When VR Meets BCI: (Un)Observable Brainwave-aware Privacy Reconstruction in the Metaverse via Unrestricted Inbuilt Motion Sensors

Tao Ni* Zehua Sun† Qingchuan Zhao‡(✉) Wei-Bin Lee§¶ Cong Wang‡(✉)
*King Abdullah University of Science and Technology †National University of Singapore
‡City University of Hong Kong §Feng Chia University
¶Information Security Research Center, Hon Hai Research Institute

***Abstract*—Metaverse devices, such as virtual reality (VR), have seen substantial development and widespread applications in numerous areas. Although recent studies have revealed privacy leakages in VR, these vulnerabilities were limited in the scope of observable behaviors in virtual scenes (*e.g.*, *what a user is seeing*). In this work, we uncover the feasibility of going beyond the scope of observable user behaviors to unobservable brain EEG-correlated representations (*e.g.*, *what a user is perceiving*) by leveraging unrestricted motion sensors in VR headsets to reconstruct brain EEG signals, a seemingly neglected but promising vector. The insight is that the inbuilt motion sensors (*e.g.*, accelerometers) in the VR headset can capture subtle vibrations induced by pupillary responses, which are highly correlated with users' visual stimuli and in-brain perceptions.**

Therefore, we design and implement BRAVESPY to systematically investigate and demonstrate the feasibility of this severe privacy leakage originating from brain EEG-correlated representations reconstructed from variations of inbuilt motion sensors. Our extensive evaluation results from different VR devices show that BRAVESPY, for the first time in the Metaverse, can reveal unobservable privacy, where we successfully unveiled perceptive images in the brain with 52.0%–67.2% accuracy. In particular, we also find that BRAVESPY outperforms the current approaches that are limited to coarse-grained inference of observable behaviors and achieves over 85.0% accuracy in inferring user activity-related sensitive information, such as fingerprinting websites, apps, and streaming videos, and over 96.0% accuracy in user de-anonymization, gaze movement tracking, and virtual keystroke inference.

1. Introduction

Recent years have witnessed exponential advancements in Metaverse technology, such as virtual reality (VR), which aims to create a simulated environment with lifelike scenes and objects to give users the sensation of being fully immersed in their surroundings through a VR headset. Typically, it consists of a head-mounted display (HMD) and inbuilt sensors that allow VR applications to measure user's motions and postures and to support real-time interactions by adjusting the visual content accordingly. Thus, massive sensor data access has become inevitable in VR apps, and most existing operating systems enforce limited constraints

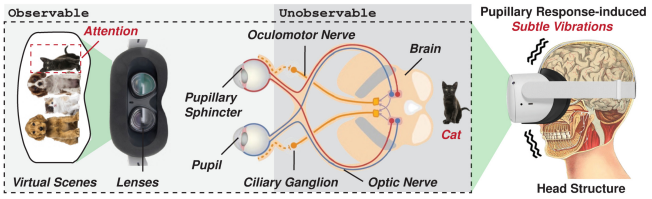


Figure 1: An illustration of brain perception in virtual scenes.

on this data access. For example, according to a recent empirical study, almost 42% VR apps from the Meta Oculus Quest store [1] collect motion sensor data and have no permission requests in the manifest files of the apps [2].

Unfortunately, the inevitable access and collection of sensor data in VR applications is expected to result in severe leakage of observable user behaviors, such as keystrokes [3]–[5], the launch of websites/apps [6]–[8], and speeches [9]–[11], because a large body of research on the smartphone platform has well studied and unveiled the vulnerabilities through which sensor data can leak such sensitive privacy. Following that line of research, recent security studies specific to the VR platform have also found that its sensor data could reveal observable and measurable privacy issues, including virtual keyboard inputs (*e.g.*, [12]–[18]), 360° streaming videos [19], user avatar [20]–[22], speech [23], [24], and further reconstructing user biometric information, *i.e.*, body fat ratio [25], respiration rates and heartbeats [26], [27] or even blood pressure [28]. Indeed, these findings have successfully complemented the understanding of the attack surface in the VR platform alongside other attack vectors reported in VR, such as GPU profiles [29], network traffic [18], [30], [31], camera-recorded videos [16], [19], [32], [33], and controllers' button-clicking sounds [34] or leaked infrared signals [35]. However, these recent security studies on the VR platform have not fully investigated the unique features of the VR platform and are thus limited in the scope of observable privacy leakages.

In this work, we found a unique but not fully explored feature of sensors in VR headsets, which are positioned much closer to a user's brain, is promising not only for further exploring much higher granularity of *what a user is seeing or doing* but also in going beyond this scope to *what a user is perceiving*, *a.k.a.*, observable behaviors to unobservable behaviors. This transition is believed to be

DOI 10.1109/SP63933.2026.00117

critical to fully understanding the potential of user privacy leakage in VR due to the inconsistency between user visual perception and brain perception [36], [37], which is a well-established theoretical and empirical foundation in the field of brain science and cognitive neuroscience. When it comes to the VR platform, it is noted that brain perception is selective as the brain constantly decides what information from the HMD is important enough to reach our consciousness. Thus, a large part of the sensory information from virtual scenes that constantly arrives through our senses is never consciously processed. Complex mechanisms in the brain filter the incoming sensory information and shape the representation of the world in human minds [38], [39]. For example, Figure 1 shows that when the VR user views a picture with multiple objects (*e.g.*, different cats and dogs) in the virtual scene, his/her brain selectively processes only the object of attention (*e.g.*, black domestic shorthair cat) [40].

Our proposed hypothesis to leverage VR sensors to reveal users' unobservable and fine-grained observable behaviors is believed to be plausible and reasonable because previous theories have validated that (1) unobservable behaviors (*e.g.*, brain image perceptions) reflected in the eye responses [41]–[43] and brain EEG-correlated reactions [44]–[46], (2) brain EEG variations or eye responses, in turn, control the user's body reactions and movements [47], [48], and (3) these two signals are concentrated in the brain's forehead but are too weak to be measured as the distance increases. However, the VR headset is tightly mounted on the forehead scalp around the eyes, ensuring close proximity to measure high-quality data. As such, the main challenge lies in the feasibility of bridging the relationship between VR sensors and these two pieces of information. Our key insight is the observation that the pupillary responses of the human eye under different visual stimuli are correlated with the dilation and contraction of the pupillary sphincter [49], [50] that leads to *subtle facial vibrations* around the eyes. In particular, these induced subtle vibrations are related to brain EEG signals [41]–[43] and could propagate through the VR headset to vibrate inbuilt motion sensors. Hence, data fluctuations of these motion sensors could reflect pupillary responses and in-depth brain EEG-correlated variations.

Therefore, we design and implement an attack system, BRAVESPY, to systematically explore the potential privacy leakages resulting from the unrestricted motion sensors built into the VR headset. By exploiting the statistical correlation between pupillary response-induced vibrations and brain-level visual perception, BRAVESPY reconstructs EEG-correlated representations from motion sensor data to infer both observable and unobservable user activities. As depicted in Figure 2, we propose a privacy leakage model that captures the correlations between visual stimuli and pupillary responses at three levels:

- **Observable UI-level Privacy:** This privacy refers to the content displayed on the HMD. Related works have only studied single coarse-grained privacy, such as keystrokes [12]–[14], [16]–[18], [30]–[32], [34], [35], [51], [52], apps [29] and 360° videos [19]. We comprehensively explore more fine-grained UI-level privacy, including fingerprinting websites, and in-app activities such as streaming videos that have not yet been attempted.

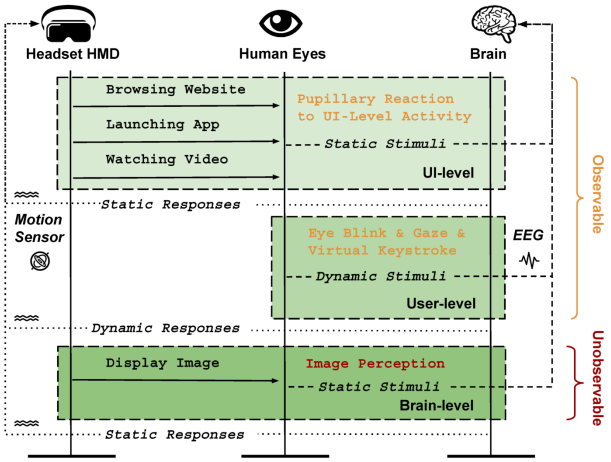


Figure 2: Three levels of observable (UI-level and user-level) behaviors and unobservable (brain-level) perceptions in VR.

- **Observable User-level Privacy:** This privacy concern is mainly related to the biometric features of VR users who wear the HMD. Previous works have focused only on user profiles, such as users' identity and gender [20], [21], [23], [24], [53], [54], vital signals such as respiration and heartbeat rates [25]–[27], and blood pressures [28]. Our research goes further by exploring user de-anonymization, gaze movements and virtual keystroke inference.
- **Unobservable Brain-level Privacy:** This privacy is relevant to the information the brain selects and reaches the user's consciousness, such as brain image perceptions [44]–[46]. To our knowledge, BRAVESPY becomes the *first* attempt to reveal such unobservable brain-level perceptions using the reconstructed EEG-correlated representations from the VR headset's motion sensors.

We have evaluated BRAVESPY on four popular VR headsets: Meta Oculus Quest 2, Meta Oculus Quest, PICO 4 All-in-One, and HTC Vive Pro. Our evaluation results show that BRAVESPY can effectively infer sensitive user information at the aforementioned three levels, and also presents high robustness and transferability when considering conditions with varying impact factors. In addition, we introduce effective countermeasures to prevent privacy leakage from inbuilt unrestricted motion sensors.

Contributions. We summarize our contributions as follows:

- We identify a novel attack vector in which adversaries can infer EEG-correlated perceptions using unrestricted motion sensors in VR headsets, which poses significant threats to user privacy within the emerging Metaverse.
- We systematically investigate the variations of brain EEG patterns with pupillary responses under visual stimuli corresponding to user activities in virtual scenes. Furthermore, we model three levels of observable and unobservable privacy with reconstructed EEG-correlated representations from visual presentation to brain perceptions.
- We design and implement an end-to-end system

BRAVESPY to validate the feasibility and practicality of the proposed attack vector, which could not only infer the privacy of user level (*e.g.*, browsing websites, running VR apps, playing streaming video) and user level (*e.g.*, identity, gaze, keystroke), but also in-brain image perceptions, which outperform all previous related studies.

2. Preliminary

2.1. Unrestricted Motion Sensors in VR Headset

In order to enable immersive interactions in the virtual environment, VR headsets (*e.g.*, Meta Oculus Quest 2) integrate with various embedded sensors such as motion sensors (*e.g.*, accelerometer, gyroscope), which are critical for estimating the orientation of the headset. However, mainstream VR SDKs and APIs, including OpenXR [55], Oculus Mobile SDK [56], and WebXR Device API [57], do not impose restrictions on accessing the motion sensor readings inside the VR headset. This openness has enabled research, as evidenced by recent studies (*e.g.*, [12]–[17], [23]–[26], [28], [58]), which demonstrate that adversaries can potentially deploy malicious apps or entice users into accessing malicious web pages, allowing stealthy and continuous logging of VR sensor data in the background without the user's permission. The data from these unrestricted motion sensors contain a substantial amount of sensitive information that could be maliciously exploited to violate sensitive user privacy.

Due to head-mounted characteristics, the unrestricted motion sensors of VR headsets can detect subtle facial vibrations, revealing sensitive biometric data such as gender and identity (*e.g.*, [23], [25], [59]) as well as vital signs such as respiration rate, heartbeat [26], and blood pressure [28]. Furthermore, because these headsets are positioned around the eyes, they can capture subtle facial vibrations triggered by pupillary responses to visual stimuli. These subtle vibrations around the eyes are linked to forehead EEG signals (*e.g.*, [41]–[43]), which reflect brain activity. This capability underscores the potential for VR technology to provide insight into a user's physiological and neurological state.

2.2. Pupillary Responses under Visual Stimuli

Due to the head-mounted design of HMD headsets, the users' eyes become the primary interface for receiving information within the virtual environment and transmitting it to the brain. Specifically, eyes visualize virtual scenes by receiving and processing different visual stimuli from various user interfaces and content displays (*a.k.a.*, pupillary light reflex [60], [61]), and the responded visual stimuli are then transmitted to the brain via the optic nerves. Furthermore, as the pupils adjust to focus on the displayed content, they present information that varies with their gaze. These pupil responses and movements influence brain perceptions, which results in different brainwave signals and is widely confirmed by previous studies [49], [50], [62]. In addition, pupillary responses also induce muscle movements (*e.g.*, dilation and contraction) that cause subtle vibrations, which can be detected by the inbuilt unrestricted motion sensors of the HMD headset that is tightly mounted on the forehead scalp. Therefore, pupillary responses bridge the correlations between brainwave variations and subtle vibrations on the motion sensors embedded in commercial VR headsets.

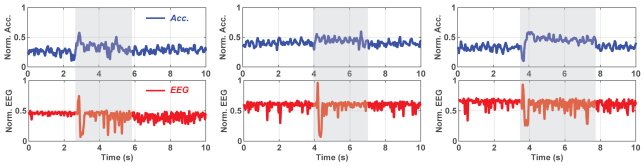
TABLE 1: Comparison with prior studies relevant to VR privacy leakage at two *observable* levels: (**L1**) UI-level visual presentation inference (🚗 → 👁); (**L2**) User-level identification and gaze recognition (👁 → 🫁), and one *unobservable* level: (**L3**) Brain-level perception inference (🚗 → 👁 → 🫁).

