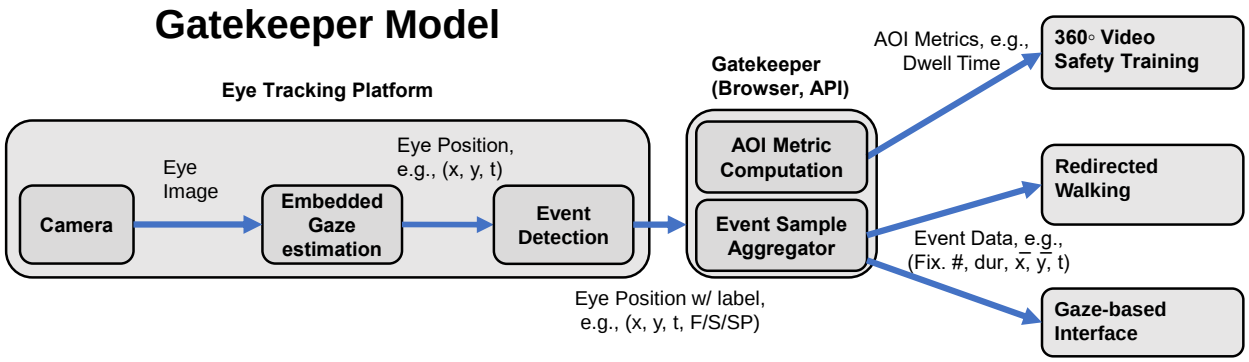

Figure 3-1. Illustration of the Gatekeeper framework. The Gatekeeper protects identity by delivering relevant data at different levels directly through the API, while withholding raw gaze samples that contain biometric features. This approach cannot be used directly with applications that require raw gaze samples.

application would be able to compute fixations in an AOI, time to first saccade into the AOI, and dwell time in the AOI. Providing raw gaze data also allows for computation of gaze velocity and other features commonly used for biometric identification [93, 86, 222]. Allowing for raw gaze access in an untrusted context, such as the web, allows arbitrary apps to re-identify users.

However, we can modify the gaze API to be privacy-preserving by acting as a Gatekeeper. Privacy vulnerabilities are caused by the design assumption that the application is benign and the data is used only for the purpose for which it is collected. As discussed previously, applications may not be benign, and connecting user data across devices will allow for richer inferences about that user. This threat motivates our proposed Gatekeeper design. An added benefit of our proposed design is that the Gatekeeper model provides desired metrics directly to applications instead of requiring them to process streamed user gaze data and calculate the metrics themselves.

3.5.3.1 Enabling AOI Metrics

Advertisers and other AOI applications are interested in the number of fixations and the fixation dwell time in a predetermined AOI [194]. Under the Gatekeeper framework, an API allows requests for metrics instead of passing along raw gaze positions. For example, a `getFixations` method takes a rectangular area and returns a list of fixations that had occurred in that area, and a `getDwellTime` method takes as input a fixation and returns in milliseconds the dwell time of the fixation. Additionally, we design a `getSaccades` method that would return a list

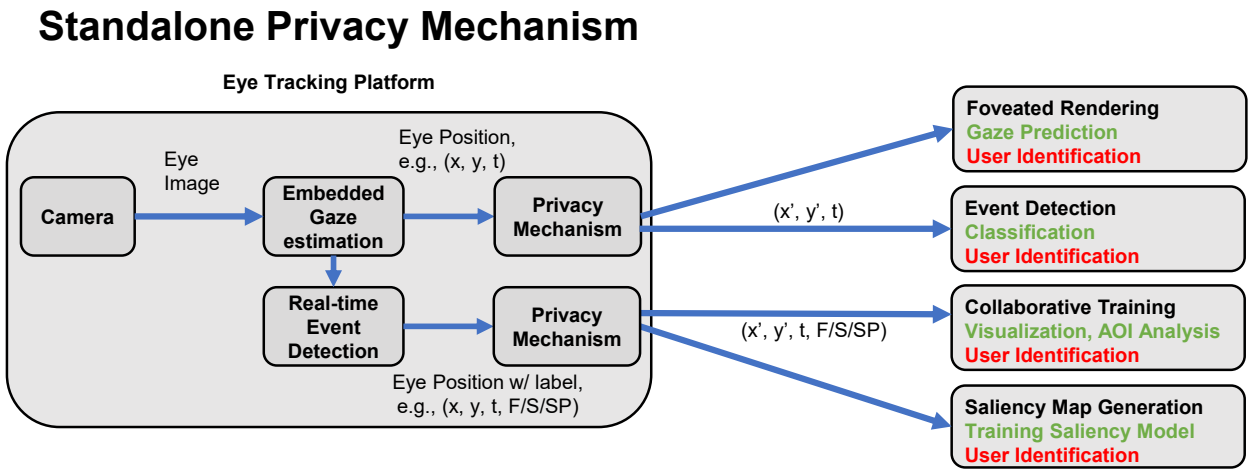

Figure 3-2. Illustration of standalone privacy mechanisms for eye-tracking data. In scenarios where a Gatekeeper API cannot be implemented, we instead apply a privacy mechanism to raw gaze samples to serve applications that use gaze samples or event data directly. Noise and downsampling are applied during eye movement events to reduce the risk of user re-identification from the stream of released gaze samples.

of saccades into the AOI. Saccades are a strong classifier feature for identity when raw gaze points are included; however, we mitigate this risk by providing only lower-dimensional summary data.

3.5.3.2 Enabling Real-Time Event Data

In situations such as gaze-based interfaces and redirected walking, applications need to be notified when a new fixation or saccade occurs instead of querying for all fixations or saccades. In this scenario, we can use an `EventListener` model instead of a query-based model. When a new event occurs, the `EventListener` will be notified and given the event data, `(x, y, t)` and a boolean indicating if it is a fixation, saccade, or smooth pursuit. More complex eye movements are difficult to detect in real-time with the sampling rate of mixed reality eye-tracking devices and typically are not implemented in real-time applications.

Our model for streaming event data sends an event when the eye movement has concluded. In a gaze-based interface, the application needs to be notified that a smooth pursuit occurred and where it landed [21]. In applications such as redirected walking, it is critical to know when a saccade begins to take advantage of saccadic blindness [236, 134, 135]. In this case, one mode of the `EventListener` would indicate when a saccade event has started instead when it has finished.

3.5.3.3 Discussion

It is important to note that the Gatekeeper API is explicitly designed to provide AOI metrics and summary data of eye movement events. The streamed metrics and event data are identical to the original data implying that there would be no loss in utility when using the Gatekeeper, and the risk of re-identification is eliminated as gaze samples are never shared outside of the eye-tracking platform. This covers a range of gaze applications for XR, ranging from AOI analysis in virtual training environments to redirected walking and gaze-based interaction. The Gatekeeper API does not scale to address applications such as foveated rendering or short-term gaze prediction, which requires raw gaze samples with low latency for their utility.

3.5.4 Standalone Privacy Mechanisms

While a set of applications will be able to function with the API mentioned above, core mixed-reality applications will require sample-level data. In these scenarios, the eye-tracking platform must stream sample-level data, and it is impossible to simply abstract data using a privacy-preserving API. Therefore, we proposed using a privacy mechanism to manipulate gaze samples as they are streamed to increase privacy.

Privacy is increased when the accuracy of user identification is reduced. Eye movement biometrics are based on features derived from common eye events, such as fixations and saccades. Thus, we proposed that privacy mechanisms can be deployed in conjunction with real-time event detection to modify the gaze samples that are released during events (Figure 3-2). Implementing privacy mechanisms in this manner preserved the event boundaries of the gaze data and enabled measuring the utility lost as a result of the different privacy mechanisms and parameters.

We considered the following privacy mechanisms: adding Gaussian noise to raw gaze data, temporal downsampling, and spatial downsampling.

The data received by the privacy mechanisms are defined to be a time series where each tuple is comprised of horizontal and vertical gaze positions (x, y), a time stamp t, and the event label e assigned to the sample. For example, $X = \{(x_1, y_1, t_1, e_1), (x_2, y_2, t_2, e_2), ..., (x_G, y_G, t_G, e_G)\}$ is a set of G gaze positions in our representation of the data. This data is processed via a privacy

mechanism and the processed output is a time series X'. Additional variables for the privacy mechanisms are defined in Table 3-2.

Additive Gaussian noise: Noise is sampled from a Gaussian distribution of zero mean and standard deviation σ defined in visual degree and added to the gaze positions. Noise is independently sampled for horizontal and vertical gaze positions as $X' = \{(x_1 + N(0,\sigma), y_1 + N(0,\sigma), t_1, e_1), (x_2 + N(0,\sigma), y_2 + N(0,\sigma), t_2, e_2), ..., (x_G + N(0,\sigma), y_G + N(0,\sigma), t_G, e_G)\}$.

Temporal downsampling: Temporal downsampling reduces the temporal resolution of the eye-tracking data stream. Downsampling is implemented by streaming the data at a frequency equal to the original sampling rate divided by a scaling parameter K. The output time series is defined as $X' = \{(x_{(K \cdot p)+1}, y_{(K \cdot p)+1}, t_{(K \cdot p)+1}, e_{(K \cdot p)+1}), ...\}$ for all integers $p \in [0, \frac{G}{K}]$. For example, with a scaling parameter of two, the private gaze positions are defined as $X' = \{(x_1, y_1, t_1, e_1), (x_3, y_3, t_3, e_3), (x_5, y_5, t_5, e_5), ...\}$, retaining only every other gaze sample. For a scaling parameter of three, $X' = \{(x_1, y_1, t_1, e_1), (x_4, y_4, t_4, e_4), (x_7, y_7, t_7, e_7), ...\}$.

Spatial downsampling: Spatial downsampling reduces the resolution of eye-tracking data down to a discrete set of horizontal and vertical gaze positions. The scene is divided into a grid and each gaze sample is approximated by the grid cell that it lies within. Spatial downsampling is performed by defining a target equirectangular domain spanning 180° vertically and 360° horizontally with M rows and N columns. For smaller values of M and N there are less possible positions, and thus reduced spatial resolution. Raw gaze positions $(x \in [0, 360^\circ), y \in [0, 180^\circ), t)$ are transformed by first computing the horizontal step size $\delta_y = \frac{180}{M}$ and vertical step size $\delta_x = \frac{360}{N}$. Downsampled gaze positions are then computed as $(\lfloor \frac{x}{\delta_x} \rfloor \cdot \delta_x, \lfloor \frac{y}{\delta_y} \rfloor \cdot \delta_y, t)$, where $\lfloor \cdot \rfloor$ represents the floor function that rounds down to the nearest integer.

Spatial downsampling is parameterized as a factor L relative to an equirectangular domain of $M = 2160$ and $N = 3840$, mapping to a domain of $M = \frac{2160}{L}$ and $N = \frac{3840}{L}$. For example, an input downsampling factor of L equals two will result in $M = 1080$ and $N = 1920$, a factor of L equals three will result in a resolution of $M = 720$ and $N = 1280$, and so on.

Table 3-2. Standalone privacy mechanism variable definitions.

Variable	Description
x	Horizontal gaze position
y	Vertical gaze position
t	Timestamp
e	Event label: Fix. (F), Sacc. (S), Smooth Pursuit (SP)
X	Input time series of gaze samples
G	Number of gaze positions in time series
X'	Output privacy-enhanced time series
K	Temporal downsample factor relative to sampling rate
L	Spatial downsample factor relative to 3840×2160
M	Number of rows in equirectangular projection
N	Number of columns in equirectangular projection
δ_x	Horizontal step size: $\frac{360}{N}$
δ_y	Vertical step size: $\frac{180}{M}$

3.6 Study: Evaluating Standalone Privacy Mechanisms

3.6.1 Research Questions

Privacy-utility trade-offs are explored for each privacy mechanism across several VR datasets to identify the best performing of the three mechanisms and assess the current risk of user re-identification for each dataset.

Formally we propose the following research questions:

- $RQ_{3.1}$: *Which privacy mechanism had the lowest impact on utility while reducing identification rate the most for each application?*
- $RQ_{3.2}$: *Is the optimal privacy mechanism able to reduce identification rates from an RBFN model to chance for all datasets?*

3.6.2 Implementation

The performance of eye movement biometrics depends on the amount of data, the task being performed, and the features used for classification. State-of-the-art approaches leverage statistical features extracted from fixation and saccade events to perform identity classification with an accuracy as high as 95% [93, 222]. We consider an attempt at re-identification by collecting eye movements for a specific stimulus or task and computing a sequence of features that can be used for identification. Privacy mechanisms are applied to both the training and the testing set. Noisy data is used for both sets of data to satisfy the assumption that there is a trusted

platform for processing data before streaming it to colluding third-party applications. The identification protocol relies on eye-tracking features generated during a specific task from a set of VR stimuli to train the identification model. The model is then tested on data from the same task but a different set of stimuli to emulate two apps colluding to recognize a user.

3.6.3 Protocol

In order to evaluate the privacy mechanisms on how effectively they prevented an adversary from re-identifying the user, we selected five existing VR datasets. Table 3-3 presents characteristics of each dataset included in analysis. Datasets were selected to have diversity in the number of participants, the number of stimuli presented, and the task being performed.

The evaluation protocol for the RBF-based biometric (Sec. 3.5.1), illustrated in Figure 3-3, is derived from [222], where a stream of gaze data collected from multiple participants viewing numerous static images is used for training and testing the identity classification. The size of the training and testing sets are defined by the number of stimuli from which gaze data is used. For example, with a train/test split of 50%/50%, gaze data from half of the dataset is selected randomly and used for training and the other half for testing. Each participant is present in both the training set and the testing set.

3.6.4 Metrics

Biometric identification was performed by classifying streams of gaze data from multiple stimuli into one user identity. The identification rate was then computed as the number of correct matches divided by the number of classifications. Identification rates reported are the average over ten runs with random stimuli selected as part of the training and test set to account for variance in stimuli content.

3.6.4.1 Gaze Prediction

Using the DGaze architecture and dataset, we evaluated the ability to predict ground-truth gaze position 100 ms into the future when gaze data output from a privacy mechanism is used as the testing data and as both the training and testing data. The utility was measured as angular gaze prediction error for each input gaze sample, with lower values indicating higher accuracy.

Table 3-3. Characteristics of VR eye-tracking datasets.

Dataset	# Ppts.	# Stimuli	Avg. # stimuli	Stimuli duration	Stimuli type
ET-DK2 (ours)	18	50	50	25s	Images
VR-Saliency [228]	130	23	8	30s	Images
VR-EyeTracking [250]	43	208	148	20s-70s	Videos
360_em [10]	13	14	14	38s-85s	Videos
DGaze [114]	43	5	2	180s-350s	3D scene

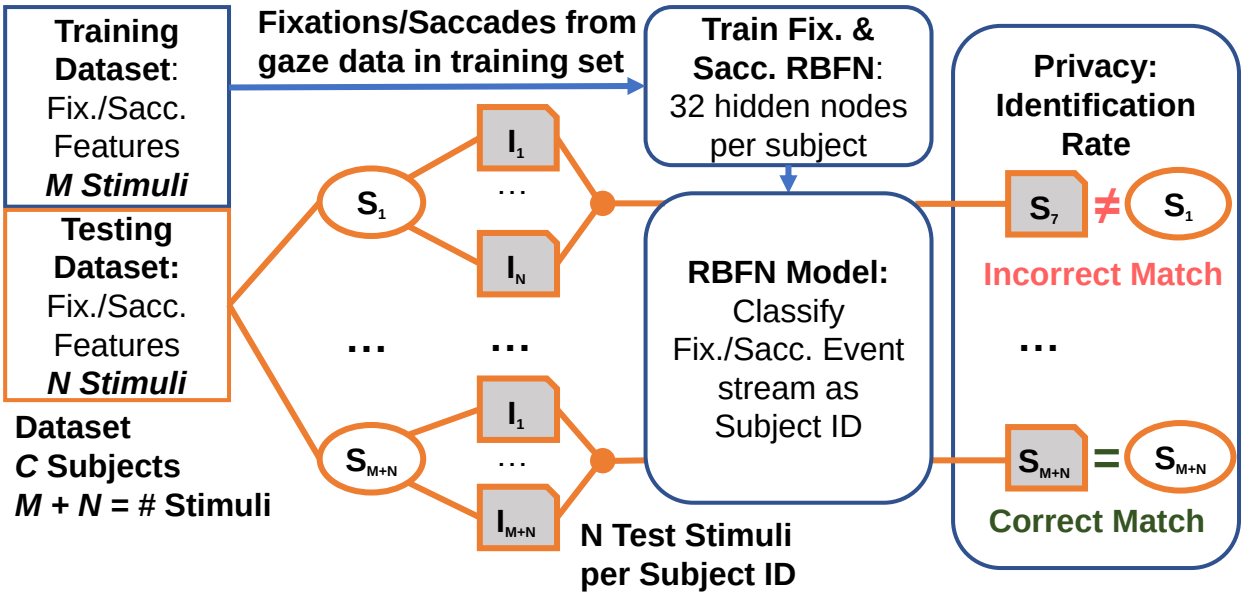

Figure 3-3. Evaluation protocol for computing Identification Rate with the gaze-based biometric classifier.

3.6.4.2 AOI Analysis

The most common form of eye-tracking analysis is performed using static AOIs defined within image content [148, 188]. AOI analysis is used to study gaze behavior during social interaction [29], while viewing websites [252], and to evaluate content placement in 3D environments [13], among many other applications. A key AOI metric that is robust to fixation detection parameters is dwell time [188]. Dwell time measures how long a viewer's gaze fell within an AOI, and allows for comparison between which AOIs attracted the most attention. We evaluated the loss in utility between ground truth and modified gaze data by computing the Root Mean Squared Error (RMSE) between AOI dwell times. AOI utility was measured for the ET-DK2 dataset, as two rectangular AOIs were marked within each image. The AOIs corresponded with salient objects, such as people or natural landmarks within the scene.

3.6.4.3 Saliency Maps

Eye-tracking data are also used to generate saliency maps, which represent a probability distribution over visual content that highlights regions most likely to be looked at by a viewer [148]. Saliency maps are generated from aggregate eye-tracking data across many viewers and are used to train and evaluate deep learning models for saliency and scanpath prediction [19, 49]. Saliency metrics are computed for both 360° images (VR-Saliency), and 360° video (VR-EyeTracking and 360_em). We compute KL-Divergence [148] to measure the deviation between saliency maps from raw and modified data. Among other saliency map metrics, KL-Divergence is most sensitive to false positives and provides a differentiable loss function that is useful for training deep networks [46].

3.6.5 Results

3.6.5.1 Privacy

Figure 3-4 presents the mean and standard deviation of identification rates for each dataset, along with baseline rates corresponding to random guessing. For all datasets, identification rates were highest when there was more training data than testing data, i.e., a 75%/25% split. The ET-DK2 dataset with 18 individuals produced the highest identification rate of 85%, while DGaze with 43 individuals produced the lowest identification rate of 2%.

Figure 3-5 presents the mean and standard deviations achieved for privacy mechanisms applied to each dataset. A training/testing split of 75%/25% was used to generate these results. We observe that Gaussian noise achieves the most privacy, reducing the identification rate of ET-DK2 from 85% to 30% on average. Temporal downsampling is not recommended, as it had the least observed impact on identification rate, and event detection is degraded at sampling rates less than 120Hz [253].

3.6.5.2 Utility

The utility of eye-tracking data depends on the context of the application. Thus, we evaluated the impact of our privacy mechanisms at three different scales: sample-level gaze points, individual-level gaze behavior, and aggregate-level gaze behavior over many individuals.

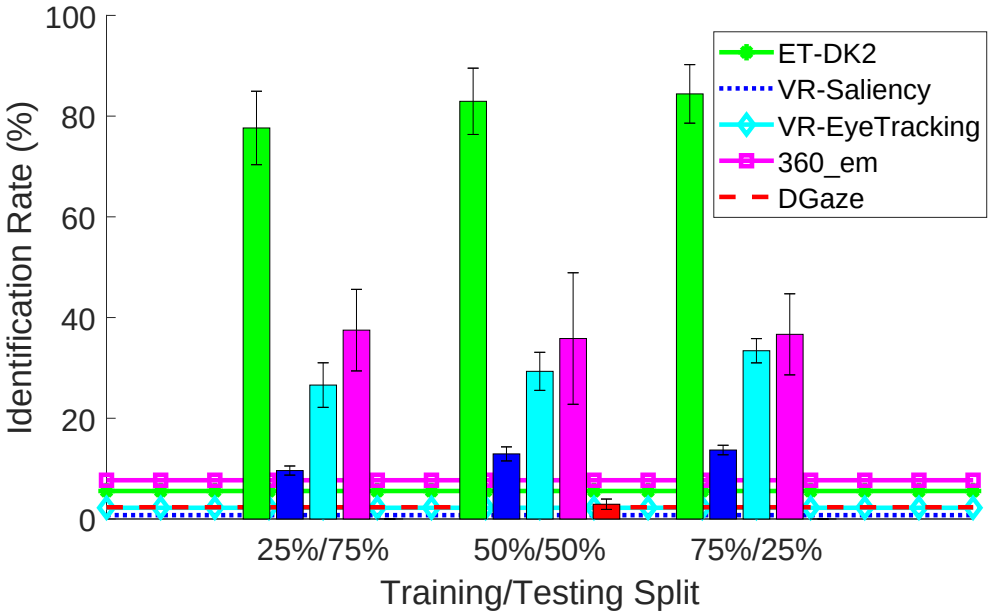

Figure 3-4. Mean and standard deviations of identification rates across datasets. Datasets included 360° images (ET-DK2, VR-Saliency), 360° videos (VR-EyeTracking, 360_em), and 3D rendered scenes (DGaze). Lines for each dataset indicate a baseline of random guessing for the given number of subjects.

First, we evaluate sample-level utility by computing gaze prediction error using the DGaze neural network architecture, then, individual-level utility by computing dwell time for AOIs defined in the ET-DK2 dataset, and finally, we compute aggregate-level utility measures for generating saliency heatmaps of 360° images and video by computing KL-Divergence for the VR-Saliency, VR-EyeTracking, and 360_em datasets.

Gaze prediction: Evaluating gaze prediction accuracy involved configuring the DGaze neural network to predict gaze position 100ms into the future, which as a baseline produces an average gaze prediction error of 4.30°. The DGaze prediction model combines fully connected networks that take saliency information as input with 1D convolutional layers that process sequences of head, gaze, and object positions to predict future gaze positions. Data from the past 500 milliseconds are used as input for the prediction model.

Gaze prediction error using the pre-trained DGaze model was as high as 9.50° for the Gaussian mechanism, more than double the baseline gaze prediction error reported in [114]. Next, we evaluated performance by re-training the DGaze model from scratch and applying

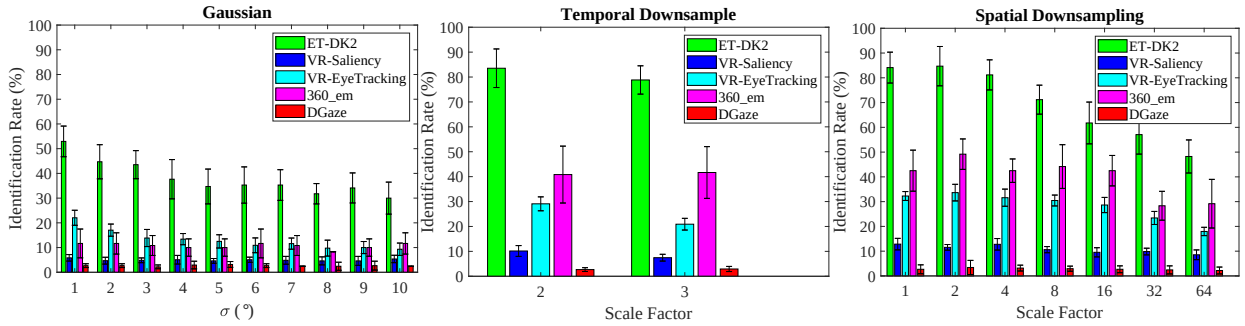

Figure 3-5. Mean and standard deviation of identification rate for each privacy mechanism with different internal parameters. Gaussian noise generates the lowest observed identification rates across all datasets, while temporal downsampling has the least impact.

privacy mechanisms to both the training and testing dataset. Applying noise to training data resulted in much lower prediction errors, with results as low as 5.44°, comparable to the 4.30° reported in [114].

Introducing the privacy mechanism to training and testing data implies that raw gaze data is not shared with any party during model training and deployment. Our experiments indicated that learning a reasonable gaze prediction model is still possible without access to the raw gaze data. Withholding raw gaze data from the training dataset is desirable, as it removes the need to safeguard additional data and alleviates the risk of membership inference attacks [50]. Future gaze prediction models will improve in prediction performance, evidenced by a 11.7% decrease in error between DGaze and the successor FixationNet [113]. Future advancements will further decrease the absolute gaze prediction error when using gaze data output from the privacy mechanisms and retain higher utility.

AOI analysis: RMSE in dwell time computation for additive Gaussian noise and temporal downsampling was found to be below 40ms, which is insignificant for applying AOI metrics, as a fixation typically lasts 200ms [219, 254]. However, for spatial downsampling, an RMSE of 247ms is introduced, which is greater than the length of one visual fixation. While being a few fixations off on average may not have a large effect on AOI applications such as user experience design, it may be noticeable in scenarios with multiple small AOIs close together or when a stimulus is only viewed for a short period [185].

Saliency map generation: The spatial errors introduced by the privacy mechanism may cause regions highlighted by the saliency map to shift or spread out, leading to larger KL-Divergence values. A recent survey revealed that the best performing model in predicting human fixations produced a KL-Divergence of 0.48 for the MIT300 dataset, with baseline models producing values of 1.24 or higher [38]. We observed that spatial downsampling produces the largest KL-Divergence on average of 0.1293, while Gaussian and temporal downsampling mechanisms produce much smaller values of 0.0367 and 0.0019, respectively. Spatial downsampling introduced errors of approximately a fourth of the existing gap in fixation prediction. Errors of this magnitude will cause saliency maps generated from spatially downsampled gaze data to deviate from ground truth and negatively impact the performance of models that use the maps for training.

3.7 Discussion

We proposed a Gatekeeper model to alleviate biometric authentication by apps that need AOI metrics or event-specific data for their utility. This model provides API calls that return desired metrics and summary information of fixation and saccades to applications without providing streams of raw gaze data, which suffices for certain classes of mixed reality use cases. However, streaming gaze data is required for use cases such as foveated rendering. We propose that in this case, privacy mechanisms can be applied to the raw data stream to reduce the identification rate while maintaining the utility needed for the given application.

We evaluated three privacy mechanisms that reduce the risk of identification: additive Gaussian noise, temporal downsampling, and spatial downsampling to answer $RQ_{3.1}$. Our evaluation found additive Gaussian noise performed best compared to the other evaluated methods in reducing identification rate from 85% to 30% while retaining practical data utility for AOI analysis, gaze prediction, and saliency map generation. Table 3-4 summarizes the impact of standalone privacy mechanisms and highlights the applications enabled by the Gatekeeper API. The computed identification rates were above chance for all evaluated privacy mechanisms and parameters, answering $RQ_{3.2}$ and indicating that additional exploration of mechanisms is needed

to achieve biometric privacy for eye-tracking datasets.

The presented methods provide a template for developing a privacy-preserving system for individual sensors with XR applications. Privacy-by-design is an ideal approach for protecting user privacy; however, such designs are at odds with applications that require accurate low-level data. Our initial exploration shows that we can add privacy to real-time eye-tracking systems while understanding the impact on XR applications. Standalone privacy mechanisms do not require information about the scene being viewed by the user as inputs to the algorithm, in contrast to the Kalɛido streaming approach [151]. Practitioners can deploy our methods to let users control their data privacy level, and explore the implications on additional applications or when considering different XR sensors.

Table 3-4. Summary of utility loss and impact on identification rates for standalone privacy mechanisms and data applications. Check marks indicate that the application is enabled by a Gatekeeper API and does not require a standalone mechanism.

Application	Gatekeeper	Standalone Mechanism	Utility loss	Identification rate
AOI Analysis	✓	Gaussian	36 ms	85% → 30%
AOI Analysis	✓	Spatial downsample	247 ms	85% → 48%
AOI Analysis	✓	Temporal downsample	6 ms	85% → 79%
Saliency maps	✓	Gaussian	KLD = .036	47% → 14%
Saliency maps	✓	Spatial downsample	KLD = .129	47% → 29%
Saliency maps	✓	Temporal downsample	KLD = .002	47% → 42%
Gaze prediction	×	Gaussian	1.14°	3% → 2%
Gaze prediction	×	Spatial downsample	0.51°	3% → 3%
Gaze prediction	×	Temporal downsample	0.22°	3% → 3%

3.8 Limitations

The key limitation of the Gatekeeper approach is that it only applies to aggregate and event-level eye-tracking applications. Such applications do not require gaze samples, and the threat model considered in this chapter is based on withholding gaze samples as they can later be processed for user identification. The standalone mechanisms applied when streaming gaze samples serve as a set of baseline methods. Further work should explore combinations of mechanisms or propose alternatives that can reduce the identification rate further.

The current threat model assumes a trusted platform that can implement the Gatekeeper or standalone mechanisms without the risk of leaking sample data. In cases where the platform

processing data itself cannot be trusted, there is a need for alternative solutions, such as hardware modifications controlled by the user that impact gaze estimation and protect raw samples.

Our characterization of identification from gaze features is based on one biometric authentication approach (RBF). As newer identification approaches are developed [154], we will need to continuously consider new privacy mechanisms and evaluate the identification risk across applications and datasets.

Our data utility characterization employed offline model accuracy analysis in terms of gaze prediction error. The computed error results provide a quantitative measure of how well private data can be used to train and test a gaze prediction model. However, the trained model was not evaluated with a user study to understand the impact on perceived differences between the model trained on private data and the original model.

