PipID: Light-Pupillary Response Based User Authentication for Virtual Reality

Muchen Pan
Shanghai Jiao Tong University
Shanghai, China
pmc296289396@sjtu.edu.cn

Yan Meng*
Shanghai Jiao Tong University
Shanghai, China
yan_meng@sjtu.edu.cn

Yuxia Zhan
New York University
New York, NY, USA
yz10951@nyu.edu

Guoxing Chen
Shanghai Jiao Tong University
Shanghai, China
guoxingchen@sjtu.edu.cn

Haojin Zhu
Shanghai Jiao Tong University
Shanghai, China
zhu-hj@sjtu.edu.cn

Abstract

During the use of Virtual Reality (VR) applications such as gaming, education, and military training, sensitive information may be generated or collected by VR sensors, raising user concerns about potential data leakage. This highlights the critical need for effective user authentication to prevent unauthorized access. Existing authentication methods for VR are often either cumbersome (e.g., entering passwords via handheld controllers), reliant on specialized hardware (e.g., iris recognition), or vulnerable to credential replay attacks. In this study, we propose PipID, a lightweight VR authentication approach that leverages commercial off-the-shelf (COTS) eye trackers integrated into VR headsets. PipID is based on the fact that users' pupillary responses to visual stimuli vary uniquely. Thus, by displaying lights of randomly selected colors (i.e., wavelengths) on the VR screen, PipID can utilize pupil diameter responses to these wavelengths as the basis for authentication. For pupil data collected by precision-limited COTS eye trackers, PipID mitigates the impact of unrelated eye movements (e.g., blinks) and leverages pupillary response differences between the left and right eyes to further enhance the granularity of authentication features. Additionally, the randomized sequence of light colors helps prevent replay attacks. We implemented PipID on a COTS VR headset and tested it with 52 participants. Experimental results show that PipID achieves an accuracy of 98.65% and maintains robust performance under various conditions (e.g., keeping 98% and 91% accuracy after 7 and 14 days respectively).

CCS Concepts

• **Security and privacy → Biometrics**; **Privacy protections**.

Keywords

Virtual Reality, User Authentication, Eye Tracking, Pupillary Response

ACM Reference Format:
Muchen Pan, Yan Meng, Yuxia Zhan, Guoxing Chen, and Haojin Zhu. 2025. PipID: Light-Pupillary Response Based User Authentication for Virtual Reality. In *Proceedings of the 2025 ACM SIGSAC Conference on Computer and Communications Security (CCS '25), October 13–17, 2025, Taipei, Taiwan.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3719027.3744850

1 Introduction

Virtual Reality (VR) has reshaped users' digital experiences through immersive simulated environments, making it extremely popular and widely applicable. The number of VR users in the United States has reached 77 million in 2024 [15], and with growing user interest, the global market size of VR is projected to reach USD 520 billion by 2031 [10]. By utilizing various devices like head-mounted displays (HMDs) and handheld controllers, users enjoy an immersive experience in VR applications. However, this process also exposes them to the risk of data breaches, especially in sensitive contexts such as financial transactions [4, 48], private meetings [7, 40], or military simulations [30, 52]. Unauthorized access to VR devices can lead to privacy leaks, economic fraud, theft of virtual assets, and business disruptions, posing significant threats to individual livelihoods and communities.

Unfortunately, text-based password solutions, the current mainstream authentication methods adopted by VR devices, suffer from the cumbersome use of physical hand controllers or virtual keyboards [24, 62]. According to statistics in [78], it takes an average of 10.5 seconds to enter a 4-digit PIN in VR. These methods are also vulnerable to side-channel attacks [16, 32, 41] and impose an additional memorization burden on users. In addition to traditional password entry, there is growing interest in exploring biometrics for VR user authentication. With advanced tracking sensors, a rich source of biometric data, including both *static* and *dynamic* characteristics, can be utilized for this purpose.

Biometric authentication in VR is still immature. Static biometrics, such as fingerprints, face recognition [31], and iris scanning [6], have long been considered reliable for user authentication. Nevertheless, the sophisticated measuring devices increases their deployment cost in VR scenarios. Moreover, as irreversible personal identifiers, data breaches could leave individuals without an

[*] Yan Meng is the corresponding author.