VR Attack	Venue	Attack Vector	M1	M2	M3	L1	L2	L3
VR-Spy [30]	IEEE VR'21	Wi-Fi CSI	✗	✓	–	●	○	○
Hologlogger [12]	IEEE VR'22	Motion sensor	✓	✗	✓	●	○	○
TyPose [13]	Security'23	Motion sensor	✓	✗	✓	●	○	○
Wu *et al.* [14]	S&P'23	Motion sensor	✓	✗	✓	●	○	○
SnoopFinger [17]	S&P'25	Motion sensor	✓	✗	✓	●	○	○
ImmerSpy [24]	NDSS'25	Motion sensor	✓	✗	✓	○	●	○
Gopal *et al.* [32]	Security'23	Recorded video	✗	✓	–	●	○	○
Intrude [19]	Security'24	Recorded video	✗	✓	–	●	○	○
LineTalker [52]	TIFS'24	Charging current	✗	✓	–	●	○	○
Heimdall [34]	NDSS'24	Controllers' sound	✗	✓	–	●	○	○
Su *et al.* [31]	Security'24	Unencrypted network	✗	✓	–	●	○	○
VRecKey [35]	NDSS'25	Infrared signals	✗	✓	–	●	○	○
TwiST [18]	NDSS'26	Wi-Fi packets	✗	✓	–	●	○	○
Zhang *et al.* [15]	Security'23	Sensor&System load	✓	✗	✗	●	○	○
OVRWatcher [29]	NDSS'26	GPU profile	✓	✗	✗	●	○	○
AvatarHunter [20]	INFOCOM'23	Users' avatar	✓	✗	✗	○	●	○
Yang *et al.* [22]	Security'24	Users' avatar	✓	✗	✗	●	○	○
GAZEploit [51]	CCS'24	User's avatar	✓	✗	✗	●	●	○
Face-Mic [23]	MobiCom'21	Motion sensor	✓	✗	✓	●	●	○
FaceReader [26]	CCS'23	Motion sensor	✓	✗	✓	●	●	○
BPSniff [28]	S&P'25	Motion sensor	✓	✗	✓	●	●	○
BRAVESPY (Our work)		Motion sensor	✓	✗	✓	●	●	●

[1] Three comparison metrics: **M1**–Malware installation; **M2**–Require external equipment (*e.g.*, cameras, smartphones); and **M3**–Unrestricted sensor policy.
[2] "●" and "○" indicate "Yes" and "No" that privacy leakage is explored.

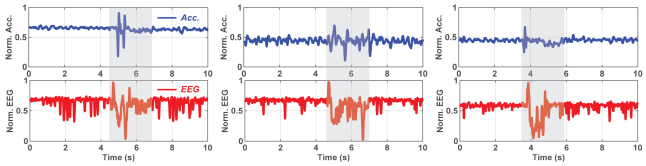
As depicted in Figure 1, when users wear the VR headset, the integrated lenses display digital content within the virtual environment, triggering visual stimuli that lead to various pupillary responses. *(i) Static Pupillary Response:* When the eyes are stationary and focused on the displayed content (*e.g.*, running apps, watching videos), the primary pupillary responses involve the dilation and contraction of the pupillary sphincter [49], [50]. In this scenario, pupillary responses to the content displayed in the VR headset are transmitted through the optic nerve to the brain for processing. *(ii) Dynamic Pupillary Response:* In contrast, when the users' eyes move during interactive activities in virtual scenes, such as searching for keys on a virtual keyboard or reading the text on the website, pupillary responses include both static changes and dynamic gaze information, which are detected by the ciliary ganglion and transmitted to the brain through the oculomotor nerve [62].

Both types of pupillary responses involve two distinct but correlated mechanisms. First, visual stimuli trigger pupillary sphincter dilation and contraction, producing subtle mechanical vibrations around the eyes [28] that propagate through the VR headset and are captured by the inbuilt motion sensors. Second, these pupillary responses independently correlate with brain EEG signals, as both are driven by the brain's visual processing pathway (*e.g.*, [41]–[43], [49], [50]). We note that the accelerometer measures me-



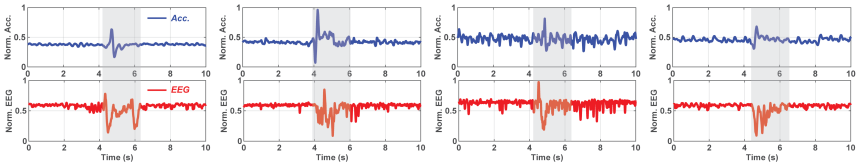
(a) Website. (b) App. (c) Video.

Figure 3: UI-level visual presentations.



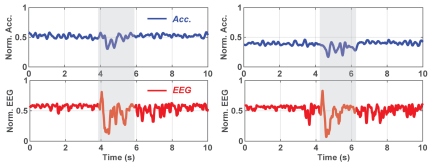
(a) Man #1. (b) Woman #1. (c) Man #2.

Figure 4: User-level identity recognition.



(a) Up. (b) Down. (c) Left. (d) Right.

Figure 5: UI-level reactions of gaze movements.



(a) Cat image. (b) Dog image.

Figure 6: Brain-level image perception.

chanical vibrations rather than electrical neural activity directly, whereas the well-established statistical correlation between pupillary responses and EEG signals makes it feasible to reconstruct EEG-correlated representations from the captured vibrations. Based on this insight, as shown in Figure 2, we model the correlations between the subtle vibrations in VR motion sensors and brain EEG variations, and systematically explore three levels of privacy inference targeting observable user behaviors and unobservable brain perceptions, including UI-level visual presentation, user-level identification and reaction, and brain-level image perceptions.

3. Threat Model

3.1. Attack Vectors and Adversary Capability

Following the recent works [12]–[14], [17], [23], [26], [28] listed in Table 1, we consider the potential attackers to be intentional VR application (app) developers who aim to acquire *only* inbuilt unrestricted motion sensor data from the VR headset to infer sensitive user information inside the virtual scenes. Unlike other related works that rely on external equipment (*e.g.*, Wi-Fi interceptor [18], [30], hidden cameras [19], [32], charging cables [52], and microphones [34]) or require privileged permission, such as access to system memory [15] or virtual cameras [20], [22], [51], our assumption about the capability of the adversary is reasonable and practical. This is because there are no deviations from the official VR app development guidelines specified by the associated SDKs (*e.g.*, OpenXR SDK [55], Oculus Mobile SDK [56]) that VR apps could access unrestricted motion sensor data without limits, including the inbuilt accelerometers and gyroscopes, from a user's headset in the background. In the *online deployment phase*, the adversary can only remotely obtain the motion sensor data from malicious code snippets embedded in these VR apps or third-party plugins released in open-source repositories. As such, our hypothesized intentional app behaves in the same way as benign apps that are acceptable in official VR app stores. Note that we do not assume that the adversary can access eye-tracking sensors to obtain gaze information directly because (1) most eye-tracking sensors are under higher restrictions and (2) only 17% VR apps incorporate functionalities with eye movements [2].

3.2. Attack Surface and Objectives

Unlike previous research, this work explores a new attack surface by measuring EEG signals derived from pupillary response-induced variations with unrestricted motion sensors, despite previous studies using these motion sensors primarily for virtual keystroke inference [12]–[15], [23]. Furthermore, the attack objectives of this work not only involve unveiling a much finer granularity of a user's observable behaviors and privacy, which advances the current approaches' sole ability in coarse-grained privacy inference, but also extend beyond the scope of observable behaviors to unobservable behaviors. Based on our three-level privacy model, the objectives are detailed as follows:

(L1) Observable UI-level Privacy. Based on the static pupillary response, the adversary can use the reconstructed EEG signals from the victim's brain to infer sensitive information related to various user-VR interactions, *i.e.*, launching VR apps, browsing websites, and watching videos inside streaming apps like NETFLIX. In particular, existing studies [12]–[14], [16] cannot achieve UI-level privacy inference because they are based solely on motion sensor data, which lacks the necessary granularity to fingerprint specific apps, websites, and video content due to associated privacy sensitivities. Figure 3 illustrates the correlation between motion sensor data (*e.g.*, accelerometer, denoted as *acc.*) and EEG signals during activities such as launching VR apps like NETFLIX, browsing websites such as TWITCH.TV, and watching videos like 3 BODY PROBLEM on the Meta Oculus Quest 2 headset. The synchronous changes observed in both accelerometer and EEG signals confirm the feasibility of reconstructing brainwaves and inferring user-interface privacy.

(L2) Observable User-level Privacy. The adversary can exploit alterations in brain EEG signals to breach user-level privacy, such as identifying users to initiate de-anonymization attacks, a threat that has been widely recognized threats by prior studies [20], [23], [26]. For example, Figure 4 displays reconstructed EEG signals from three different VR users (two men and one woman). Moreover,

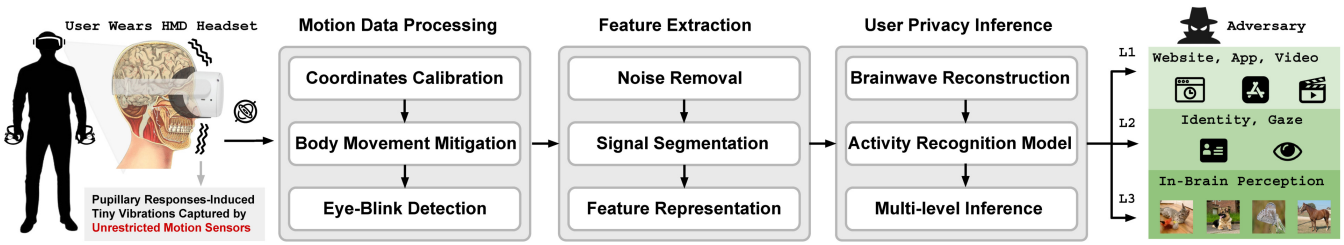


Figure 7: Overview of BRAVESPY.

based on the dynamic pupillary response (§2.2), adversaries can analyze reconstructed brainwaves to track the gaze movements of VR users. This capability allows them to connect user attention data to more accurately determine UI-level privacy concerns, such as virtual keystrokes, by monitoring user-level gaze movements toward specific keys on a virtual keyboard. Figure 5 illustrates the alterations in accelerometer and EEG signals with gaze movements: up, down, left, and right, respectively. As a result, these examples demonstrate the feasibility of using reconstructed EEG signals for user-level identity recognition and gaze movement extraction.

(L3) Unobservable Brain-level Privacy. The adversary leverages data collected from unrestricted motion sensor to reconstruct EEG signals for inferring brain-level visual perceptions, such as decoding perceptive images within the user's brain, which is a well-established brain-level task and has been widely validated by previous EEG-to-image studies (*e.g.*, [44]–[46], [63], [64]). Specifically, the adversary uses the reconstructed EEG signals to fine-tune advanced generative models, *i.e.*, DreamDiffusion [44], which interpret the EEG data to match the images perceived in the user's brain. Figure 6 exhibits the accelerometer and EEG signals when the VR user visualizes images in the mind, demonstrating that distinct perceptive images generate unique EEG patterns, which highlight the potential to uncover brain-level visual perceptions within the current threat model.

4. Attack Design

4.1. Overview of BRAVESPY

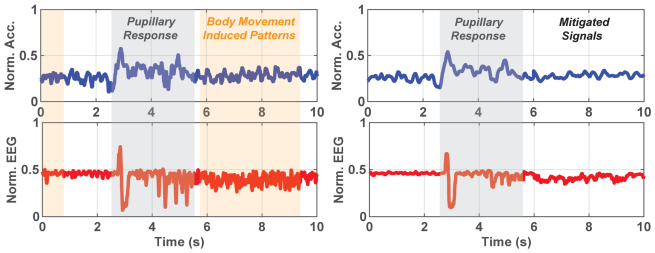
Figure 7 shows the end-to-end system overview of BRAVESPY. The adversary first obtains the unrestricted motion sensor data from the victim's VR headset and then conducts the following three stages: *(1) Motion Data Processing:* The adversary leverages the acquired motion data to calibrate the initial coordinates, mitigate the influence resulting from body movement, and detect the eye-blink events; *(2) Feature Extraction:* Then, the adversary applies filters to remove high-frequency noise, divide the informative segments from the signals, and extract feature representations; and *(3) User Privacy Inference:* The extracted features are then used to reconstruct brain wave signals (EEG signals), which are then used to infer various user privacy at different granularity levels. Finally, the adversary could obtain the output to infer observable sensitive information at the UI level (*e.g.*, website history, app usage, video preference), user level (*e.g.*, identity, gaze movements, and keystrokes), and unobservable brain level (*e.g.*, in-brain perceptive images).

4.2. Motion Data Processing

Coordinates Calibration. To mitigate the impact of varying initial body postures, we first perform coordinate calibration on the collected motion sensor data. Since interacting with a VR headset involves inherently three-dimensional movements, the 3D motion data captured by inbuilt accelerometers, which are recorded independently at different body postures, lack spatial coordinate calibration, making it unsuitable for direct use in brainwave reconstruction. To resolve this, we transform the acceleration values associated with different head postures into a common body reference coordinate system that is independent of orientation and location. We define the world coordinate system by the axes north, east, and down, or in the direction of gravity, and refer to the local coordinate system of the VR headset as (X, Y, Z). The plane of motion for the headset is defined as the front-side plane, which is perpendicular to gravity, with the side pointing toward the right side of the user's forehead Assuming the linear acceleration signals along the three orthogonal directions of the VR headset are Acc_X, Acc_Y, and Acc_Z, we compute the linear acceleration in the reference system as follows:

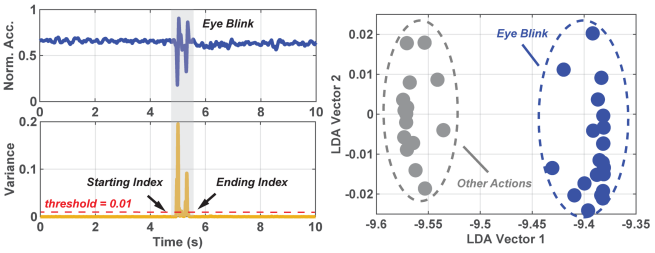
$$\begin{bmatrix} Acc_{X'} \\ Acc_{Y'} \\ Acc_{Z'} \end{bmatrix} = R_b^w \cdot R_w^v \cdot \begin{bmatrix} Acc_X \\ Acc_Y \\ Acc_Z \end{bmatrix} \quad (1)$$

Specifically, $Acc_{X'}$, $Acc_{Y'}$, and $Acc_{Z'}$ are linear accelerations along the direction of gravity, the front direction, and the side direction. The transformation matrices R_h^w and R_w^v, which represent the rotation from the world coordinate system to the body coordinate system and from the VR headset coordinate system to the world coordinate system, can be derived using angular data from the inbuilt motion sensors, based on established methods in previous work [65]. Note that accurately determining the absolute head direction of a VR user is not feasible using only the inbuilt accelerometer of a VR headset. Therefore, our focus is primarily on coordinate calibration rather than precise real-world head posture. Once the acceleration data are transformed into the body coordinate system, we use $Acc_{X'}$, $Acc_{Y'}$, and $Acc_{Z'}$ as calibrated motion sensor data to proceed with further analysis. In addition, as shown above, the $Acc_{Y'}$ presents similar signal patterns to EEG signals in response to user activities at different levels, which we select to reconstruct brain EEG signals for further analysis.