CHAPTER 4
PRIVACY FOR EYE-TRACKING DATASETS

4.1 Introduction

Re-identification attacks in literature have been extensively explored for social networks [181], location data [196], and medical data [75]. Real-world re-identification attacks have been demonstrated to learn the medical prescriptions of a politician [237] or reveal the Netflix preferences of half of a million users [180]. Re-identification attacks have real implications for user privacy. For example, in the Netflix dataset attack, a woman sued the company over the risk that her leaked viewing patterns would reveal her sexual orientation to her family [227]. In eye tracking, there are multiple algorithms that can authenticate and identify a user based on eye movement data alone [93, 230, 222, 154]. Numerous datasets of eye-tracking data for XR applications are publicly available [250, 228, 9, 232, 62, 76, 112]. There is a clear trend towards increased datasets in public repositories and large-scale data collection to support deep learning model training and deployment within the major XR platforms. The risk of making available or leaking large-scale datasets, coupled with the deployment of integrated eye trackers in the next wave of consumer XR devices [79], motivates the need for methods that retain utility for developing eye-tracking applications while addressing and reducing privacy risks for users.

Surveys by Adams et al. [8] and Steil et al. [232] have established that both users and developers have privacy concerns over XR and eye-tracking data collection and how they are applied to make inferences about the user. For example, XR developers have cited that they are aware of privacy concerns for users and share their sentiments; however, most developers are not experts in these fields and thus lack the tools to address topics like ethics or privacy issues. For users, survey participants have indicated that they would be willing to accept beneficial XR applications that collect eye-tracking data if they are sharing the data with trusted governmental health agencies or with a university for research purposes. The same users also responded that they would not share their data publicly or with private services unless there were constraints on how the data would be used. Privacy-preserving mechanisms provide a straightforward solution by adding privacy at the data level. Legal approaches such as GDPR address risks by restricting high-level access control and data sharing.

Privacy laws and regulations protect traditional biometric identifiers, such as iris patterns and face scans [109]. The established privacy laws protect users with a framework that limits how long biometrics are stored before being discarded and manage data access by requiring user consent before biometric data are shared or sold to other private entities. However, legal scholars have pointed out that these privacy laws rarely hold up in court and would not apply to behavioral data streams, as they feature ambiguous wording over what data is considered a biometric [210]. A lack of enforceable privacy laws and data release standards implies that XR platforms could store or sell identities through eye-tracking and behavioral data captured alongside demographics, typically used for personalized ads on the web [55]. Beyond personalized ads, identifying users could lead to more harmful consequences such as stalking and cyber-bullying in the context of XR platforms and communities [8, 158].

Prior work has addressed re-identification attacks on eye-tracking datasets through mechanisms that add privacy noise to released data. Privacy mechanisms for eye-tracking saliency maps, feature data, and gaze samples achieve differential privacy (DP) at different stages in the eye-tracking pipeline. DP is a formal privacy guarantee that provides a mathematically proven bound on the amount of information leaked by the presence of a single element in the dataset. DP guarantees' implications depend on the type of data protected and what can be inferred from such data.

For example, a DP mechanism for saliency maps protects the contribution of any individual's fixation locations on the released saliency distribution, even in a worst-case scenario where data from all users had leaked. Fixation locations of an individual could reveal personal interests within web pages or behavioral diagnoses in an educational environment [153]. When considering gaze-based features extracted from fixations, saccades, blinks, and pupil responses, sensitive attributes about the user such as age [255], sexual orientation [203], and personality traits [30] can be inferred. DP in this context protects the feature values extracted from the gaze data recorded from each user, obscuring the ability to train models that infer such attributes [232].

For the case of gaze samples, i.e., time series of where the user is looking at any point in

time, DP applies to protect against singling out what specific users were paying attention to [151]. For samples, a spatial DP bound provides privacy within the content being viewed. Li et al. vary the spatial bound based on the width of an object or AOI in the scene, such as a human face. The method limits the adversary's ability to distinguish what was looked at as long as gaze positions fall within the spatial bound. Practically, this means what is gazed upon within the span of an AOI region could not be inferred. For the example of face AOIs, an adversary cannot detect whether a user had increased eye contact that could reveal whether they are familiar with the other person or not [169]. In contrast, a spatial DP bound based on the distance between AOIs would protect saccadic shifts in the user's gaze between the different regions, protecting against being able to determine which AOI was looked at most. When considering transitions between AOIs that correspond to faces, shifts could reveal how attentive a user is at maintaining eye contact, showing behavioral markers for autism [226], or which human face they were most attracted too [144].

DP has a standard definition, though how it is applied to data determines how to interpret the implications of the privacy guarantee. With DP protection comes an inevitable loss to data utility [136]. The negative impact on data utility must be determined for each application to understand where the DP privacy-utility trade-off is practical. Prior to this work, DP methods were the only formal privacy guarantee for eye-tracking data.

We developed privacy mechanisms for eye-tracking datasets that serve as an alternative to DP. The proposed mechanisms achieve *k*-anonymity and plausible deniability (PD) for both eye-tracking feature and sample datasets. The ability of our privacy mechanisms to protect against re-identification while enabling eye-tracking applications create a trade-off between privacy and utility for XR applications. Our results demonstrate that eye-tracking feature mechanisms mitigate re-identification attacks while retaining the dataset's utility for training a document type recognition model (Study 1, Sec. 4.4). For datasets of sample data, synthesis methods and DP noise can prevent re-identification while supporting utility for training activity recognition models and gaze prediction models (Study 2, Sec. 4.5). An earlier version of this work was published at ETRA 2022 [60]. The ETRA 2022 paper presents the feature mechanisms

and results from Study 1. The methods and results from Study 2 are intended to be submitted to the IEEE TVCG journal by the end of August 2022.

4.2 Related Work

4.2.1 Privacy Guarantees for Eye-tracking Data

Mechanisms that achieve formal privacy guarantees for protecting eye-tracking data against re-identification attacks apply to gaze samples [151] and features extracted from gaze data [232, 41]. Table 4-1 lists existing mechanisms that achieve formal privacy guarantees for eye-tracking data, the type of input data, and how the mechanism was adapted to eye tracking. The only formal privacy guarantee that has been previously explored was DP. While DP is popular in the privacy community due to the robust definition, there is an inevitable trade-off between increased DP privacy and lower data utility [136].

The two most prominent data types listed in Table 4-1 are statistical features and gaze samples. Statistical features refer to statistics extracted during a fixed time window, such as the count of small, medium, or large amplitude saccades or the average fixation duration [44]. Alternatively, statistical features can be extracted from each identified fixation and saccade event, capturing information like the average gaze velocity during a saccade or the spatial dispersion of gaze points during a fixation [93]. Statistical features summarize eye movement behavior, and a wide array of features have been explored for biometric identification of users (Section 3.4.2). In addition, the same sets of features can also be used to train a classifier for sensitive attributes, such as age or gender [232, 41].

DP applied to feature datasets protects the feature values of each individual from being released. Specifically, DP bounds how much the output distribution of data from the mechanism changes in the case where any one feature vector in the dataset is omitted or included. DP is traditionally applied to output aggregate data, such as the average heights of a demographic or the average daily power consumption of homes on a specific block. In the context of eye-tracking datasets, released data matches the same format as the input with separate files for each individual. The result of the DP guarantee means that the variation across individuals is reduced,

Table 4-1. Privacy mechanisms for eye-tracking data with formal privacy guarantees. Shaded rows indicate our mechanisms. Key: DP = Differential Privacy, PD = Plausible Deniability.

Mechanism	Guarantee	Data type	Mechanism input	Adaption to eye tracking
Gaussian [153]	ε, δ-DP	Saliency maps	User fixation map	Adapt DP noise mechanism [70] to protect fixation counts over image pixels
Exponential-DP [232]	ε-DP	Statistical features	Gaze features extracted over window of time t	Adapt DP Noise mechanism [72] applied to features independently
DCFPA [41]	ε-DP	Statistical features	Gaze features extracted over window of time t	Adapt Fourier DP mechanism [200] to include difference and chunking of sliding windows over time
k-same-select sequence (ours)	k-anonymity	Statistical features	Gaze features extracted over window of time t	Randomly group features and apply k-same-select [98] over a sequence
Task-based Marginals (ours)	k, γ-PD	Statistical features	Gaze features extracted over window of time t	Apply Marginals Generative Model and PD test [34] to each task
Kalεido [151]	ε, w, r-DP	Gaze samples	Window of w gaze positions, spatial bound r	Adapt spatial DP mechanism [15] to incorporate a sequence [133] of gaze positions relative to ROIs in scene
k-same-synth (ours)	k-anonymity	Gaze samples	Gaze positions with event labels	Apply k-same-select sequence to parameters of models that generate event gaze positions
Event-synth-PD (ours)	k, γ-PD	Gaze samples	Gaze positions with event labels	Sample generative model for event gaze positions and apply PD test to each task

forcing the released data to be more homogenous and making individual differences harder to detect. Thus, identifying individual users or detecting group traits such as age, gender, or ethnicity becomes more difficult as privacy noise is added and masks individual differences. Masking individual differences enforces that less information is leaked by the presence of any particular element released in the dataset. However, the privacy noise needed to achieve such a guarantee also reduces the utility of the released data by masking valuable insights.

The only mechanism for DP applied to sample data has adapted existing definitions to

protect the spatial-temporal trace of gaze positions. DP in the context of spatial data ensures a bound on how much output locations change within a spatial radius around the original data. The only DP guarantee achieved for gaze samples is from the kalεido mechanism. The kalεido approach protects gaze positions at most a distance of r away from each other within windows of size w.

The algorithm runs in real-time and allows for streaming of gaze samples with a DP guarantee; however, the context of this DP guarantee does not provide theoretical protection against re-identification. It has been demonstrated that the amount of privacy noise added with kalεido does reduce the risk of re-identification to chance for the 360_em dataset. Yet, there is no analytical method to directly link the DP parameter ε with the theoretical risk of re-identifying the user from features that will be extracted from the released gaze positions.

4.2.2 Alternative Privacy Guarantees

Many formal privacy guarantees exist to protect against different types of privacy risks. We pursued k-anonymity and k, γ-plausible deniability (PD) as alternatives to DP, as they directly protect against re-identification attacks. First, we explored k-anonymity to provide intuitive protection in that individual data cannot be distinguished from k-1 others. The k-same [183, 98, 205] approach is common to achieve k-anonymity for numerical data and works by averaging data together in groups of size k and releasing duplicate values. The duplicate values have equal contribution to the released data, establishing an upper bound of $\frac{1}{k}$ on the probability of individuals being re-identified. k-same is typically used to protect identity within facial images, as the numeric pixel values can easily be averaged across individuals. From an eye-tracking perspective, releasing duplicate data is not a satisfying solution.

Limiting output data to k copies of the same data led us to k, γ-PD, which extends a similar intuition applied to synthetically generated data [33, 34]. PD retains the intuition of the k parameter in terms of privacy for linking synthetic data to the real dataset. The γ parameter is used to threshold the probability that $k-1$ real data inputs could have generated the synthetic output before it can be released, allowing control over the level of privacy for data synthesized by

a generative model. PD has been applied in the domain of spatial-temporal data in the form of location traces [33], motivating an application to spatial-temporal gaze data. Synthetic location traces retained utility for location-based services while protecting real individuals from being re-identified and leaking the specific location of their home, doctor's office, or work locations.

Adaptions of existing mechanisms for k-anonymity and k, γ-PD (Table 4-1) process feature data directly and allow for the protection of datasets that only release eye-tracking features. The mechanism's guarantee holds for the released feature data, as the only source of identification are the released feature values. In contrast, datasets of eye-tracking samples are difficult to protect against re-identification with a formal guarantee. The feature set an attacker may use for identification might not be known at the time of dataset release, preventing the privacy mechanism from providing a robust guarantee against future attacks. As described above, even the DP methods do not offer a direct theoretical guarantee against re-identification. They would require empirical analysis to determine if data is safe for release against a given feature set and model. To address this limitation, we consider generative models that can synthesize gaze positions during the most common eye-movement events, fixation and saccades. Synthesizing new gaze samples within each detected event preserves realistic eye movement behavior, compared to kalεido in which DP privacy noise is added in a manner that would no longer retain event boundaries within the released data. Modeling eye movements from generative models allows the assignment of probabilities on whether actual data from individuals could have generated a set of synthetic gaze positions. Mechanisms can achieve privacy guarantees for synthetic sample data and the models that generate them with respect to re-identification.

4.2.3 Synthesizing Gaze Data

Synthesizing eye-tracking data has been explored in the eye-tracking community to drive saliency-based applications [198, 160, 256], and for training deep network models [145]. For these applications, generative models of gaze data are trained with the intention of deploying the model on new unseen data. For example, a deep model that predicts a fixation scanpath can take as input an unseen image and predict the most relevant regions to optimize while streaming the

content. Deep data synthesis models typically take the stimulus as input and predict the eye movement behavior of a viewer, which is considered synthetic data. In contrast, our proposed approach considers eye movements at the event level, synthesizing data from a saccade of a particular amplitude and direction or a fixation with a specific horizontal and vertical spread.

Historical work on modeling eye movements have developed both simple [243, 16, 95] and complex models [139, 140, 128] for events that vary in the number of model parameters and whether they are based on heurisitics [22] or physical simulation [128]. Models based on statistical distributions are amenable to modeling the probabilities that relate distributions from each individual with samples of real gaze data. Modeling the probability that an individual produced a set of samples or a particular event is key to preventing re-identification from features that generalize common behavioral trends as features.

We considered applications that process individual events or raw samples in our work. Such applications benefit from modeling eye-tracking data at a low level compared to the high-level prediction and synthesis models discussed above. For example, in training a real-time gaze model, the output could predict where the user will look in 100 milliseconds. This task requires high utility of data at the sample level, which is achieved by modeling each event detected in a sequence of eye-tracking data. Predicting a fixation scanpath to synthesize sample data may miss minor details in the time frame of 100 milliseconds and lose utility in the characteristics of fixation and saccade events.

In the context of privacy, researchers have also turned to machine-learning methods to learn how to manipulate eye-tracking data and balance privacy and utility. Fuhl et al. [84] deployed a reinforcement learning model that integrates privacy terms into the optimization. The model learns to protect specific attributes like identity by adding error terms to the cost optimization function. Implementing these terms depends on a trained model to be evaluated against, optimizing the representation of the released data to achieve high privacy for a known identification model while retaining utility for the task at hand. The limitation of such an approach is that the dataset is released assuming that an attacker will use a similar attack model.

However, with rapid advancements in deep learning biometrics, this assumption may not hold for the lifespan of the dataset [155]. Thus, formal privacy guarantees are the preferred approach when considering large-scale datasets and re-identification attacks [191].

4.3 Methodology

Applying formal privacy guarantees to eye-tracking datasets depends on the format of data being processed and determines whether feature or sample mechanisms are used. Figure 4-1 illustrates the flow of raw data collected from an eye tracker as images of the eye to 3D gaze positions represented as a time series. It is assumed that the time series of eye-tracking data is segmented into separate files by identity and stimulus. Gaze data are then processed to extract events that are available to feature and sample mechanisms. For feature datasets, computed features are specified before data processing and define the input to the feature mechanisms.

We considered three formal definitions of privacy that are applied to eye-tracking data: k-anonymity, k, γ-plausible deniability, and ε-differential privacy. Algorithms that achieve these definitions for feature and sample data are defined in Section 4.4 and Section 4.5, respectively. Section 4.3.2 motivates the use of privacy mechanisms on eye-tracking data to protect against re-identification with a threat scenario. Section 4.3.3 provides the assumptions of the threat model considered in this chapter.

4.3.1 Privacy Definitions

This section defines three privacy definitions that can be applied to re-identification attacks on eye-tracking data. First, we discuss k-anonymity as the seminal definition of anonymity for a released dataset. Second, we present the definition of plausible deniability, which leverages the intuition of k-anonymity for synthetically generated data. Last, we provide the definition and practical implications of ε-differential privacy.

4.3.1.1 k-anonymity

k-anonymity is a seminal definition of privacy within a dataset proposed by Samarati et al. [220]. In 2002, k-anonymity was used by Sweeney to protect against re-identification attacks for public medical datasets [237]. Sweeney wanted to show that the public medical dataset of

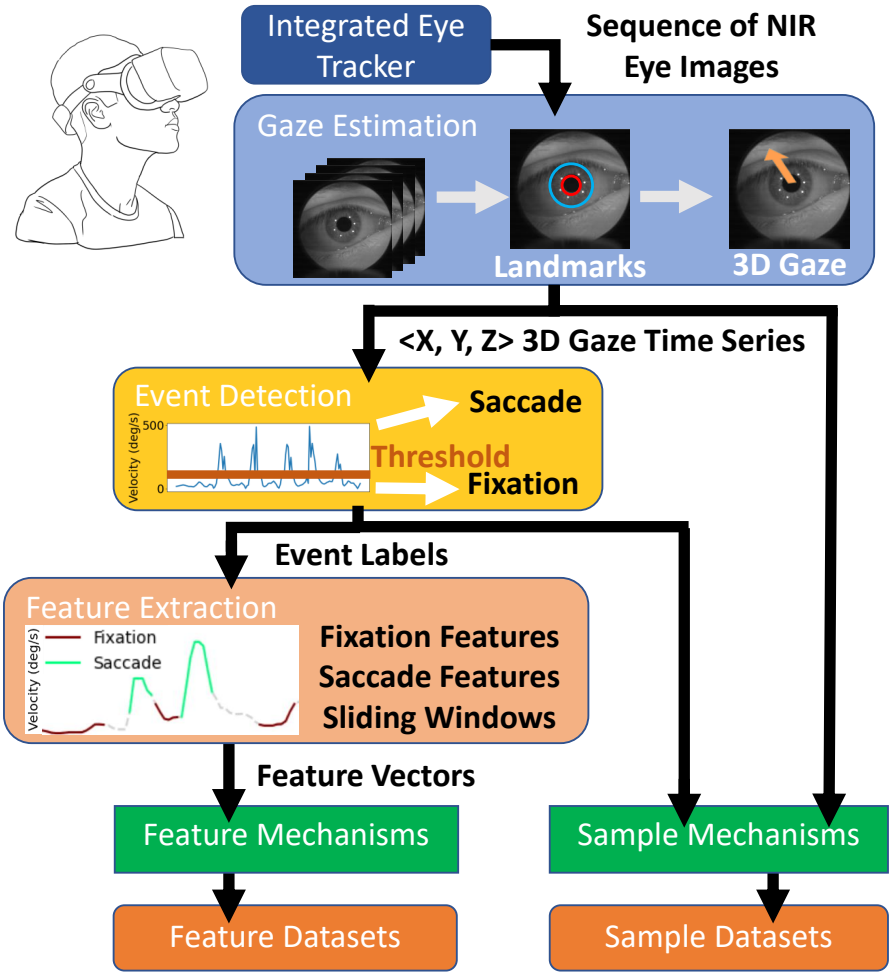

Figure 4-1. Data flow for sample and feature-based privacy mechanisms. Feature sets extract information from gaze samples over windows of time. Features can be extracted over time windows of fixed length, or extracted from fixation and saccade events of variable duration.

Massachusetts state employees did not protect privacy by simply removing their names, in contrast to claims made by the then Governor of Massachusetts. Sweeney demonstrated that the public medical dataset paired with voter registration data could identify the prescriptions of the Governor of Massachusetts himself, as he was in the dataset of state employees. The re-identification attack was made possible as the medical dataset contained the gender, date of birth, and zip code of patients along with a list of prescriptions. By purchasing the state voter records for $20, Sweeney could match the Governor's zip code and birth date to a single record within the medical dataset. Sweeney showed that re-identification would be prevented if more

than one record matched the given demographics. The definition of k-anonymity guarantees that at least k released data records would match any unique combination of such feature values. Formally,

Definition 1. *k-anonymity*

Given a person-specific dataset D, a de-identified dataset D' is k-anonymized by privacy process $\mathscr{P} : D \mapsto D'$ *if all released features* $\Gamma_d = \mathscr{P}(\Gamma) \in D'$ *cannot be recognized as* Γ *with probability higher than* $\frac{1}{k}$.

We say a dataset has k-anonymity if the above condition is true for all unique combinations of feature values, including zip code and birth date in the given scenario. In other words, this means that at least $k-1$ other data records would have matched the demographics of the Governor, and the attacker would not be sure which one corresponds to the Governor. Using k-anonymity to protect the privacy of individuals within a dataset is useful as it provides an upper bound of $\frac{1}{k}$ on the probability of re-identification. However, the implications on privacy depend on what value of k is used to release the dataset. Selecting an appropriate value of k is not always intuitive and varies depending on what knowledge we assume that the adversary has access to. In the case of the Governor of Massachusetts, the risk of re-identification resulted from a specific external dataset of voter records. Knowing an adversary's information or features is critical in determining a value of k that prevents re-identification. In practice, a value of k between five and fifteen is common in releasing medical datasets as a rule of thumb [74]. There is no standard approach for determining k across fields, as the optimal value of k depends on the type of data, and what probability of re-identification would make an individual safe from attacks. In the context of eye-tracking data and our threat model (Sec. 4.3.3), we consider a k-anonymous dataset privacy-preserving if the upper-bound on probability of re-identification $\frac{1}{k}$ is near chance rates.

4.3.1.2 Plausible Deniability

Plausible deniability (PD) was first defined by Bindschaedler and Shokri in the context of location traces [33] and later extended to general data formats [34]. PD prevents re-identification by utilizing the generation of synthetic data to achieve privacy. A synthetic dataset is released that

captures the original characteristics without leaking the identity of those that contributed to the original dataset. PD provides a guarantee that there are at least k individual records that could have plausibly generated a synthetic data output.

PD has two privacy parameters: k, an integer greater than or equal to one, and γ, a real number greater than or equal to one. Let $\mathbf{M}$ be a probabilistic generative model that takes as input a data record d and generates synthetic records y with probability equal to $Pr\{y = \mathbf{M}(d)\}$.

Definition 2. *Plausible Deniability*

For any dataset D where $|D| \geq k$, and any record y generated by a probabilistic generative model $\mathbf{M}$ *such that $y = \mathbf{M}(d_1)$ for $d_i \in D$, we state that y is releasable with (k,γ)-plausible deniability if there exist at least $k-1$ unique records $d_2, ..., d_k \in D \setminus \{d_1\}$, such that*

$$\gamma^{-1} \leq \frac{Pr\{y = \mathbf{M}(d_i)\}}{Pr\{y = \mathbf{M}(d_j)\}} \leq \gamma$$

where $k \geq 1$ is an integer and $\gamma \geq 1$ is a real number.

The level of privacy is controlled by parameters k and γ. Large values of k and values of γ that are closer to one imply higher privacy. In practice, PD ensures that at least $k-1$ plausible seeds, i.e., inputs, to the model $\mathbf{M}$ could have plausibly produced the synthetic output record y. The parameter γ bounds how close together the probabilities are to determine that they are plausible. Privacy-preserving datasets are generated by only releasing synthetic records y if they pass the PD privacy test.

PD privacy test:

1. Let i ≥ 0 be the only integer that fits the inequality $\gamma^{-i-1} < Pr\{y = \mathbf{M}(d)\} \leq \gamma^{-i}$
2. Let k' be the count of records $d_a \in D$ such that $\gamma^{-i-1} < Pr\{y = \mathbf{M}(d_a)\} \leq \gamma^{-i}$
3. If $k' \geq k$: return PASS, else return FAIL

Step 1 is formulated as there is only one integer that satisfies the inequality when $\gamma \geq 1$, as the range of values covered by the set $(\gamma^{-i-1}, \gamma^{-i}]$ represent disjoint sections of the real number

line for different integer values of i. Therefore, $Pr\{y = \mathbf{M}(d)\}$ can only fall within one such range. Step 2 is a sufficient condition to achieve (k,γ)-PD when both $Pr\{y = \mathbf{M}(d)\}$ and $Pr\{y = \mathbf{M}(d_a)\}$ fall within the range of $(\gamma^{-i-1},\gamma^{-i}]$. Please see the Appendix for a proof of the sufficient condition.

Implementing the Privacy Test requires a method to compute probability values of the form $Pr\{y = \mathbf{M}(d)\}$ that represent the probability that random mechanism $\mathbf{M}$ would generate the synthetic output y for a given input. Bindschaedler et al. model the conditional probabilities of tabular data with a Multinomial Dirichlet distribution (MDD) [242]. An MDD assumes that feature values are a set of discrete values and cannot be applied to continuous values. Continuous features can be made discrete by binning values into buckets over a range of values.

The process of estimating conditional probabilities using the MDD depends on the number of data records in the dataset that have a particular value for each feature, i.e., the histogram counts of each value. Conditional probabilities from the MDD model are used to compute the probability that synthesized feature values match the set of feature values from actual data. The conditional distributions are then used to calculate $Pr\{y = \mathbf{M}(d)\}$, and execute the Privacy Test.

Plausible deniability provides a definition of privacy that is intuitive against re-identification in terms of k, similar to k-anonymity. Using synthetic data achieves privacy while retaining data utility if the generated data captures the characteristics of the original dataset. In the context of eye-tracking data and our threat model (Sec. 4.3.3), we compare achieved privacy of PD against re-identification attacks on eye-tracking features for different values of k and γ.

4.3.1.3 Differential Privacy

Differential privacy (DP) is a theoretical definition of privacy that has quickly become a standard in the privacy community [71]. First proposed by Dwork in 2006 [70], DP is popular as it provides a theoretical bound on the output data distribution. The privacy guarantee applies even in the worst-case scenario where all other entries from the original dataset have been leaked. The privacy parameters for DP are defined to quantify how much information an adversary gains when they access data released by the privacy mechanism. Compared to k-anonymity, selecting the

privacy parameters for DP does not require an assumption on what type of external information an adversary has. An assumption on adversary knowledge is unnecessary as the DP guarantee applies to any two datasets that differ by at most one element.

Formally, ε-differential privacy is defined as,

Definition 3. *ε-Differential Privacy*

A mechanism $\mathbf{M}$ *provides* ε*-differential privacy if for all databases* D, D' *that differ in at most one element and for every* $O \subseteq Range(\mathbf{M})$*, we have*

$$Pr[\mathbf{M}(D) \in O] \leq e^{\varepsilon} \cdot Pr[\mathbf{M}(D') \in O]$$

It is important to note that the definition of ε-Differential Privacy (ε-DP) applies to the mechanism $\mathbf{M}$, and not the database D or D'. A guarantee on the mechanism ensures that the formal guarantee generalizes to all possible datasets. The e^{ε} term in the privacy guarantee bounds the probability that an adversary can detect a difference if a given data element was or was not contained within the original dataset.

A major benefit of DP mechanisms is the ability to generalize the information that an adversary has. Unlike other privacy definitions, it does not depend on assumptions of external information that would influence the initial suspicion that a record is in the dataset. As a result, the privacy parameter ε is easier to interpret across specific datasets and applications when applied to the same data type. In the context of eye-tracking data and our threat model (Sec. 4.3.3), we consider a differentially private dataset privacy-preserving if the probability of successful re-identification attacks is near chance rates.

4.3.2 Threat Scenario

Publicly released datasets risk leaking the identity of individuals that contributed to the dataset. Using the classic example described in Section 4.3.1.1, the Governor of Massachusetts's medical prescriptions were leaked as a result of releasing gender, date of birth, and zip code [237]. For a high-profile politician, the risk of re-identification could reveal an undisclosed medical condition that fuels negative propaganda. Extending this example, consider a

hypothetical dataset from a medical study that releases XR data to classify a medical condition. The released data is de-identified by removing participant names and date of birth, but age and gender were included along with eye-tracking data. Given past re-identification attacks, the sponsored study is required to implement k-anonymity with k greater than or equal to four.[1] k-anonymity can easily be achieved for the released age and gender data by generalizing the values of the age data into ranges of age values [220]. However, if k-anonymity is not also achieved for the released eye-tracking data, then the risk of a successful re-identification attack is no longer bounded above by $\frac{1}{k}$. We demonstrate this risk with an example using the ET-DK2 and 360_em datasets (see Section 4.4.3 for more details on these specific datasets).

In this scenario, it is assumed that the adversary can select the identities that match the demographics of their target and then train and apply a model to the subset of identities. For example, suppose only four identities in the dataset have an age between 18 and 20 and identify as Male. In that case, the adversary can train the model and predict which of the four identities is predicted to be the target.

A prototype of a gaze-based re-identification attack is conducted by combining the ET-DK2 and 360_em eye-tracking datasets with age and gender demographics. The combined dataset in total includes 24 identities. A standard method of data generalization is used to achieve k-anonymity on age and gender data by releasing ranges of values instead of exact values (see the Appendix for the generalized k-anonymous groupings). Figure 4-2 demonstrates the success rate of re-identification attacks with and without the k-same-select sequence mechanism applied to the eye-tracking data (Section 4.4.2.1). The biometric identification evaluation detailed in Section 4.4.5.1 was applied to perform the re-identification attack. Attack success remained above 80% for all values of k if only the age and gender demographics were k-anonymous. In contrast, attack success remained less than the theoretical $\frac{1}{k}$ bound when the eye-tracking data was also made to be k-anonymous prior to release.

[1] El Emam et al. [74] discussed the values of k typically seen in medical datasets, with a value of k equals three considered a minimum and a value of five being most common. Values of k as large as fifteen are rarely seen in the context of medical datasets.

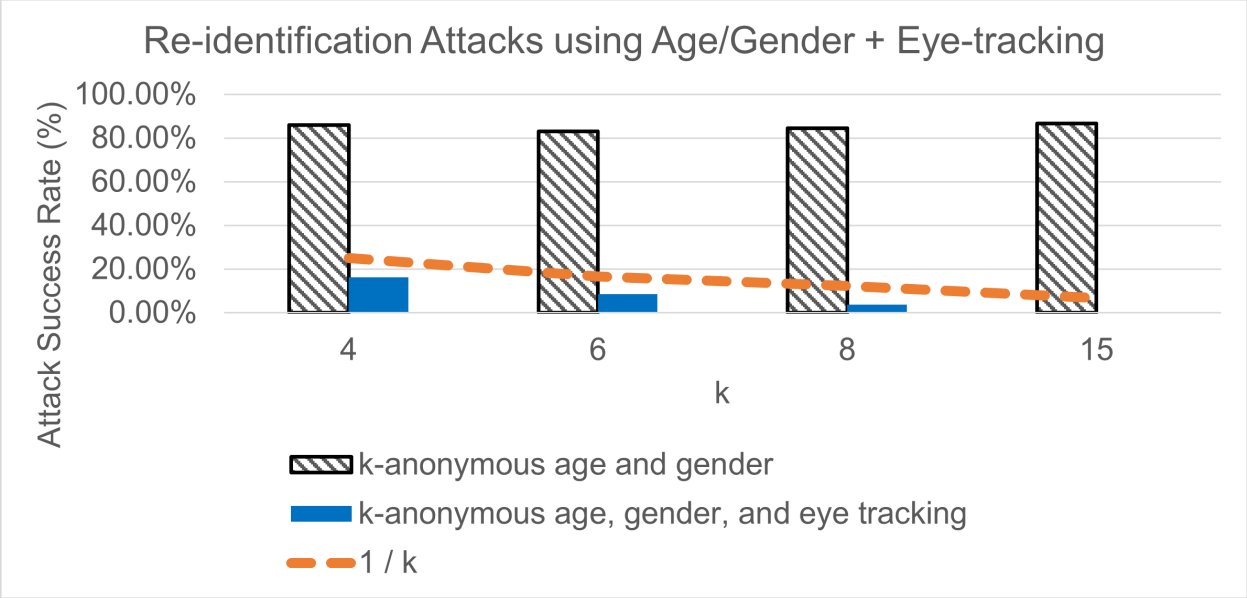

Figure 4-2. Success rate of re-identification attacks on ET-DK2 and 360_em using age, gender, and eye-tracking data with k-anonymity. Solid bars indicate results where age and gender are k-anonymous, while bars with lines indicate results when eye-tracking data is also made k-anonymous. The orange dashed line plots $\frac{1}{k}$, the theoretical upper bound on re-identification. For all values of k, re-identifications attacks remain below the theoretical bound only if the released eye-tracking data are also made k-anonymous.