(a) Raw signals. (b) Body movement mitigation.

Figure 8: Recorded acceleration and EEG signals before and after body movement mitigation.



(a) Eye-blink segmentation. (b) LDA distribution.

Figure 9: Eye-blink detection, including variance-based signal segmentation and LDA distributions

Body Movement Mitigation. In practical settings, users generally interact with the VR headset with inevitably subtle body movements. Hence, the embedded motion sensors in the VR headset could capture diverse and unpredictable motion-related patterns, *i.e.*, spontaneous head movements and instinctive body rotations, which introduce extra interference in the collected motion sensor data due to the overlap in the frequency range associated with accelerations. To effectively reduce these noise patterns, we have implemented a generalized *Short-Time Energy* (STE) approach to realize adaptive time-series filtering [26], [66]. We apply a two-second sliding window to the obtained acceleration signal $\mathcal{S}(t)$ and then calculate the total energy of this segment, which is used to determine the presence of significant body motions in the analyzed signals. Specifically, the obtained segment energy is compared with a power threshold, and an adaptive filter is applied when the detected segment energy exceeds the empirically defined power threshold $\mathcal{T}$ (0.5). To obtain the suitable configuration of the adaptive filter and to acquire the filtered sensor signals, assuming the reference signal $\mathcal{S}_R(t)$ is collected when the user's head is absolutely static, we have to solve this optimization problem by establishing the optimal adaptive weight vector $\mathcal{V}_w$ as follows:

$$\begin{aligned} \underset{\mathcal{V}_w}{\text{argmin}} \quad & \sum_{t\in\tau} D_{KL}(\mathcal{S}_R(t), \mathcal{S}(t)) = \sum_{t\in\tau} \mathcal{S}_R(t) \log\left(\frac{\mathcal{S}_R(t)}{\mathcal{S}(t)}\right) \\ \text{subject to} \quad & \sum_{t\in\tau} \mathcal{V}_W(t) = \alpha \sum_{t\in\tau} \mathcal{V}_W(t) + \delta \sum_{t\in\tau} \mathcal{E}(t)\cdot\mathcal{S}^{'}(t), \\ & \sum_{t\in\tau} \mathcal{S}(t) = \sum_{t\in\tau} \mathcal{V}_W(t)\cdot\mathcal{S}^{'}(t), \sum_{t\in\tau} \mathcal{S}^{'}(t)^2 \le \mathcal{T}. \end{aligned} \tag{2}$$

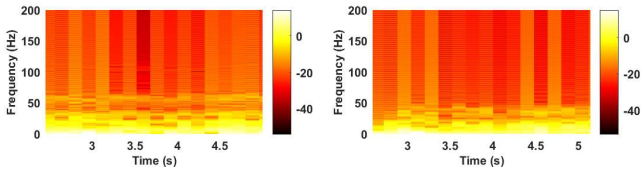
In particular, $\mathcal{S}^{'}(t)$ and $\mathcal{E}(t)$ represent the filtered signals after body movement mitigation and the error function, and τ, α, and δ are the indices of the time-series sensor signal, hyper-parameters, and the optimizing step size, respectively. In addition, we utilize the *Kullback-Leibler* (KL) divergence [67] as the error function to measure discrepancies between the unfiltered and reference signals and to optimize the weights of the adaptive filter because the KL divergence presents a robust ability to quantify differences between the distributions of two series of signals. In particular, Figure 8a and Figure 8b illustrate the acceleration and EEG signals before and after applying the body movement mitigation. We can observe that the patterns of interference resulting from subtle body movements are significantly mitigated.

Eye-Blink Detection. During user interactions with a VR headset, involuntary eye blinks commonly occur. Previous studies (*e.g.*, [41], [42]) have shown that these eye blinks could induce dynamic pupillary responses, leading to detectable vibrations in motion sensor data and EEG signals. These EEG responses to eye blinks vary between individual users because of differences in biometric structure, depicting a potential method for user de-anonymization. To address this, we developed an eye-blink detector that identifies eye blinks from the acceleration data captured by the VR headset's unrestricted motion sensors. Figure 9a illustrates the significant changes in acceleration signals triggered by eye blinks while wearing a Meta Oculus Quest 2 headset. Typically, one eye blink lasts between 0.2 s and 2 s [42]. Hence, in practice, we determine the starting and ending indices of an eye-blink event by applying the moving variance window with a length of 0.2 s and an empirical threshold of 0.01. Furthermore, we calculate the *Linear Discriminant Analysis* (LDA) [68] values of the amplitude within segments of eye-blink and other activities, and use a k-NN detector to accurately identify eye blinks, as detailed in Figure 9b, which achieves 100% accuracy in determining eye blinks from the obtained acceleration signals.

4.3. Feature Extraction

Noise Removal. The accelerometers embedded in headsets are susceptible to high-frequency interference from human speaking frequencies (*e.g.*, 90–255 Hz) and high-frequency electromagnetic radiation at kHz and MHz levels. Therefore, to mitigate the influence of this interference, we apply a *Savitzky-Golay* (S-G) [69] filter to the collected time-series signals, which could remove this high-frequency noise without distorting the signal shapes. In practice, we set the window size to 0.1 of the sampling rate and the polynomial order to 3. Then, considering the different starting amplitude values of acceleration signals, we calculate the average value of the first one-second data as the static acceleration data and then deduct this offset from the motion sensor data to correct for any biases.

Signal Segmentation. After obtaining the filtered signals, we then divide the signal segments that contain relevant user activities and information. As discussed in § 4.2, burst fluctuations in acceleration signals due to pupillary responses can be identified through variance analysis. To this end, we utilize a moving variance window with a threshold of 0.05



(a) Accelerometer. (b) Brain EEG.

Figure 10: Spectrograms of acceleration and EEG signals.

applied to the filtered signals. This approach allows us to detect pupillary responses at the appearance of the first peak. Subsequently, the segment of acceleration data between the first and second peaks in the moving variance signal will be extracted because it contains critical user activity information. In addition, due to the variability in the duration of different pupillary responses (*e.g.*, eye blink: 0.2–2 s, opening a VR app: 1–5 s), we normalize the processed signals of each attempt to the same length of time (*e.g.*, 0.5 s) by exploiting up-sampling (*e.g.*, interpolation [70]) or down-sampling (*e.g.*, decimation factor [71]) algorithms.

Feature Representation. To explore the interconnected features between the acceleration and EEG signal segments comprehensively, we apply the *Short-Term Fourier Transform* (STFT) to these time-series signals, extracting time-frequency spectrogram images as feature representations. We assume that the sampling frequency for the inbuilt accelerometer is f_{acc} and that the processed signal segments are $\overline{S'}(t)$. The FFT of the acceleration signals is given by:

$$\overline{S'}(f) = STFT(\overline{S'}(t)) = \int_0^{t_s} S(t) \cdot \omega(t-\sigma) \cdot e^{-i2\pi ft} dt \quad (3)$$

Specifically, we use a hamming window of size $0.5 f_{acc}$ with an overlapping rate of 50% to convert the segment into discrete expressions. The terms ω and σ represent the window function and the time around the center of the window, respectively. For ground truth EEG signals captured at a different sampling frequency f_{eeg}, we adjust the sliding window size to $0.5 f_{eeg}/f_{acc}$ and convert the time-series EEG signals to a spectrogram with the same time frame length. Figure 10a and Figure 10b illustrate the spectrograms for the acceleration and EEG signals, respectively. These spectrogram images, particularly those related to pupillary response segments shown in Figure 8b, exhibit similar pattern distributions, validating the potential of using acceleration signals to reconstruct brain EEG signals. Hence, these feature representations from both signal types are utilized as inputs and outputs for the brainwave reconstruction model, which will be illustrated in subsequent sections.

4.4. User Privacy Inference

Brainwave Reconstruction. To reconstruct the brainwave, we first build a *Conditional Generative Adversarial Network* (cGAN) model to generate the brain EEG spectrogram $\mathcal{P}_E$ using the given acceleration spectrogram $\mathcal{A}$ as conditions. The cGAN model extracts the mapping relationship from the acceleration spectrogram to the corresponding EEG spectrogram and converts it to time-series brain EEG signals to further uncover VR user privacy. In practice, we leverage the pix2pix [72] image-to-image generative model as the backbone, which has shown remarkable capabilities in transformations between paired training data. Specifically, the cGan model consists of two primary components: a generator $\mathcal{G}_\theta$ and a discriminator $\mathcal{D}_\phi$. The generator is tasked with generating brain EEG spectrogram $\mathcal{P}_E$ that is indistinguishable from the acceleration spectrogram $\mathcal{P}_A$ given the inbuilt accelerometer data from the VR headset. On the other hand, the discriminator tries to distinguish between $\mathcal{P}_E$ and $\mathcal{P}_A$. The generator $\mathcal{G}_\theta$ takes the acceleration spectrogram $\mathcal{A}$ as input and generates a reconstructed spectrogram $\mathcal{P}_E = \mathcal{G}_\theta(\mathcal{A})$. The discriminator $\mathcal{D}_\phi$ receives the groundtruth image pair $(\mathcal{P}_A, \mathcal{A})$ and a reconstructed image pair $(\mathcal{P}_E, \mathcal{A})$ and attempts to fit the EEG signals from the brainwave sensors. The training process involves alternating between updating the parameters and weights of the generator $\mathcal{G}_\theta$ and the discriminator $\mathcal{D}_\phi$. In particular, $\mathcal{D}_\phi$ is trained to maximize its ability to correctly fit the real and generated EEG spectrograms, while $\mathcal{G}_\theta$ is trained against $\mathcal{D}_\phi$ to obtain optimal efficacy.

To realize the training of this cGAN model, we design and implement the objective loss function $\mathcal{L}(\theta, \phi)$ considering both adversarial loss and L1 loss. In particular, adversarial loss ensures that the generated EEG spectrograms are close to those of the EEG signals from brainwave sensors, while L1 loss ensures fidelity to the input acceleration profiles. Hence, the objective loss function is shown as:

$$\mathcal{L}(\theta, \phi) = \mathcal{L}_{adv}(\mathcal{G}_\theta, \mathcal{D}_\phi) + \lambda \mathcal{L}_{L_1}(\mathcal{G}_\theta) \quad (4)$$

where $\mathcal{L}_{adv}$, $\mathcal{L}_{L1}$, and λ are the adversarial loss, L1 loss, and the hyperparameter for balancing the two losses, respectively. Specifically, the adversarial loss $\mathcal{L}_{adv}$ and L1 loss $\mathcal{L}_{L1}$ can be formulated by considering the acceleration and EEG spectrograms as follows:

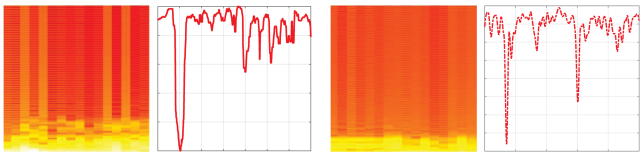
$$\begin{cases} \mathcal{L}_{adv}(\mathcal{G}_\theta, \mathcal{D}_\phi) = \mathbb{E}_{\mathcal{P}_A, \mathcal{A} \sim p(\mathcal{P}_A, \mathcal{A})} \log \mathcal{D}_\phi(\mathcal{P}_A, \mathcal{A}) \\ \qquad\qquad + \mathbb{E}_{\mathcal{A} \sim p(\mathcal{A}), \mathcal{P}_E \sim \mathcal{G}_\theta(\mathcal{A})} \log(1 - \mathcal{D}_\phi(\mathcal{P}_E, \mathcal{A})) \\ \mathcal{L}_{L1}(\mathcal{G}_\theta) = \mathbb{E}_{\mathcal{P}_A, \mathcal{A} \sim p(\mathcal{P}_A, \mathcal{A}), \mathcal{P}_E \sim \mathcal{G}_\theta(\mathcal{A})} |\mathcal{P}_A - \mathcal{P}_E| \end{cases} \quad (5)$$

where we set the batch size to one and used the Adam optimizer with 0.5 momentum rate in both the generator $\mathcal{G}_\theta$ and the discriminator $\mathcal{D}_\phi$, with an initial learning rate of 0.0001 and an initial balance parameter $\lambda = 100$. In addition, we utilize the U-Net [73] as the backbone in $\mathcal{G}_\theta$ and resample both the acceleration and EEG spectrograms to a size of 256×256 and trained the cGAN model for 1000 epochs.

Figure 11b shows an example of reconstructed EEG spectrograms using acceleration spectrograms. Specifically, the reconstructed EEG spectrograms are leveraged to train different activity recognition models to infer UI-level, user-level activity, and privacy. In addition, we apply the *Inverse Short-Term Fourier Transform* (ISTFT) to reconstruct the EEG signals to fit the real brainwave signal as:

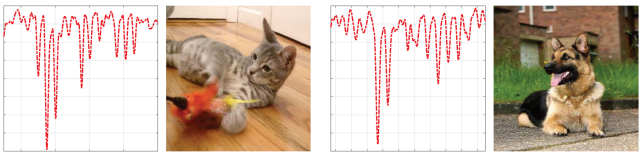
$$\overline{S''}(t) = \frac{1}{\int_0^{t_s} |\omega(t)|^2 \, dt} \int_0^{t_s} \int_0^{t_s} \overline{S'}(f) \cdot \omega(t-\sigma) \cdot e^{i2\pi ft} df \, d\sigma \quad (6)$$

In addition, Figure 11 also presents an example of original and reconstructed EEG signals. We can observe



(a) Original EEG. (b) Reconstructed EEG.

Figure 11: Original and reconstructed time-frequency EEG spectrograms and time-series EEG signals.



(a) Reconstructed EEG of Cat. (b) Reconstructed EEG of Dog.

Figure 12: Reconstructed EEG signals and brain-level visual perception of a "Cat" image and a "Dog" image.

that while most informative patterns are preserved, some information loss occurs due to the signal processing and compression steps involved in STFT and ISTFT. Despite this, the reconstructed EEG signal remains effective as input to fine-tune EEG-to-image models, *i.e.*, DreamDiffusion [44], to infer the brain-level visual perceptions of VR users.