In this example attack, a public dataset does not meet the k-anonymity privacy guarantee required by the research sponsor, impacting the researchers and institution that released the dataset. Furthermore, the scenario puts participants' privacy at risk, with successful re-identification attacks allowing the adversary to identify medical conditions or other sensitive information about victims. Applying k-anonymity to eye-tracking data would enable the researchers to satisfy their data privacy requirements while still contributing a public dataset to the research community.

4.3.3 Threat Model

Assumptions for the considered re-identification attack include an adversary who has a target identity that they want to identify within the dataset. The adversary has read access to the public dataset. The adversary knows the demographics of their target user and has access to eye-tracking data from them performing the same task as the dataset. The adversary can then

build a model trained on the public dataset that predicts which identity most closely matches the input data. If the prediction is correct, the target is successfully re-identified.

In this chapter, we considered a threat model where a privacy mechanism has processed the public dataset while the testing data used to re-identify individuals is unmodified. It is reasonable to assume that an adversary could gain access to raw tracking data through unauthorized code or by logging data streamed to third-party applications [241]. The explored privacy mechanisms also rely on processing the entire dataset at once and would not apply to eye-tracking data collected in the wild.

4.4 Study 1: Privacy Guarantees for Feature Datasets

Datasets of eye-tracking features without corresponding gaze samples are helpful for analyzing human behavior during experimental tasks or training machine learning models when developing naturalistic interfaces. The risk behind releasing or storing eye-tracking features that may be leaked to adversaries is the ability to use the features directly to identify individuals. Figure 4-3 outlines the process for taking a dataset of eye-tracking features as input and releasing them in the same format. Features from the input dataset are assumed to be segmented in separate files across C individuals viewing a set of M stimuli. Three distinct privacy mechanisms are explored: k-same-select sequence, Marginals Generative Model (PD), and Exponential-DP.

4.4.1 Research Questions

Formal privacy guarantees for eye-tracking datasets have been limited to the definition of DP. We explored how to achieve additional privacy guarantees to protect against re-identification attacks and evaluated the privacy-utility trade-off when using feature data to train a document type classifier.

Formally we propose the following research questions:

- $RQ_{4.1}$: *Can we protect eye-tracking feature data against re-identification through k-anonymity and plausible deniability?*
- $RQ_{4.2}$: *Which privacy mechanism achieves the highest level of utility for document type recognition while reducing identification rate to chance?*

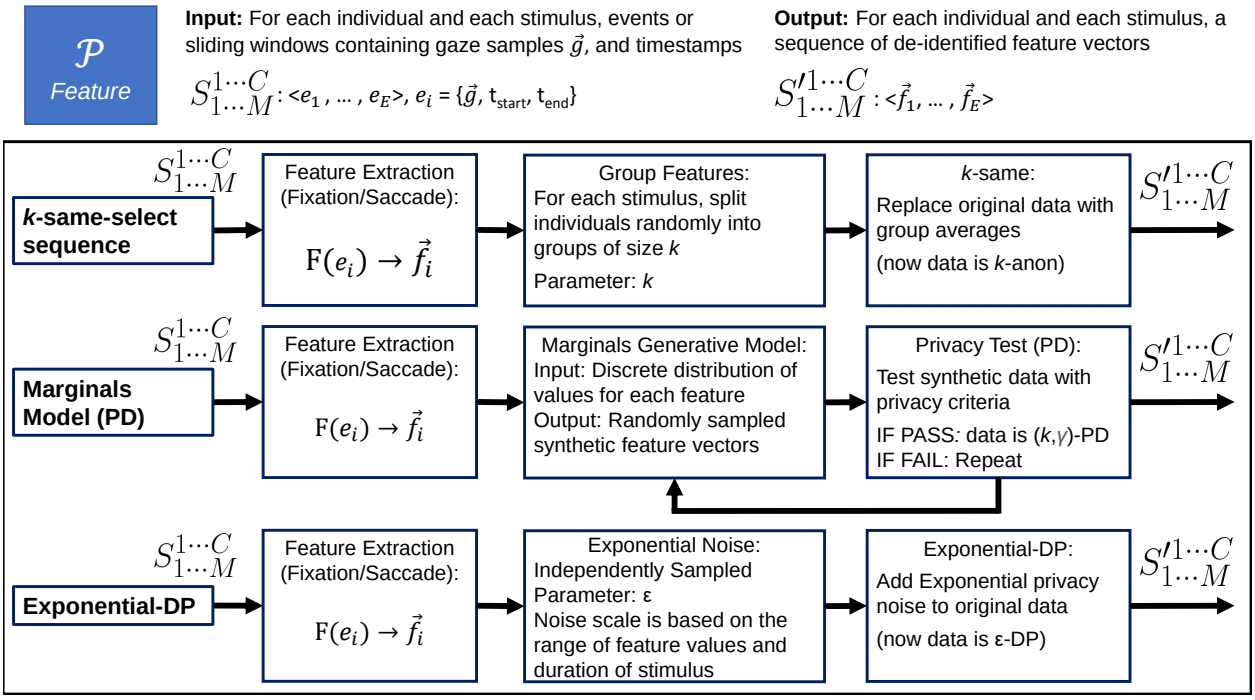

Figure 4-3. Privacy mechanisms for releasing eye-tracking feature data.

4.4.2 Implementation

We contribute two privacy mechanisms, one that satisfies *k*-anonymity and one that satisfies plausible deniability. We provide pseudocode for ease of re-implementation and publicly release code for *k*-same-select sequence.[2] Both mechanisms are adaptations of prior work to consider sequences of eye-tracking features. For completeness, we also provide pseudocode and code for our implementation of the DP-oriented mechanism defined by Steil et al. [232].

4.4.2.1 *k*-same-select sequence

The *k*-same family of mechanisms [183, 98] accomplish *k*-anonymity by first splitting individual data into groups of size *k*. Each group is averaged to produce a value which is then released *k* times in the released dataset (Figure 4-4). Releasing duplicates enforces the upper bound on re-identification probabilities, as *k* of the identities from the original dataset will contribute equally to the privatized data.

[2] www.doi.org/10.5281/zenodo.6463849

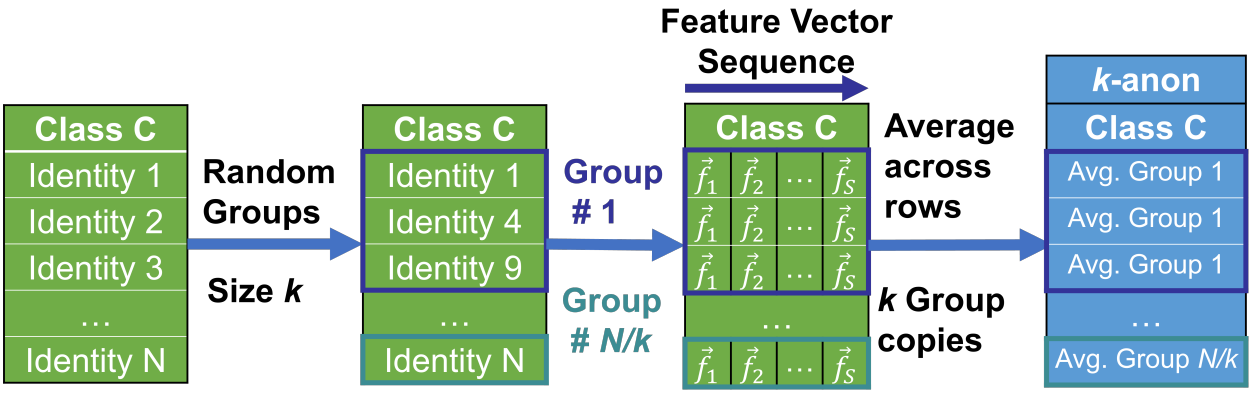

Figure 4-4. The k-same-select sequence mechanism processes sequences of feature vectors from each individual within the target utility class (stimulus). First, the identities are grouped into $\lceil \frac{N}{k} \rceil$ groups of size k. The sequences of feature vectors are then aligned temporally and averaged within the assigned groups. The average feature sequence of each group is published for each individual assigned to that group. Releasing k copies in this manner establishes k-anonymity for the released sequence of feature vectors, bounding the probability that any individual in the group could be identified from the $k-1$ others.

The implementation of k-same depends on the format of data being released. For example, k-same can be applied directly to face images by clustering and releasing averages [183]. For eye-tracking data, the computed feature vectors are grouped and averaged to satisfy k-anonymity. We adapted the k-same-select mechanism by separately processing the sequence of feature vectors generated for each task in the dataset. The data from all individuals are processed sequentially, i.e., the first feature vector of all individuals viewing a specific stimulus within a given task are randomly placed into groups of size k to compute average values for release. The mechanism assumes that data from at least k individuals is available for grouping. The same groupings of individuals are used within each stimulus to achieve k-anonymity across the entire sequence of feature vectors. The adapted sequence mechanism is generalized by processing feature vectors in sequence; however, there is no guarantee that each individual has the same number of feature vectors per stimulus. Data are padded to repeat the last feature vector in the sequence for individuals with fewer features.

1: **procedure** k-SAME-SELECT SEQUENCE(k, feature_data)

2: **Parameters:** k - k-anonymity parameter

feature_data - sequences of feature vectors, indexed by stimulus m and identity i

for $m = 1$ to num_task **do** ▷ Process features from each task independently

  $curr\_data \leftarrow feature\_data[m, :]$

  $G \leftarrow$ Randomize N individuals into H groups of size k

  for $i = 1$ to $num\_feature\_vectors$ **do** ▷ Loop over sequence of feature vectors within task m

    $curr\_features \leftarrow curr\_data[i, :]$ ▷ i-th feature vector from all individuals

    $avg\_features \leftarrow avg\_groups(curr\_features, G)$ ▷ Average feature vectors for each group

    $curr\_data[i, :] \leftarrow avg\_features$

  $feature\_data[m] \leftarrow curr\_data$ ▷ Update feature vector data for task m

return $feature\_data$

4.4.2.2 Marginals Generative Model (PD)

As defined in Section 4.3.1.2, PD is not a condition of a privacy mechanism but a privacy criterion that is checked before a data record can be released [34]. Thus, various approaches can be applied to generate data satisfying PD.

To achieve PD for eye-tracking features, we applied the Marginals generative model approach proposed in the original paper with the publicly available code [34].[3] Marginals builds a distribution of discrete values for each feature column in the dataset and releases synthetic data by randomly sampling each feature independently. The Marginals approach was applied to model feature distributions representative of each stimulus or task. The resulting distributions are used to synthesize data by document type and retain utility.

[3] https://vbinds.ch/node/69

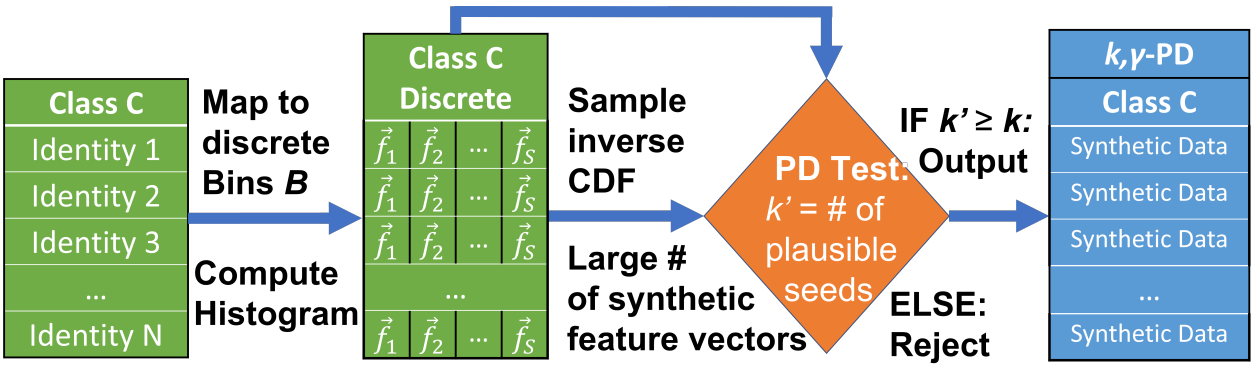

Figure 4-5. The Marginals generative model with PD criterion uses feature vectors from each target utility class or stimulus to build a distribution of values for each feature. The feature values are first binned to build a discrete probability distribution, or histogram with B bins for each feature. Then, a large dataset of synthetic feature vectors is generated by randomly sampling each feature distribution independently. The generated synthetic feature vectors are then subject to the PD Privacy test, and those that pass are retained as candidates for release. The k, γ-PD synthetic feature vectors are then randomly assigned to the different individual identities contained in the dataset. The number of synthetics per individual and task is stratified to match the number in the real dataset.

We adapted this approach by binning each continuous feature into $B = 30$ uniformly sized buckets over the range of feature values. Each bucket corresponds to a discrete value that the feature could take. The generated synthetic feature vectors consist of values corresponding to the buckets used to discretize each feature. The counts of data points that fall into each bucket define a probability mass function for each feature. To map synthetic data back into a set of continuous feature values, we sample values between the min and max range from the corresponding bucket using a random uniform distribution. The synthetic dataset is stratified to contain the same number of feature vectors for each individual and stimulus as the original dataset. The guarantee of the resulting data differs from that of k-anonymity in that PD guarantees $k-1$ other features from the original dataset could have generated the synthetic output, while k-anonymity guarantees that $k-1$ other individuals could have generated a sequence of output features.

1: **procedure** TASK-BASED MARGINALS MODEL(k, γ, B, num_samples, feature_data)

2: **Parameters:** k, γ - plausible deniability parameters, B - # of bins for discrete feature data

3: num_synthetics - sequences of feature vectors, indexed by stimulus m and identity i

feature_data - sequences of feature vectors, indexed by stimulus m and identity i

$bin_feature_data \leftarrow$BinData($feature\_data, B$) ▷ Map each feature into B uniformly sized bins

for $m = 1$ to num_task **do** ▷ Process features from each task independently

  M ←MarginalsDist(bin_feature_data[m]) ▷ Learn distribution of discrete feature values

  $synth_data \leftarrow$Generate(**M**,num_synthetics) ▷ Randomly generate synthetic dataset

  $private\_data \leftarrow$ PD Privacy Test(synth_data) ▷ Retain synthetics that pass PD privacy test

  $bin\_feature\_data[m] \leftarrow private\_data$ ▷ Update feature vectors with synthetic data

$feature_data \leftarrow$BinToContinuous($bin\_feature\_data$) ▷ Map back to continuous features

return $feature\_data$

4.4.2.3 Exponential-DP Mechanism

The Exponential-DP noise mechanism was proven to be ε-DP by Steil et al. [232] and applies to each individual feature in the feature set.[4] Exponential noise is sampled independently for each feature vector extracted during stimulus viewing and depends on the range of each feature and the stimulus duration. The first step in applying Exponential-DP is to compute the range δ_i for each feature i as the maximum value minus the minimum value. The maximum number of feature vectors t_{max} from any individual viewing the stimulus is used for padding the data from other individuals. The mechanism repeats the last feature vector recorded for an individual to ensure that each individual has t_{max} total feature vectors. For each feature a value y is sampled from an Exponential distribution with a scale of $\frac{1}{\lambda}$, where $\lambda = \frac{\varepsilon}{2 \cdot t_{max} \cdot \delta_i}$. The additive noise is then computed as $r = \pm \frac{log_e(y)}{\lambda \cdot t_{max}}$ and the positive or negative sign is randomly assigned. The additive noise values of r are computed for every feature from the stimulus and are added to the original data to produce noisy feature vectors for release.

procedure EXPONENTIAL-DP(ε, feature_data: structure indexing data by identity and task)

  $\delta \leftarrow Range(feature\_data)$ ▷ Maximum value for each feature minus minimum value

  for $m = 1$ to $num\_stimuli$ **do** ▷ Process feature vectors from each individual and task independently

    Compute t_{max} for task m and pad individual data

[4] Note that due to The Composition Theorem, the Exponential mechanism achieves an overall guarantee of ε times the number of feature columns in the dataset. For consistency with [232], we reference ε as the privacy parameter for each feature and not the composed guarantee.

$\lambda \leftarrow \frac{\varepsilon}{2 \cdot t_{max} \cdot \delta}$ ▷ λ computed using δ, and t_{max} from task

$Exp \leftarrow Exponential(scales = \frac{1}{\lambda})$ ▷ Define Exponential distribution for each feature based on λ

for $i = 1$ to $num\_feature\_vectors$ **do** ▷ Loop over sequence of feature vectors within task m

$y \leftarrow Sample(Exp)$ ▷ Sample synthetic value from Exponential distribution for each feature

$r \leftarrow \frac{log_e(y)}{\lambda \cdot t_{max}}$ ▷ Compute additive noise value

$feature\_data[m,i] \leftarrow feature\_data[m,i] \pm r$ ▷ Randomly flip noise sign

return $feature\_data$

4.4.3 Datasets

We evaluated the above-detailed privacy mechanisms on publicly available VR datasets of eye-tracking features. The datasets varied based on the number of individuals, amount of data available, task, and type of stimulus. Table 4-2 summarizes the characteristics of dataset included in our evaluation.

4.4.3.1 MPIIDPEye

The MPIIDPEye dataset contains eye-tracking features extracted from sliding windows of fixation, saccade, blink, and pupil statistics for 20 individuals [232]. The purpose of the dataset was to benchmark privacy-preserving mechanisms for eye-tracking data. The dataset's utility is recognizing what type of document a user reads in VR between Comic, Newspaper, or Textbook.

4.4.3.2 ET-DK2

The ET-DK2 dataset contains eye-tracking data from 18 individuals freely viewing 50 different 360° images in VR. The purpose of the dataset was to generate saliency maps [124].

4.4.3.3 VR-Saliency

The VR-Saliency dataset contains eye-tracking data from 130 individuals freely viewing eight different 360° images in VR [228]. The purpose of the dataset was to explore visual saliency and human behavior in VR viewing.

4.4.3.4 360_em

The 360_em dataset contains eye-tracking data from 13 individuals freely viewing eight different 360° videos in VR [10]. The purpose of the dataset was to explore classification methods for eye movement events.

4.4.3.5 VR-EyeTracking

The VR-EyeTracking dataset contains eye-tracking data from 43 individuals freely viewing 360° videos in VR [250]. The purpose of the dataset was to compute saliency maps and provide input to a deep neural network-based gaze prediction model.

4.4.3.6 OpenEDS

The OpenEDS dataset contains eye-tracking data from 44 individuals interacting with objects and exploring a 3D rendered VR environment [76]. The dataset was part of the Facebook OpenEDS 2021 gaze prediction challenge.

4.4.3.7 EHTask

The EHTask dataset contains eye-tracking data from 30 individuals viewing 360° videos in VR while performing four different tasks (Free Viewing, Search, Saliency, Track) [112]. The purpose of the dataset was to train a deep neural network-based model that classifies tasks based on eye and head movements.

Table 4-2. Characteristics of VR eye-tracking datasets.

Dataset	# Ppts.	Chance rate	# Stim.	Data per ppt.	Stimuli type	Task
MPIIDPEye [232]	20	1/20 = 5.0%	3	30 mins	Documents	VR reading
ET-DK2 [62]	18	1/18 = 5.5%	50	21 mins	360° images	Free viewing
VR-Saliency [228]	130	1/130 = 0.8%	8	4 mins	360° images	Free viewing
360_em [10]	13	1/13 = 7.7%	14	17 mins	360° videos	Free viewing
VR-EyeTracking [250]	43	1/43 = 2.3%	208	Avg: 88 mins	360° videos	Free viewing
OpenEDS [76]	44	1/4 = 2.3%	2	10 mins	3D scene	Scene exploration
EHTask [112]	30	1/30 = 3.3%	15	30 mins	360° videos	Free viewing, search, saliency, track

4.4.4 Feature Sets

Five of the datasets listed in Table 4-2 release raw gaze sample data, while the MPIIDPEye dataset included both raw samples and a set of pre-computed sliding windows of gaze-based features [44]. To maintain consistency with past results from MPIIDPEye, we used their feature set in our analysis of this dataset. For all other datasets, we extracted features from fixation and saccade events detected using the I-S^{5}T algorithm with default parameters [62]. The features extracted from the stream of fixation and saccades events leverage common statistics such as duration and amplitude, as well as the velocity and acceleration of gaze positions during the event [93]. A feature set is generated for each type of event, and a separate classification model is trained for each feature set.

4.4.5 Metrics

4.4.5.1 Privacy

Biometric identification was performed by classifying streams of gaze data from multiple stimuli into one user identity, as previously described in Section 3.5.1. The only difference from the previously described re-identification approach is the use of a training set processed by the privacy mechanism and testing data left unmodified. The identification rate metric was computed as the total number of correct matches divided by the total number of classifications. Reported identification rates are averaged over ten runs with random stimuli selected as part of the training and test set to account for variance in stimuli content.

As described in Section 4.4.4, most datasets included in our evaluation use features extracted from both fixation and saccade events, requiring an RBF network trained independently on both features [93]. The output identification scores are averaged within each type of event, and then a final classification is made with a weighted average between fixation and saccade scores. A weight of 0.4 was applied to the fixation scores, with a weight of 0.6 for saccade scores, as saccade features provided a slightly higher accuracy. In the case of MPIIDPEye, the prediction scores from all inputs within a stimulus are averaged before classifying identity.

4.4.5.2 Utility

Releasing a useful privacy-preserving dataset relies on achieving a practical level of utility. The MPIIDPEye dataset classified sliding windows of eye-tracking features to determine what type of document was being read by the user [232]. We evaluated the utility of each privacy mechanism applied to the MPIIDPEye dataset by computing the accuracy of a classification model trained with the parameters presented in the original paper. The utility evaluation differed from identification in that we classify each feature vector independently and compute performance instead of classifying a sequence of features from each stimulus.

Steil et al. [232] first evaluated MPIIDPEye using an SVM model to classify document type as either Comic, Newspaper, or Textbook. The SVM used an RBF kernel, bias parameter set to one, and expressivity parameter set to one divided by the number of features. The model was trained on data from each individual during the first half of the reading processed by the privacy mechanism, and tested on data from the second half. Classifier performance is based on the true positive, false positive, true negative, and false negative predictions for each feature vector classified in the test dataset using accuracy, computed as $\frac{TP+TN}{TP+FP+TN+FN}$.

4.4.6 Results

Re-identification risk for eye-tracking data is evaluated by splitting eye-tracking features into training sets processed by privacy mechanisms and testing sets of unmodified data. Identification rates higher than chance, one divided by the number of individuals in a dataset, indicate the risk of re-identification from released data. Figure 4-6 presents the identification rates for each dataset and mechanism. The ET-DK2 dataset produced the highest identification rate of all datasets, with 100% identification with the original data. All datasets generally produced identification rates higher than chance before privacy mechanisms were applied.

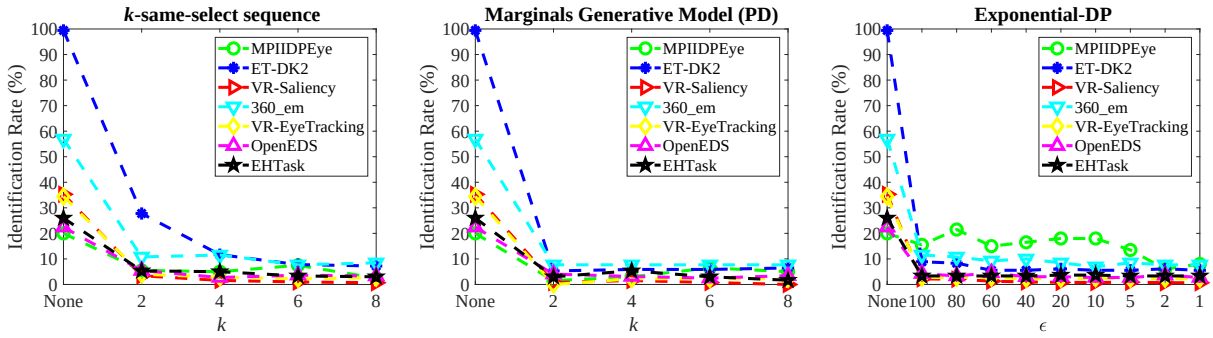

Figure 4-6. Privacy evaluation for identification rate from eye-tracking features. Privatizing the dataset with our presented mechanisms lowers all identification rates to chance for $k = 8$ in k-same and Marginals ($\gamma = 1$), and $\varepsilon = 2$ for Exponential-DP. Chance identification rates demonstrate that identity is protected within a group of individuals. The different datasets contain eye-tracking data on tasks performed within a variety of VR environments (reading documents, 360° images, 360° videos, and 3D rendered scenes). Chance rates (1/#Ppts.) vary for each dataset based on the number of identities and are listed in Table 4-2.

When privacy mechanisms were applied, the identification rates of all datasets dropped to chance rates. The Exponential-DP and Marginals methods degraded the identification rates to chance across all parameter values. The only exception was the MPIIDPEye dataset for Exponential-DP, which required a parameter value of ε equals two for an identification rate of 6%, compared to a chance rate of 5%. k-same also reduces identification rates to chance, with a larger value of k needed to bring ET-DK2 to chance (5.6%). Our results suggested that privacy mechanisms protect against re-identification attacks on eye-tracking features using a standard biometric identification model.

Figure 4-7 presents model accuracy results for each mechanism. Each plot demonstrates utility relative to the original data and chance rate of guessing (33%). The Exponential-DP mechanism reduced accuracy to chance or near chance rates. For Exponential-DP, accuracy started at 80% for $\varepsilon = 100$, and fell to chance at $\varepsilon = 20$. For Marginals, low utility was retained as accuracy remained near 53% for all parameters. The k-same approach was stable across parameter values, with slightly lower accuracy for higher levels of k. k-same across all parameters maintained performance greater than 72%. This level of accuracy would be practical for an assistive reading interface that must identify the correct document type most of the time [240].

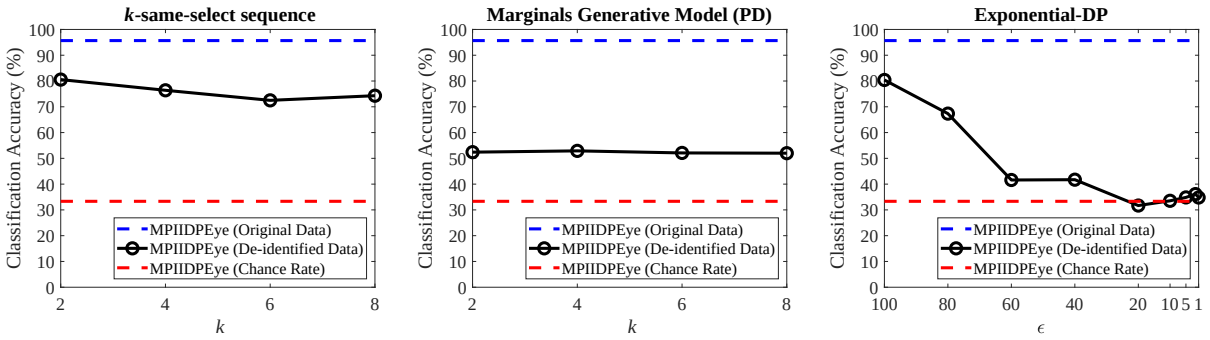

Figure 4-7. Utility evaluation for accuracy of document type classification with an SVM model. Privatizing the dataset with our k-same mechanism retains the utility of the dataset for its intended application. In comparison, the Marginals Generative method does not retain utility above 53%, and the Exponential-DP mechanism rapidly leads to utility loss at the parameter range where MPIIDPEye identification rates fell below chance.

4.5 Study 2: Privacy Guarantees for Sample Datasets

Datasets of eye-tracking samples are useful for analyzing low-level human behavior during experimental tasks [143, 223, 178] or for training deep machine learning models from large-scale datasets [137, 115, 114, 112, 76]. The risk of releasing or leaking eye-tracking samples is that adversaries can apply any feature extraction method for biometric identification.

Figure 4-8 outlines the process for taking a dataset of eye-tracking samples as input and releasing synthetic gaze samples output from a privacy mechanism. The privacy mechanisms have access to event labels (Fixation/Saccade) for every sample, as our proposed methods process sample data separately for each event.

4.5.1 Research Questions

Formal privacy guarantees for eye-tracking sample datasets have been limited to the definition of DP. We explored how to achieve alternative formal privacy guarantees against re-identification for gaze samples. The meaning of formal privacy guarantees on samples depends on where they are implemented in the pipeline and the risk being addressed. Our approach applies privacy guarantees to samples within fixation and saccade events that comprise most data and are typical for re-identification.

We evaluated our k-anonymity and plausible deniability mechanisms against re-identification attacks and compared results with the established kalɛido DP mechanism. Utility

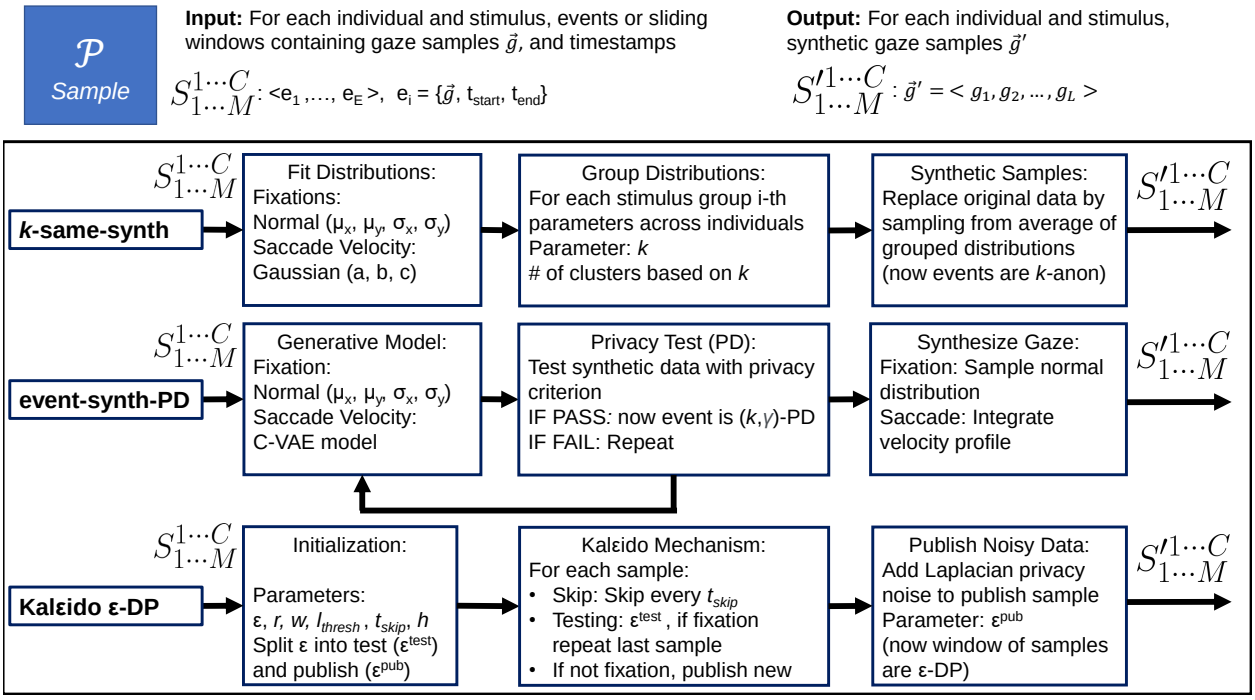

Figure 4-8. Privacy mechanisms for releasing eye-tracking sample data.

for synthetic data was using datasets to train deep learning models for activity recognition and gaze prediction tasks.

Formally we propose the following research questions:

- $RQ_{4.3}$: *Can we protect eye-tracking sample data against re-identification through k-anonymity and plausible deniability?*
- $RQ_{4.4}$: *Which privacy mechanism achieves the highest level of utility for activity type recognition and gaze prediction while reducing identification rate to chance?*

4.5.2 Implementation

This section introduces two new privacy mechanisms for eye-tracking sample datasets that achieve *k*-anonymity and PD through generative models. First, we describe the generative models used to synthesize gaze samples during fixation and saccade events. Second, we describe and provide pseudocode for our proposed methods and the existing kalɛido DP mechanism.

4.5.2.1 Synthesis Models

Privacy for sample-level data is achieved by synthesizing new gaze samples (Fig. 4-9). Data synthesis methods use the real data distributions to capture characteristics of the dataset that

preserve utility. The approach to gaze synthesis is first to identify fixation and saccade events and then replace gaze samples during the events with synthetic data. For fixations, we fit a standard 2D Normal distribution for sampling new gaze positions spatially. For saccades, we fit a three parameter Gaussian well established in literature [243] to the velocity profile and train a conditional variational auto-encoder (C-VAE).

Sampling new gaze samples based on model parameters enabled computing the probability that a particular model generated a given set of synthetic samples from a fixation or saccade event. The sampling approach has the benefit of assigning conditional probabilities to gaze samples for use with the PD Privacy Test.

4.5.2.1.1. Fixations Fixations are low-velocity eye movements best described as clusters of gaze positions that center around a fixation point. We applied a simple model that fits an anisotropic 2D Normal distribution with parameters μ_x, μ_y, σ_x, and σ_y for each fixation cluster and generated synthetic gaze samples by sampling from this distribution. The generated synthetic samples maintain the centroid of the fixation and preserve utility while degrading features that are extracted from the spatial distribution of gaze samples around the centroid.

To determine the probability that a set of t gaze samples were sampled from a given 2D Normal distribution,

$$Pr\{y = \{(x_1, y_1), \cdots, (x_i, y_i), \cdots (x_t, y_t)\} \leftarrow N(\mu_x, \mu_y, \sigma_x, \sigma_y)\}, \tag{4-1}$$

we considered the joint probability that all of the points come from the Normal distribution $N(\mu_x, \mu_y, \sigma_x, \sigma_y)$. The joint probability for independently sampled points is computed as a product of probabilities that each point came from the same distribution

$$\prod_{i=1}^{t} Pr\{(x_i, y_i) = N(\mu_x, \mu_y, \sigma_x, \sigma_y)\}.$$

Gaze positions in this context are considered a continuous random variable defined by $N(\mu_x, \mu_y, \sigma_x, \sigma_y)$. Continuous random variables do not have an analytical probability mass

function to compute $Pr\{(x_i, y_i) = N(\mu_x, \mu_y, \sigma_x, \sigma_y)\}$ directly as a function of (x_i, y_i). Thus, we computed the individual probabilities by considering the cumulative distribution function (CDF) for the Normal distribution. The CDF returns probabilities that the random variable falls within a range of values a and b in the form $Pr\{(a_x, a_y) < N(\mu_x, \mu_y, \sigma_x, \sigma_y) \leq (b_x, b_y)\}$. We approximated $Pr\{(x_i, y_i) = N(\mu_x, \mu_y, \sigma_x, \sigma_y)\}$ as $Pr\{(x_i - \partial, y_i - \partial) < N(\mu_x, \mu_y, \sigma_x, \sigma_y) \leq (x_i + \partial, y_i + \partial)\}$, where $\partial = .01$ represents a sufficiently small region around the gaze position to consider. Estimating the probability from the CDF between $(x_i - \partial, y_i - \partial)$ and $(x_i + \partial, y_i + \partial)$ is interpreted as the probability that a value near (x_i, y_i) comes from the distribution $N(\mu_x, \mu_y, \sigma_x, \sigma_y)$. The probability that a value near that (x_i, y_i) came from $N(\mu_x, \mu_y, \sigma_x, \sigma_y)$ enabled computation of the probability term in EQ. 4-1, thus allowing the PD Privacy Test to be applied at the fixation level. Figure 4-9 (Left) shows a set of synthetic gaze samples (orange) generated from a 2D Normal distribution fit to real data (blue). Synthetic data preserves the mean location of the fixation, while spatially changing the gaze sample positions.

4.5.2.1.2. Saccades A three parameter Gaussian function is fit to the profile of instantaneous saccade velocities computed from the raw gaze samples [95]. The velocities for each gaze sample v_i are computed as $d((x_{i-1}, y_{i-1}), (x_i, y_i))$ where d is the shortest angular distance between two points on a sphere of uniform radius computed with the haversine formula [211]. The Gaussian function used to model the velocity profile of a saccade is defined as

$$G(a, b, c, t) = a * e^{-\frac{(t-b)^2}{c}},$$

where a, b, and c control the shape of the velocity profile and $t \in [0, 1]$ represents time steps for each sample within the saccade. Raw velocity values are re-sampled uniformly so that each saccade consists of a fixed number of values between the start and landing point of the saccade. We fixed the number of saccade points to 30 samples for all saccade profiles. The parameters a, b, and c are determined by solving for the values that minimize the sum of squared errors for each saccade profile, $\sum_{i=1}^{30}(v_i - G(a, b, c, t_i))^2$, where $t_i = (i-1) \cdot \frac{1}{29}$.[5] The least squares optimization is

[5] A step size of $\frac{1}{29}$ is used for t to uniformly split the interval $[0, 1]$ and enforce $t_1 = 0$ and $t_{30} = 1.0$.

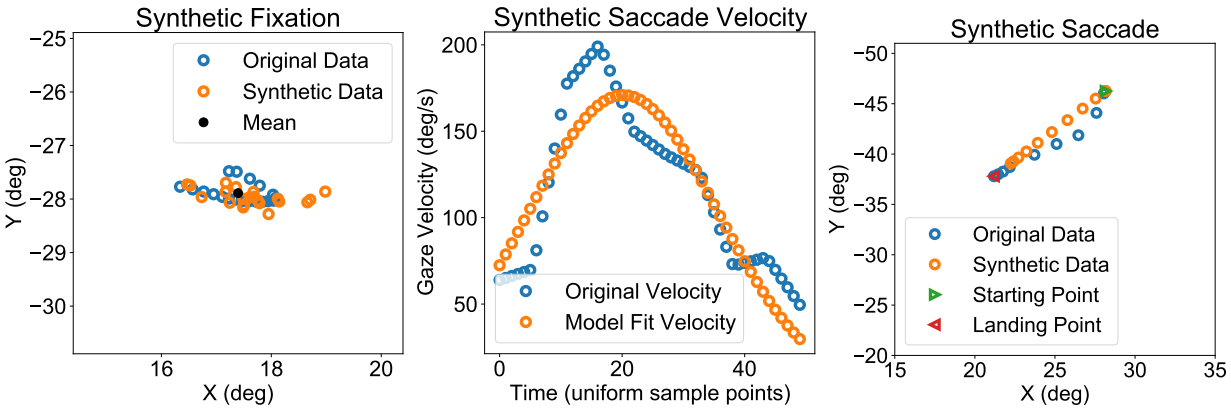

Figure 4-9. Illustration of synthesizing gaze samples for fixations and saccades. Left: Samples from a fixation modeled using a 2D Normal distribution. Blue circles represent real data used to compute mean and standard deviations and orange circles are new positions sampled around the center of the distribution. Center: Saccade velocity profile modeled using a three-parameter Gaussian model. Blue circles represent actual data used to fit the velocity profile, and orange circles represent the output of the fit model. The model fit provides a smooth profile throughout the saccade with a smaller peak velocity. Right: To generate gaze positions for a saccade the velocity profile must be integrated to generate displacements between consecutive points. The displacements are added to the saccade start point along a straight line in the direction of the landing point. Blue circles represent the real gaze positions, and orange circles represent those generated from the model.

performed using the scipy.optimize.least_squares function to return the optimal parameters that fit the Gaussian curve to the raw velocity data.

The Gaussian fit profiles specify the shape of a curve and are input to the *k*-same-synth approach by averaging shape parameters. The fit profiles are not a probability distribution or generative model that can be sampled, and cannot be used to model conditional probabilities directly. Instead, we used a generative model with an encoder and decoder with a latent feature space that is modeled as a probability distribution.

The encoder learns to take a uniformly sampled saccade velocity profile as input and outputs a lower-dimensional feature vector in latent space. The latent space feature vector is input to a decoder model that learns how to map the low dimensional feature values back into a velocity profile. The encoder acts as the generative model by manipulating values of the latent feature vectors and passing them into the decoder to produce synthetic data in the same domain as the

original data. To preserve utility, generative models should provide accurate data reconstructions and enable users to control the output with additional parameters, known as conditions.

A C-VAE model is employed as it provides a deep generative model that can reconstruct velocity profiles using conditions of identity label, amplitude, and duration as input. The C-VAE model benefits from including conditional values as they provide control over the generated synthetic data as parameters.

Probabilities that an individual in the dataset produced a given synthetic saccade profile is required to implement the PD test for C-VAE generated data. The C-VAE inputs and outputs must have the same dimension of 30 values sampled across the saccade profile. The saccade velocity values for each uniform time step i of the real dataset are used to define probability mass functions by mapping continuous values to discrete bins, similar to the step described in Section 4.4.2.2. Fifty bins discretized the velocity values for each time step. The velocity bins uniformly covered a range of velocity values between 0 and a maximum velocity of 1000 degrees per second. A histogram of velocity values for each individual counted the number of values that fell into each bin at every time step i. The counts are divided by the total number of saccades from each individual, so the sum of all probabilities is equal to one. The resulting values provide a joint probability distribution over the likelihood of producing a specific velocity value across the saccade duration. The probabilities across time points are summed to compute the likelihood that an individual generated the set of values in a synthetic saccade profile.

Figure 4-9 (Center) shows a saccade velocity profile from real and synthetic data in the Gaussian model form. The generated velocity profile is smoother than the raw data and retains the starting and landing positions of the saccade. The synthetic saccade positions are then generated from the velocity profile by first computing the displacement between each sample. The displacement A represents the distance between subsequent gaze positions at each time step. The displacement is used to generate to generate a new point $(x_{i+1}, y_{i+1}) = (x_i, y_i) + A \cdot T$ where T is a normalized vector from the saccade start position towards the original landing point (Figure 4-9 (Right)).

4.5.2.2 k-same-synth

In this context, the meaning of k-anonymity is that a released event has an equal probability of coming from the original identity and $k-1$ others. Gaze samples from each detected fixation and saccade event are used to fit model parameters for each event in the dataset. Fixation or saccade events are processed sequentially in the order in which they occurred. The k-same-select sequence mechanism for features is applied directly to the model parameters. The k-anonymous model parameters are then used to sample synthetic data points for fixations and saccades.

For fixations, the μ_x, μ_y, σ_x, and σ_y parameters are processed by the mechanism to modify the centroid position of the fixation using other individuals' data and varying the spatial spread of the samples. The absolute position of the fixation within the stimulus could be shifted to a completely different region as a result. For saccades, the parameters of a Gaussian function model are averaged and used to construct a k-anonymous velocity profile.

procedure k-SAME-SYNTH(k, sample_data, fix_event_params, sacc_event_params)
Parameters: k - k-anonymity parameter
    sample_data - time series of gaze samples, indexed by stimulus m, identity i, and fix/sacc events e
    fix_event_params - Fixation Gaussian parameters, indexed by stimulus m, identity i, and event e
    sacc_event_params - Velocity profile parameters, indexed by stimulus m, identity i, and event e
    $fix\_event\_params \leftarrow k$-same-select sequence($k$,$fix\_event\_params$) ▷ Make fix. params k-anonymous
    $sacc\_event\_params \leftarrow k$-same-select-sequence($k$,$sacc\_event\_params$) ▷ Make sacc. params k-anonymous
    for $m = 1$ to $num\_stimuli$ **do** ▷ Process events from each stimulus independently
        for $i = 1$ to $num\_identities$ **do** ▷ Process samples for each identity
            $fix\_data\_params \leftarrow fix\_event\_params[m,i,:]$ ▷ List of fixation event parameters
            for $e = 1$ to $num\_fixations$ **do**
                $\mu_x, \mu_y, \sigma_x, \sigma_y, t \leftarrow fix\_data\_params[e]$
                $sample\_data[m,i,e] \leftarrow SynthFixation(\mu_x, \mu_y, \sigma_x, \sigma_y, t)$ ▷ Synthesize samples for fixation e
            $sacc\_data\_params \leftarrow sacc\_event\_params[m,i,:]$ ▷ List of saccade event parameters
            for $e = 1$ to $num\_saccades$ **do**
                $a, b, c, t \leftarrow sacc\_data\_params[e]$
                $sample\_data[m,i,e] \leftarrow SynthSaccade(a,b,c,t)$ ▷ Synthesize samples for saccade e
    return $sample\_data$

4.5.2.3 event-synth-PD

Plausible deniability is achieved for samples by generating synthetic gaze positions for fixation and saccade events. The feature vector extracted from the events must pass the privacy criteria (EQ. 2) before being released, as illustrated in Figure 4-8. For fixations, gaze samples are generated by randomly sampling the Gaussian distribution defined by μ, σ_x, and σ_y parameters until the privacy criteria are met. For saccades, gaze samples are generated by synthesizing new velocity profiles with the C-VAE model until the criteria are met.

We defined a PD Event Privacy Test that determines if a synthetic fixation or saccade is k, γ-PD as an alternative to the original privacy test defined in Sec. 4.3.1.2. For each synthetic, the modified privacy test loops over event parameters from other individuals. After identifying an event that passes the test for an individual, k' is incremented and the loop moves on to the next individual. The last step returns pass or fail based on whether $k' \geq k-1$. A fixation or saccade event that passes the privacy test has the guarantee that at least $k-1$ other individuals could have plausibly generated it.

The difference in the guarantee achieved by the PD and PD Event privacy tests is that the k parameter refers to either data records or individuals, respectively. The original PD Privacy Test counts k' based on the number of events that satisfy the PD criterion and could provide a passing result even though all of the records that incremented k' were from the same individual. For Event PD, k' is only incremented once per individual.

procedure PD EVENT PRIVACY TEST(k,γ,Pr_d,$\mathbf{M}$,D))
Parameters: k, γ - plausible deniability parameters, Pr_d - Probability of real seed for y, $Pr\{y \leftarrow \mathbf{M}(d)\}$
        $\mathbf{M}$ - generative model that synthesized y, D - data records from identities other than input
    $i' \leftarrow$ unique integer i', $s.t. \gamma^{-i'-1} < Pr_d \leq \gamma^{-i'}$
    $k' \leftarrow 0$
    for $i = 1$ to *num_identities* **do**
        $D_i \leftarrow D[i]$
        for $d_a \in D_i$ **do**
            if $\gamma^{-i'-1} < Pr\{y = \mathbf{M}(d_a)\} \leq \gamma^{-i'}$ **then**
                $k' \leftarrow k' + 1$

11: **Break**

12: **if** $k' \geq k-1$ **then return** Pass

13: **else return** Fail

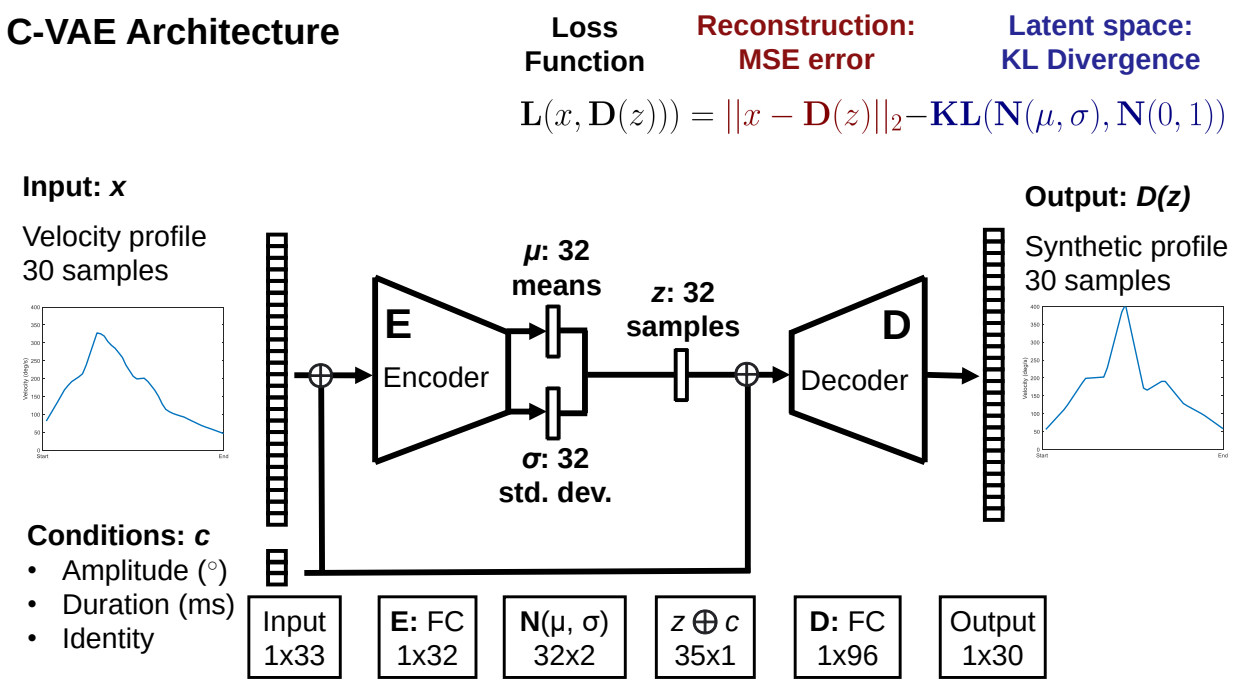

Figure 4-10. C-VAE architecture deployed to generate synthetic saccade velocity profiles. The encoder network maps input velocity profiles and saccade conditions into a latent feature space of normal distributions. The latent feature space is sampled to produce a noise vector that is input to the decoder network along with the saccade conditions. The decoder reconstructs a synthetic velocity profile that appears similar to the input.

The deployed C-VAE model for saccades is used to output synthetic velocity profiles. Synthetic velocity profiles are then used to generate new gaze position samples between a saccade starting and ending point. The decoder network **D** of the C-VAE takes a randomly sampled noise vector z along with the conditions of a real saccade event, i.e., the saccade amplitude, duration, and individual identity as input, and outputs a corresponding synthetic profile. The synthetic profile captures the characteristics of the original saccade to preserve utility while also introducing random variability that will allow the extracted feature vector to pass the privacy criterion.

As shown in Figure 4-10, the C-VAE takes as input a velocity profile x of 30 samples concatenated with conditions c that characterize the saccade. The encoder **E** consists of a one layer fully connected (FC) network with 32 nodes and a ReLU activation layer. The encoder

outputs 64 parameters defining a latent space of Normal distributions μ and σ. The Normal distributions defined by μ and σ parameters are then sampled independently using inverse transform sampling to produce a noise vector z with 64 elements. The decoder $\mathbf{D}(z)$ is a one-layer FC network with 96 nodes and a linear activation layer that takes the noise vector concatenated with c, and outputs the reconstructed synthetic profile. See the Appendix for a detailed description of model training and optimization of parameters.

procedure EVENT-SYNTH-PD(k, γ, sample_data, fix_event_params, sacc_vel_profiles, $CVAE_{enc}$, $CVAE_{dec}$)
Parameters: k, γ - plausible deniability parameters
    sample_data - time series of gaze sample, indexed by stimulus m, identity i, and fix/sacc events e
    sacc_vel_profiles - saccade velocities and conditions, indexed by stimulus m, identity i, and event e
    $CVAE_{enc}$ - Encoder network of C-VAE, maps input to latent space distributions defined by μ and σ
    $CVAE_{dec}$ - Decoder network of C-VAE, maps input random samples $z \bigoplus c$ to synthetic velocities
  for $m = 1$ to $num\_stimuli$ **do** ▷ Process events from each stimulus independently
    for $i = 1$ to $num\_identities$ **do** ▷ Process samples for each identity
      $fix\_data\_params \leftarrow fix\_event\_params[m,i,:]$ ▷ List of fixation event parameters
      for $e = 1$ to $num\_fixations$ **do** ▷ Synthesize fixation samples until PD criterion is met
        $d = (\mu_x, \mu_y, \sigma_x, \sigma_y, t) \leftarrow fix\_data\_params[e]$ ▷ Params for fixation e
        $\mathbf{M}_{fix} \leftarrow N(x,y)$ ▷ 2D Normal distribution that returns t values
        $result \leftarrow False$
        while $result == False$ **do**
          $y \leftarrow \mathbf{M_{fix}}(d)$ ▷ Generate t samples from distribution with curr params
          $Pr_d \leftarrow Pr\{y \leftarrow \mathbf{M_{fix}}(d)\}$ ▷ Probability real seed generated synthetic samples y
          $result \leftarrow$ PD Event Privacy Test($k, \gamma, Pr_d, M_{fix}, fix\_event\_params[m, \neq i, :]$))
        $sample\_data[m,i,e] \leftarrow y$
      $sacc\_data \leftarrow sacc\_vel\_profiles[m,i,:]$ ▷ List of real data saccade profiles
      for $e = 1$ to $num\_saccades$ **do** ▷ Synthesize fixation samples until PD criterion is met
        $d = (\mu_1, \sigma_1, \cdots, \mu_L, \sigma_L) \leftarrow C - VAE_{enc}(sacc\_data[e])$
        $\mathbf{M}_{sacc} \leftarrow N_1, \cdots, N_L$ ▷ Define $\mathbf{M}$ as L independent Normal distributions
        $result \leftarrow False$
        while $result == False$ **do**
          $y = (z_1, \cdots, z_L) \leftarrow \mathbf{M}_{sacc}(d)$
          $Pr_d \leftarrow Pr\{y \leftarrow \mathbf{M_{sacc}}(d)\}$ ▷ Probability real seed generated synthetic samples y

27: $result \leftarrow$ PD Event Privacy Test$(k, \gamma, Pr_d, M_{sacc}, sacc_vel_profiles[m, \neq i, e])$

28: $sample\_data[m, i, e] \leftarrow y$

return $sample\_data$

4.5.2.4 Kalεido

The kalεido privacy mechanism is a composition of multiple DP definitions for processing streams of eye-tracking samples in a real-time manner [151]. The original algorithm processed gaze samples as 2D pixel locations, while our evaluation considered 3D gaze positions. The 3D gaze positions are represented as horizontal and vertical gaze angles mapped on a unit 3D sphere. The distance between gaze positions is computed using the haversine formula as the shortest distance between positions on a sphere [211]. Kalεido combines the concept of (ε,r)-geo-indistinguishability [15] and w-event DP [133]. Geo-indistinguishability is a spatial DP guarantee that is applied to a set of spatial locations. When the pair-wise distance between any two locations x and x' are all less than or equal to r, then $Pr[\mathbf{M}(x) \in O] \leq e^{\varepsilon} \cdot Pr[\mathbf{M}(x') \in O]$. This definition is analogous to the traditional DP guarantee (EQ. 3), except the inputs to the mechanism M are replaced by spatial positions within a distance of r from each other.

The concept of w-event DP is applied to protect sequences of events within a data stream. In the context of gaze data, events refer to gaze positions. The windows of gaze positions are considered neighboring if they overlap for all samples besides one. The guarantee establishes $Pr[\mathbf{M}(S_t) \in O] \leq e^{\varepsilon} \cdot Pr[\mathbf{M}(S'_t) \in O]$ holds for any w-neighboring stream prefixes. In this context, a prefix S_t is all of the gaze samples that occur prior to a specific timestamp, and O represents the output of mechanism $\mathbf{M}$. The guarantee from w-event-DP establishes that any consecutive sequence of w gaze positions before the current time point has an e^{ε} bound on the change in output probabilities of the Kalεido mechanism. Larger values of w mean longer sequences are protected, using more spatial noise to satisfy DP for the same value of ε.

The privacy guarantee of kalεido combines geo-indistinguishability with w-event DP to define (ε,w,r)-DP. Combining the definitions allows users to publish differentially private streams of gaze data within a spatial bound of r on sets of w gaze positions.

Formally, this is defined as

Definition 4. *(ε,w,r)-DP for gaze stream prefixes*

A mechanism $\mathbf{M} : S^g \rightarrow C^g$ *where* S^g *is the domain of all stream prefixes, satisfies (ε,w,r)-DP if for all pairs (w,r)-neighboring gaze stream prefixes* $\{S_t^g, S_t^{g'}\} \in S^g \times S^g$*, we have*

$$\forall O \in C^g, \forall t, Pr[\mathbf{M}(S_t^g) = O] \leq e^{\varepsilon} \cdot Pr[\mathbf{M}(S_t^{g'}) = O],$$

where S_t^g and S_t^g are neighboring sequences of w gaze positions prior to timestamp t, and C^g is the output set of private gaze positions.

The kalεido algorithm ensures the privacy budget ε is distributed over each window of w consecutive samples within the data stream. The algorithm relies on splitting ε into a testing budget ε^{test} and a publishing budget ε^{pub}. The testing budget generates random noise that is added to l_{thresh} and determines whether a new gaze sample should be released based on whether it is part of the current fixation. The spatial threshold plus noise acts as a fixation detector by determining if the current gaze position is close enough to the previous position to skip publishing a new position. If the distance between gaze positions is less than the threshold, then the previous gaze position is repeated in the data stream. Repeating published samples essentially lowered the temporal resolution of the released gaze positions.

The publishing budget influences the scale of spatial noise added when a new gaze sample must be released. The parameter h defines the ratio of testing budget to publishing budget, providing a trade-off between skipping samples more randomly and adding more spatial noise to the released data. The parameters l_{thresh} and h were determined empirically and scaled based on values of r and are not explicitly specified by Liu et al. [151]. Determining optimal parameters for the adaptive budget algorithm does not impact DP privacy, as ε does not change, but impacts utility by determining how often gaze positions are repeated.

The pseudocode below details the kalεido approach for a stream of n_{raw} gaze samples $g_{1,\cdots,n_{raw}}$, window size w, privacy parameter ε, sample distance threshold l_{thresh}, sample skipping parameter t_{skip}, spatial parameter r, and ratio of testing to publishing privacy budget h.

1: **procedure** KALεIDO DP($g_{1,\cdots,n_{raw}}$, w, ε, l_{thresh}, t_{skip}, r, h)

Parameters: $g_{1,\cdots,n_{raw}}$ - stream of gaze positions, w - window size (# samples), ε - DP privacy level

l_{thresh} - Distance threshold for testing, t_{skip} - # of samples to skip over during testing

r - Privacy radius for DP, h - ratio of privacy budget used for testing

$n_{test} \leftarrow \lceil w/t_{skip} \rceil$ ▷ Number of points to test for each window

$\varepsilon_{test} \leftarrow \varepsilon/(h \cdot n_{test})$ ▷ Privacy budget allocated to test each sample

$i_{test} \leftarrow null$ ▷ Index of the last tested gaze position.

$i_{pub} \leftarrow null$ ▷ Index of the last published gaze position.

$g'_i \leftarrow zeros(n_{raw})$ ▷ Published gaze position for sample i, initialized to zeros.