Activity Recognition Model. In particular, we design and implement different activity recognition models to infer the aforementioned three-level user privacy when using the VR headset. First, to infer observable UI-level (*e.g.*, website history, app usage) and user-level (*e.g.*, identification, gaze movements) privacy, we design a transformer-based activity recognition model that takes the reconstructed EEG spectrogram images as input and recognizes the corresponding user activities. Specifically, the model architecture is designed based on the *Vision Transformer* (ViT) [74], which consists of an input layer that resizes and normalizes the input image, a patch embedding layer to split the image into 16×16 patches, a positional encoding to add fixed or learnable positional embeddings to each patch for positional information retention, a transformer encoder with 12 layers of multi-head self-attention and feed-forward networks, and a classification layer that calculates the average pooling over the output and uses a fully-connected layer to map the transformer output to the class with the highest probability. We implement the activity recognition model in PyTorch 2.0 with an initial learning rate of 0.001 and cross-entropy loss as the loss function. The output shape of the last fully-connected layer depends on the corresponding task (*e.g.*, the number of websites, apps, and videos).

Second, to rebuild unobservable brain-level image perceptions, we leverage the open-sourced MOABB dataset [75], which represents the largest available open-sourced EEG-image dataset, to train an EEG encoder that generates valid input for fine-tuning the pre-trained Stable Diffusion model proposed in DreamDiffusion [44]. Specifically, we replace data in 40 common classes with the EEG signals and images collected by our experimental devices (§ 5.1) and train the EEG encoder with a mask ratio of 0.75 over 800 epochs, which we validate as a classification task rather than pixel-faithful image reconstruction. Next, the cross-attention heads and the trained EEG encoder will be jointly optimized with EEG-image pairs to accomplish fine-tuning. The loss function in the Stable Diffusion fine-tuning process can be described as follows:

$$\mathcal{L}_{SD} = \mathbb{E}_{x,\epsilon\sim\mathcal{N}(0,1),t} \left\| \epsilon - \epsilon^{'}(x_t, t, \tau^{'}(y) \right\|_2^2 \quad (7)$$

where x is the given image, y is the output of the EEG encoder, $\tau^{'}$ is the projector to obtain the embedding EEG representations, and $\epsilon^{'}$ is the denoising function implemented as the U-Net.

In the end, the fine-tuned Stable Diffusion model will take the reconstructed EEG signal as input and generate the image prediction that represents the visual perceptions inside the user's brain. For instance, Figure 12a and Figure 12b present the reconstructed EEG signals and the corresponding brain-level image perception of "Cat" and "Dog", respectively. Note that due to the versatile and complex characteristics of human eye perception and brain thought, we take perceptive images as an example to demonstrate the EEG-to-image perception capability of BRAVESPY in this paper, whose feasibility has been widely validated and recognized by most previous relevant studies (*e.g.*, [44]–[46], [63], [64]).

Multi-level Inference. Ultimately, the adversary can exploit BRAVESPY to identify three levels of user activities while wearing a VR headset, thus exposing user privacy and sensitive information. At the UI level, well-documented privacy violations, such as fingerprinting on the website and app [76] and detailed streaming video leakage [77] can lead to the disclosure of metadata related to political and financial affiliations, potentially resulting in credential leakage. Furthermore, BRAVESPY facilitates UI-level activity recognition, which previous efforts could not achieve using only unrestricted motion sensors in the VR headset (*e.g.*, [12]–[14]). At the user level, BRAVESPY can utilize the reconstructed EEG signals to de-anonymize the VR user's identity and recognize gaze movements that might cause credential breaches, such as keystroke inference [78]–[81]. Finally, BRAVESPY is the first work that reveals brainwave-aware user privacy, *i.e.*, unveiling visual perceptive images formed in the brain from reconstructed EEG signals.

5. Evaluation

5.1. Experimental Setup

Experiment Devices. BRAVESPY is evaluated on four commercial VR devices, *i.e.*, Meta Oculus Quest 2, Meta Oculus Quest, PICO 4 All-in-One, and HTC Vive Pro. In particular, the Meta Oculus Quest 2 and Meta Oculus Quest are equipped with a motion sensor board (*i.e.*, 330-00193-03 1PASF8K [26]) originally designed by Meta. The PICO 4 All-in-One adopts the TDK ICM42688 motion tracking module [82], and the HTC Vive Pro contains a G-sensor [83], which consists of an accelerometer and a

(a) Participant setup. (b) Brainwave sensor headband.

Figure 13: Experiment setup for data collection.

gyroscope. In practice, we set the sampling frequency of the motion sensors in the four VR headsets as 500 Hz, which is the most stable sampling frequency for these headsets. We developed a tool to collect motion sensor data in the background from Meta Oculus Quest 2 and Meta Oculus Quest through the function *ovr_GetTrackingState()* on Oculus Mobile SDK [56], as well as implemented a tool for data collection in PICO 4 All-in-One and HTC Vive Pro through the function *getDeviceToAbsoluteTrackingPose()* in OpenVR SDK [84] or *getViewerPose()* in the WebXR Device API [57]. Figure 23 in the Appendix depicts the code snippets implemented in the data collection tools on different VR platforms.

To obtain real brain EEG signals from participants as groundtruth data for training reconstruction models, we utilize ThinkGear brainwave sensor headbands [85], [86] as shown in Figure 13b, which consist of a steel electrode to detect the EEG signals, an ear clip to obtain the reference signals and mitigate interference, and a TGAM board with a Bluetooth module to receive and transmit real-time EEG signals to the paired Android smartphone app. Note that both the collected acceleration and EEG data samples from VR headsets' motion sensors are processed on a desktop computer. cGAN and transformer-based activity recognition models were trained on a NVIDIA RTX A6000 GPU.

Data Collection Process. As shown in Figure 13a, all participants were informed that motion sensor data from VR headsets and EEG signals from the sensors in the headband would be recorded. Specifically, we recruit 25 participants and evaluate BRAVESPY by collecting data samples under the aforementioned three levels of attack surfaces[1]. *(1) UI-level Data Collection:* In this scenario, each participant is required to open 50 websites in default browsers, launch 50 VR apps, and watch 50 streaming videos within immersive VR apps such as NETFLIX. Specifically, the actions of opening websites and apps are repeated for 100 times to collect sufficient data from both the VR's inbuilt motion sensors and the EEG headband. In particular, we collect data samples in the first minute of performing the above activities to guarantee the distinctive patterns in the collected signals induced by visual stimuli. The complete lists of websites, VR apps, and videos are listed in Table 4 in the Appendix. *(2) User-level Data Collection:* For validating the user-level identity recognition, we collect data samples from the 25 participants when starting the VR devices to make their eyes receive the same stimulus and repeat the collection procedure for 100 times. To extract user-level gaze information, we ask each participant to perform eight types of gaze movements (*i.e.*, U: up, D: down, L: left, R: right, UL: up left, UR: up right, DL: down left, and DR: down right) and collect the motion and EEG data in the same procedure. *(3) Brain-level Data Collection:* In this scenario, we collect data samples to demonstrate the feasibility of leveraging the reconstructed brainwave to infer brain-level perceptions, *i.e.*, perceptive images in the human brain. In particular, we show each participant images from common 40 classes exploited in prior visual-to-brain works (*e.g.*, [44]–[46], [63], [64]) while repeatedly collecting the corresponding motion sensor data from VR and EEG signals from the brainwave sensors. Specifically, the data collected from the EEG headband are also utilized to fine-tune the generative models derived from the pre-trained Stable Diffusion models. We note that the requirement of collecting paired training data in the *offline phase* does not violate our threat model, as the adversary can independently recruit participants to build training datasets in advance, while only requiring unrestricted motion sensor data from the victim during the *online deployment* phase.

Evaluation Metrics. We use two metrics to quantify the effectiveness of BRAVESPY in reconstructing different levels of user privacy within virtual scenes, including UI-level website/app/video fingerprinting, user-level identification de-anonymization and gaze movement recognition, as well as brain-level perceptive images inference. *(i) Accuracy:* We select accuracy as the metric for different levels of user privacy recognition, which is defined as the ratio of classes that are correctly identified. *(ii) Confusion matrix:* We evaluate the performance in recognizing user identities and gaze movements with confusion matrices, which visualize the actual target values alongside those predicted by the trained models. In practice, we divide each collected data set into a training set and a testing set with a ratio of $8:2$. Training sets are used to train recognition models, and testing sets are used for performance evaluation.

5.2. Evaluation of UI-level Privacy Inference

Website Fingerprinting Results. Figure 14 shows the evaluation results of BRAVESPY in recognizing the 50 different browsing websites, where it achieves an overall 89.3% accuracy rate. Among the evaluated websites, we observed that websites such as TWITCH.TV (100%), TICKTOCK.COM (100%), and CNN.COM (98%) are more recognizable when being browsed because these websites automatically load video streaming or many high-resolution images, leading to strong visual stimuli for the eyes that cause distinctive pupillary response patterns in reconstructed EEG signals. On the contrary, BRAVESPY presents an accuracy of only 70% and 77% in recognizing websites such as GOOGLE.COM and DUCKDUCKGO.COM, which are widely used search engines with only a simple search bar and static page layout to render. Hence, these websites cause weak pupillary responses and present similar patterns that are easily misclassified.

1. To support the Open Science policy, we make all relevant source code and supporting scripts publicly available on the Zenodo platform via the following permanent access: https://doi.org/10.5281/zenodo.18957102

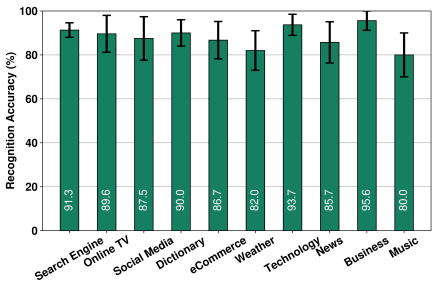


Figure 14: UI-level website fingerprinting results with 50 websites from 10 categories.

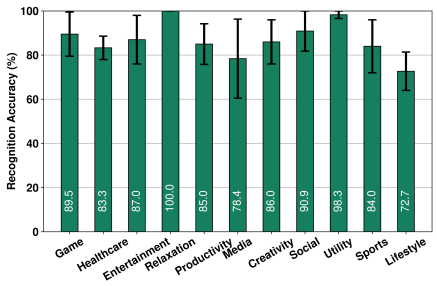


Figure 15: UI-level VR app fingerprinting results with 50 VR apps from 11 categories.

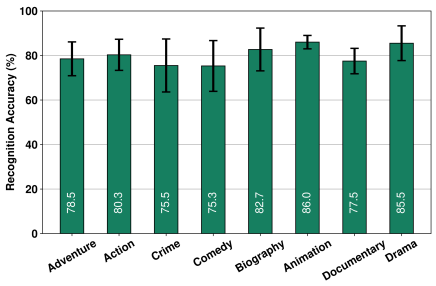


Figure 16: UI-level Netflix video fingerprinting results with 50 videos from 8 categories.

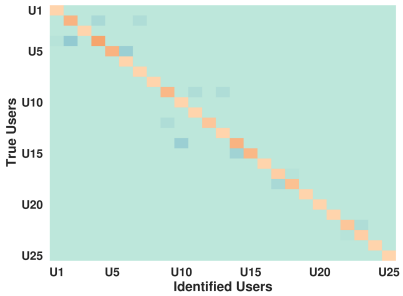


Figure 17: User-level user de-anonymization. U_n: the nth participants.

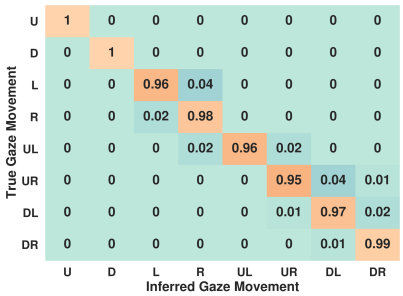


Figure 18: User-level gaze movement recognition with different directions.

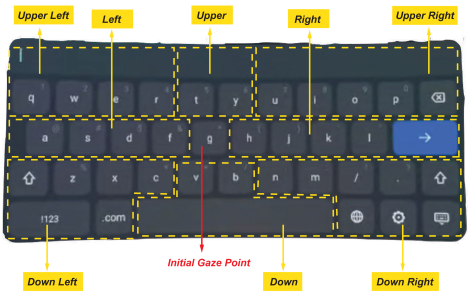


Figure 19: Gaze-based keystroke inference by dividing virtual keyboard to eight zones.

VR App Fingerprinting Results. Furthermore, Figure 15 demonstrates that BRAVESPY achieves an overall 85.8% accuracy in fingerprinting 50 VR apps from different categories. BRAVESPY achieves the highest performance (100%) in fingerprinting VR apps from the game category (*e.g.*, POPULATION: ONE and CARDS & TANKARDS) and the media and streaming category (*e.g.*, NETFLIX and PLUTO TV), as these apps involve dynamic animations in the launch stage, resulting in strong visual stimuli to the user's eyes and causing drastic fluctuations in the reconstructed EEG signals. However, BRAVESPY performs worse in the recognition of apps that adopt the default app launch interface, *i.e.*, a static white background with a simple app logo, such as social apps like MULTIVERSE (63%) and lifestyle apps such as MALOKA (51%), whose reconstructed EEG patterns are plain due to limited visual stimuli.

Netflix Streaming Video Fingerprinting Results. Prior studies (*e.g.*, [19], [88], [89]) highlight the importance of investigating video fingerprinting that may leak user privacy, such as the political or religious interests of certain organizations or regions. Furthermore, in Figure 16, we also present the evaluation results of fingerprinting 50 streaming apps within NETFLIX. In particular, BRAVESPY achieves an average accuracy of 80.0% in video fingerprinting, showcasing the highest performance when the VR user is watching streaming videos from the action category, like BANDIDOS (95%) and THE WAGES OF FEAR (93%), as they contain more dynamic content of activities and quick scene switching, making intense visual stimuli for the human eyes and strong brainwave responses. In contrast, streaming videos such as MEA CULPA (57%) and A KILLER PARADOX (54%) exhibit the worst performance because these two videos contain a lot of bland characters dialog and dark scenes, such that they cannot provide strong stimulation to the human eyes.

5.3. Evaluation of User-level Identity Recognition and Gaze Movement Extraction

User Re-identification Results. User identity in virtual environments is highly sensitive to privacy concerns, which becomes susceptible to identity leakage caused by de-anonymization attacks [20]. Hence, we evaluate BRAVESPY's performance in recognizing the 25 VR participants' identities (denoted as U_1–U_{25}) by collecting the eye-blink clips to reconstruct the corresponding brain EEG signals. Figure 17 shows the evaluation results of BRAVESPY in user-level identity re-identification, where it achieves an overall 98.4% accuracy rate. Specifically, we observe that BRAVESPY performs better in user-level re-identification than in UI-level privacy inference because eye blink results in apparent patterns that contain distinctive biometric characteristics due to the varying facial architecture of the participants. The results also highlight BRAVESPY's potential as a novel user authentication approach, leveraging biometric EEG signals reconstructed by the unrestricted motion sensors within the VR headset.