$\varepsilon_i^{pub} \leftarrow zeros(n_{raw})$ ▷ List of privacy budget consumed for sample i, initialized to zeros.

for $i = 1$ to num_raw **do** ▷ Process each window of raw gaze samples

  if $i_{test} \neq null$ AND $t(i) - t(i_{test}) < t_{skip}$ **then** ▷ check if should skip based on t_{skip} parameter

    $g'_i \leftarrow g'_{i_{pub}}$

    $\varepsilon_i^{pub} \leftarrow 0$

    Continue

  $i_{test} = i$

  $l_{dis} = d(g_i, g'_{i_{pub}})$ ▷ Distance between gaze sample i and last published

  $\eta \sim Lap(1/\varepsilon_{test})$ ▷ Sample from Laplace distribution, small values of ε_{test} introduce more noise

  if $l_{dis} \neq null$ AND $l_{dis} \leq l_{thresh} + \eta$ **then** ▷ Test if current gaze is close enough to last published to repeat

    $g'_i \leftarrow g'_{i_{pub}}$

    $\varepsilon_i^{pub} \leftarrow 0$

    Continue

  $i_{pub} \leftarrow i$ ▷ We will publish a new gaze sample, update index of last published

  $\varepsilon_{rem} \leftarrow \varepsilon - \varepsilon/h - \sum_{k=i-n_{raw}+1}^{i-1} \varepsilon_k^{pub}$ ▷ Compute remaining privacy budget for this window

  $\varepsilon_i^{pub} \leftarrow \varepsilon_{rem}/2$

  $g'_i \leftarrow PlanarLap(g_i, \varepsilon_i^{pub}/r)$

return g'

The adaptive algorithm includes several parameters that allow for privacy budget savings while processing the gaze sample at each timestamp. First, a fixed time duration $t_{skip} = 50ms$ is used to skip gaze samples that arrive within t_{skip} of the last published gaze position. Next, after t_{skip} has passed since the last published gaze point, the algorithm moves on to the testing phase. If the current gaze position is within the fixation threshold determined by l_{thresh} and ε^{test}, then the previously published position is re-used, and only ε^{test} of the budget for the current time window

is consumed. The algorithm enters the publishing phase if the new gaze position is farther than the threshold. A noisy gaze position is generated using the ε^{pub} budget with a Planar Laplacian mechanism [15]. The amount of the ε^{pub} budget used decreases adaptively to preserve as much utility as possible while maintaining ε-DP guarantee within each time window. This process is repeated for each time window, and any leftover ε^{pub} budget is recycled into the next window. A complete description of the algorithm and a proof that each window consumes at most ε of the privacy budget is available in the original paper [151].

Li et al. evaluated window sizes of 0.5 seconds and 2 seconds and propose a novel approach for setting the spatial bound parameter r based on the ROIs contained within the stimulus content. We reproduced window sizes for the 100Hz eye-tracking data in our evaluation by setting w to 50 and 200 samples, respectively. For a fair comparison with k-same-synth and event-synth-PD, which do not consider stimulus content, we fix the value of r as either the typical spatial dispersion of fixations or the amplitude of saccade events during viewing tasks. Fixations are typically contained within two visual degrees, and the median saccade amplitude for the EHTask dataset was ten visual degrees during the viewing task [112].

The privacy protection of kalεido is unique in that it can be applied in real-time to arbitrary length streams of eye-tracking data. The DP guarantee through noisy gaze position bounds the probability that the attacker learns what the user was looking at by e^{ε}. Kalεido is the only existing mechanism applied to generate differentially private gaze data at the sample level.

4.5.3 Datasets

We evaluate the above-detailed sample privacy mechanisms on publicly available VR datasets for activity recognition and gaze prediction. The public datasets were released to enable researchers to explore deep learning models and evaluate progress towards better understanding their respective tasks.

4.5.3.1 EHTask

The EHTask [112] dataset includes VR gaze data collected at 100Hz from 30 participants viewing three 360° videos. Participants viewed each video four times, performing different tasks:

free viewing, visual search, saliency, and tracking. Free viewing involved the participant freely exploring the video, visual search had the participant search for and count shapes, saliency asked participants to estimate if the top half or bottom half of the environment was more salient, and tracking required the user to look at and track the nearest moving object. A deep network was trained to process gaze and head data to classify windows of samples into the four activity classes. Activity recognition has applications in adaptive XR interfaces that leverage context when presenting digital content [126]. For example, detecting when a user is performing a visual search in a grocery store could trigger visual labels that guide the user to the cheapest item or an item from their shopping list [92].

4.5.3.2 DGaze

The DGaze [114] dataset included VR gaze data collected at 100Hz from 43 participants that explored and navigated two 3D rendered scenes. Within each environment, multiple animals dynamically move around, attracting the visual attention of the participant. Gaze data trained the DGaze deep learning model for gaze prediction. DGaze processed saliency of scene content, tracked objects, and current gaze position to predict a future gaze position. Gaze prediction by DGaze has been demonstrated in the context of foveated rendering and can help account for latency in the eye-tracking and rendering pipeline [192, 16, 114].

4.5.4 Metrics

4.5.4.1 Identification Rate

Identification rate is computed by extracting features and applying an RBFN as described in Section 4.4.5.1. Privacy mechanisms are applied to sample data before events are detected and biometrics features are extracted. The mechanism impacts the resulting feature values, reducing the risk of re-identification.

4.5.4.2 EHTask

Utility for EHTask is based on classifying windows of samples into one of four tasks. Performance is computed as the accuracy of window classification $\frac{TP+TN}{TP+FP+TN+FN}$. The chance rate of guessing for EHTask is equal to 25% as there are four possible activities.

The EHTask classification model computes features from 1D convolutional layers applied to sequences of eye-in-head, head-in-world, and gaze-in-world samples fed into bidirectional GRU layers with outputs that are concatenated into a fully-connected network used to predict the activity being performed. The classification model considers data from the past ten seconds for inputs to the model.

Classification accuracy is computed with test data from 25% of the total data for each task by segmenting users into train and test sets. The EHTask model is a deep architecture based on 1D CNNs, processing gaze-in-world, eye-in-head, and head-in-world movement signals. The gaze privacy mechanisms only modify the eye-in-head and gaze-in-world data streams to affect the task classification utility.

4.5.4.3 DGaze

Gaze prediction accuracy is measured as the angular error for predicted gaze position 100ms in the future compared to the actual gaze position for each input to the gaze prediction model, as described in Section 3.6.4.1.

4.5.5 Results

4.5.5.1 EHTask

Utility results from each privacy mechanism are presented in Figure 4-14. Visual results for each mechanism applied to EHTask can be seen in Figures 4-11, 4-12, and 4-13. Gaze positions from an individual in EHTask viewing a stimulus for 150 seconds is plotted for each privacy mechanism across various parameters.

Table 4-3. EHTask results for the k-same-synth privacy mechanism. EHTask classification rates showed a sharp decline for large values of k, lowering rates to 29.1% at k=8. Identification rates fell to as low as 8.5%, remaining above the chance rate of 3%.

Params	RBFN identification rate % (↓)	EHTask classification accuracy % (↑)
No mechanism	28.0	82.8
$k = 2$	9.7	61.8
$k = 4$	8.7	45.8
$k = 6$	8.5	40.6
$k = 8$	7.5	29.1

Table 4-4. Privacy and utility results for the event-synth-PD privacy mechanism. Re-identification rates ranged from 9.2% to 15.0%, while classification rates covered a range of 66.7% to 73.0%.

Params	RBFN identification rate % (↓)				EHTask classification accuracy % (↑)			
No mechanism	28.0				82.8			
k	$\gamma = 1.0$	$\gamma = 1.5$	$\gamma = 2.0$	$\gamma = 3.0$	$\gamma = 1.0$	$\gamma = 1.5$	$\gamma = 2.0$	$\gamma = 3.0$
$k = 2$	12.5	13.5	11.7	13.8	71.9	74.0	72.2	71.1
$k = 4$	9.2	12.2	15.0	14.2	73.0	73.5	70.1	66.7

The k-same-synth mechanism enforces k-anonymity on the parameters that define fixation and saccade events from each individual. Figure 4-11 shows the k-same-synth mechanism and the effect of privacy parameter k on the output gaze sample positions. Horizontal gaze positions are impacted more than vertical positions as typically, viewers keep their head level during image viewing [59]. Large shifts in the synthetic horizontal gaze positions are observed as early as k equals two, as fixation positions are updated based on other individuals' μ_x and μ_y values.

Table 4-3 provides re-identification rates and classification accuracy for several values of k. The mechanism produced re-identification rates in the range of 7.5% to 9.7%, lower than 28.0% from unmodified data. Computed re-identification rates remained less than $\frac{1}{k}$. The lowest re-identification rate was 7.5% at k equals eight, which is higher than chance (1/30 = 3.3%). Rates above chance result from the type of biometric features extracted from the synthesized data. For example, the fixation feature set includes the duration of fixations and the spatial dispersion within a fixation. The fixation synthesis method we deployed does not directly modify fixation duration but does modify the spatial distribution that influences dispersion. Updating the samples may impact the event boundaries detected in the synthetic data but overall have a smaller impact on the features that are temporal in nature. Thus, any individual trend uniquely influenced by temporal features is not guaranteed to be removed by k-same-synth.

Table 4-5. EHTask results for the kalεido privacy mechanism. The input parameters ε, w, r determine the amount of privacy noise added to the released gaze samples. Lower values of ε and larger values for r and w indicate higher privacy. DP parameters of ($r = 10^\circ$, $w = 200$ samples, $\varepsilon = 1$) produced re-identification rates near chance while lowering accuracy on the utility classification task to 73.3% (10% less than the original data).

Params	RBFN identification rate % (↓)				EHTask classification accuracy % (↑)			
No mechanism	28.0%				82.8			
ε	$w = 50$ $r = 2^\circ$	$w = 50$ $r = 10^\circ$	$w = 200$ $r = 2^\circ$	$w = 200$ $r = 10^\circ$	$w = 50$ $r = 2^\circ$	$w = 50$ $r = 10^\circ$	$w = 200$ $r = 2^\circ$	$w = 200$ $r = 10^\circ$
$\varepsilon = 10$	10.2	5.7	8.0	6.7	71.6	70.9	70.6	74.9
$\varepsilon = 5$	7.3	4.0	6.0	7.2	43.9	76.6	72.2	73.5
$\varepsilon = 2$	7.7	10.5	9.0	10.3	64.8	73.5	55.1	61.0
$\varepsilon = 1$	7.8	9.3	8.3	6.0	72.0	69.7	68.1	73.3

The k-anonymous dataset was used to train an activity classification model for utility. k-anonymity introduced a loss in utility, dropping the classification accuracy of gaze windows in the test set from 82.8% to as low as 29.1%. A large drop in classification accuracy impacts interfaces or models that depend on recognizing the user's activity. Reasonable classification accuracy of 61.8% was achieved at k equals two, but quickly falls off and reaches chance rates at k equals eight (29.1%).

The event-synth-PD mechanism guarantees that fixation positions and the generated saccade velocity profiles are k, γ-PD. The interpretation of this guarantee is that there are $k-1$ other events that could have plausibly synthesized the released gaze samples. Figure 4-12 shows the event-synth-PD mechanism and the effect of privacy parameters k and γ on the output gaze sample positions. Synthetic horizontal and vertical gaze positions are largely unaffected for both values k, and all values of γ.

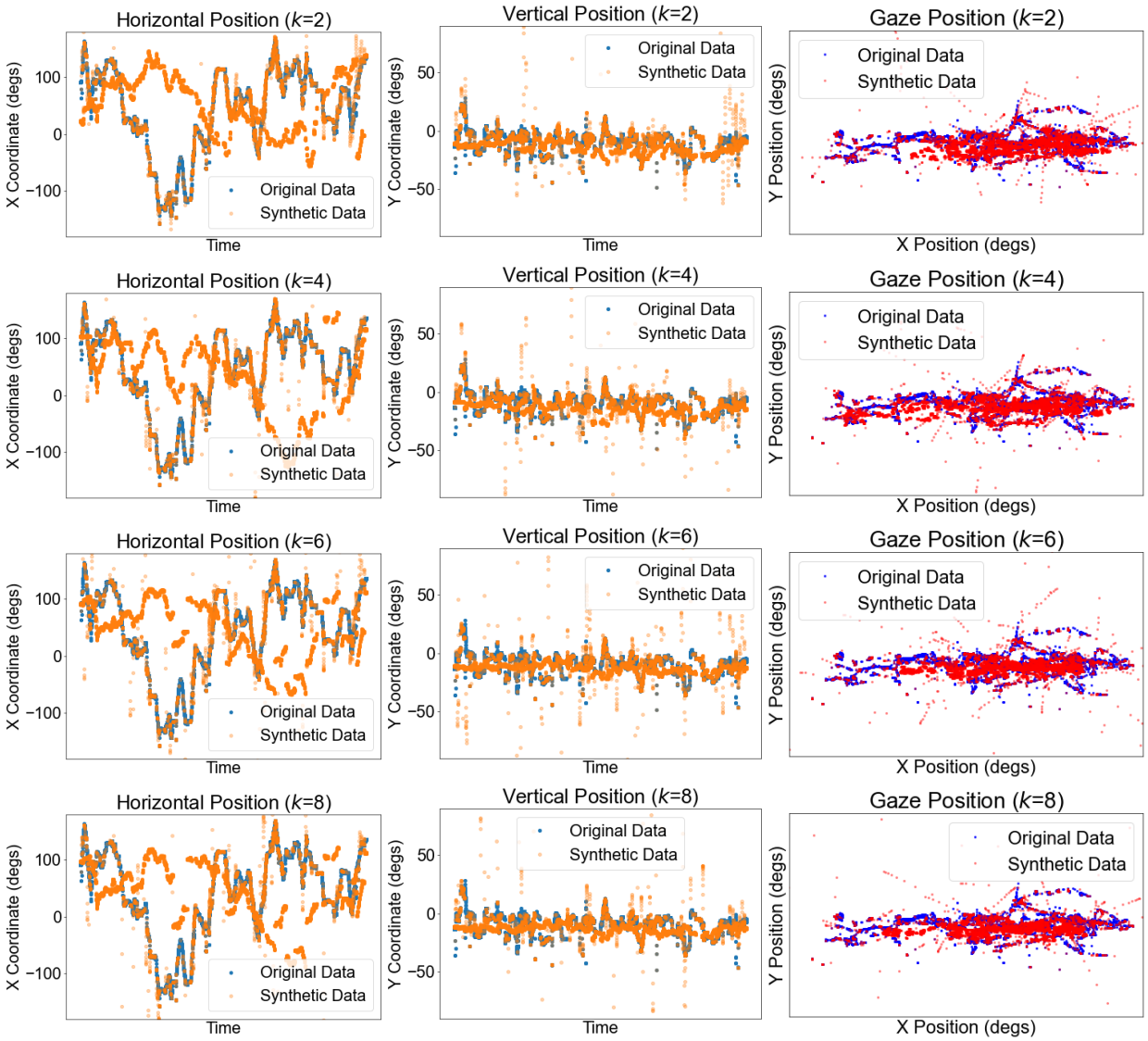

Figure 4-11. Real and synthetic gaze positions for the *k*-same-synth mechanism from Identity 1 of EHTask performing the viewing task on stimulus 1. Left Column: Horizontal Gaze position time series. Middle: Column: Vertical Gaze position time series. Right Column: 2D Gaze positions in equirectangular format. Higher privacy is achieved at larger values of *k* where more noise is observed in the trace of gaze positions.

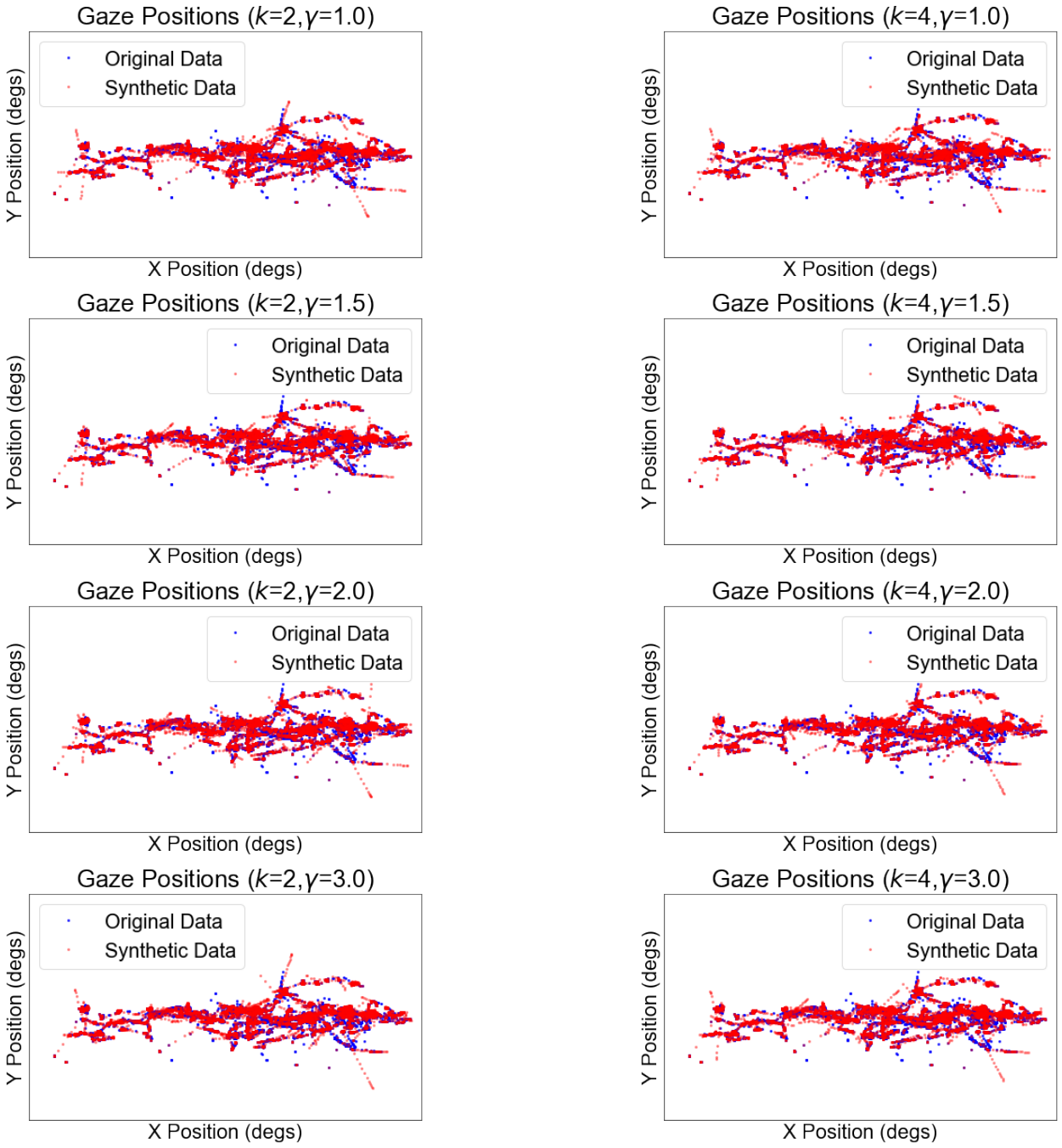

Figure 4-12. Real and synthetic gaze positions for the event-synth-PD mechanism from Identity 1 of EHTask performing the viewing task on stimulus 1. Synthetic gaze positions are consistent across privacy parameters, suggesting that similar synthetic gaze positions are passing the privacy test in all configurations. The generated points do not deviate much from the original data and re-identification attacks are still possible on the generated data (Table 4-4).

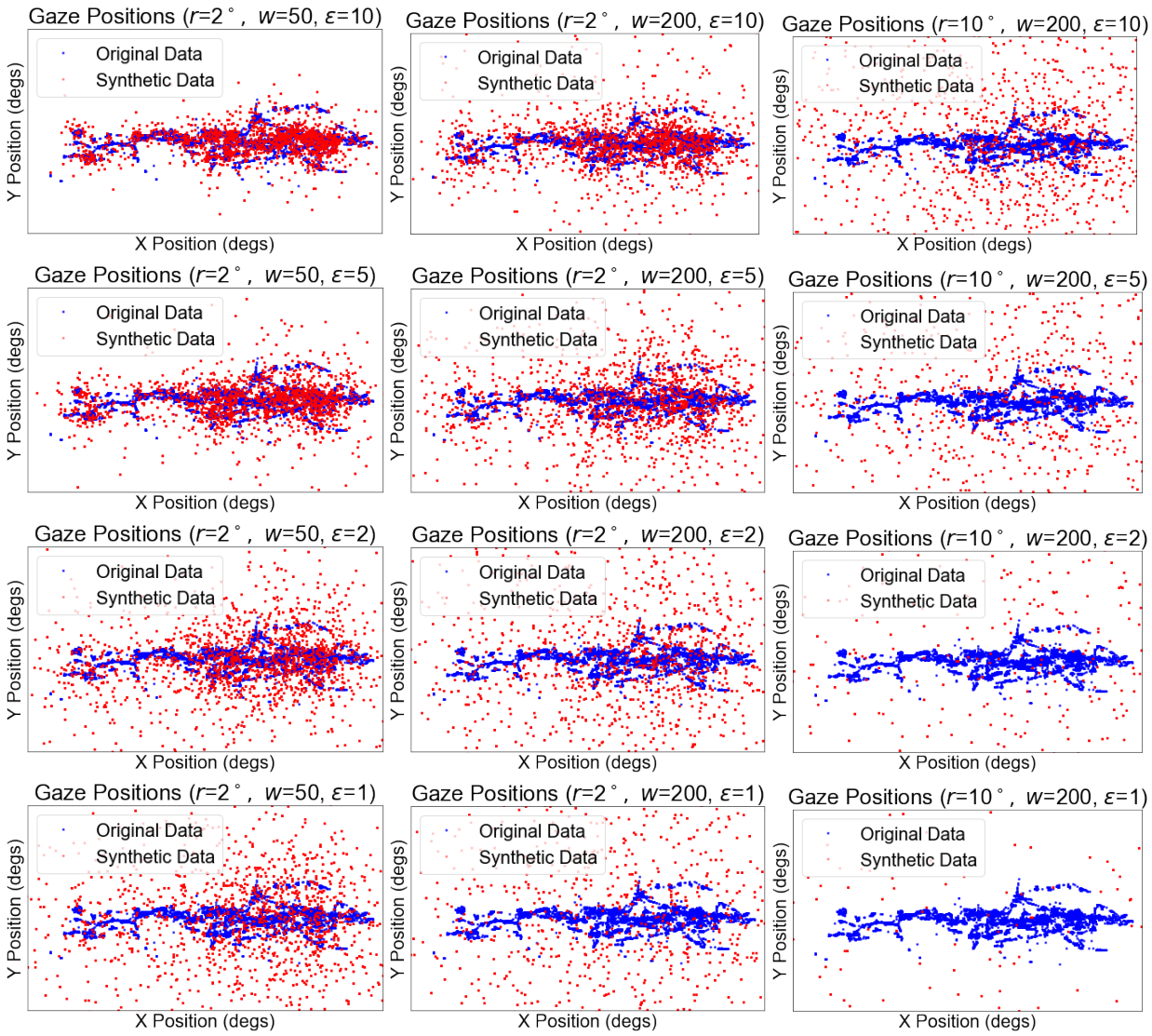

Figure 4-13. Real and synthetic gaze positions for the kalεido mechanism from Identity 1 of EHTask performing the viewing task on stimulus 1. Higher levels of privacy increase spatial noise in the data and force the data to spread out away from the original gaze positions, losing utility of the gaze data at the cost of the spatial DP guarantee.

Table 4-4 provides re-identification rates and classification accuracy for values of k equal to two and four, and several values of γ. The lowest re-identification rate was 9.2%, which was the strongest level of PD privacy evaluated (k=2 and γ=1.0). The produced rates were higher than chance (1/30 = 3.3%), suggesting that event-synth-PD does not provide complete protection against re-identification attacks at this privacy level.

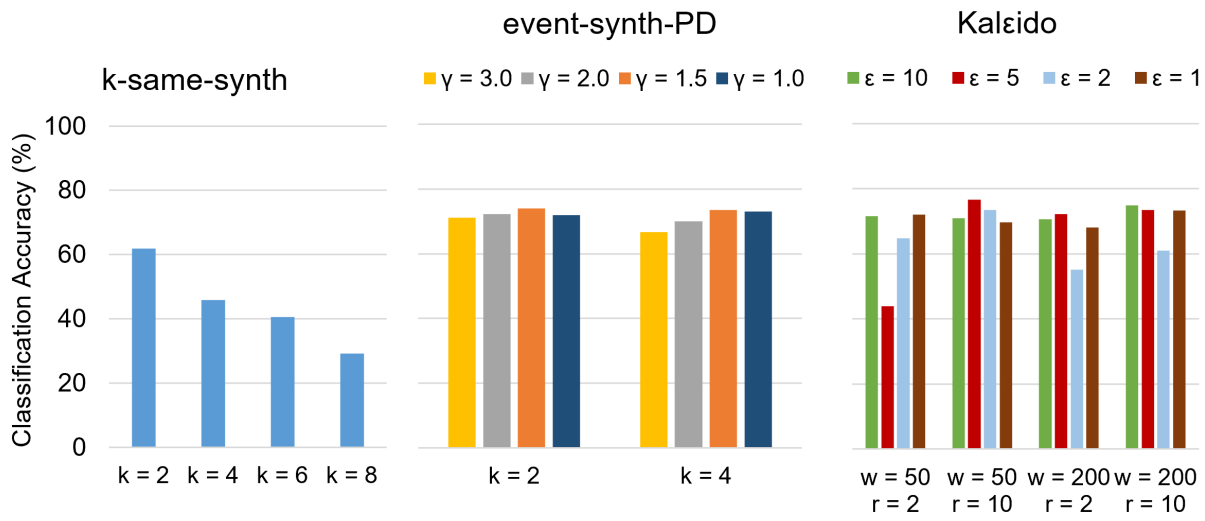

Figure 4-14. Classification rates presented a downward trend for k-same-synth across values of k with classification rates starting near 60% and dropping to 29% with more privacy. Classification rates of event-synth-PD presented a uniform trend near 72% for all values of k and γ. Classification rates for kalεido ranged from 42% to 76%; with an accuracy of 69% for parameters that achieved strong DP privacy ($r = 10°$, $w = 200$, and $\varepsilon = 1$).

The k, γ-PD dataset was used to train an activity classification model for utility. The mechanism introduced a loss in utility, dropping the classification accuracy of gaze windows in the test set from 82.8% to as low as 66.7%. A 16% drop in classification accuracy may have an impact on interfaces or models that depend on recognizing the user's activity and was similar in scale to the utility achieved by k-same-synth at k equals 2. The utility was preserved with a small impact on performance while reducing the risk of re-identification.

Figure 4-13 shows the kalεido DP data and the effect of mechanism parameters r, w and ε on the output gaze positions. Visualizations show that more spatial displacements are introduced for higher privacy achieved at larger values of r and w and smaller values of ε. Eventually, parameter values introduce so much noise that the gaze positions deviate and form a cloud of positions around the real data at the highest level of privacy ($r = 10°$, $w = 200$, $\varepsilon = 1$). Overall, data generated at ε equals two or less does not retain the structure of the data impacting the ability to both re-identify users and classify activities.

Table 4-5 provides re-identification rates and classification accuracy across parameter values. The lowest re-identification rate was near chance (3.3%) at the highest level of privacy. Even at lower privacy settings, the produced re-identification rates were at most 10.5%. Rates were comparable to or lower than the re-identification rates from all other mechanisms. Kalεido was the only mechanism that achieved chance re-identification rates.

The ε-DP dataset was successful at training an activity classification model. The mechanism introduced a loss in utility, dropping the classification accuracy of gaze windows in the test set from 82.8% to as low as 43.9%. However, the 43.9% classification rate was not produced at the highest privacy parameter, and generally rates remained near 70%. The kalεido mechanism repeats gaze positions at higher privacy levels as part of the testing stage as the algorithm skips samples more frequently for lower values of the privacy budget ε. The EHTask model also takes head movement data as input. Repeating gaze samples modifies the training dataset to contain less information about eye movements and prioritizes optimization around the head movement data, which remains sampled at 100Hz. The resulting EHTask model learns to classify tasks based on noisy gaze samples at a low sampling rate and head movements relative to them, achieving comparable utility at both $\varepsilon = 10$ and $\varepsilon = 1$. The classification rates achieved with kalεido were less than other mechanisms; however, the small impact on utility with the ability to lower identification rates to chance produced the best privacy-utility trade-off.

4.5.5.2 DGaze

Figure 4-16 presents utility results from each privacy mechanism. Visual results for each mechanism applied to the DGaze prediction task can be seen in Figure 4-15. The last gaze position of the input sequence to the model is plotted with an arrow drawn to the ground-truth gaze position 100ms in the future. DGaze predictions from training with and without mechanisms demonstrate how well each approach preserved utility relative to the baseline performance and the actual gaze position.

Table 4-6. DGaze results for the k-same-synth privacy mechanism. Identification rates are near chance (1/43 = 2.3%) for unmodified data, and remain at chance for the mechanism. Accuracy for gaze estimation 100ms into the future is decreased across all values of k, introducing up to 2.2° of additional error on average compared to the results with no mechanism applied.

Params	RBFN identification rate % (↓)	DGaze prediction error (°) (↓)
No mechanism	2.3	4.3
$k = 2$	2.0	5.4
$k = 4$	2.3	6.0
$k = 6$	2.1	6.3
$k = 8$	1.1	6.5

Table 4-7. DGaze utility results for the event-synth-PD privacy mechanism. Identification rates are near chance (1/43 = 2.3%) for unmodified data, and remain at chance for the mechanism. Gaze prediction error ranged between 6.8° and 9.1°, with lowest errors produced for higher values of γ and k=2. Classification rates did vary with higher utility at PD parameters with less privacy.

Params	RBFN identification rate % (↓)				DGaze prediction error (°) (↓)			
No mechanism	2.3				4.3			
k	$\gamma = 1.0$	$\gamma = 1.5$	$\gamma = 2.0$	$\gamma = 3.0$	$\gamma = 1.0$	$\gamma = 1.5$	$\gamma = 2.0$	$\gamma = 3.0$
$k = 2$	1.2	1.3	1.9	1.5	8.4	8.9	6.8	7.0
$k = 4$	1.3	1.2	1.5	1.5	8.8	8.7	9.1	9.0

Table 4-8. DGaze results for the kalεido privacy mechanism. Identification rates are near chance (1/43 = 2.3%) for unmodified data, and remain at chance for the mechanism.