Gaze Movement Recognition Results. As illustrated in § 5.1, we divide gaze movements into eight different directions to validate BRAVESPY's ability to extract gaze, including upper (U), down (D), left (L), right (R), upper left (UL), upper right (UR), down left (DL), and down right (DR). Overall, Figure 18 depicts that BRAVESPY achieves 97.6% accuracy in recognizing the eight directions of gaze movement. Since gaze movements have been shown to reveal sensitive information (*e.g.*, [78]–[80]),

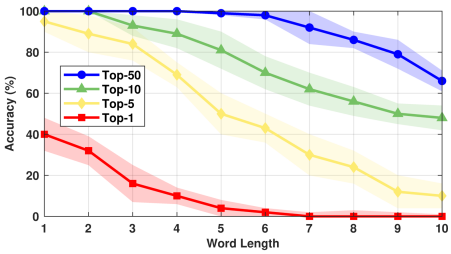


Figure 20: Gaze-based keystroke inference results under top-1, 5, 10, 50 settings.

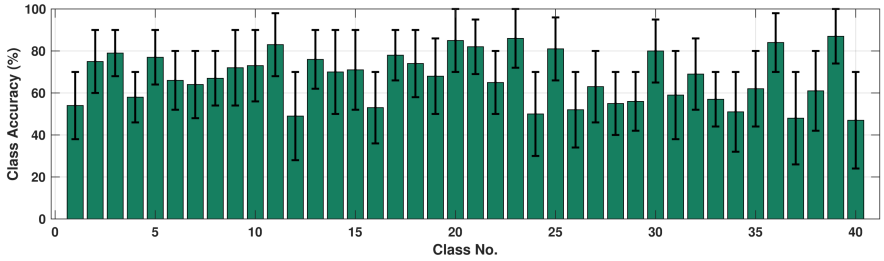


Figure 21: Brain-level visual perception inference, including the image perceptions of 40 classes in the ImageNet dataset [87] of Stable Diffusion.

such as UI-level keystrokes, we further investigate whether BRAVESPY can also infer keystrokes on virtual keyboards from gaze data. Specifically, we designate the key "G" as the focal point and divide the virtual keyboard, as shown in Figure 19, into eight zones corresponding to the directions of gaze movement. This approach stems from the observation that eyes actively search for keys during the typing process. By analyzing the extracted gaze movements, we are then able to generate the top-k candidate combinations (*e.g.*, $k = 1, 5, 10, 50$) for each sequence of keys and subsequently infer the virtual keystrokes.

Figure 20 shows the results of BRAVESPY in inferring 100 words with lengths of one to ten using the extracted gaze movements. The words are randomly selected from the Cambridge English vocabulary list [90]. We can observe that BRAVESPY achieves an average 10.4% top-1 accuracy, 50.6% top-5 accuracy, 74.9% top-10 accuracy, and 92.0% top-50 accuracy, respectively. Despite previous VR keystroke inference attacks (*e.g.*, [12], [14], [15], [34]) that present higher recognition accuracy, our method, BRAVESPY, introduces a novel orthogonal approach to realize UI-level virtual keystroke inference by reconstructing the brain EEG signals. Note that the virtual keystroke inference accuracy of BRAVESPY could be further enhanced if incorporated with previous state-of-the-art gaze analysis methods [78]–[80] and we only investigate its feasibility in this study.

5.4. Evaluation of Brain-level Image Perception

In this section, we further evaluate the performance of BRAVESPY in recognizing brain-level visual perceptions. Due to the non-deterministic characteristics of the Stable Diffusion model, the output perceptual images from the fine-tuned model vary, whereas images in the same class often present similar patterns [44]. Specifically, we utilize *Class Accuracy* (CA) as the evaluation metric to assess BRAVESPY's performance. That is, if the output image and the ground truth image belong to the same class (*e.g.*, "Cat", "Dog", *etc.*), we consider it a successful trial of recognizing brain-level activity.

In Figure 21, the class accuracy of perceptive image inference across the aforementioned 40 classes is presented, demonstrating that BRAVESPY achieves an overall class accuracy rate of 67.2%. In particular, accuracy rates vary significantly between different classes due to image heterogeneity within each class. For example, BRAVESPY achieves high accuracy rates of 79% for the "Butterfly" class (No. 3) and 83% for the "Kayak" class (No. 11), mainly because most of the images in these classes consist of similar and easily recognizable elements, such as green leaves and red kayaks floating on water surfaces, respectively. In contrast, classes with heterogeneous images, such as "Cat" (54%) and "Cellular telephone" (49%), demonstrate lower performance and limited channel capacity, attributed to the diverse variations within these categories. Despite these challenges, BRAVESPY establishes itself as the pioneering work in reconstructing brain-level image perceptions from the inbuilt accelerometer in the VR headset, showcasing competitive performance.

5.5. Ablation Study

To validate the necessity of EEG reconstruction and rule out potential evaluation artifacts, we conduct ablation studies on the Meta Oculus Quest 2 in two dimensions: *(1) Input Type:* We compare raw accelerometer spectrograms against reconstructed EEG spectrograms, and *(2) Evaluation Protocol:* We compare the standard random 8:2 split against a session-disjoint split, in which training and testing data are drawn from separate collection sessions at least one week apart to control for session-specific artifacts, such as headset placement and user posture. Table 3 presents the empirical results of all four combinations, which show that: (1) EEG reconstruction consistently improves performance, with the most significant gain at the brain level, *i.e.*, +26.2% under random split, +23.8% under session-disjoint, which rules out the possibility that the model only predicts class labels from head-motion cues, (2) The session-disjoint protocol introduces only moderate drops, *i.e.*, 3.6%–4.4% for UI-level, 4.1%–4.4% for user-level, and 7.6% for brain-level, which confirms that BRAVESPY captures genuine stimulus-response patterns rather than session-specific correlations, and (3) Even in the strictest setting combining raw accelerometer input with session-disjoint splits, BRAVESPY achieves 74.6% website fingerprinting and 35.8% brain-level perception inference, both substantially above random baselines, *i.e.*, 2.0% and 2.5%, respectively.

6. Robustness of BRAVESPY

Different VR Headsets. Given the diverse hardware designs of various VR headsets, such as differences in accelerome-

TABLE 2: Robustness evaluation of BRAVESPY under different scenarios.

User Privacy	VR Headsets				Unseen Users				Sampling Rates				Light Intensities			
	Oculus 2	Oculus	PICO 4	HTC Vive	$User_{26}$	$User_{27}$	$User_{28}$	$User_{29}$	100 Hz	250 Hz	500 Hz	1000 Hz	25%	50%	75%	100%
Website Fingerprinting (UI)	89.3	87.3	88.7	86.5	87.2	84.2	88.6	83.0	70.5	82.0	89.3	90.0	75.2	86.5	89.3	79.3
VR App Fingerprinting (UI)	85.8	83.6	84.3	82.0	83.5	81.0	84.0	79.1	65.6	80.5	85.8	86.1	69.6	82.7	85.8	74.5
Video Fingerprinting (UI)	80.0	76.0	76.7	75.5	75.7	72.0	78.7	70.5	58.8	73.7	80.0	80.5	64.4	78.7	80.0	70.3
User Re-identification (User)	98.4	92.9	93.2	92.8	–	–	–	–	77.5	91.7	98.4	98.8	77.5	95.5	98.4	88.6
Gaze Recognition (User)	97.6	90.0	90.3	89.6	90.1	86.6	94.2	82.3	74.7	89.6	97.6	98.3	75.0	93.4	97.6	86.5
EEG-based Perception (Brain)	67.2	66.6	52.5	52.0	56.5	50.8	62.3	48.4	39.6	63.5	67.2	68.0	46.6	62.0	67.2	58.8

TABLE 3: Ablation study results: *Raw Acc.* (raw accelerometer) and *Recon. EEG* (reconstructed EEG) under *random split* and *session-disjoint* split. Δ_{EEG}: Improvement from EEG reconstruction under each split. Δ_{SD}: Drop from session-disjoint evaluation.

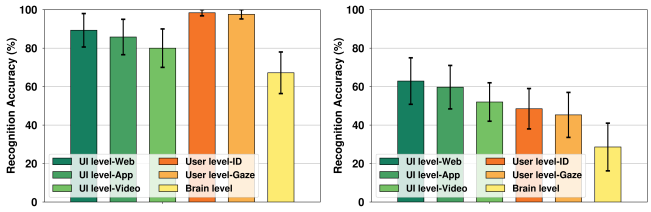
Multi-level User Privacy	Random Split			Session-Disjoint			Δ_{SD}	
	Raw Acc.	Recon. EEG	Δ_{EEG}	Raw Acc.	Recon. EEG	Δ_{EEG}	Raw Acc.	Recon. EEG
Website Fingerprinting (UI)	78.5	89.3	▲ +10.8	74.6	85.7	▲ +11.1	▼ −3.9	▼ −3.6
VR App Fingerprinting (UI)	73.2	85.8	▲ +12.6	69.1	81.4	▲ +12.3	▼ −4.1	▼ −4.4
Video Fingerprinting (UI)	67.8	80.0	▲ +12.2	63.5	75.8	▲ +12.3	▼ −4.3	▼ −4.2
User Re-identification (User)	89.6	98.4	▲ +8.8	84.8	94.0	▲ +9.2	▼ −4.8	▼ −4.4
Gaze Recognition (User)	85.3	97.6	▲ +12.3	80.7	93.5	▲ +12.8	▼ −4.6	▼ −4.1
EEG-based Perception (Brain)	41.0	67.2	▲ +26.2	35.8	59.6	▲ +23.8	▼ −5.2	▼ −7.6

ters and gyroscopes, the recorded acceleration signals may exhibit subtle variations in patterns. To evaluate the robustness and extensibility of BRAVESPY to other VR headsets, we carried out further experiments applying the Meta Oculus Quest 2 to data samples collected from three other VR devices, including the Meta Oculus Quest, PICO 4 All-in-One, and HTC Vive Pro. The evaluation results of the three-level user privacy inference based on reconstructed brain EEG signals from these three VR headsets are depicted in Table 2. Notably, BRAVESPY maintains high recognition accuracy in UI-level inference, achieving 86.5%–88.7% in fingerprinting 50 websites, 82.0%–84.3% in fingerprinting 50 VR apps, and 75.5%–76.7% in fingerprinting 50 streaming videos on Netflix, respectively. However, there is a slight decrease in BRAVESPY's performance in user-level recognition, from 98.4% to 92.9%, 93.2%, and 92.8% in user re-identification and from 97.6% to 90.0%, 90.3%, and 89.6% in gaze movement extraction and keystroke inference. In particular, there is a significant decrease in BRAVESPY's performance in brain-level perception recognition when directly transferred to PICO 4 and HTC Vive VR headsets, decreasing from 67.2% to 52.5%, 52.0%, respectively. This decrease can be attributed to the varying sensitivity of the inbuilt motion sensors in these VR headsets from different brands, leading to additional perturbations in the input acceleration signals and resulting in a performance decrease in the EEG reconstruction models. Nevertheless, this issue can be mitigated by collecting more data samples from the new VR headset to fine-tune the pre-trained recognition models.

Different VR Users. Due to the inherent variability in biometric facial characteristics among different participants, pupillary responses to identical visual stimuli can exhibit subtle differences, introducing variations in reconstructed EEG signals. To comprehensively understand the impact of participant diversity, we recruit four additional participants, denoted as U_{26}–U_{29}, and collect unseen data samples from them. We then apply the pre-trained models to evaluate BRAVESPY's robustness across these new users. Specifically, the evaluation results shown in Table 2 reveal that BRAVESPY's performance presents an average decrease of 4.4% across various UI-level inference, *i.e.*, 3.6% website fingerprinting, 3.9% app fingerprinting, and 5.8% video fingerprinting. Likewise, we observe a decrease of 9.3% in user-level gaze movement recognition and keystroke inference, along with a decrease of 12.7% in brain-level perception recognition. However, BRAVESPY maintains acceptable performance in inferring fine-grained user privacy, and the observed reduction in performance was predominantly due to user variability resulting from the limited datasets collected for model training. Thus, addressing this limitation involves collecting data samples from more extensive and diverse participants.

Different Accelerometer Sampling Rates. In § 5.1, we set the sampling rate of the inbuilt accelerometer to 500 Hz to capture acceleration signals to reconstruct brain EEG signals. As different sampling rates encode user activity information with varying granularity, we conducted additional experiments to explore their impact by collecting data samples from the accelerometer at rates of 100 Hz, 250 Hz, 500 Hz, and 1,000 Hz using the Meta Oculus Quest 2 VR headset, then assessed performance using pre-trained models at the default 500 Hz rate, respectively. The evaluation results, depicted in Table 2, reveal an increase in BRAVESPY's performance with higher sampling rates. When setting the sampling rate at 100 Hz, BRAVESPY exhibited significantly lower performance compared to the default rate, achieving only 65.0% in UI-level, 76.1% in user-level, and 39.6% in brain-level inference, respectively. In particular, as sampling rates beyond 500 Hz offer marginal increments because of the limited frequency range of pupillary responses, we conclude that 500 Hz achieves a balance between BRAVESPY's performance and stealthiness, considering the higher energy consumption associated with higher sampling frequencies, which may raise suspicions.

Different Light Intensities of the Virtual Scene. As shown in previous studies [91], [92], light intensity



(a) Before applying defense. (b) After applying defense.

Figure 22: Empirical results of BRAVESPY before and after applying sensor signal obfuscation defense method.

could impact the response time of the human eye. Since BRAVESPY relies on pupillary responses to reconstruct brain EEG signals and infer user privacy, the light intensity of virtual scenes (*a.k.a.*, brightness) becomes a significant factor influencing visual stimuli. In our previous evaluation detailed in § 5, we maintain default settings with a brightness level set to 75%. To study the impact of varying light intensities on the performance of BRAVESPY, we adjusted the brightness levels to 25%, 50%, and 100%. The evaluation results, presented in Table 2, show the optimal performance of BRAVESPY at the 50% and 75% brightness levels. We also observe that BRAVESPY's performance diminishes under low-light conditions (25% brightness) due to weak visual stimuli, which lead to less apparent pattern changes in the pupillary responses. On the other hand, at 100 % brightness, the increased light intensity decreases the eyes' sensitivity to content changes in the VR headset's head-mounted display, resulting in a decrease in pupillary responses and pattern distinctiveness associated with various user activities and lower recognition accuracy rates.

7. Countermeasures

Permission-based and Privacy-aware Management. Similar to previous works (*e.g.*, [12]–[16]), BRAVESPY requires access to unrestricted motion sensors in the VR headset. Hence, a potential countermeasure to defend against privacy leakage is to design and implement a privacy-based and permission-based sensor management scheme in the VR platform, as proposed on Android platforms [93], [94]. For instance, VR users would need to grant permission for VR apps or webpages to access specific sensors during both installation and runtime. Furthermore, to improve transparency and control over data usage, the sensor management scheme should be privacy-aware, providing comprehensive information on the types of sensors, their activation times, durations of operation, and data collection activities in both the foreground and background. Using contextual information, the privacy-aware sensor management scheme could allow users to create and update customized access control policies for all embedded sensors, enabling them to limit data access for background apps. Additionally, the framework assesses the quality of sensor data flow to VR apps and allows users to customize specific data attributes (*e.g.*, sampling frequency, resolution) for a specific VR app.

Sensor Signal Obfuscation. Another countermeasure is to apply a signal obfuscation method when accessing data from unrestricted motion sensors at a specific time. For example, the VR system could add random noise (*e.g.*, Gaussian white noise) to the motion sensor data when launching apps or opening websites, which could prevent the UI-level privacy inference from BRAVESPY. Besides, we could exploit data transformation and encryption methods [95]–[97] to mask motion sensor data to hide sensitive information and retain data that are necessary to support the specific VR app to reduce recognition effectiveness and mitigate threats from user-level de-anonymization and gaze data leakage, as well as brain-level perceptions inference. In practice, we designed an automatic, software-defined mechanism to apply Gaussian white noise to the recorded data from unrestricted motion sensors in Meta Oculus Quest 2 with a standard deviation of 0.1 and evaluated BRAVESPY's performance on UI-level, user-level, and brain-level privacy inferences. Figure 22 shows that the accuracy of the three levels of privacy inference decreases approximately 26.8%, 51.1%, and 38.6%, respectively. However, obfuscating the signals from unrestricted motion sensors could also decrease the usability of VR headsets (*e.g.*, long latency, slower response), which rely on these sensor readings to provide their real-time position and movement tracking ability. In addition, recent studies [14], [16] have demonstrated the effectiveness of defending against inference attacks by randomly shuffling the layouts of some privacy-related UI, such as virtual keyboards, when entering sensitive information like accounts and passwords.

8. Limitations

Although BRAVESPY reveals both observable and unobservable user privacy leakage, our framework still has some limitations. First, as a proof-of-concept study, we only evaluated one type of brain activity, *i.e.*, validated its class-level perception inference across 40 image categories shown in DreamDiffusion [44], and additional EEG-image pairs from other brain-level activity inference will require fine-tuning models or utilizing EEG-ImageNet [98] (*e.g.*, 80 image classes), or exploiting EEG-to-video backbones like EEG2Video [99] to expand both the scope and granularity of brain-level inference. Second, it shows limited generalizability and performs well in relatively controlled scenarios such as browsing or light interaction, and the brain-level perception inference degrades across different headsets, which could be mitigated by cross-subject domain adaptation, *i.e.*, domain-adversarial neural networks [99] and adaptive feature representation learning methods [100]. In addition, while the subtle signals of interest could potentially be overwhelmed by intensive movements, our body movement mitigation module (§ 4.2) addresses this issue and improves the signal-to-noise ratio (SNR) of the acceleration signals from approximately 11–21 dB. Moreover, prior studies have demonstrated even weaker physiological signals from VR motion sensors at comparable or lower SNR levels, such as heartbeats (5–15 dB) [26] and blood pressure-related vibrations (5–10 dB) [28], which further validate the feasibility of our approach.

9. Related Works

Attacks on VR Devices. With the rise of the Metaverse, recent studies have investigated VR attacks to steal sensitive information, *e.g.*, and keystrokes on the virtual keyboard. Most of the previous work installs malicious websites or apps to acquire readings from unrestricted motion sensors inside the VR headset and trains specific DNN models to reveal virtual keyboard input (*e.g.*, [12]–[17], [21], [26], [33], [58], [101]). Face-Mic [23] and ImmerSpy [24] show the feasibility of speech eavesdropping using zero-permission motion data. Recent studies also utilize cameras to record user videos while wearing the VR headset and infer keyboard input [16], [32] from watching videos [19] in virtual scenes and other user's keystrokes in the surrounding real environment [102]. Moreover, attackers can exploit virtual avatars to de-anonymize user identities [20] and infer real-world keyboard typing through the avatars' movements [22], [31]. VR-Spy [30], Heimdall [34] and VRecKey [35] present side-channel attacks using Wi-Fi signals, button click sounds, and leaked infrared signals from VR controllers to infer virtual keystrokes, respectively. In addition, recent research efforts have shown that attackers could use motion sensor data to infer biometric features, such as respiration rates and heartbeats [26], gender and body fat ratio [25], and blood pressure [28]. Compared with previous work, our work, BRAVESPY, shows the feasibility of brainwave reconstruction with unrestricted motion sensors in VR headsets, which could lead to various attack surfaces.

Brain Activity Inference via BCI. Most of the initial research on neural decoding and brain activity inference focuses on exploiting functional magnetic resonance imaging (fMRI) (*e.g.*, [103], [104]) or electroencephalography (ECoG) [105]. Specifically, researchers scan the human brain with professional BCI facilities (*e.g.*, fMRI scanner [106]) under different visual stimuli from real images, and then input the fMRI data into generative neural networks (*e.g.*, GAN [107], diffusion models [108]) to uncover brain activities. However, due to the difficulty of obtaining fMRI data, recent studies have demonstrated the feasibility of leveraging EEG signals extracted from human brains to decode brain activities, *i.e.*, perceptive images (*e.g.*, [44], [46], [63]), which can be extracted using non-invasive and portable BCI devices (*e.g.*, brainwave sensor headbands [109]). In particular, a recent study [110] also explored the potential attack vectors from the side of the operating, system as well as various on-device models integrated with commercial BCI platforms. Our work, BRAVESPY, becomes the *first* work to demonstrate the feasibility of bridging the gap between the Metaverse and BCI by utilizing unrestricted motion sensors in head-mounted VR headsets to reconstruct brain EEG signals.

10. Conclusion

We propose BRAVESPY, a novel system for reconstructing brain EEG signals by exploiting inbuilt unrestricted motion sensors of a VR headset. Specifically, BRAVESPY leverages the acceleration signals induced by pupillary responses under various visual stimuli to reconstruct EEG signals. Then, the adversary can exploit the reconstructed brain EEG signals to infer observable UI-level user activities (*e.g.*, browsing websites, launching apps, and streaming videos), user-level identity de-anonymization and gaze movement extraction, and unobservable brain-level image perception. The extensive evaluation suggests that BRAVESPY achieves high accuracy in recognizing the privacy of three levels of users within the virtual world and the human mind. To our knowledge, BRAVESPY is the first study to highlight unexplored avenues of unobservable activities of brain-level perception inference among all newly proposed VR-related research.

Ethics Considerations

We take ethical considerations seriously, and this study was approved by the Human Subjects Ethics Sub-Committee of City University of Hong Kong (No. HU-STA-00000169). We recruited 25 volunteers from university students and staff (15 males and 10 females, with ages ranging from 18 to 35) for data collection in this study. In particular, we observed that only 11 participants have previous VR experience, while the other 14 participants do not have knowledge of using VR devices. Hence, we asked participants to perform different activities in VR devices for one hour to become familiar with VR interactions while wearing the headband with inbuilt EEG brainwave sensors before the official data collection.

Before the experiments, each participant must sign a written consent form that allows us to collect data on human behavior for evaluation. During the experiments, participants could move slightly while sitting or standing, as they were casually playing with the VR device, and we only used our own accounts on VR platforms to browse websites, run apps, play VR games, and type keystrokes. Considering the dizziness and motion sickness associated with VR usage, we require participants to rest for 1–10 minutes before reporting their conditions after collecting data with different time durations in each trial. Note that we spent one year on the data collection, and all collected data samples are securely stored on a encrypted local server to prevent any form of privacy leakage pertaining to the volunteers, which is only accessible by the authorized principal instructor and the graduate students in this project.

Acknowledgment

We sincerely appreciate our shepherd and all anonymous reviewers for their constructive feedback and invaluable comments. This work was fully supported by the Hong Kong Research Grants Council (RGC) under Grants CityU 21219223, 11219524, 11219025, RFS2122-1S04, C1029-22G, C6015-23G, CRS_HKUST601/24, DON_RMG 9229170, and AC-202403-02-15. Any opinions, findings, and conclusions in this paper are those of the authors and are not necessarily those of the supported organizations.

References

[1] M. Quest, "Oculus Quest 2 Store: VR Games, Apps, & More," 2023, https://www.oculus.com/experiences/quest.

[2] H. Guo, H.-N. Dai, X. Luo, Z. Zheng, G. Xu, and F. He, "An empirical study on Oculus virtual reality applications: Security and privacy perspectives," in *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering (ICSE)*, 2024.

[3] A. J. Aviv, B. Sapp, M. Blaze, and J. M. Smith, "Practicality of accelerometer side channels on smartphones," in *Proceedings of the Annual Computer Security Applications Conference (ACSAC)*, 2012.

[4] L. Simon and R. Anderson, "Pin Skimmer: Inferring PINs through the camera and microphone," in *Proceedings of ACM Workshop on Security and Privacy in Smartphones and Mobile Devices*, 2013.

[5] L. Lu, J. Yu, Y. Chen, Y. Zhu, X. Xu, G. Xue, and M. Li, "Keylistener: Inferring keystrokes on QWERTY keyboard of touch screen through acoustic signals," in *Proeedings of IEEE International Conference on Computer Communications (INFOCOM)*, 2019.

[6] N. Matyunin, Y. Wang, T. Arul, K. Kullmann, J. Szefer, and S. Katzenbeisser, "Magneticspy: Exploiting magnetometer in mobile devices for website and application fingerprinting," in *Proc. of ACM Workshop on Privacy in the Electronic Society*, 2019.

[7] H. Pan, L. Yang, H. Li, C.-W. You, X. Ji, Y.-C. Chen, Z. Hu, and G. Xue, "Magthief: Stealing private app usage data on mobile devices via built-in magnetometer," in *Proceedings of IEEE International Conference on Sensing, Communication, and Networking (SECON)*, 2021.

[8] R. Ning, C. Wang, C. Xin, J. Li, and H. Wu, "Deepmag: Sniffing mobile apps in magnetic field through deep convolutional neural networks," in *Proceedings of IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2018.

[9] S. A. Anand and N. Saxena, "Speechless: Analyzing the threat to speech privacy from smartphone motion sensors," in *Proceedings of IEEE Symposium on Security and Privacy (SP)*, 2018.

[10] Z. Ba, T. Zheng, X. Zhang, Z. Qin, B. Li, X. Liu, and K. Ren, "Learning-based practical smartphone eavesdropping with built-in accelerometer," in *Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2020.

[11] P. Hu, H. Zhuang, P. S. Santhalingam, R. Spolaor, P. Pathak, G. Zhang, and X. Cheng, "AccEar: Accelerometer acoustic eavesdropping with unconstrained vocabulary," in *Proceedings of IEEE Symposium on Security and Privacy (SP)*, 2022.

[12] S. Luo, X. Hu, and Z. Yan, "Holologger: Keystroke inference on mixed reality head-mounted displays," in *Proceedings of IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2022.

[13] C. Slocum, Y. Zhang, N. Abu-Ghazaleh, and J. Chen, "Going through the motions: AR/VR keylogging from user head motions," in *Proceedings of USENIX Security Symposium*, 2023.

[14] Y. Wu, C. Shi, T. Zhang, P. Walker, J. Liu, N. Saxena, and Y. Chen, "Privacy leakage via unrestricted motion-position sensors in the age of virtual reality: A study of snooping typed input on virtual keyboards," in *Proceedings of IEEE Symposium on Security and Privacy (SP)*, 2023.

[15] Y. Zhang, C. Slocum, J. Chen, and N. Abu-Ghazaleh, "It's all in your head (set): Side-channel attacks on AR/VR systems," in *Proceedings of USENIX Security Symposium*, 2023.

[16] Ü. Meteriz-Yıldıran, N. F. Yıldıran, A. Awad, and D. Mohaisen, "A keylogging inference attack on air-tapping keyboards in virtual environments," in *Proceedings of IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2022.

[17] S. Lee and W. Choi, "Eyes on your typing: Snooping finger motions on virtual keyboards," in *Proceedings of IEEE Symposium on Security and Privacy (SP)*, 2025.

[18] C. V. J. F. H. Khalili and N. S. O. Abari, "Xr devices send wifi packets when they should not: Cross-building keylogging attacks via non-cooperative wireless sensing," in *Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2026.

[19] A. Nguyen, X. Zhang, and Z. Yan, "Penetration vision through virtual reality headsets: Identifying 360 videos from head movements," in *Proceedings of USENIX Security Symposium*, 2024.

[20] Y. Meng, Y. Zhan, J. Li, S. Du, H. Zhu, and X. S. Shen, "De-anonymization attacks on metaverse," in *Proceedings of IEEE Conference on Computer Communications (INFOCOM)*, 2023.

[21] V. Nair, W. Guo, J. Mattern, R. Wang, J. F. O'Brien, L. Rosenberg, and D. Song, "Unique identification of 50,000+ virtual reality users from head & hand motion data," in *Proceedings of USENIX Security Symposium*, 2023.

[22] Z. Yang, Z. Sarwar, I. Hwang, R. Bhaskar, B. Y. Zhao, and H. Zheng, "Can virtual reality protect users from keystroke inference attacks?" in *Proceedings of USENIX Security Symposium*, 2024.

[23] C. Shi, X. Xu, T. Zhang, P. Walker, Y. Wu, J. Liu, N. Saxena, Y. Chen, and J. Yu, "Face-Mic: Inferring live speech and speaker identity via subtle facial dynamics captured by AR/VR motion sensors," in *Proceedings of Annual International Conference on Mobile Computing and Networking*, 2021.

[24] D. Cayir, R. Mohamed, R. Lazzeretti, M. Angelini, A. Acar, M. Conti, Z. B. Celik, and S. Uluagac, "Speak up, i'm listening: Extracting speech from zero-permission VR sensors," in *Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2025.