Params	RBFN identification rate % (↓)				DGaze prediction error (°) (↓)			
No mechanism	2.3%				4.3			
ε	$w = 50$ $r = 2°$	$w = 50$ $r = 10°$	$w = 200$ $r = 2°$	$w = 200$ $r = 10°$	$w = 50$ $r = 2°$	$w = 50$ $r = 10°$	$w = 200$ $r = 2°$	$w = 200$ $r = 10°$
$\varepsilon = 10$	2.3	2.3	2.3	2.8	5.0	6.6	5.5	7.4
$\varepsilon = 5$	2.3	2.8	2.1	3.0	5.4	7.0	6.4	8.7
$\varepsilon = 2$	1.5	2.1	1.9	2.3	5.6	8.9	7.9	8.4
$\varepsilon = 1$	2.6	2.3	4.6	2.1	7.0	10.5	7.9	9.3

Figure 4-15 shows gaze prediction from a model trained using the k-same-synth mechanism with k equals four. The figure demonstrates that the predicted gaze position deviates from the unmodified model's predicted gaze. Gaze errors arise from overshooting the predicted position or shifting away from the actual gaze. Table 4-6 provides re-identification rates and average gaze

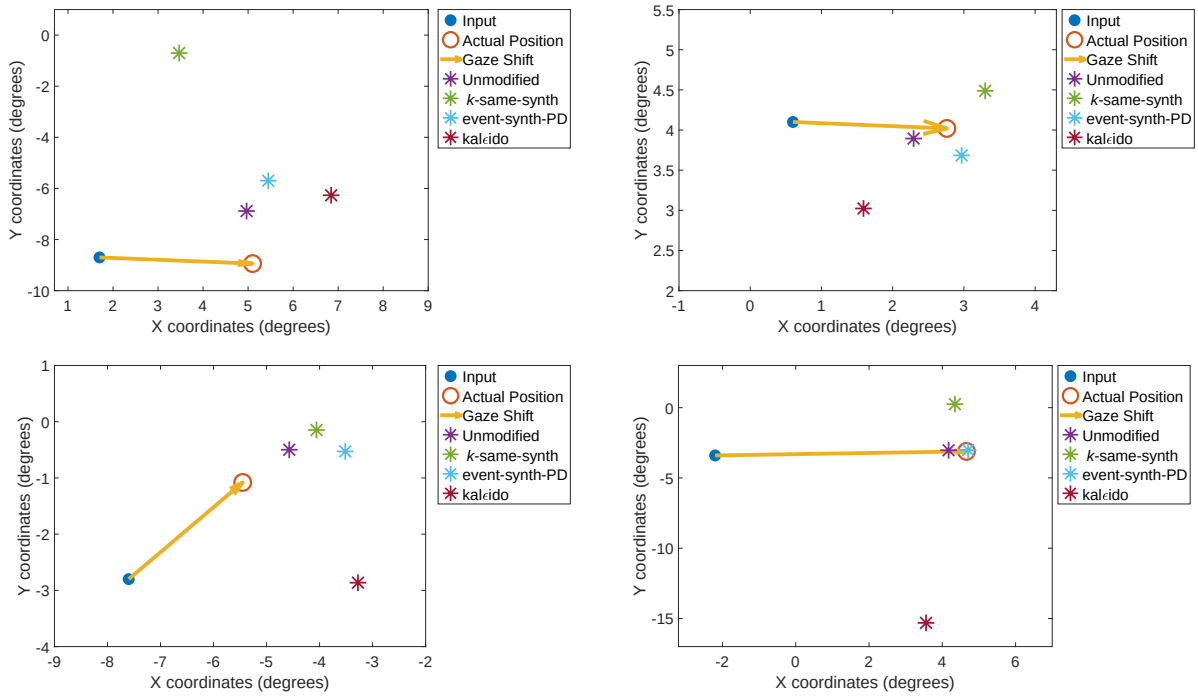

Figure 4-15. Illustration of DGaze gaze predictions from models trained on unmodified and private data. Colored stars indicate gaze predictions output from DGaze trained for each privacy mechanism and unmodified data. The blue dot indicates the gaze position at the time of prediction, with an arrow drawn to the actual gaze position 100ms into the future (orange circle). Top Left: All predictions are inaccurate. Top Right: *k*-same-synth and event-synth-PD overshoot the actual gaze position and unmodified undershoots. Bottom Left: All predictions overshoot the actual gaze position. Bottom Right: All predictions besides *k*-same-synth and Kalɛido are accurate. The kalɛido mechanism introduces the most error on average.

prediction accuracy for several values of k. Without a privacy mechanism applied, the DGaze dataset produced identification rates at chance. We hypothesized that low rates result from the DGaze dataset providing viewers only two scenes to explore, containing sparse environments where they were instructed to follow animals around by using teleporting to navigate. The combination of a prescribed task with minimal individual variation and low diversity in stimuli results in features that are not reliable for user identification. This claim is supported by prior work that demonstrated identification rates for free viewing tasks were 60% higher than that of guided training sessions within similar 360° image environments [147].

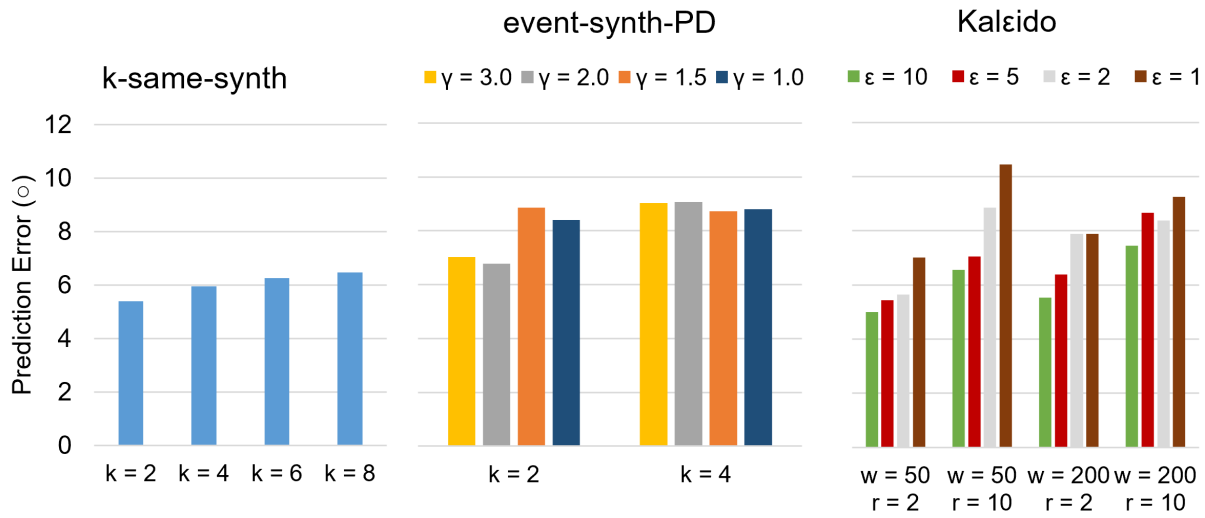

Figure 4-16. The gaze errors from k-same-synth increase from 5.4° to 6.5° across values of k, indicating a slight negative linear trend. Gaze errors from event-synth-PD was uniform at 4.6° for all k and γ. Gaze errors for kalεido presented an increasing linear trend within each combination of r and w as ε goes from lower privacy (10) to higher privacy (1).

The k-anonymous dataset introduced a loss in utility, increasing the prediction accuracy from 4.3° up to 6.5°. An increase in prediction error on the scale of 2.2° introduces a small impact on gaze prediction applications, such as foveated rendering. However, it can be compensated for with a larger foveal region parameter. For example, perceptual experiments by Guenter et al. [99] on traditional displays have found an optimal size for the foveal region between three to four visual degrees. Increasing the foveal region within this range by 2.2° to accommodate additional error would reduce the rendering speedup from a factor of ten to four. The privacy mechanism introduces a loss in rendering savings but still offers practical improvement by reducing rendering cost to a fourth of the time needed for a non-foveated system.

Figure 4-15 shows gaze prediction error from a model trained using the event-synth-PD mechanism with $k = 2$ and $\gamma = 2.0$. Table 4-7 provides re-identification rates and average gaze prediction accuracy that demonstrated a uniform trend for several values of k and γ. The k,γ-PD dataset introduced a moderate loss in utility, increasing the prediction accuracy from 4.3° up to

9.1° across parameters. The introduced error is almost double that of the unmodified data, introducing more errors than k-same-synth.

Figure 4-15 shows the prediction from DGaze trained on kalεido DP data with parameters $r = 2.0$, $w = 50$ and $\varepsilon = 1$. Visualizations show high spatial displacement from the actual gaze position for the kalεido predictions. Table 4-8 provides re-identification rates and average gaze prediction accuracy. Kalεido introduced the most error of all mechanisms at high DP privacy but introduced reasonable errors of 5.6° or less for the smallest values of r and w. Figure 4-16 demonstrates a linear trend for increased prediction error within each set of DP parameters as ε decreases and stronger privacy is achieved.

4.6 Discussion

To explore $RQ_{4.1}$ and $RQ_{4.2}$, we presented feature data mechanisms that offer alternatives to DP in the form of k-anonymity and plausible deniability. Our alternatives protect both feature and sample datasets against re-identification attacks by considering the risk of re-identification probabilities instead of a DP privacy guarantee. The k-same-select sequence mechanism produced identification rates at chance while preserving model accuracy of 72% for document type classification using the MPIIDPEye dataset. We recommend using k-same-select sequence to protect feature datasets against re-identification in a computationally efficient manner. When the feature set of the attacker is explicitly known, the method directly bounds the probability of re-identification and provides a strong defense against attacks.

To explore $RQ_{4.3}$ and $RQ_{4.4}$, we presented sample data mechanisms that achieve alternative privacy guarantees to DP. For sample data, the achieved k-anonymity, and plausible deniability guarantees are applied at the event level before features are extracted. The presented mechanisms reduce the risk of re-identification, though not all mechanisms reduce identification rate to chance. We validated $RQ_{4.3}$ by observing that re-identification rates on the EHTask dataset were lower than unmodified data; however, it was difficult to achieve chance identification rates for the EHTask dataset, as only kalεido lowered rates to 3%.

For utility evaluations, the EHTask dataset was used to train a deep model that classified

task types based on input eye and head movement sequences. The DGaze dataset trained a model for gaze prediction 100ms into the future for input gaze, head, and saliency data. The computed results answered $RQ_{4.4}$ for task classification by revealing a practical trade-off between privacy and utility for kalεido at $r = 10°$, $w = 200$ samples, and $\varepsilon = 1.0$.

For the gaze prediction task, which predicts a continuous gaze direction instead of a categorical task type, both the k-same-synth and kalεido mechanisms introduced a negative linear trend between their privacy parameters and the impact on utility. The mechanism for plausible deniability had a moderate impact on gaze prediction utility, suggesting that k-same-synth and kalεido with small r and w parameters are the best choice for this task.

The DGaze dataset provided a low risk of re-identification due to using only two 3D rendered stimuli with a prescribed search and follow task. Identification rates were already near chance for the unmodified DGaze dataset. However, other gaze prediction datasets could produce a higher risk if 360° images or videos were used with a free viewing task. Our utility results answered $RQ_{4.4}$ by indicating a linear negative impact on utility for the k-anonymous and ε-DP mechanisms, with moderate impact across all k, γ-PD parameters. For datasets outside of DGaze with a higher risk of re-identification, the kalεido mechanism is recommended, as it has a large impact on identification rates, despite a systematic negative impact on utility.

Table 4-9 summarizes the privacy guarantees achieved for eye-tracking datasets and the mechanisms that achieved practical privacy-utility trade-offs. For feature datasets used for classification tasks, we recommended the k-same sequence mechanism as it retains between class differences while averaging away individual differences. For sample datasets used on classification models, practical trade-offs were achieved by both event-synth-PD and kalεido. However, the parameters for kalεido that achieved the best trade-off resulted in a highly sparse sampling of the gaze data positions. The EHTask model we evaluated also takes head movements as input and retained utility by relying less on gaze data. The gaze positions produced by event-synth-PD resembled real data much more closely (Figures 4-12 and 4-13), and would generalize better to utilities that only rely on gaze data. For sample datasets used in gaze

prediction, a practical trade-off was only achieved for the k-same-synth mechanism that limits the introduced gaze prediction error.

Table 4-9. Summary of privacy-utility trade-offs for across privacy mechanisms and data applications. Check marks indicate a practical trade-off of applying a privacy mechanism for that application.

Mechanism	Guarantee	Data type	Utility	Practical trade-off
k-same-select sequence	k-anonymity	Features	Classification	✓
Marginals	k,γ-PD	Features	Classification	×
Exponential-DP	ε-DP	Features	Classification	×
k-same-synth	k-anonymity	Samples	Classification	×
event-synth-PD	k,γ-PD	Samples	Classification	✓
Kalεido DP	ε-DP	Samples	Classification	✓
k-same-synth	k-anonymity	Samples	Prediction	✓
event-synth-PD	k,γ-PD	Samples	Prediction	×
Kalεido DP	ε-DP	Samples	Prediction	×

4.7 Limitations

Our identification results were limited to an RBF network, although prior work explored random forest [222], SVM [170], k-NNs [41] and deep network [171, 155] models. The identification model impacts our empirical results, but it does not impact the achieved theoretical guarantees and the fundamental differences between k-anonymity, plausible deniability, and differential privacy. For feature mechanisms, we explored the seminal DP approach [232], and did not evaluate the DCFPA mechanism of Bozikir et al. [41]. While this method's DP privacy parameter ε has the same meaning, the achieved utility and specific privacy result may vary from that of the Exponential mechanism.

Our utility results depend on the application and metric applied, i.e., gaze prediction evaluated by angular error 100 ms into the future, and the model trained and tested on the data. Each utility model has multiple hyper-parameters in defining its architecture and training process. Our analysis was limited to the parameters reported to be optimal by the authors of their respective papers. While this provides a benchmark on the change in performance relative to unmodified data, de-identified data with tweaked model parameters could result in higher utility. It would be interesting to explore trends in utility performance for privacy mechanisms across different model parameters to observe if similar trends are produced compared to unmodified data.

A fundamental limitation of privacy mechanisms with the parameter k is the assumption that each stimulus or task has data from at least k individuals. For datasets with a small number of individuals, or cases where all individuals do not view every stimulus, there is an upper bound on the value of k. Thus, the desirable level of privacy or identification rates may not be achieved for the privacy mechanisms. In contrast, DP allows the data owner to increase privacy until the desired level is reached.

High utility results for classification tasks depend on the difficulty of the task and the model being used. Our results considered document type recognition (MPIIDPEye) and activity recognition (EHTask) datasets that included three and four utility classes, respectively. Datasets with a larger number of classes would be more difficult to perform accurately and may impact the generalization of our comparison of privacy mechanisms with respect to utility.

CHAPTER 5
CONCLUSIONS

An increasing amount of XR applications and datasets collect eye-tracking data that introduces privacy concerns based on what can be inferred from data, including identity. To address concerns of eye trackers collecting identifying data we have developed mechanisms to enhance privacy for eye images [122, 120, 123], designed frameworks to protect privacy while streaming eye-tracking data to third-party applications [62], and provided mechanisms that achieve formal privacy guarantees against re-identification attacks on datasets [60].

5.1 Protecting Iris Biometrics

Most XR eye trackers rely on infrared images of the eye to perform gaze estimation. Eye images also capture the iris pattern, a gold standard biometric, introducing the risk of leaking the user's identity in the data stream. Our proposed solution applies blur to images post-capture in software when the platform is trusted or pre-capture by the user using optical defocus. Blur filters away the high-frequency iris patterns and reduces the risk of using eye-tracking images to spoof the user's identity. The trade-off from introducing blur impacts the quality of gaze estimation, measured in data-level accuracy and downstream impact on different applications. We employed gaze estimation software and perceptual studies with virtual avatars to understand how much error is introduced in the data signal and model the impact of introduced errors on the perception of an avatar. We were the first to identify the risk of leaking identity through iris patterns in eye-tracking images. Our approach provides a user-controlled pre-capture mechanism to protect identity while enabling eye-tracking applications.

User-controlled privacy lets users determine what applications they trust. Users make similar decisions when they choose not to connect to public wireless networks or rely on hardware hacks such as placing tape over a laptop webcam to protect their privacy. Future XR ecosystems will collect data at home, work, and in public. For example, VR arcade machines can be outfitted with blurring optics that enable gaze estimation without collecting identifiable information about the user's eye. Our approach provides a template for identifying risks from XR sensors, in this case leaking iris biometric data, and providing a defense mechanism in a user-controlled manner.

5.2 Privacy for Streaming Eye-Tracking Data

Eye-tracking applications of saliency, redirected walking, and gaze prediction contributes to the future of user modeling and interaction in XR. Applications take raw gaze samples from the eye-tracking platform and apply pre-processing to extract relevant metrics for the corresponding utility. Sharing raw gaze samples introduces the risk of identifying users from eye movement biometrics. Features extracted from raw samples identify users performing common tasks. Tasks such as free viewing produce identification rates with high accuracy [62] and could be used to recognize users across applications. Our Gatekeeper model is an API that processes gaze sample data within the XR platform. The API serves the processed data metrics directly to the application instead of sharing raw samples.

By implementing the appropriate data metrics on the platform, the API has no impact on application utility. For applications such as gaze prediction requiring sample data, our privacy mechanisms add noise or downsample data to modify the data stream in real-time and reduce the identification rate. Real-time implementations are required for time-sensitive eye-tracking applications and limit the complexity of privacy mechanisms that can be deployed in practice. A benefit of the Gatekeeper model is that it can easily be extended to restrict permissions at different data pipeline components. If a user wants to eliminate eye-tracking risks, they can turn off foveated rendering or other core functionalities implemented by the XR platform. The exact details on how core platform components process data are known by the operating system and could be controlled by the user in their device settings. However, privacy is not always a binary setting when data leaves the platform to a third party. Once data leaves the platform, how it is saved or processed is unknown. In the scope of our evaluation, multiple variables impact identification rates, including the volume of data collected, the task being performed, the quality of data, and the selected privacy mechanism. It is not easy to extrapolate results and declare a data stream as private or not when users do not know or have control over the other variables in re-identification. Evaluations on more scenarios can provide additional insight into risk severity beyond the datasets we explored.

In security and privacy research, it is critical to be proactive instead of reactive. Current XR risks are becoming a topic of discussion among companies [39], researchers [103, 189, 212], and professional standards organizations [164]. However, no XR attacks have been heavily publicized to this point. Thus, now is the time to get ahead of risks and learn from pitfalls in the past within mobile devices. A specific example is evaluating whether mobile apps directed at children are compliant with existing privacy laws [204]. Alomar and Egelman pointed out that child privacy legislation has existed for over 25 years, but most Android apps that target children were not in compliance [14]. Through surveys and interviews with child app developers, they found that a major source of non-compliance was the third-party SDKs they integrated, including advertising services, and a lack of resources to evaluate whether their app was compliant. Larger game studios can outsource privacy audits to external services, but smaller groups assumed that their app was compliant if it was not rejected by the app store when published. While app store compliance checking can identify privacy violations, developers pointed out that the rejection notifications took up to a week for processing and did not provide sufficient details to identify and fix the offending SDK or feature. Based on these findings, we remark that the XR platforms of the future should provide app stores and useable tools that allow developers, who are not privacy experts, to evaluate apps before they reach production. Happa et al. have previously discussed the requirements of a privacy certification framework specific to XR, including the need for a common vocabulary for XR-specific threats, a multi-disciplinary standards group to exhaust the type of threats to consider, and forensics tools that can identify threats related to the societal harms introduced by XR [104]. Establishing privacy in this manner places the onus on the XR platform to implement and maintain privacy and does not require the users or developers to become privacy experts.

In the scenario where XR platforms are trustworthy in handling data, the solutions that vet third-party services can protect privacy from the sensor to the application. Our Gatekeeper fits this privacy model and provides an efficient means for privacy-by-design specific to typical applications of eye-tracking data. When the Gatekeeper does not enable a specific application,

privacy-enhanced gaze samples can be streamed using noise mechanisms implemented by the platform. Data-level privacy mechanisms introduce another issue for developers in understanding the impact of privacy mechanisms on their specific applications. A solution for developers must be context specific to understand what mechanism should be applied and the range of parameters that retain practical utility.

In a perfect world, the most private XR platform would know exactly what data sensors an application needs and what it intends to do with the data. Contextual integrity is a theoretical privacy framework applicable to such a scenario, as it relies on data flows that respect societal norms for data sharing [184, 26]. However, apps cannot always be trusted to be honest about data use. In response, contextual integrity was implemented within internet-of-things devices by detecting malicious activity using a context-aware permissions system [119]. By analyzing app behavior, the system triggers additional permissions requests when actions are performed in a new context, such as a smart device requesting to open a window of the user's house for temperature control in the middle of the night instead of during the daytime as previously permitted by the user. By asking the user before opening the window, an intrusion resulting from malicious commands sent to the device can be prevented. Additional work in platform design and technical understanding of identifying data misuse is needed to achieve privacy when streaming data with future XR devices.

5.3 Privacy for Eye-Tracking Datasets

Eye-tracking datasets risk privacy by enabling re-identification attacks on eye movement biometrics. Public datasets of eye-tracking data from common XR tasks enable re-identification at reasonable success rates, demonstrating the need to de-identify data. Our privacy mechanisms tap into established privacy definitions of k-anonymity and plausible deniability. Our mechanisms enable favorable privacy-utility trade-offs when applying k-anonymity to classification models trained on eye-tracking features and gaze prediction models from eye-tracking samples. Our mechanisms and analysis inform data owners about the impact of de-identifying eye-tracking datasets before release.

Our findings suggest that there was no one-size-fits-all privacy solution for de-identifying the eye-tracking datasets included in our evaluation. While kalɛido-DP for gaze samples lowered identification rates to chance, it did not produce practical utility for training gaze prediction models. A DP guarantee did not retain utility for classification tasks on feature datasets. However, it retained utility for sample datasets on the same task due to differences in how data is processed to achieve the guarantee and the inclusion of head tracking data in the sample dataset. Within applications of gaze prediction, we recommend *k*-anonymity for sample data achieved by *k*-same-synth guarantee but do not recommend *k*-same-synth for classification tasks with sample data. Based on these results, our recommendation is that the data owner must consider their application and what is an acceptable loss in utility to determine which mechanism and parameters to apply.

While formal guarantees against re-identification are future-proofed against attacks within the bounds of our threat model, new attacks in the future can still introduce an unexpected risk to privacy. For example, in the area of systems security, there was an attack known as Spectre that affected a vast majority of processors in 2017 [7]. The attack targeted the functionality of speculative execution, which was released approximately a decade before the attack came to fruition. An analogy in the context of XR datasets is determining that the head movement data released within the dataset also leaked identity through speech reconstruction [225], rendering a formal guarantee applied to gaze data alone ineffective. Such a risk is not known at the time of dataset release and is difficult to anticipate. Unanticipated risks are a fundamental flaw of the assumptions in threat modeling. While a robust threat model defines a clear attack surface and data inputs to evaluate defense mechanisms, there is a natural trade-off between what researchers can anticipate with existing knowledge and the threats most likely to occur at the time.

A takeaway from our work is considering the level of data utility needed when dealing with sensitive eye-tracking datasets and considering the appropriate definition of privacy to account for future threats. Researchers concerned with re-identification can rely on *k*-anonymity to defend their eye-tracking features against attacks. At the same time, an XR web browser can deploy

differential privacy with kalεido to protect user gaze data against a broader range of attacks that consider where a user is looking. Notably, the current standard today is releasing or capturing raw data. Instead, we recommend that data owners preemptively defend against attacks by releasing de-identified data with formal privacy mechanisms. Datasets for machine-learning tasks can report model performance on unmodified and de-identified data to establish the privacy-utility trade-off and release the de-identified data publicly. Researchers can use the public data for developing new models that are later submitted for evaluation on a withheld test dataset of unmodified data, a standard process for dataset challenges.[1] In cases where unmodified datasets must be shared, such as XR companies sharing large-scale user data with contractors for data analysis or model development, privacy agreements between trusted parties can protect user privacy from unknown adversaries. Our mechanisms contribute to the set of privacy tools for data owners to protect eye-tracking data. Our analysis informs the need to protect XR data preemptively before such attacks are publicized and generate harms that impact real users.

5.4 Future Directions

A next step in developing better privacy mechanisms in the short term is to explore further generative models that achieve the definition of PD for datasets. The Marginals model for generating synthetic eye-tracking features is a baseline method for discrete datasets [34], with the assumption that all feature columns in the dataset are independent of each other. The assumption of independence influences the probability distributions sampled to generate new data, resulting in output distributions that are not similar to the real data and impact data utility. Linking novel generative models for continuous features and sample data to PD will provide the opportunity to achieve better trade-offs between privacy and utility for synthetic data.

The threat models we had considered relate to identifying the user from their data stream. Attention data paired with content has the potential to violate privacy expectations concerning personalized ads, revealing biases, and identifying sexual orientation [168]. Inferences of this nature are known as biometric pyschography, referring to data that can indicate the emotional

[1] https://salient360.ls2n.fr/

state or intention of a user [109]. A recent report from Common Sense Media discussed the implications of privacy specific to children in the Metaverse highlighting that typical users do not know what is being shared when they enable XR data streams [202]. Children are at risk of over-sharing and are more susceptible to targeted advertising. Even de-identified data could leak mental health conditions [18] or neurological diseases like Autism [37, 48]. While current privacy solutions have only explored the additional threat of gender classification with DP methods [232, 41, 84], there is still a gap between the large body of work on using eye movements for individual classifications like medical conditions or emotional state in ideal lab conditions [20, 105, 229] and how frequently the scenarios that produce these risks would arise in everyday use of XR. Eliminating a privacy risk completely is difficult considering eye-tracking data are produced from different types of sensors with variations in data quality resulting from calibration, and a large diversity in demographics of users and environments. Future work must expand the understanding of which contexts introduce potential harms when sharing gaze data and prioritize privacy while developing the next generation of mixed-reality technologies.

Beyond eye-tracking data, future devices introduce risks to privacy from other biometric sensors, such as heart rate that can reveal emotion or brain-sensing EEG signals. Extending data processing methods to additional data streams is essential, as XR interfaces and platforms will integrate multi-modal sensor fusion [31]. Data risks grow when more data sources are introduced and can be cross-analyzed. Researchers have already demonstrated that VR devices share these data streams with third-party apps while violating privacy policies [241]. Privacy mechanisms or formal guarantees only applied to one sensor may not achieve satisfactory levels of privacy. Enforcing standalone privacy mechanisms for each sensor and extending a more comprehensive Gatekeeper interface for sharing data after sensor fusion could allow developers to develop novel interactions and experiences while limiting the impact on user privacy.

APPENDIX A
PROOF OF SUFFICIENT CONDITION FOR PD

A.1 Theorem

For real $\gamma \geq 1$, if

$$\gamma^{-i-1} < Pr\{y = \mathbf{M}(d_i)\} \leq \gamma^{-i} \text{ and } \gamma^{-i-1} < Pr\{y = \mathbf{M}(d_j)\} \leq \gamma^{-i}$$

are true for the only integer $i > 0$, then

$$\gamma^{-1} \leq \frac{Pr\{y = \mathbf{M}(d_i)\}}{Pr\{y = \mathbf{M}(d_j)\}} < \gamma,$$

which satisfies (k,γ)-PD.

A.2 Proof

Assume that

$$\gamma^{-i-1} < Pr\{y = \mathbf{M}(d_i)\} \leq \gamma^{-i} \text{ and } \gamma^{-i-1} < Pr\{y = \mathbf{M}(d_j)\} \leq \gamma^{-i}$$

for the only integer $i > 0$. Starting with

$$\gamma^{-i-1} < Pr\{y = \mathbf{M}(d_i)\} \leq \gamma^{-i},$$

divide all terms by $Pr\{y = \mathbf{M}(d_j)\}$ to get

$$\frac{\gamma^{-i-1}}{Pr\{y = \mathbf{M}(d_j)\}} < \frac{Pr\{y = \mathbf{M}(d_i)\}}{Pr\{y = \mathbf{M}(d_j)\}} \leq \frac{\gamma^{-i}}{Pr\{y = \mathbf{M}(d_j)\}}.$$

Because $Pr\{y = \mathbf{M}(d_j)\} \leq \gamma^{-i}$, we have that

$$\frac{\gamma^{-i-1}}{Pr\{y = \mathbf{M}(d_j)\}} \geq \frac{\gamma^{-i-1}}{\gamma^{-i}} = \gamma^{-1}.$$

Because $Pr\{y = \mathbf{M}(d_j)\} > \gamma^{-i-1}$, we have that

$$\frac{\gamma^{-i}}{Pr\{y = \mathbf{M}(d_j)\}} < \frac{\gamma^{-i}}{\gamma^{-i-1}} = \gamma$$

therefore,

$$\gamma^{-1} \leq \frac{\gamma^{-i-1}}{Pr\{y = \mathbf{M}(d_j)\}} < \frac{Pr\{y = \mathbf{M}(d_i)\}}{Pr\{y = \mathbf{M}(d_j)\}} < \frac{\gamma^{-i}}{Pr\{y = \mathbf{M}(d_j)\}} < \gamma$$

which satisfies

$$\gamma^{-1} \leq \frac{Pr\{y = \mathbf{M}(d_i)\}}{Pr\{y = \mathbf{M}(d_j)\}} < \gamma. \qquad \square$$

APPENDIX B
SECTION 4.3.2 THREAT SCENARIO *K*-ANONYMITY DETAILS

Age and gender demographics are generalized by grouping values into ranges to achieve *k*-anonymity. The number of data rows for each unique combination of age and gender ranges must be *k* or greater to maintain the privacy guarantee. The combined dataset of ET-DK2 and 360_em consists of 24 individuals with age and gender values listed in Table B-1.

Table B-1. Age and Gender demographics for ET-DK2 and 360_em datasets. Note that Subject ID 1 from both datasets were excluded from analysis due to data loss and subject sickness during data collection, respectively.