[25] T. Zhang, C. Shi, P. Walker, Z. Ye, Y. Wang, N. Saxena, and Y. Chen, "Passive vital sign monitoring via facial vibrations leveraging AR/VR headsets," in *Proceedings of Annual International Conference on Mobile Systems, Applications and Services (MobiSys)*, 2023.

[26] T. Zhang, Z. Ye, A. T. Mahdad, M. M. R. R. Akanda, C. Shi, Y. Wang, N. Saxena, and Y. Chen, "FaceReader: Unobtrusively mining vital signs and vital sign embedded sensitive info via AR/VR motion sensors," in *Proceedings of ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2023.

[27] T. Zhang, Q. Ji, M. M. Rahman, R. Akanda, Z. Ye, A. T. Mahdad, C. Shi, Y. Wang, N. Saxena, and Y. Chen, "Harnessing vital sign vibration harmonics for effortless and inbuilt xr user authentication," in *Proceedings of ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2025.

[28] Z. Ye, A. T. Mahdad, Y. Wang, C. Shi, Y. Chen, and N. Saxena, "BPSniff: Continuously surveilling private blood pressure information in the metaverse via unrestricted inbuilt motion sensors," in *Proceedings of IEEE Symposium on Security and Privacy (SP)*, 2025.

[29] S. Son, C. Mukherjee, R. M. Aburas, B. Gulmezoglu, and Z. B. Celik, "Side-channel inference of user activities in AR/VR using GPU profiling," in *Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2026.

[30] A. Al Arafat, Z. Guo, and A. Awad, "VR-Spy: A side-channel attack on virtual key-logging in VR headsets," in *Proc. of IEEE VR*, 2021.

[31] Z. Su, K. Cai, R. Beeler, L. Dresel, A. Garcia, I. Grishchenko, Y. Tian, C. Kruegel, and G. Vigna, "Remote keylogging attacks in multi-user VR applications," in *Proceedings of USENIX Security Symposium*, 2024.

[32] S. R. K. Gopal, D. Shukla, J. D. Wheelock, and N. Saxena, "Hidden Reality: Caution, your hand gesture inputs in the immersive virtual world are visible to all!" in *Proceedings of USENIX Security Symposium*, 2023.

[33] H. Khalili, A. Chen, T. Papaiakovou, T. Jacques, H.-J. Chien, C. Liu, A. Ding, A. Hass, S. Zonouz, and N. Sehatbakhsh, "Virtual keymysteries unveiled: Detecting keystrokes in VR with external side-channels," in *Proceedings of IEEE Symposium on Security and Privacy (SP) Workshops*, 2024.

[34] S. Luo, A. Nguyen, H. Farooq, K. Sun, and Z. Yan, "Eavesdropping on controller acoustic emanation for keystroke inference attack in virtual reality," in *Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2024.

[35] T. Ni, Y. Du, Q. Zhao, and C. Wang, "Non-intrusive and unconstrained keystroke inference in VR platforms via infrared side channel," in *Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2025.

[36] J. Smythies, "How the brain decides what we see," *Journal of the Royal Society of Medicine*, 2005.

[37] I. E. Dror, "Perception is far from perfection: the role of the brain and mind in constructing realities," *Behavioral and Brain Sciences*, 2005.

[38] A. Dwarakanath, V. Kapoor, J. Werner, S. Safavi, L. A. Fedorov, N. K. Logothetis, and T. I. Panagiotaropoulos, "Bistability of prefrontal states gates access to consciousness," *Neuron*, 2023.

[39] S. Abbas, N. Okdeh, R. Roufayel, H. Kovacic, J.-M. Sabatier, Z. Fajloun, and Z. Abi Khattar, "Neuroarchitecture: How the perception of our surroundings impacts the brain," *Biology*, 2024.

[40] D. E. L. Lockhofen and C. Mulert, "Neurochemistry of visual attention," *Frontiers in neuroscience*, 2021.

[41] L.-W. Ko, O. Komarov, W.-K. Lai, W.-G. Liang, and T.-P. Jung, "Eyeblink recognition improves fatigue prediction from single-channel forehead EEG in a realistic sustained attention task," *Journal of neural engineering*, 2020.

[42] M. Shahbakhti, M. Beiramvand, I. Rejer, P. Augustyniak, A. Broniec-Wójcik, M. Wierzchon, and V. Marozas, "Simultaneous eye blink characterization and elimination from low-channel prefrontal EEG signals enhances driver drowsiness detection," *IEEE Journal of Biomedical and Health Informatics (JBHI)*, 2021.

[43] M. M. Gusso, K. L. Christison-Lagay, D. Zuckerman, G. Chandrasekaran, S. I. Kronemer, J. Z. Ding, N. C. Freedman, P. Nohama, and H. Blumenfeld, "More than a feeling: scalp EEG and eye signals in conscious tactile perception," *Consciousness and Cognition*, 2022.

[44] Y. Bai, X. Wang, Y.-p. Cao, Y. Ge, C. Yuan, and Y. Shan, "Dreamdiffusion: Generating high-quality images from brain EEG signals," *arXiv preprint arXiv:2306.16934*, 2023.

[45] Y.-T. Lan, K. Ren, Y. Wang, W.-L. Zheng, D. Li, B.-L. Lu, and L. Qiu, "Seeing through the brain: image reconstruction of visual perception from human brain signals," *arXiv preprint arXiv:2308.02510*, 2023.

[46] K. M. Davis, C. de la Torre-Ortiz, and T. Ruotsalo, "Brain-supervised image editing," in *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[47] P. M. Reis, F. Hebenstreit, F. Gabsteiger, V. von Tscharner, and M. Lochmann, "Methodological aspects of eeg and body dynamics measurements during motion," *Frontiers in Human Neuroscience*, 2014.

[48] M. F. Land, "Eye movements and the control of actions in everyday life," *Progress in Retinal and Eye Research*, 2006.

[49] G. Slanzi, J. A. Balazs, and J. D. Velásquez, "Combining eye tracking, pupil dilation and EEG analysis for predicting web users click intention," *Information Fusion*, 2017.

[50] S. Park and M. Whang, "Infrared camera-based non-contact measurement of brain activity from pupillary rhythms," *Frontiers in physiology*, 2018.

[51] H. Wang, Z. Zhan, H. Shan, S. Dai, M. Panoff, and S. Wang, "GAZEploit: Remote keystroke inference attack by gaze estimation from avatar views in VR/MR devices," in *Proceedings of ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2024.

[52] J. Li, Y. Meng, Y. Zhan, L. Zhang, and H. Zhu, "Dangers behind charging VR devices: Hidden side channel attacks via charging cables," *IEEE Transactions on Information Forensics and Security (TIFS)*, 2024.

[53] T. Zhang, Q. Ji, Z. Ye, M. M. R. R. Akanda, A. T. Mahdad, C. Shi, Y. Wang, N. Saxena, and Y. Chen, "SAFARI: Speech-associated facial authentication for AR/VR settings via robust vibration signatures," in *Proceedings of ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2024.

[54] H. Zhu, M. Xiao, D. Sherman, and M. Li, "Soundlock: A novel user authentication scheme for VR devices using auditory-pupillary response." in *Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2023.

[55] Khronos, "OpenX software development kit (SDK) sources project," https://github.com/KhronosGroup/OpenXR-SDK-Source, 2023.

[56] M. Quest, "Oculus mobile SDK," https://developer.oculus.com/downloads/package/oculus-mobile-sdk/, 2021.

[57] Khronos, "OpenX software development kit (SDK) sources project," https://immersiveweb.dev/, 2023.

[58] Z. Ling, Z. Li, C. Chen, J. Luo, W. Yu, and X. Fu, "I know what you enter on gear VR," in *Proceedings of IEEE Conference on Communications and Network Security (CNS)*, 2019.

[59] V. Nair, G. M. Garrido, and D. Song, "Exploring the unprecedented privacy risks of the metaverse," *arXiv:2207.13176*, 2022.

[60] R. Kardon, "Pupillary light reflex," *Current opinion in ophthalmology*, 1995.

[61] K. Nozaki, K. Kamiya, Y. Matsue, N. Hamazaki, R. Matsuzawa, S. Tanaka, E. Maekawa, T. Kishi, A. Matsunaga, T. Masuda *et al.*, "Pupillary light reflex as a new prognostic marker in patients with heart failure," *Journal of Cardiac Failure*, 2019.

[62] J. Zhang, S. Park, A. Cho, and M. Whang, "Recognition of empathy from synchronization between brain activity and eye movement," *Sensors*, 2023.

[63] I. Kavasidis, S. Palazzo, C. Spampinato, D. Giordano, and M. Shah, "Brain2image: Converting brain signals into images," in *Proceedings of ACM International Conference on Multimedia (MM)*, 2017.

[64] P. Tirupattur, Y. S. Rawat, C. Spampinato, and M. Shah, "Thoughtviz: Visualizing human thoughts using generative adversarial network," in *Proceedings of ACM International Conference on Multimedia (MM)*, 2018.

[65] N. Mohssen, R. Momtaz, H. Aly, and M. Youssef, "It's the human that matters: accurate user orientation estimation for mobile computing applications," in *Proceedings of International Conference on Mobile and Ubiquitous Systems (MobiQuitous)*, 2014.

[66] X. Li, S. Liang, S. Yan, J. Ryu, and Y. Wu, "Adaptive detection of Ahead-sEMG based on short-time energy of local-detail difference and recognition in advance of upper-limb movements," *Biomedical Signal Processing and Control*, 2023.

[67] Y. Huang, Y. Zhang, and J. A. Chambers, "A novel Kullback–Leibler divergence minimization-based adaptive student's t-filter," *IEEE Transactions on Signal Processing*, 2019.

[68] P. Xanthopoulos, P. M. Pardalos, T. B. Trafalis, P. Xanthopoulos, P. M. Pardalos, and T. B. Trafalis, "Linear discriminant analysis," *Robust data mining*, 2013.

[69] J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh, "A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter," *Remote sensing of Environment*, 2004.

[70] O. Rukundo and H. Cao, "Nearest neighbor value interpolation," *arXiv preprint arXiv:1211.1768*, 2012.

[71] D. M. Kreindler and C. J. Lumsden, "The effects of the irregular sample and missing data in time series analysis," in *Nonlinear Dynamical Systems Analysis for the Behavioral Sciences Using Real Data*, 2016.

[72] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[73] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015.

[74] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[75] V. Jayaram and A. Barachant, "MOABB: trustworthy algorithm benchmarking for BCIs," *Journal of neural engineering*, 2018.

[76] A. Panchenko, F. Lanze, J. Pennekamp, T. Engel, A. Zinnen, M. Henze, and K. Wehrle, "Website fingerprinting at internet scale." in *Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2016.

[77] R. Schuster, V. Shmatikov, and E. Tromer, "Beauty and the burst: Remote identification of encrypted video streams," in *Proceedings of USENIX Security Symposium*, 2017.

[78] Y. Wang, W. Cai, T. Gu, and W. Shao, "Your eyes reveal your secrets: An eye movement based password inference on smartphone," *IEEE Transactions on Mobile Computing*, 2019.

[79] Y. Chen, T. Li, R. Zhang, Y. Zhang, and T. Hedgpeth, "Eyetell: Video-assisted touchscreen keystroke inference from eye movements," in *Proceedings of IEEE Symposium on Security and Privacy (SP)*, 2018.

[80] Y. Abdrabou, J. Schütte, A. Shams, K. Pfeuffer, D. Buschek, M. Khamis, and F. Alt, "Your eyes tell you have used this password before: Identifying password reuse from gaze and keystroke dynamics," in *Proceedings of ACM CHI Conference on Human Factors in Computing Systems*, 2022.

[81] T. Ni, X. Zhang, C. Zuo, J. Li, Z. Yan, W. Wang, W. Xu, X. Luo, and Q. Zhao, "Uncovering user interactions on smartphones via contactless wireless charging side channels," in *Proceedings of IEEE Symposium on Security and Privacy (SP)*, 2023.

[82] T. InvenSense, "High precision 6-axis mems motiontracking device," https://invensense.tdk.com/wp-content/uploads/2020/04/ds-000347_icm-42688-p-datasheet.pdf, 2023.

[83] HTC, "VIVE Pro Specs & User Guide," 2020, https://developer.vive.com/resources/hardware-guides/vive-pro-specs-user-guide/.

[84] V. Corporation, "OpenVR," https://partner.steamgames.com/doc/features/steamvr/openvr, 2021.

[85] S. Studio, "ThinkGear AM - brainwave sensor," 2024, https://www.seeedstudio.com/ThinkGear-AM-Brainwave-Sensor-p-1441.html.

[86] M. F. Mridha, S. C. Das, M. M. Kabir, A. A. Lima, M. R. Islam, and Y. Watanobe, "Brain-computer interface: Advancement and challenges," *Sensors*, 2021.

[87] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.

[88] J. Gu, J. Wang, Z. Yu, and K. Shen, "Traffic-based side-channel attack in video streaming," *IEEE/ACM Transactions on Networking (ToN)*, 2019.

[89] M. Enev, S. Gupta, T. Kohno, and S. N. Patel, "Televisions, video privacy, and powerline electromagnetic interference," in *Proceedings of ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2011.

[90] T. U. of Cambridge, "Cambridge English Vocabulary List," 2012, https://www.cambridgeenglish.org/images/84669-pet-vocabulary-list.pdf.

[91] X. Shen, H. Jiang, D. Liu, K. Yang, F. Deng, J. C. Lui, J. Liu, S. Dustdar, and J. Luo, "Pupilrec: leveraging pupil morphology for recommending on smartphones," *IEEE Internet of Things Journal (IoTJ)*, 2022.

[92] G. Lan, B. Heit, T. Scargill, and M. Gorlatova, "Gazegraph: Graph-based few-shot cognitive context sensing from human visual behavior," in *Proceedings of ACM Conference on Embedded Networked Sensor Systems (SenSys)*, 2020.

[93] W. Han, C. Cao, H. Chen, D. Li, Z. Fang, W. Xu, and X. S. Wang, "sendroid: Auditing sensor access in Android system-wide," *IEEE Transactions on Dependable and Secure Computing (TDSC)*, 2017.

[94] R. Li, W. Diao, Z. Li, J. Du, and S. Guo, "Android custom permissions demystified: From privilege escalation to design shortcomings," in *Proceedings of IEEE Symposium on Security and Privacy (SP)*, 2021.

[95] V. Rastogi and S. Nath, "Differentially private aggregation of distributed time-series with transformation and encryption," in *Proceedings of International Conference on Management of Data (SIGMOD)*, 2010.