Dataset	Subject ID	Age	Gender
ET-DK2	2	M	43
ET-DK2	3	F	27
ET-DK2	4	M	29
ET-DK2	5	M	32
ET-DK2	6	F	28
ET-DK2	8	M	26
ET-DK2	9	F	23
ET-DK2	10	M	30
ET-DK2	11	F	28
ET-DK2	12	M	26
ET-DK2	13	M	52
ET-DK2	14	M	26
ET-DK2	15	M	35
ET-DK2	16	M	50
ET-DK2	17	M	33
ET-DK2	18	M	31
ET-DK2	19	M	32
ET-DK2	20	M	36
360_em	2	M	38
360_em	3	M	29
360_em	4	F	23
360_em	5	F	31
360_em	6	M	27
360_em	7	M	31
360_em	8	F	23
360_em	9	M	24
360_em	10	M	23
360_em	11	M	27
360_em	12	M	23
360_em	13	M	23
360_em	14	M	32

Ranges were selected for each value of k that maximized the total number of groups while ensuring each group had at least k rows matching the ranges of age and gender. The ranges of age and gender used to establish k-anonymity are listed in Table B-2.

Table B-2. Gender and age ranges used to generalize the ET-DK2 and 360_em demographics for k-anonymity. For each value of k the data rows are mapped into the listed ranges based on actual values. For example, (Male, 23-31) would be assigned to all Males between the age of 23 and 31. Male/Female refers to the data rows not specifying either value for Gender.

k	Gender & Age Generalization
4	(Female, 23-31), (Male, 23-27), (Male, 29-31), (Male, 32-33), (Male, 35-52)
6	(Female, 23-31), (Male, 23-27), (Male, 29-33), (Male, 35-52)
8	(Male/Female, 23-27), (Male/Female, 28-31), (Male/Female, 32-52)
15	(Male/Female, 23-28), (Male/Female, 29-52)

APPENDIX C
C-VAE MODEL TRAINING PROCEDURE

The C-VAE model for generating synthetic saccade profiles (Sec. 4.5.2.3) was trained using tensorflow version 1.13.1. Models were trained independently for each dataset using data from all individuals and stimuli. Training was performed using 75% of the available data with the remaining 25% used as a validation set. An experiment to optimize model hyper-parameters is described in Appendix D.

All models were trained with an ADAM optimizer [138] using tensorflow's Model compile and fit functions. The loss function was defined as

$$\mathbf{L}(x,\mathbf{D}(z)) = ||x-\mathbf{D}(z)||_2 - \mathbf{KL}(\mathbf{N}(\mu,\sigma),\mathbf{N}(0,1)),$$

where the first term is Mean Squared Error for the reconstructed synthetic profile and the second terms employs KL Divergence to enforce latent space sampling that follows a normal distribution with zero mean.

APPENDIX D
C-VAE MODEL HYPER-PARAMETER OPTIMIZATION

Hyper-parameters were tuned using the EHTask dataset as it contained a longer duration of data compared to the DGaze dataset. Grid search optimization was performed over the following sets of values, with optimal parameters in bold:

- Learning Rate: 0.001, **0.01**
- Batch Size: 20, 60, **100**
- Number of Epochs: 10, **20**, 30
- Encoder Hidden Layer with ReLU activation function: **32**, 64, 96 Nodes
- Latent Space Dimension: 32, **64**, 96
- Decoder Hidden Layer with linear activation function: 32, 64, **96** Nodes

The optimal parameters produced an average loss of 0.33 on the validation set.

REFERENCES

[1] *Biometrics for education*, https://www.iritech.com/biometric-education-Kenya, Accessed: 2019-08-22.

[2] *Keeping an eye on security*, https://www.schiphol.nl/en/page/how-the-iris-scan-works/, Accessed: 2019-08-22.

[3] *Louisiana hospital adopts iris-based patient id system*, https://findbiometrics.com/la-hospital-adopts-iris-based-patient-id-system-26203/, Accessed: 2019-08-22.

[4] *Microsoft unveils hololens 2: Twice the field of view, eye tracking*, https://arstechnica.com/gadgets/2019/02/microsoft-unveils-hololens-2-twice-the-field-of-view-eye-tracking/, Accessed: 2019-08-29.

[5] *Somaliland election saw iris id technology deployed*, https://www.biometricupdate.com/201801/somaliland-election-saw-iris-id-technology-deployed, Accessed: 2019-08-22.

[6] Danielle Abril, *Future of work: Smart glasses, holograms and AI-equipped robots will change our jobs*, https://www.washingtonpost.com/technology/2022/03/17/qualcomm-cristiano-amon-future-of-work/, 2022, Accessed: 2022-05-23.

[7] Nael Abu-Ghazaleh, Dmitry Ponomarev, and Dmitry Evtyushkin, *How the spectre and meltdown hacks really worked*, IEEE Spectrum **56** (2019), no. 3, 42–49.

[8] Devon Adams, Alseny Bah, Catherine Barwulor, Nureli Musaby, Kadeem Pitkin, and Elissa M Redmiles, *Ethics emerging: the story of privacy and security perceptions in virtual reality*, Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018), 2018, pp. 427–442.

[9] Ioannis Agtzidis, Mikhail Startsev, and Michael Dorr, *360-degree video gaze behaviour: A ground-truth data set and a classification algorithm for eye movements*, Proceedings of the 27th ACM International Conference on Multimedia, 2019, pp. 1007–1015.

[10] _____, *A ground-truth data set and a classification algorithm for eye movements in 360-degree videos*, arXiv preprint arXiv:1903.06474 (2019).

[11] Tousif Ahmed, Patrick Shaffer, Kay Connelly, David Crandall, and Apu Kapadia, *Addressing physical safety, security, and privacy for people with visual impairments*, Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016), 2016, pp. 341–354.

[12] Amgad AA Algawhari and Yongfeng Huang, *Iris recognition under unconstrained conditions*, 2018 International Conference on Image and Video Processing, and Artificial Intelligence, vol. 10836, International Society for Optics and Photonics, 2018, p. 108360D.

[13] Rawan Alghofaili, Michael S Solah, and Haikun Huang, *Optimizing visual element placement via visual attention analysis*, 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, 2019, pp. 464–473.

[14] Noura Alomar and Serge Egelman, *Developers say the darnedest things: Privacy compliance processes followed by developers of child-directed apps*, Proceedings on Privacy Enhancing Technologies **4** (2022), 250–273.

[15] Miguel E Andrés, Nicolás E Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi, *Geo-indistinguishability: Differential privacy for location-based systems*, Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security, 2013, pp. 901–914.

[16] Elena Arabadzhiyska, Okan Tarhan Tursun, Karol Myszkowski, Hans-Peter Seidel, and Piotr Didyk, *Saccade landing position prediction for gaze-contingent rendering*, ACM Transactions on Graphics (TOG) **36** (2017), no. 4, 1–12.

[17] William A Arbaugh, David J Farber, and Jonathan M Smith, *A secure and reliable bootstrap architecture*, Proceedings of the 1997 IEEE Symposium on Security and Privacy, 1997, pp. 65–71.

[18] Thomas Armstrong and Bunmi O Olatunji, *Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis*, Clinical psychology review **32** (2012), no. 8, 704–723.

[19] Marc Assens, Xavier Giro-i Nieto, Kevin McGuinness, and Noel E O'Connor, *Pathgan: visual scanpath prediction with generative adversarial networks*, Proceedings of the European Conference on Computer Vision (ECCV), 2018.

[20] Ofer Avital, *Method and system of using eye tracking to evaluate subjects*, October 8 2015, US Patent App. 14/681,083.

[21] Mihai Bâce, Vincent Becker, Chenyang Wang, and Andreas Bulling, *Combining gaze estimation and optical flow for pursuits interaction*, ACM Symposium on Eye Tracking Research and Applications, 2020, pp. 1–10.

[22] A Terry Bahill and Lawrence Stark, *The trajectories of saccadic eye movements*, Scientific American **240** (1979), no. 1, 108–117.

[23] Reynold Bailey, Ann McNamara, Aaron Costello, Srinivas Sridharan, and Cindy Grimm, *Impact of subtle gaze direction on short-term spatial information recall*, Proceedings of the Symposium on Eye Tracking Research and Applications, 2012, pp. 67–74.

[24] Reynold Bailey, Ann McNamara, Nisha Sudarsanam, and Cindy Grimm, *Subtle gaze direction*, ACM Transactions on Graphics (TOG) **28** (2009), no. 4, 1–14.

[25] Yair Bar-Haim, Talee Ziv, Dominique Lamy, and Richard M Hodes, *Nature and nurture in own-race face processing*, Psychological science **17** (2006), no. 2, 159–163.

[26] Adam Barth, Anupam Datta, John C Mitchell, and Helen Nissenbaum, *Privacy and contextual integrity: Framework and applications*, 2006 IEEE symposium on security and privacy (S&P'06), IEEE, 2006, pp. 15–pp.

[27] Behnam Bastani, Eric Turner, Carlin Vieri, Haomiao Jiang, Brian Funt, and Nikhil Balram, *Foveated pipeline for AR/VR head-mounted displays*, Information Display **33** (2017), no. 6, 14–35.

[28] Roman Bednarik, Hana Vrzakova, and Michal Hradis, *What do you want to do next: a novel approach for intent prediction in gaze-based interaction*, Proceedings of the symposium on eye tracking research and applications, ACM, 2012, pp. 83–90.

[29] Justin K Bennett, Srinivas Sridharan, Brendan John, and Reynold Bailey, *Looking at faces: autonomous perspective invariant facial gaze analysis*, Proceedings of the ACM Symposium on Applied Perception, ACM, 2016, pp. 105–112.

[30] Shlomo Berkovsky, Ronnie Taib, Irena Koprinska, Eileen Wang, Yucheng Zeng, Jingjie Li, and Sabina Kleitman, *Detecting personality traits using eye-tracking data*, Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–12.

[31] Guillermo Bernal, *Developing Galea: An open source tool at the intersection of VR and neuroscience*, https://www.media.mit.edu/posts/galea/, 2021, Accessed: 2022-06-06.

[32] Guillermo Bernal, Nelson Hidalgo, Conor Russomanno, and Pattie Maes, *Galea: A physiological sensing system for behavioral research in virtual environments*, 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, 2022, pp. 66–76.

[33] Vincent Bindschaedler and Reza Shokri, *Synthesizing plausible privacy-preserving location traces*, 2016 IEEE Symposium on Security and Privacy (SP), IEEE, 2016, pp. 546–563.

[34] Vincent Bindschaedler, Reza Shokri, and Carl A Gunter, *Plausible deniability for privacy-preserving data synthesis*, Proceedings of the VLDB Endowment **10** (2017), no. 5.

[35] Benjamin Bolte and Markus Lappe, *Subliminal reorientation and repositioning in immersive virtual environments using saccadic suppression*, IEEE transactions on visualization and computer graphics **21** (2015), no. 4, 545–552.

[36] Thomas Booth, Srinivas Sridharan, Ann McNamara, Cindy Grimm, and Reynold Bailey, *Guiding attention in controlled real-world environments*, Proceedings of the ACM Symposium on Applied Perception, 2013, pp. 75–82.

[37] Zillah Boraston and Sarah-Jayne Blakemore, *The application of eye-tracking technology in the study of autism*, The Journal of physiology **581** (2007), no. 3, 893–898.

[38] Ali Borji, *Saliency prediction in the deep learning era: Successes and limitations*, IEEE TPAMI (2019).

[39] Andrew Bosworth and Nick Clegg, *Building the Metaverse Responsibly*, https://about.fb.com/news/2021/09/building-the-metaverse-responsibly/, 2022, Accessed: 2022-07-16.

[40] Efe Bozkir, Onur Günlü, Wolfgang Fuhl, Rafael F Schaefer, and Enkelejda Kasneci, *Differential privacy for eye tracking with temporal correlations*, arXiv preprint arXiv:2002.08972 (2020).

[41] _____, *Differential privacy for eye tracking with temporal correlations*, Plos one **16** (2021), no. 8, e0255979.

[42] Efe Bozkir, Ali Burak Ünal, Mete Akgün, Enkelejda Kasneci, and Nico Pfeifer, *Privacy preserving gaze estimation using synthetic images via a randomized encoding based framework*, ACM Symposium on Eye Tracking Research and Applications, 2020, pp. 1–5.

[43] Jessica Bradshaw, Frederick Shic, Anahita N Holden, Erin J Horowitz, Amy C Barrett, Tamsin C German, and Ty W Vernon, *The use of eye tracking as a biomarker of treatment outcome in a pilot randomized clinical trial for young children with autism*, Autism Research **12** (2019), no. 5, 779–793.

[44] Andreas Bulling, Jamie A Ward, Hans Gellersen, and Gerhard Tröster, *Eye movement analysis for activity recognition using electrooculography*, IEEE transactions on pattern analysis and machine intelligence **33** (2010), no. 4, 741–753.

[45] Alisa Burova, John Mäkelä, Jaakko Hakulinen, Tuuli Keskinen, Hanna Heinonen, Sanni Siltanen, and Markku Turunen, *Utilizing vr and gaze tracking to develop ar solutions for industrial maintenance*, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–13.

[46] Zoya Bylinskii, Tilke Judd, Aude Oliva, Antonio Torralba, and Frédo Durand, *What do different evaluation metrics tell us about saliency models?*, IEEE transactions on pattern analysis and machine intelligence **41** (2018), no. 3, 740–757.

[47] Aayush Kumar Chaudhary and Jeff B Pelz, *Privacy-preserving eye videos using rubber sheet model*, ACM Symposium on Eye Tracking Research and Applications, 2020, pp. 1–5.

[48] Katarzyna Chawarska and Frederick Shic, *Looking but not seeing: Atypical visual scanning and recognition of faces in 2 and 4-year-old children with autism spectrum disorder*, Journal of autism and developmental disorders **39** (2009), no. 12, 1663.

[49] Dongwen Chen, Chunmei Qing, Xiangmin Xu, and Huansheng Zhu, *Salbinet360: Saliency prediction on 360° images with local-global bifurcated deep network*, 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, 2020, pp. 92–100.

[50] Min Chen, Zhikun Zhang, Tianhao Wang, Michael Backes, Mathias Humbert, and Yang Zhang, *When machine unlearning jeopardizes privacy*, Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security, 2021, pp. 896–911.

[51] Shaoyu Chen, Budmonde Duinkharjav, Xin Sun, Li-Yi Wei, Stefano Petrangeli, Jose Echevarria, Claudio Silva, and Qi Sun, *Instant reality: Gaze-contingent perceptual optimization for 3d virtual reality streaming*, IEEE Transactions on Visualization and Computer Graphics **28** (2022), no. 5, 2157–2167.

[52] Mauro Conti, Nicola Dragoni, and Viktor Lesyk, *A survey of man in the middle attacks*, IEEE Communications Surveys & Tutorials **18** (2016), no. 3, 2027–2051.

[53] Andrei Costin, *Security of cctv and video surveillance systems: Threats, vulnerabilities, attacks, and mitigations*, Proceedings of the 6th international workshop on trustworthy embedded devices, 2016, pp. 45–54.

[54] Kim M Dalton, Brendon M Nacewicz, Tom Johnstone, Hillary S Schaefer, Morton Ann Gernsbacher, Hill H Goldsmith, Andrew L Alexander, and Richard J Davidson, *Gaze fixation and the neural circuitry of face processing in autism*, Nature neuroscience **8** (2005), no. 4, 519–526.

[55] Amit Datta, Michael Carl Tschantz, and Anupam Datta, *Automated experiments on ad privacy settings*, Proceedings on privacy enhancing technologies **2015**, no. 1, 92–112.

[56] John Daugman, *New methods in iris recognition*, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) **37** (2007), no. 5, 1167–1175.

[57] John Daugman, *Chapter 25 - how iris recognition works*, The Essential Guide to Image Processing (Al Bovik, ed.), Academic Press, Boston, 2009, pp. 715–739.

[58] John G Daugman, *High confidence visual recognition of persons by a test of statistical independence*, IEEE transactions on pattern analysis and machine intelligence **15** (1993), no. 11, 1148–1161.

[59] Erwan J David, Jesús Gutiérrez, Antoine Coutrot, Matthieu Perreira Da Silva, and Patrick Le Callet, *A dataset of head and eye movements for 360 videos*, Proceedings of the 9th ACM Multimedia Systems Conference, ACM, 2018, pp. 432–437.

[60] Brendan David-John, Kevin Butler, and Eakta Jain, *For your eyes only: Privacy-preserving eye-tracking datasets*, ACM Symposium on Eye Tracking Research and Applications, 2022, pp. 1–6.

[61] Brendan David-John, Diane Hosfelt, Kevin Butler, and Eakta Jain, *Let's soup up xr: Collected thoughts from an ieee vr workshop on privacy in mixed reality*, VR4Sec: 1st International Workshop on Security for XR and XR for Security, 2021, pp. 1–6.

[62] ———, *A privacy-preserving approach to streaming eye-tracking data*, IEEE Transactions on Visualization and Computer Graphics (2021).

[63] Brendan David-John, Candace Peacock, Ting Zhang, T Scott Murdison, Hrvoje Benko, and Tanya R Jonker, *Towards gaze-based prediction of the intent to interact in virtual reality*, ACM Symposium on Eye Tracking Research and Applications, 2021, pp. 1–7.

[64] Jaybie A De Guzman, Kanchana Thilakarathna, and Aruna Seneviratne, *Security and privacy approaches in mixed reality: A literature survey*, ACM Computing Surveys (CSUR) **52** (2019), no. 6, 1–37.

[65] Mina Deng, Kim Wuyts, Riccardo Scandariato, Bart Preneel, and Wouter Joosen, *A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements*, Requirements Engineering **16** (2011), no. 1, 3–32.

[66] Tamara Denning, Zakariya Dehlawi, and Tadayoshi Kohno, *In situ with bystanders of augmented reality glasses: Perspectives on recording and privacy-mediating technologies*, Proceedings of the 32nd annual ACM conference on Human factors in computing systems, ACM, 2014, pp. 2377–2386.

[67] Yujie Dong and Damon L Woodard, *Eyebrow shape-based features for biometric recognition and gender classification: A feasibility study*, 2011 International Joint Conference on Biometrics (IJCB), IEEE, 2011, pp. 1–8.

[68] Yuzhu Dong, Sophie Jörg, and Eakta Jain, *Is the avatar scared? pupil as a perceptual cue*, Computer Animation and Virtual Worlds **33** (2022), no. 2, e2040.

[69] Andrew T Duchowski, Vinay Shivashankaraiah, Tim Rawls, Anand K Gramopadhye, Brian J Melloy, and Barbara Kanki, *Binocular eye tracking in virtual reality for inspection training*, Proceedings of the 2000 symposium on Eye tracking research & applications, 2000, pp. 89–96.

[70] Cynthia Dwork, *Differential privacy*, Automata, Languages and Programming, Springer Berlin Heidelberg, 2006, pp. 1–12.

[71] Cynthia Dwork, *Differential privacy: A survey of results*, International conference on theory and applications of models of computation, Springer, 2008, pp. 1–19.

[72] Cynthia Dwork, Aaron Roth, et al., *The algorithmic foundations of differential privacy.*, Found. Trends Theor. Comput. Sci. **9** (2014), no. 3-4, 211–407.

[73] Simon Eberz, Giulio Lovisotto, Kasper B Rasmussen, Vincent Lenders, and Ivan Martinovic, *28 blinks later: Tackling practical challenges of eye movement biometrics*, Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, 2019, pp. 1187–1199.

[74] Khaled El Emam and Fida Kamal Dankar, *Protecting privacy using k-anonymity*, Journal of the American Medical Informatics Association **15** (2008), no. 5, 627–637.

[75] Khaled El Emam, Elizabeth Jonker, Luk Arbuckle, and Bradley Malin, *A systematic review of re-identification attacks on health data*, PloS one **6** (2011), no. 12, e28071.

[76] Kara J Emery, Marina Zannoli, James Warren, Lei Xiao, and Sachin S Talathi, *Openneeds: A dataset of gaze, head, hand, and scene signals during exploration in open-ended vr environments*, ACM Symposium on Eye Tracking Research and Applications, 2021, pp. 1–7.

[77] Anton Mølbjerg Eskildsen and Dan Witzner Hansen, *Analysis of iris obfuscation: Generalising eye information processes for privacy studies in eye tracking.*, ACM Symposium on Eye Tracking Research and Applications, 2021, pp. 1–10.

[78] Anna Maria Feit, Shane Williams, Arturo Toledo, Ann Paradiso, Harish Kulkarni, Shaun Kane, and Meredith Ringel Morris, *Toward everyday gaze input: Accuracy and precision of eye tracking and implications for design*, Proceedings of the 2017 Chi conference on human factors in computing systems, 2017, pp. 1118–1130.

[79] Jamie Feltham, *Project cambria: Everything we know about meta's next headset*, UploadVR, https://uploadvr.com/project-cambria-everything-we-know/ (Jan. 26, 2022).

[80] Xianglong Feng, Viswanathan Swaminathan, and Sheng Wei, *Viewport prediction for live 360-degree mobile video streaming using user-content hybrid motion tracking*, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **3** (2019), no. 2, 43.

[81] Darrell G Ford and Marty Ludlum, *Employee privacy outside the workplace*, Southern Law Journal **26** (2016), no. 2, 321.

[82] Yu Fu, Yan Hu, and Veronica Sundstedt, *A systematic literature review of virtual, augmented, and mixed reality game applications in healthcare*, ACM Transactions on Computing for Healthcare (HEALTH) **3** (2022), no. 2, 1–27.

[83] Wolfgang Fuhl, Efe Bozkir, and Enkelejda Kasneci, *Reinforcement learning for the privacy preservation and manipulation of eye tracking data*, arXiv preprint arXiv:2002.06806 (2020).

[84] ———, *Reinforcement learning for the privacy preservation and manipulation of eye tracking data*, International Conference on Artificial Neural Networks, Springer, 2021, pp. 595–607.

[85] Wolfgang Fuhl, Gjergji Kasneci, and Enkelejda Kasneci, *Teyed: Over 20 million real-world eye images with pupil, eyelid, and iris 2d and 3d segmentations, 2d and 3d landmarks, 3d eyeball, gaze vector, and eye movement types*, 2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), IEEE, 2021, pp. 367–375.

[86] Chiara Galdi, Michele Nappi, Daniel Riccio, and Harry Wechsler, *Eye movement analysis for human authentication: a critical survey*, Pattern Recognition Letters **84** (2016), 272–283.

[87] Aparna G Gale and SS Salankar, *A review on advance methods of feature extraction in iris recognition system*, IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE) e-ISSN (2014), 2278–1676.

[88] Abhishek Gangwar, Akanksha Joshi, Ashutosh Singh, Fernando Alonso-Fernandez, and Josef Bigun, *Irisseg: A fast and robust iris segmentation framework for non-ideal iris images*, ICB 2016, IEEE, 2016, pp. 1–8.

[89] Maia Garau, Mel Slater, Simon Bee, and Martina Angela Sasse, *The impact of eye gaze on communication using humanoid avatars*, Proceedings of the SIGCHI conference on Human factors in computing systems, ACM, 2001, pp. 309–316.

[90] Stephan J Garbin, Yiru Shen, Immo Schuetz, Robert Cavin, Gregory Hughes, and Sachin S Talathi, *Openeds: Open eye dataset*, arXiv preprint arXiv:1905.03702 (2019).

[91] Charlotte Garden, *Labor organizing in the age of surveillance*, Louis ULJ **63** (2018), 55.

[92] Christoph Gebhardt, Brian Hecox, Bas van Opheusden, Daniel Wigdor, James Hillis, Otmar Hilliges, and Hrvoje Benko, *Learning cooperative personalized policies from gaze data*, Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, 2019, pp. 197–208.

[93] Anjith George and Aurobinda Routray, *A score level fusion method for eye movement biometrics*, Pattern Recognition Letters **82** (2016), 207–215.

[94] Reiko Graham, Alison Hoover, Natalie A Ceballos, and Oleg Komogortsev, *Body mass index moderates gaze orienting biases and pupil diameter to high and low calorie food images*, Appetite **56** (2011), no. 3, 577–586.

[95] Henry Griffith, Samantha Aziz, and Oleg Komogortsev, *Prediction of oblique saccade trajectories using learned velocity profile parameter mappings*, 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), IEEE, 2020, pp. 0018–0024.

[96] Steve Grogorick, Georgia Albuquerque, and Marcus Magnor, *Gaze guidance in immersive environments*, 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), March 2018, pp. 563–564.

[97] Steve Grogorick, Michael Stengel, Elmar Eisemann, and Marcus Magnor, *Subtle gaze guidance for immersive environments*, Proceedings of the ACM Symposium on Applied Perception, 2017, pp. 1–7.

[98] Ralph Gross, Edoardo Airoldi, Bradley Malin, and Latanya Sweeney, *Integrating utility into face de-identification*, International Workshop on Privacy Enhancing Technologies, Springer, 2005, pp. 227–242.

[99] Brian Guenter, Mark Finch, Steven Drucker, Desney Tan, and John Snyder, *Foveated 3d graphics*, ACM Transactions on Graphics (TOG) **31** (2012), no. 6, 164.

[100] Richard Guest, *Information technology–biometric data interchange formats–19794-part 7: Signature/sign time series data*, (2014).

[101] Simon NB Gunkel, Rick Hindriks, Karim M El Assal, Hans M Stokking, Sylvie Dijkstra-Soudarissanane, Frank ter Haar, and Omar Niamut, *Vrcomm: an end-to-end web system for real-time photorealistic social vr communication*, Proceedings of the 12th ACM Multimedia Systems Conference, 2021, pp. 65–79.

[102] Daniel Hanley and Sally Hubbard, *Eyes everywhere: Amazon's surveillance infrastructure and revitalizing worker power*, Open Markets Institute (2020).

[103] Jassim Happa, Mashhuda Glencross, and Anthony Steed, *Cyber security threats and challenges in collaborative mixed-reality*, Frontiers in ICT **6** (2019), 5.

[104] Jassim Happa, Anthony Steed, and Mashhuda Glencross, *Privacy-certification standards for extended-reality devices and services*, 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE, 2021, pp. 397–398.

[105] Katarzyna Harezlak and Pawel Kasprowski, *Application of eye tracking in medicine: A survey, research issues and challenges*, Computerized Medical Imaging and Graphics **65** (2018), 176–190.

[106] Rakibul Hasan, Eman Hassan, Yifang Li, Kelly Caine, David J Crandall, Roberto Hoyle, and Apu Kapadia, *Viewer experience of obscuring scene elements in photos to enhance privacy*, Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, ACM, 2018, p. 47.

[107] Scott Hayden, *'Rec Room' Studio Estimates Facebook Sold 2-3M Oculus Quest 2 Headsets in Q4*, https://www.roadtovr.com/facebook-rec-room-2-3m-oculus-quest-2-sold/, 2021, Accessed: 2021-05-26.

[108] Hamish Hector, *Oculus Quest 2's amazing 2021 sales figures spell doom for PSVR 2*, https://www.techradar.com/news/meta-quest-2s-2021-success-could-spell-doom-for-psvr-2, 2022, Accessed: 2022-05-23.

[109] Brittan Heller, *Watching androids dream of electric sheep: Immersive technology, biometric psychography, and the law*, Vanderbilt Journal of Entertainment & Technology Law **23** (2020), no. 1, 1.

[110] Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, and Joost Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*, OUP Oxford, 2011.

[111] Roberto Hoyle, Robert Templeman, Steven Armes, Denise Anthony, David Crandall, and Apu Kapadia, *Privacy behaviors of lifeloggers using wearable cameras*, Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ACM, 2014, pp. 571–582.

[112] Zhiming Hu, Andreas Bulling, Sheng Li, and Guoping Wang, *Ehtask: Recognizing user tasks from eye and head movements in immersive virtual reality*, IEEE Transactions on Visualization and Computer Graphics (2021).

[113] _____, *Fixationnet: Forecasting eye fixations in task-oriented virtual environments*, IEEE Transactions on Visualization and Computer Graphics **27** (2021), no. 5, 2681–2690.

[114] Zhiming Hu, Sheng Li, Congyi Zhang, Kangrui Yi, Guoping Wang, and Dinesh Manocha, *Dgaze: Cnn-based gaze prediction in dynamic scenes*, IEEE transactions on visualization and computer graphics **26** (2020), no. 5, 1902–1911.

[115] Zhiming Hu, Congyi Zhang, Sheng Li, Guoping Wang, and Dinesh Manocha, *SGaze: A data-driven eye-head coordination model for realtime gaze prediction*, IEEE Transactions on Visualization and Computer Graphics **25** (2019), no. 5, 2002–2010.

[116] James Hutson, *Social virtual reality: Neurodivergence and inclusivity in the metaverse*, Societies **12** (2022), no. 4, 102.

[117] Gazi Karam Illahi, Thomas Van Gemert, Matti Siekkinen, Enrico Masala, Antti Oulasvirta, and Antti Ylä-Jääski, *Cloud gaming with foveated video encoding*, ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) **16** (2020), no. 1, 1–24.

[118] Anil K Jain, Arun Ross, and Salil Prabhakar, *An introduction to biometric recognition*, IEEE Transactions on circuits and systems for video technology **14** (2004), no. 1, 4–20.

[119] Yunhan Jack Jia, Qi Alfred Chen, Shiqi Wang, Amir Rahmati, Earlence Fernandes, Zhuoqing Morley Mao, Atul Prakash, and SJ Unviersity, *Contexlot: Towards providing contextual integrity to appified iot platforms.*, NDSS, vol. 2, San Diego, 2017, pp. 2–2.

[120] Brendan John, Sophie Jörg, Sanjeev Koppal, and Eakta Jain, *The security-utility trade-off for iris authentication and eye animation for social virtual avatars.*, IEEE transactions on visualization and computer graphics (2020).

[121] Brendan John, Sriram Kalyanaraman, and Eakta Jain, *Look out! a design framework for safety training systems a case study on omnidirectional cinemagraphs*, 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE, 2020, pp. 147–153.

[122] Brendan John, Sanjeev Koppal, and Eakta Jain, *EyeVEIL: degrading iris authentication in eye tracking headsets*, Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications, ACM, 2019, p. 37.

[123] Brendan John, Ao Liu, Lirong Xia, Sanjeev Koppal, and Eakta Jain, *Let it snow: Adding pixel noise to protect the user's identity*, ACM Symposium on Eye Tracking Research and Applications, 2020, pp. 1–3.

[124] Brendan John, Pallavi Raiturkar, Olivier Le Meur, and Eakta Jain, *A benchmark of four methods for generating 360° saliency maps from eye tracking data*, International Journal of Semantic Computing **13** (2019), no. 03, 329–341.

[125] Brennan Jones, Yaying Zhang, Priscilla NY Wong, and Sean Rintel, *Belonging there: Vroom-ing into the uncanny valley of xr telepresence*, Proceedings of the ACM on Human-Computer Interaction **5** (2021), no. CSCW1, 1–31.

[126] Tanya R Jonker, Ruta Desai, Kevin Carlberg, James Hillis, Sean Keller, and Hrvoje Benko, *The role of ai in mixed and augmented reality interactions*, CHI2020 ai4hci Workshop Proceedings. ACM, 2020.

[127] AT Kahlil and FEM Abou-Chadi, *Generation of iris codes using 1d log-gabor filter*, The 2010 International Conference on Computer Engineering & Systems, IEEE, 2010, pp. 329–336.

[128] Alex Karpov, Jacob Liberman, Dillon Lohr, and Oleg Komogortsev, *Parallel oculomotor plant mathematical model for large scale eye movement simulation*, arXiv preprint arXiv:2007.09884 (2020).

[129] Pawel Kasprowski and Katarzyna Harężlak, *The second eye movements verification and identification competition*, IEEE International Joint Conference on Biometrics, IEEE, pp. 1–6.

[130] Paweł Kasprowski, Oleg V Komogortsev, and Alex Karpov, *First eye movement verification and identification competition at btas 2012*, 2012 IEEE fifth international conference on biometrics: theory, applications and systems (BTAS), IEEE, 2012, pp. 195–202.

[131] Moritz Kassner, William Patera, and Andreas Bulling, *Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction*, Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication, ACM, 2014, pp. 1151–1160.

[132] Sam Kavanagh, Andrew Luxton-Reilly, Burkhard Wuensche, and Beryl Plimmer, *A systematic review of virtual reality in education*, Themes in Science and Technology Education **10** (2017), no. 2, 85–119.

[133] Georgios Kellaris, Stavros Papadopoulos, Xiaokui Xiao, and Dimitris Papadias, *Differentially private event sequences over infinite streams*, Proceedings of the VLDB Endowment **7** (2014), no. 12, 1155–1166.

[134] Maryam Keyvanara and Robert Allison, *Transsaccadic awareness of scene transformations in a 3d virtual environment*, ACM Symposium on Applied Perception 2019, 2019, pp. 1–9.

[135] _____, *Effect of a constant camera rotation on the visibility of transsaccadic camera shifts*, Proceedings of the Symposium on Eye Tracking Research and Applications, 2020, pp. 1–8.

[136] Daniel Kifer and Ashwin Machanavajjhala, *No free lunch in data privacy*, Proceedings of the 2011 ACM SIGMOD International Conference on Management of data, 2011, pp. 193–204.

[137] Joohwan Kim, Michael Stengel, Alexander Majercik, Shalini De Mello, David Dunn, Samuli Laine, Morgan McGuire, and David Luebke, *Nvgaze: An anatomically-informed dataset for low-latency, near-eye gaze estimation*, Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–12.

[138] Diederik P Kingma and Jimmy Ba, *Adam: A method for stochastic optimization*, ICLR (Poster), 2015.

[139] Oleg Komogortsev, Corey Holland, Sampath Jayarathna, and Alex Karpov, *2d linear oculomotor plant mathematical model: Verification and biometric applications*, ACM Transactions on Applied Perception (TAP) **10** (2013), no. 4, 1–18.

[140] Oleg V Komogortsev and Alex Karpov, *Liveness detection via oculomotor plant characteristics: Attack of mechanical replicas*, 2013 international conference on biometrics (ICB), IEEE, 2013, pp. 1–8.

[141] Oleg V Komogortsev, Alexey Karpov, and Corey D Holland, *Attack of mechanical replicas: Liveness detection with eye movements*, IEEE Transactions on Information Forensics and Security **10** (2015), no. 4, 716–725.

[142] Jacob Leon Kröger, Otto Hans-Martin Lutz, and Florian Müller, *What does your gaze reveal about you? on the privacy implications of eye tracking*, IFIP International Summer School on Privacy and Identity Management, Springer, 2019, pp. 226–241.

[143] Thomas C Kübler, Wolfgang Fuhl, Elena Wagner, and Enkelejda Kasneci, *55 rides: attention annotated head and gaze data during naturalistic driving*, ACM Symposium on Eye Tracking Research and Applications, 2021, pp. 1–8.

[144] Dylan G Kwart, Tom Foulsham, and Alan Kingstone, *Age and beauty are in the eye of the beholder*, Perception **41** (2012), no. 8, 925–938.

[145] Guohao Lan, Tim Scargill, and Maria Gorlatova, *Eyesyn: Psychology-inspired eye movement synthesis for gaze-based activity recognition*, Proceedings of ACM/IEEE IPSN, 2022.

[146] Eike Langbehn, Frank Steinicke, Markus Lappe, Gregory F Welch, and Gerd Bruder, *In the blink of an eye: Leveraging blink-induced suppression for imperceptible position and orientation redirection in virtual reality*, ACM Transactions on Graphics (TOG) **37** (2018), no. 4, 66.

[147] Karina LaRubbio, Jeremiah Wright, Brendan David-John, Andreas Enqvist, and Eakta Jain, *Who do you look like? gaze-based authentication for workers in vr*, 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE, 2022.

[148] Olivier Le Meur and Thierry Baccino, *Methods for comparing scanpaths and saliency maps: strengths and weaknesses*, Behavior research methods **45** (2013), no. 1, 251–266.

[149] R John Leigh and David S Zee, *The neurology of eye movements*, Oxford USA Press, 2015.

[150] David Li, Ruofei Du, Adharsh Babu, Camelia D. Brumar, and Amitabh Varshney, *A log-rectilinear transformation for foveated 360-degree video streaming*, IEEE Transactions on Visualization and Computer Graphics (2021), 1–1.

[151] Jingjie Li, Amrita Roy Chowdhury, Kassem Fawaz, and Younghyun Kim, *Kalεido: Real-time privacy control for eye-tracking systems*, 29th USENIX Security Symposium (USENIX Security 20), 2020.

[152] Tianxing Li, Qiang Liu, and Xia Zhou, *Ultra-low power gaze tracking for virtual reality*, Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems, ACM, 2017, p. 25.

[153] Ao Liu, Lirong Xia, Andrew Duchowski, Reynold Bailey, Kenneth Holmqvist, and Eakta Jain, *Differential privacy for eye-tracking data*, Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications, ACM, 2019, p. 28.

[154] Dillon Lohr, Henry Griffith, and Oleg V Komogortsev, *Eye know you: Metric learning for end-to-end biometric authentication using eye movements from a longitudinal dataset*, arXiv preprint arXiv:2104.10489 (2021).

[155] Dillon Lohr and Oleg V Komogortsev, *Eye know you too: A densenet architecture for end-to-end biometric authentication via eye movements*, arXiv preprint arXiv:2201.02110 (2022).

[156] Dillon J Lohr, Samantha Aziz, and Oleg Komogortsev, *Eye movement biometrics using a new dataset collected in virtual reality*, ACM Symposium on Eye Tracking Research and Applications, 2020, pp. 1–3.

[157] Pietro Lungaro, Rickard Sjöberg, Alfredo José Fanghella Valero, Ashutosh Mittal, and Konrad Tollmar, *Gaze-aware streaming solutions for the next generation of mobile VR experiences*, IEEE Transactions on Visualization and Computer Graphics **24** (2018), no. 4, 1535–1544.

[158] Divine Maloney, Guo Freeman, and Andrew Robb, *A virtual space for all: Exploring children's experience in social virtual reality*, Proceedings of the Annual Symposium on Computer-Human Interaction in Play, 2020, pp. 472–483.

[159] Claudio Marforio, Hubert Ritzdorf, Aurélien Francillon, and Srdjan Capkun, *Analysis of the communication between colluding applications on modern smartphones*, Proceedings of the 28th Annual Computer Security Applications Conference, 2012, pp. 51–60.

[160] Daniel Martin, Ana Serrano, Alexander W Bergman, Gordon Wetzstein, and Belen Masia, *Scangan360: A generative model of realistic scanpaths for 360 images*, IEEE Transactions on Visualization & Computer Graphics (2022), no. 01, 1–1.

[161] Libor Masek and Peter Kovesi, *Matlab source code for a biometric identification system based on iris patterns, ms thesis*, 2013.

[162] Lucas Matney, *Magic Leap raises $461 million in fresh funding from the Kingdom of Saudi Arabia*, https://techcrunch.com/2018/03/07/magic-leap-raises-461-million-in-fresh-funding-from-the-kingdom-of-saudi-arabia/?guccounter=1, 2018, Accessed: 2022-05-23.

[163] Nora McDonald, *Online data could be used against people seeking abortions if Roe v. Wade falls*, https://theconversation.com/online-data-could-be-used-against-people-seeking-abortions-if-roe-v-wade-falls-182830, 2022, Accessed: 2022-05-23.

[164] Mark McGill, *Xr and the erosion of anonymity and privacy*, The IEEE Global Initiative on Ethics of Extended Reality Report, IEEE, 2021.

[165] Chao Mei, Bushra T Zahed, Lee Mason, and John Ouarles, *Towards joint attention training for children with ASD-a VR game approach and eye gaze exploration*, 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, 2018, pp. 289–296.

[166] Xiaoxu Meng, Ruofei Du, and Amitabh Varshney, *Eye-dominance-guided foveated rendering*, IEEE transactions on visualization and computer graphics **26** (2020), no. 5, 1972–1980.

[167] Xiaoxu Meng, Ruofei Du, Matthias Zwicker, and Amitabh Varshney, *Kernel foveated rendering*, Proceedings of the ACM on Computer Graphics and Interactive Techniques **1** (2018), no. 1, 1–20.

[168] Abraham Hani Mhaidli and Florian Schaub, *Identifying manipulative advertising techniques in xr through scenario construction*, Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 2021, pp. 1–18.

[169] Ailsa E Millen, Lorraine Hope, Anne P Hillstrom, and Aldert Vrij, *Tracking the truth: the effect of face familiarity on eye fixations during deception*, Quarterly Journal of Experimental Psychology **70** (2017), no. 5, 930–943.

[170] Mark Roman Miller, Fernanda Herrera, Hanseul Jun, James A Landay, and Jeremy N Bailenson, *Personal identifiability of user tracking data during observation of 360-degree vr video*, Scientific Reports **10** (2020), no. 1, 1–10.

[171] Robert Miller, Natasha Kholgade Banerjee, and Sean Banerjee, *Using siamese neural networks to perform cross-system behavioral authentication in virtual reality*, 2021 IEEE Virtual Reality and 3D User Interfaces (VR), IEEE, 2021, pp. 140–149.

[172] John V Monaco, *Classification and authentication of one-dimensional behavioral biometrics*, IEEE International Joint Conference on Biometrics, IEEE, 2014, pp. 1–8.

[173] Donald M Monro, Soumyadip Rakshit, and Dexin Zhang, *Dct-based iris recognition*, IEEE TPAMI (2007), no. 4, 586–595.

[174] James Morris, Stephen Smalley, and Greg Kroah-Hartman, *Linux security modules: General security support for the linux kernel*, Proceedinsg of the 2002 USENIX Security Symposium, 2002.

[175] Joerg H Mueller, Philip Voglreiter, Mark Dokter, Thomas Neff, Mina Makar, Markus Steinberger, and Dieter Schmalstieg, *Shading atlas streaming*, SIGGRAPH Asia 2018 Technical Papers, ACM, 2018, p. 199.

[176] Peter Mundy, *A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder*, European Journal of Neuroscience **47** (2018), no. 6, 497–514.

[177] DP Munoz, JR Broughton, JE Goldring, and IT Armstrong, *Age-related performance of human subjects on saccadic eye movement tasks*, Experimental brain research **121** (1998), no. 4, 391–400.

[178] T Scott Murdison, Gunnar Blohm, and Frank Bremmer, *Saccade-induced changes in ocular torsion reveal predictive orientation perception*, Journal of Vision **19** (2019), no. 11, 10–10.

[179] Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, and Norman Sadeh, *Privacy expectations and preferences in an IoT world*, Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017), 2017, pp. 399–412.

[180] Arvind Narayanan and Vitaly Shmatikov, *Robust de-anonymization of large sparse datasets*, 2008 IEEE Symposium on Security and Privacy, IEEE, 2008, pp. 111–125.

[181] _____, *De-anonymizing social networks*, 2009 30th IEEE symposium on security and privacy, IEEE, 2009, pp. 173–187.

[182] Carman Neustaedter, Saul Greenberg, and Michael Boyle, *Blur filtration fails to preserve privacy for home-based video conferencing*, ACM Transactions on Computer-Human Interaction (TOCHI) **13** (2006), no. 1, 1–36.

[183] Elaine M Newton, Latanya Sweeney, and Bradley Malin, *Preserving privacy by de-identifying face images*, IEEE transactions on Knowledge and Data Engineering **17** (2005), no. 2, 232–243.

[184] Helen Nissenbaum, *Privacy as contextual integrity*, Wash. L. Rev. **79** (2004), 119.

[185] Marcus Nyström, Richard Andersson, Kenneth Holmqvist, and Joost Van De Weijer, *The influence of calibration method and eye physiology on eyetracking data quality*, Behavior research methods **45** (2013), no. 1, 272–288.

[186] Lubos Omelina, Jozef Goga, Jarmila Pavlovicova, Milos Oravec, and Bart Jansen, *A survey of iris datasets*, Image and Vision Computing **108** (2021), 104109.

[187] Jason Orlosky, Yuta Itoh, Maud Ranchet, Kiyoshi Kiyokawa, John Morgan, and Hannes Devos, *Emulation of physician tasks in eye-tracked virtual reality for remote diagnosis of neurodegenerative disease*, IEEE Transactions on Visualization and Computer Graphics **23** (2017), no. 4, 1302–1311.

[188] Jacob L Orquin, Nathaniel JS Ashby, and Alasdair DF Clarke, *Areas of interest as a signal detection problem in behavioral eye-tracking research*, Journal of Behavioral Decision Making **29** (2016), no. 2-3, 103–115.

[189] Jessica Outlaw, *"Don't Track My Life" Virtual and Augmented Reality Consumer Data & Privacy Survey*, https://www.extendedmind.io/survey, 2021, Accessed: 2022-05-23.

[190] Yun Suen Pai, Benjamin I Outram, Benjamin Tag, Megumi Isogai, Daisuke Ochi, and Kai Kunze, *Gazesphere: Navigating 360-degree-video environments in VR using head rotation and eye gaze*, ACM SIGGRAPH 2017 Posters, ACM, 2017, p. 23.

[191] Nicolas Papernot, Patrick McDaniel, Arunesh Sinha, and Michael Wellman, *Towards the science of security and privacy in machine learning*, arXiv preprint arXiv:1611.03814 (2016).

[192] Anjul Patney, Marco Salvi, Joohwan Kim, Anton Kaplanyan, Chris Wyman, Nir Benty, David Luebke, and Aaron Lefohn, *Towards foveated rendering for gaze-tracked virtual reality*, ACM Transactions on Graphics (TOG) **35** (2016), no. 6, 179.

[193] Kevin A Pelphrey, James P Morris, and Gregory McCarthy, *Neural basis of eye gaze processing deficits in autism*, Brain **128** (2005), no. 5, 1038–1048.

[194] Jean Pfiffelmann, Nathalie Dens, and Sébastien Soulez, *Personalized advertisements with integration of names and photographs: An eye-tracking experiment*, Journal of Business Research **111** (2020), 196–207.

[195] Francesco Pittaluga and Sanjeev J Koppal, *Privacy preserving optics for miniature vision sensors*, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 314–324.

[196] Vincent Primault, Antoine Boutet, Sonia Ben Mokhtar, and Lionel Brunie, *The long road to computational location privacy: A survey*, IEEE Communications Surveys & Tutorials **21** (2018), no. 3, 2772–2793.

[197] Hugo Proenca, Silvio Filipe, Ricardo Santos, Joao Oliveira, and Luis A Alexandre, *The ubiris. v2: A database of visible wavelength iris images captured on-the-move and at-a-distance*, IEEE TPAMI **32** (2009), no. 8, 1529–1535.

[198] Yashas Rai, Jesús Gutiérrez, and Patrick Le Callet, *A dataset of head and eye movements for 360 degree images*, Proceedings of the 8th ACM on Multimedia Systems Conference, ACM, 2017, pp. 205–210.

[199] Yasmeen Rashidi, Tousif Ahmed, Felicia Patel, Emily Fath, Apu Kapadia, Christena Nippert-Eng, and Norman Makoto Su, *"You don't want to be the next meme": College students' workarounds to manage privacy in the era of pervasive photography*, Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018), 2018, pp. 143–157.

[200] Vibhor Rastogi and Suman Nath, *Differentially private aggregation of distributed time-series with transformation and encryption*, Proceedings of the 2010 ACM SIGMOD International Conference on Management of data, 2010, pp. 735–746.

[201] Joel Reardon, Álvaro Feal, Primal Wijesekera, Amit Elazari Bar On, Narseo Vallina-Rodriguez, and Serge Egelman, *50 ways to leak your data: An exploration of apps' circumvention of the android permissions system*, 28th USENIX Security Symposium (USENIX Security 19), 2019, pp. 603–620.

[202] Nelson Reed and Katie Joseff, *Kids and the Metaverse: What Parents, Policymakers, and Companies Need to Know*, https://www.commonsensemedia.org/sites/default/files/featured-content/files/metaverse-white-paper-1.pdf, 2022, Accessed: 2022-07-17.

[203] Patrice Renaud, Joanne L Rouleau, Luc Granger, Ian Barsetti, and Stéphane Bouchard, *Measuring sexual preferences in virtual reality: A pilot study*, CyberPsychology & Behavior **5** (2002), no. 1, 1–9.

[204] Irwin Reyes, Primal Wijesekera, Joel Reardon, Amit Elazari Bar On, Abbas Razaghpanah, Narseo Vallina-Rodriguez, Serge Egelman, et al., *"won't somebody think of the children?" examining coppa compliance at scale*, The 18th Privacy Enhancing Technologies Symposium (PETS 2018), 2018.

[205] Slobodan Ribaric, Aladdin Ariyaeeinia, and Nikola Pavesic, *De-identification for privacy protection in multimedia content: A survey*, Signal Processing: Image Communication **47** (2016), 131–151.

[206] Gerulf Rieger, Brian M Cash, Sarah M Merrill, James Jones-Rounds, Sanjay Muralidharan Dharmavaram, and Ritch C Savin-Williams, *Sexual arousal: The correspondence of eyes and genitals*, Biological Psychology **104** (2015), 56–64.

[207] Ioannis Rigas, Evgeniy Abdulin, and Oleg Komogortsev, *Towards a multi-source fusion approach for eye movement-driven recognition*, Information Fusion **32** (2016), 13–25.

[208] Ioannis Rigas and Oleg V Komogortsev, *Gaze estimation as a framework for iris liveness detection*, IEEE International Joint Conference on Biometrics, IEEE, 2014, pp. 1–8.

[209] _____, *Current research in eye movement biometrics: An analysis based on bioeye 2015 competition*, Image and Vision Computing **58** (2017), 129–141.

[210] Sharon Roberg-Perez, *The future is now: Biometric information and data privacy*, Antitrust **31** (2016), 60.

[211] C Carl Robusto, *The cosine-haversine formula*, The American Mathematical Monthly **64** (1957), no. 1, 38–40.

[212] Franziska Roesner and Tadayoshi Kohno, *Security and privacy for augmented reality: Our 10-year retrospective*, VR4Sec: 1st International Workshop on Security for XR and XR for Security, 2021.

[213] Theodore Rostow, *What happens when an acquaintance buys your data: a new privacy harm in the age of data brokers*, Yale J. on Reg. **34** (2017), 667.

[214] Sylvia Rothe, Felix Althammer, and Mohamed Khamis, *Gazerecall: Using gaze direction to increase recall of details in cinematic virtual reality*, Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia, 2018, pp. 115–119.

[215] Sylvia Rothe, Daniel Buschek, and Heinrich Hußmann, *Guidance in cinematic virtual reality-taxonomy, research status and challenges*, Multimodal Technologies and Interaction **3** (2019), no. 1, 19.

[216] Kerstin Ruhland, Sean Andrist, Jeremy Badler, Christopher Peters, Norman Badler, Michael Gleicher, Bilge Mutlu, and Rachel Mcdonnell, *Look me in the eyes: A survey of eye and gaze animation for virtual agents and artificial systems*, Eurographics state-of-the-art report, 2014, pp. 69–91.

[217] Kerstin Ruhland, Christopher E Peters, Sean Andrist, Jeremy B Badler, Norman I Badler, Michael Gleicher, Bilge Mutlu, and Rachel McDonnell, *A review of eye gaze in virtual agents, social robotics and HCI: Behaviour generation, user interaction and perception*, Computer graphics forum, vol. 34, Wiley Online Library, 2015, pp. 299–326.

[218] Reiner Sailer, Xiaolan Zhang, Trent Jaeger, and Leendert Van Doorn, *Design and Implementation of a TCG-based Integrity Measurement Architecture*, Proceedings of the 2004 USENIX Security Symposium, 2004.

[219] Dario D Salvucci and Joseph H Goldberg, *Identifying fixations and saccades in eye-tracking protocols*, Proceedings of the 2000 symposium on Eye tracking research & applications, 2000, pp. 71–78.

[220] Pierangela Samarati and Latanya Sweeney, *Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression*, (1998).

[221] Negar Sammaknejad, Hamidreza Pouretemad, Changiz Eslahchi, Alireza Salahirad, and Ashkan Alinejad, *Gender classification based on eye movements: A processing effect during passive face viewing*, Advances in cognitive psychology **13** (2017), no. 3, 232.

[222] Christoph Schröder, Sahar Mahdie Klim Al Zaidawi, Martin HU Prinzler, Sebastian Maneth, and Gabriel Zachmann, *Robustness of eye movement biometrics against varying stimuli and varying trajectory length*, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–7.

[223] Immo Schuetz, T Scott Murdison, Kevin J MacKenzie, and Marina Zannoli, *An explanation of fitts' law-like performance in gaze-based selection tasks using a psychophysics approach*, Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–13.

[224] Friedhelm Schwenker, Hans A Kestler, and Günther Palm, *Three learning phases for radial-basis-function networks*, Neural networks **14** (2001), no. 4-5, 439–458.

[225] Cong Shi, Xiangyu Xu, Tianfang Zhang, Payton Walker, Yi Wu, Jian Liu, Nitesh Saxena, Yingying Chen, and Jiadi Yu, *Face-mic: inferring live speech and speaker identity via subtle facial dynamics captured by ar/vr motion sensors*, Proceedings of the 27th Annual International Conference on Mobile Computing and Networking, 2021, pp. 478–490.

[226] Frederick Shic, Katarzyna Chawarska, Jessica Bradshaw, and Brian Scassellati, *Autism, eye-tracking, entropy*, 2008 7th IEEE International Conference on Development and Learning, IEEE, 2008, pp. 73–78.

[227] Ryan Singel, *Netflix spilled your brokeback mountain secret, lawsuit claims*, Threat Level (blog), Wired (2009).

[228] Vincent Sitzmann, Ana Serrano, Amy Pavel, Maneesh Agrawala, Diego Gutierrez, Belen Masia, and Gordon Wetzstein, *Saliency in VR: How do people explore virtual environments?*, IEEE Transactions on Visualization and Computer Graphics **24** (2018), no. 4, 1633–1642.

[229] Vasileios Skaramagkas, Giorgos Giannakakis, Emmanouil Ktistakis, Dimitris Manousos, Ioannis Karatzanis, Nikolaos Tachos, Evanthia Eleftherios Tripoliti, Kostas Marias, Dimitrios I Fotiadis, and Manolis Tsiknakis, *Review of eye tracking metrics involved in emotional and cognitive processes*, IEEE Reviews in Biomedical Engineering (2021).

[230] Ivo Sluganovic, Marc Roeschlin, Kasper B Rasmussen, and Ivan Martinovic, *Analysis of reflexive eye movements for fast replay-resistant biometric authentication*, ACM Transactions on Privacy and Security (TOPS) **22** (2018), no. 1, 1–30.

[231] Srinivas Sridharan, Reynold Bailey, Ann McNamara, and Cindy Grimm, *Subtle gaze manipulation for improved mammography training*, Proceedings of the Symposium on Eye Tracking Research and Applications, 2012, pp. 75–82.

[232] Julian Steil, Inken Hagestedt, Michael Xuelin Huang, and Andreas Bulling, *Privacy-aware eye tracking using differential privacy*, 11th ACM Symposium on Eye Tracking Research & Applications, ACM, 2019.

[233] Julian Steil, Marion Koelle, Wilko Heuten, Susanne Boll, and Andreas Bulling, *Privaceye: privacy-preserving head-mounted eye tracking using egocentric scene image and eye movement features*, Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications, ACM, 2019, p. 26.

[234] Frank Steinicke, Gerd Bruder, Jason Jerald, Harald Frenz, and Markus Lappe, *Estimation of detection thresholds for redirected walking techniques*, IEEE transactions on visualization and computer graphics **16** (2009), no. 1, 17–27.

[235] William Steptoe, Anthony Steed, Aitor Rovira, and John Rae, *Lie tracking: social presence, truth and deception in avatar-mediated telecommunication*, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2010, pp. 1039–1048.

[236] Qi Sun, Anjul Patney, Li-Yi Wei, Omer Shapira, Jingwan Lu, Paul Asente, Suwen Zhu, Morgan Mcguire, David Luebke, and Arie Kaufman, *Towards virtual reality infinite walking: dynamic saccadic redirection*, ACM Transactions on Graphics (TOG) **37** (2018), no. 4, 67.

[237] Latanya Sweeney, *k-anonymity: A model for protecting privacy*, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **10** (2002), no. 05, 557–570.

[238] Dean Takahashi, *Magic Leap 2 mixed reality headsets for enterprise will debut for $3,300 on September 30*, https://venturebeat.com/2022/07/12/magic-leap-2-mixed-reality-headsets-for-enterprise-will-debut-for-3300-on-september-30/, 2022, Accessed: 2022-07-15.

[239] Tobii, *Tobii XR SDK*, https://vr.tobii.com/sdk/develop/unity/, 2022, Accessed: 2022-07-17.

[240] Takumi Toyama, Andreas Dengel, Wakana Suzuki, and Koichi Kise, *Wearable reading assist system: Augmented reality document combining document retrieval and eye tracking*, 2013 12th International Conference on Document Analysis and Recognition, IEEE, 2013, pp. 30–34.

[241] Rahmadi Trimananda, Hieu Le, Hao Cui, Janice Tran Ho, Anastasia Shuba, and Athina Markopoulou, *Ovrseen: Auditing network traffic and privacy policies in oculus vr*, arXiv preprint arXiv:2106.05407 (2021).

[242] Stephen Tu, *The dirichlet-multinomial and dirichlet-categorical models for bayesian inference*, Computer Science Division, UC Berkeley **2** (2014).

[243] AJ Van Opstal and JAM Van Gisbergen, *Skewness of saccadic velocity profiles: a unifying parameter for normal and slow saccades*, Vision research **27** (1987), no. 5, 731–745.

[244] Pranav Venuprasad, Tushal Dobhal, Anurag Paul, Tu NM Nguyen, Andrew Gilman, Pamela Cosman, and Leanne Chukoskie, *Characterizing joint attention behavior during real world interactions using automated object and gaze detection*, Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications, ACM, 2019, p. 21.

[245] Matias Volonte, Andrew Robb, Andrew T Duchowski, and Sabarish V Babu, *Empirical evaluation of virtual human conversational and affective animations on visual attention in inter-personal simulations*, 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, 2018, pp. 25–32.

[246] Human Rights Watch, *#Outlawed the love that dare note speak its name*, Web: http://internap.hrw.org/features/features/lgbt_laws/ (accessed April 11, 2021).

[247] Eric Whitmire, Laura Trutoiu, Robert Cavin, David Perek, Brian Scally, James Phillips, and Shwetak Patel, *Eyecontact: scleral coil eye tracking for virtual reality*, Proceedings of the 2016 ACM International Symposium on Wearable Computers, ACM, 2016, pp. 184–191.

[248] Wenjun Xiong and Robert Lagerström, *Threat modeling–a systematic literature review*, Computers & security **84** (2019), 53–69.

[249] Mai Xu, Chen Li, Shanyi Zhang, and Patrick Le Callet, *State-of-the-art in 360 video/image processing: Perception, assessment and compression*, IEEE Journal of Selected Topics in Signal Processing **14** (2020), no. 1, 5–26.

[250] Yanyu Xu, Yanbing Dong, Junru Wu, Zhengzhong Sun, Zhiru Shi, Jingyi Yu, and Shenghua Gao, *Gaze prediction in dynamic 360° immersive videos*, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 5333–5342.

[251] Wencheng Yang, Song Wang, Jiankun Hu, Guanglou Zheng, and Craig Valli, *Security and accuracy of fingerprint-based biometrics: A review*, Symmetry **11** (2019), no. 2, 141.

[252] Chaitra Yangandul, Sachin Paryani, Madison Le, and Eakta Jain, *How many words is a picture worth? attention allocation on thumbnails versus title text regions*, Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications, 2018, pp. 1–5.

[253] Raimondas Zemblys and Oleg Komogortsev, *Developing photo-sensor oculography (PS-OG) system for virtual reality headsets*, Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications, ACM, 2018, p. 83.

[254] Raimondas Zemblys, Diederick C Niehorster, Oleg Komogortsev, and Kenneth Holmqvist, *Using machine learning to detect events in eye-tracking data*, Behavior research methods **50** (2018), no. 1, 160–181.

[255] A Tianyi Zhang and Olivier Le Meur, *How old do you look? inferring your age from your gaze*, 2018 25th IEEE International Conference on Image Processing (ICIP), IEEE, 2018, pp. 2660–2664.

[256] Yunzhan Zhou, Tian Feng, Shihui Shuai, Xiangdong Li, Lingyun Sun, and Henry Been-Lirn Duh, *Edvam: a 3d eye-tracking dataset for visual attention modeling in a virtual museum*, Frontiers of Information Technology & Electronic Engineering **23** (2022), no. 1, 101–112.

BIOGRAPHICAL SKETCH

Brendan David-John received a Ph.D. in computer science from the University of Florida, where he was advised by Dr. Eakta Jain. While pursuing his doctoral degree, Brendan completed an internship at Facebook Reality Labs Research and was awarded a Google Ph.D. Fellowship to support his work on eye-tracking privacy. Prior to his doctoral studies, he attended the Rochester Institute of Technology, where he received a B.S. in computational mathematics and M.S. in computer science.