[96] M. Malekzadeh, R. G. Clegg, A. Cavallaro, and H. Haddadi, "Privacy and utility preserving sensor-data transformations," *Pervasive and Mobile Computing*, 2020.

[97] T. Ni, G. Lan, J. Wang, Q. Zhao, and W. Xu, "Eavesdropping mobile app activity via radio-frequency energy harvesting," in *Proceedings of the USENIX Security Symposium*, 2023.

[98] S. Zhu, Z. Ye, Q. Ai, and Y. Liu, "Eeg-imagenet: An electroencephalogram dataset and benchmarks with image visual stimuli of multi-granularity labels," 2024.

[99] X.-H. Liu, Y.-K. Liu, Y. Wang, K. Ren, H. Shi, Z. Wang, D. Li, B.-L. Lu, and W.-L. Zheng, "Eeg2video: Towards decoding dynamic visual perception from eeg signals," in *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

[100] Z. Chen, J. Qing, and J. H. Zhou, "Cinematic mindscapes: High-quality video reconstruction from brain activity," in *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

[101] T. Ni, "Sensor security in virtual reality: Exploration and mitigation," in *Proceedings of the Annual International Conference on Mobile Systems, Applications and Services (MobiSys)*, 2024.

[102] Z. Yang, Y. Chen, Z. Sarwar, H. Schwartz, B. Y. Zhao, and H. Zheng, "Towards a general video-based keystroke inference attack," in *Proceedings of USENIX Security Symposium*, 2023.

[103] Y. Takagi and S. Nishimoto, "High-resolution image reconstruction with latent diffusion models from human brain activity," in *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[104] S. Lin, T. Sprague, and A. K. Singh, "Mind reader: Reconstructing complex images from brain activities," in *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

[105] T. Costecalde, T. Aksenova, N. Torres-Martinez, A. Eliseyev, C. Mestais, C. Moro, and A. L. Benabid, "A long-term BCI study with ecog recordings in freely moving rats," *Neuromodulation: Technology at the Neural Interface*, 2018.

[106] J. C. Snow, C. E. Pettypiece, T. D. McAdam, A. D. McLean, P. W. Stroman, M. A. Goodale, and J. C. Culham, "Bringing the real world into the fMRI scanner: Repetition effects for pictures versus real objects," *Scientific reports*, 2011.

[107] T. Fang, Y. Qi, and G. Pan, "Reconstructing perceptive images from brain activity by shape-semantic GAN," in *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

[108] Y. Lu, C. Du, Q. Zhou, D. Wang, and H. He, "Minddiffuser: Controlled image reconstruction from human brain activity with semantic and structural diffusion," in *Proceedings of ACM International Conference on Multimedia (MM)*, 2023.

[109] I. W. Store, "TGAM module development kit brainwave smart headband EEG sensor," https://www.aliexpress.com/item/1005003264615528.html, 2024.

[110] Z. Tarkhani, L. Qendro, M. O. Brown, O. Hill, C. Mascolo, and A. Madhavapeddy, "Enhancing the security & privacy of wearable brain-computer interfaces," *arXiv preprint arXiv:2201.07711*, 2022.

Appendix A.
Supplementary Figures and Tables

```
ovr_Create(&session, &luid);
...
DisplayTime = ovr_GetPredictedDisplayTime(session, frame)
motionData = ovr_GetTrackingState(session, 0.0, ovrFalse);
...                                  Unrestricted Permission
x = motionData.HeadPose.ThePose.Position.x;
y = motionData.HeadPose.ThePose.Position.y;
z = motionData.HeadPose.ThePose.Position.z;
```

(a) Oculus Mobile SDK.

```
vrObj = openvr.init(openvr.VRApplication_Other)
...                                  Unrestricted Permission
motionRawData = vrObj.getDeviceToAbsoluteTrackingPose(
                    openvr.TrackingUniverseStanding, 0,
                    openvr.k_unMaxTrackedDeviceCount)

hmdPose = motionRawData[openvr.k_unTrackedDeviceIndex_Hmd]
motionArr = hmdPose.mDeviceToAbsoluteTracking
x, y, z = motionArr[0][3], motionArr[1][3], motionArr[2][3]
```

(b) OpenVR SDK.

```
const onXRFrame = (time, frame) => {
                                     Unrestricted Permission
    motionData = frame.getViewerPose(RefSpace);

    hmdPose = motionData.transform.position;
    x, y, z = hmdPose.x, hmdPose.y, hmdPose.z;

    session.requestAnimationFrame(onXRFrame)
};
```

(c) WebXR Device API.

Figure 23: Code snippets depict the accessing built-in motion sensor data with unrestricted permission on three VR APIs/SDKs.

TABLE 4: Full list of 50 VR apps, 50 Netflix videos, and 50 websites used in the evaluation (§ 5.1) as of May 2025.

VR Apps				Netflix Videos				Websites			
#	App	Category	# Ratings	#	Video	Category	# Votes	#	Website	Category	Monthly Visits
A_1	Population: One	Games	14,761	V_1	Society of the Snow	Adventure	129,537	W_1	google.com	Search engines	162,500,000,000
A_2	FitXR	Health & Fitness	9,436	V_2	The Gentlemen	Action	82,777	W_2	youtube.com	Online TV	106,500,000,000
A_3	Gun Raiders	Games	4,919	V_3	3 Body Problem	Adventure	81,461	W_3	facebook.com	Social media networks	17,500,000,000
A_4	Epic Roller Coasters	Games	3,800	V_4	Damsel	Action	75,905	W_4	twitter.com	Social media networks	9,600,000,000
A_5	YouTube VR	Entertainment	3,233	V_5	Avatar: The Last Airbender	Action	65,931	W_5	wikipedia.org	Dictionaries & encyclopedias	9,200,000,000
A_6	TRIPP	Relaxation	2,898	V_6	Fool Me Once	Crime	52,113	W_6	instagram.com	Social media networks	8,000,000,000
A_7	Bait	Games	2,868	V_7	One Day	Comedy	39,797	W_7	reddit.com	Social media networks	7,400,000,000
A_8	Meta Quest Browser	Productivity	2,769	V_8	Lift	Action	37,603	W_8	amazon.com	eCommerce & shopping	4,300,000,000
A_9	Blaston	Games	2,667	V_9	Griselda	Biography	36,651	W_9	duckduckgo.com	Search engines	4,200,000,000
A_{10}	Netflix	Media & Streaming	2,579	V_{10}	Spaceman	Adventure	30,652	W_{10}	yahoo.com	Search engines	4,100,000,000
A_{11}	Gods of Gravity	Games	1,636	V_{11}	The Brothers Sun	Action	21,952	W_{11}	tiktok.com	Online TV	3,500,000,000
A_{12}	Cards & Tankards	Games	1,374	V_{12}	Boy Swallows Universe	Crime	16,855	W_{12}	fandom.com	Dictionaries & encyclopedias	3,200,000,000
A_{13}	Gravity Sketch	Creativity & Design	1,073	V_{13}	Ripley	Crime	14,743	W_{13}	weather.com	Weather	3,100,000,000
A_{14}	Multiverse	Social	979	V_{14}	Irish Wish	Comedy	13,188	W_{14}	whatsapp.com	Social media networks	3,100,000,000
A_{15}	HOLOFIT	Health & Fitness	946	V_{15}	Orion and the Dark	Animation	12,150	W_{15}	bing.com	Search engines	3,000,000,000
A_{16}	Half+Half	Games	862	V_{16}	Code 8: Part II	Action	11,125	W_{16}	openai.com	Computer technology	2,600,000,000
A_{17}	Meta Quest Move	Utility	805	V_{17}	The Greatest Night in Pop	Documentary	10,873	W_{17}	yandex.ru	Social media networks	2,500,000,000
A_{18}	VZFit	Sports	601	V_{18}	Scoop	Biography	9,878	W_{18}	linkedin.com	Social media networks	2,000,000,000
A_{19}	Ultimechs	Games	591	V_{19}	Badland Hunters	Action	9,058	W_{19}	microsoft.com	Computer technology	1,900,000,000
A_{20}	Liminal	Lifestyle	543	V_{20}	The Abyss	Action	6,986	W_{20}	twitch.tv	Online TV	1,900,000,000
A_{21}	Spatial	Social, Utilities	521	V_{21}	House of Ninjas	Action	5,750	W_{21}	live.com	Online TV	1,800,000,000
A_{22}	Neverboard	Games	476	V_{22}	Parasyte: The Grey	Action	5,661	W_{22}	quora.com	Social media networks	1,700,000,000
A_{23}	Alcove	Social	458	V_{23}	Baby Reindeer	Biography	5,647	W_{23}	netflix.com	Online TV	1,700,000,000
A_{24}	ForeVR	Games	392	V_{24}	Amar Singh Chamkila	Biography	5,493	W_{24}	office.com	Computer technology	1,600,000,000
A_{25}	Maloka	Lifestyle	390	V_{25}	The Signal	Drama	5,336	W_{25}	tsyndicate.com	Computer technology	1,500,000,000
A_{26}	Sphere Toon	Media & Streaming	266	V_{26}	Mea Culpa	Crime	5,190	W_{26}	bit.ly	Computer technology	1,500,000,000
A_{27}	Hoame	Health & Fitness	189	V_{27}	A Killer Paradox	Comedy	4,840	W_{27}	globo.com	News & media publishers	1,400,000,000
A_{28}	MLB VR	Media & Streaming	186	V_{28}	The Tearsmith	Drama	4,333	W_{28}	imdb.com	Social media networks	1,300,000,000
A_{29}	Noda	Productivity	166	V_{29}	The Beautiful Game	Drama	3,683	W_{29}	vk.com	Social media networks	1,200,000,000
A_{30}	Instagram	Social	155	V_{30}	Supersex	Biography	3,068	W_{30}	cnn.com	News & media publishers	1,200,000,000
A_{31}	vTime	Social	140	V_{31}	Einstein and the Bomb	Documentary	3,063	W_{31}	manganato.com	News & media publishers	1,100,000,000
A_{32}	Immerse	Lifestyle	120	V_{32}	What Jennifer Did	Documentary	2,840	W_{32}	x.com	Social media networks	1,100,000,000
A_{33}	VR Workout	Health & Fitness	104	V_{33}	Testament: The Story of Moses	Documentary	2,562	W_{33}	pinterest.com	Social media networks	1,000,000,000
A_{34}	Facebook	Social	103	V_{34}	Through My Window: Looking at You	Comedy	2,525	W_{34}	doubleclick.net	Business	957,100,000
A_{35}	REMIO	Productivity	87	V_{35}	My Name is Loh Kiwan	Drama	2,464	W_{35}	aliexpress.com	eCommerce & shopping	952,900,000
A_{36}	MeetinVR	Social	78	V_{36}	The Wages of Fear	Action	2,390	W_{36}	ebay.com	eCommerce & shopping	872,700,000
A_{37}	Flipside Studio	Productivity	72	V_{37}	Furies	Action	2,389	W_{37}	discord.com	Social media networks	864,100,000
A_{38}	Prisms MATH	Social	64	V_{38}	Delicious in Dungeon	Animation	2,254	W_{38}	sharepoint.com	Business	842,600,000
A_{39}	Zoe	Social	60	V_{39}	Turning Point: The Bomb & Cold War	Documentary	2,066	W_{39}	zoom.us	Computer technology	838,100,000
A_{40}	Arthur	Social	59	V_{40}	Bandidos	Action	1,921	W_{40}	spotify.com	Music	835,700,000
A_{41}	VirtualSpeech	Productivity	41	V_{41}	Shirley	Biography	1,831	W_{41}	indeed.com	Business	833,600,000
A_{42}	Pluto TV	Media & Streaming	39	V_{42}	Crooks	Action	1,173	W_{42}	github.com	Computer technology	831,000,000
A_{43}	Messenger	Utilities	21	V_{43}	Heart of the Hunter	Action	1,117	W_{43}	msn.com	Social media networks	803,100,000
A_{44}	Smartsheet	Productivity	19	V_{44}	Love, Divided	Comedy	1,063	W_{44}	canva.com	Business	794,000,000
A_{45}	Nanome	Productivity	14	V_{45}	The Antisocial Network	Documentary	1,026	W_{45}	bbc.com	News & media publishers	791,000,000
A_{46}	Resolve	Productivity	11	V_{46}	Woody Woodpecker Goes to Camp	Adventure	662	W_{46}	foxnews.com	News & media publishers	748,500,000
A_{47}	Softspace	Productivity	8	V_{47}	The Hijacking of Flight 601	Drama	580	W_{47}	taboola.com	Business	720,000,000
A_{48}	Monday.com	Productivity	5	V_{48}	The Tourist	Action	454	W_{48}	espn.com	News & media publishers	705,900,000
A_{49}	LastPass	Utilities	5	V_{49}	Nuevo rico, nuevo pobre	Comedy	262	W_{49}	paypal.com	Computer technology	703,900,000
A_{50}	STARZ	Media & Streaming	2	V_{50}	The Cartel: The Origin	Drama	254	W_{50}	samsung.com	Computer technology	684,200,000

Appendix B.
Meta-Review

The following meta-review was prepared by the program committee for the 2026 IEEE Symposium on Security and Privacy (S&P) as part of the review process as detailed in the call for papers.

B.1. Summary

The paper presents BRAVESPY, a VR side-channel attack that uses unrestricted motion sensors in commercial VR headsets to reconstruct the user's brain-wave signal and extract privacy-sensitive information. BRAVESPY leverages subtle pupillary-response-induced vibrations captured through motion sensors and translates them into EEG signals to infer user activity and cognition.

B.2. Scientific Contributions

- 5. Identifies an impactful vulnerability.
- 6. Provides a valuable step forward in an established field.

B.3. Reasons for Acceptance

1) The paper identifies an impactful privacy risk in VR systems. BRAVESPY showed that unrestricted motion sensors can be translated into brain-induced EEG signals that expose substantially more privacy information.
2) The paper provides a valuable step forward in an established field. The paper goes beyond prior VR motion-sensor privacy work by attempting to connect subtle headset vibrations with higher-level perceptual inference. The experimental effort is substantial, involving multiple devices, multiple users, and several inference tasks.
3) The paper presents a broad effectiveness and accuracy evaluation. The evaluation spans multiple privacy-inference tasks across UI-, user-, and brain-level settings and covers multiple VR headsets, users, sampling rates, and environmental conditions. The reported results indicate strong performance in several scenarios, especially for UI-level inference, while also showing measurable accuracy for the more challenging brain-level task.

B.4. Noteworthy Concerns

None