CCS '25, Taipei, Taiwan

ACM ISBN 979-8-4007-1525-9/2025/10
https://doi.org/10.1145/3719027.3744850

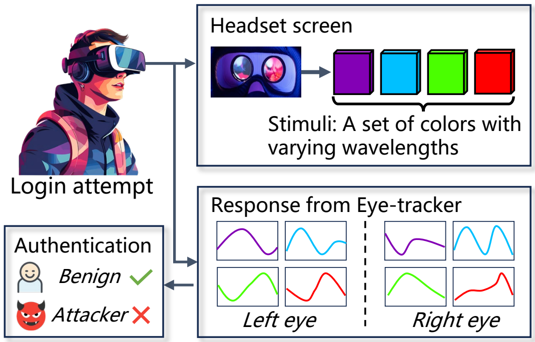


Figure 1: PipID utilizes pupil diameter responses to the VR headset screen for user authentication.

alternative credential to prevent replay attacks [54, 67]. In contrast, dynamic biometrics rely on users' movement features, such as facial expressions [80], eye movements [81, 82], and gait information [57], derived from performing specific actions [42, 47, 55, 57] or responding to audio/motion commands [82]. However, these methods require active user participation, which adds an extra burden on users, such as walking in [57] or concentrating on external music in [82]. Considering above-mentioned issues when deploying existing works in VR, it is desirable to design a novel authentication mechanism which meets following requirements: 1) *User-effortless:* eliminating the need for any active user actions; 2) *Commercial-deployable:* fully utilizing built-in VR headset devices without requiring specialized equipment; 3) *Anti-replay:* resistance to replay attacks.

Motivations of PipID. In this study, we propose PipID, a light-pupillary response based authentication system. As shown in Figure 1, PipID leverages the user's pupillary response to lights of different wavelengths displayed on an HMD's virtual screen for authentication. The chosen light-pupillary response offers two advantages: (1) Current VR headsets, such as the PICO 4 Pro [50], Meta Quest Pro [63], and Varjo XR-3 [69], are now equipped with commercial-off-the-shelf (COTS) eye trackers. This eliminates the need for additional specialized devices, enabling the widespread application of PipID. (2) As the light changes, the distinct spectral sensitivities and photon fluxes of photoreceptor cells in the retina produce different neurological reflexes [5, 33, 39], resulting in corresponding pupil constriction or dilation. According to research in the biological field [3, 72, 76], the variability of this uncontrolled, spontaneous eye movement among individuals is caused by the unique biological characteristics of neural pathways and iris muscle structure, making it nearly impossible to mimic through voluntary eye movement control. Additionally, it has been shown that left and right pupil diameters are different even under the same lighting conditions [51], which further enhances the granularity of the authentication process.

Despite the advantages of light-pupillary response, achieving the design goals of PipID faces the following challenges. **RQ1**: how can we select appropriate light stimuli for the VR screen that are user-friendly while generating pupillary responses significantly different among users? **RQ2**: it is well-established that data measured by commercial eye trackers in VR have lower accuracy compared to specialized devices[21, 56, 65, 71]. Thus, how can we capture effective features from data with limited accuracy? **RQ3**: to address the anti-replay goal, how can we mitigate replay attacks where adversaries exploit outdated pupillary responses?

In this study, to answer **RQ1**, we first conduct a series of case studies which validate that variations in pupil diameter under different light wavelengths can effectively authenticate users. By adjusting light colors, PipID enables users to freely look around the virtual environment without needing to maintain a fixed gaze or perform specific actions. Furthermore, we use the Kolmogorov-Smirnov (KS) test to select the most suitable visual stimuli for each user during login. To address **RQ2**, we propose deriving efficient time and frequency-domain features from the raw data and enhancing our feature set by leveraging the binocular tracking function of the VR eye tracker to obtain the differences in pupil diameter between left and right eyes. Finally, to tackle **RQ3**, PipID alters the wavelength sequence within the visual stimuli during each login attempt to generate a randomized challenge. This approach ensures that credentials (i.e., the pupillary responses) is hard to predict with high randomness.

To evaluate the system's performance, we conduct experiments with 52 volunteers using commercial VR headsets. For each volunteer, we present 10 lights with different wavelengths, each lasting 5 seconds, and repeat this process 20 times, constructing a dataset of 10,400 light-pupillary response profiles. PipID achieves an overall accuracy of 98.65% and demonstrates robust performance across various factors, including registration cost, number of light wavelengths, and stimulus duration. Besides, PipID demonstrates efficiency against replay attacks by reducing the attack success rate to 2.40% when the attackers fail to predict the challenge. Finally, the longitudinal study shows PipID's consistent performance, achieving 98% and 91% accuracy after 7 and 14 days, respectively.

The main contributions of this study are:

- *New user authentication in VR.* We propose PipID, a lightweight and user-effortless authentication mechanism that leverages the light-pupillary response in both eyes under various colors to verify user identity.
- *Effective features.* We propose effective statistical, temporal, and frequency-domain features, and further enhance them by utilizing inter-ocular differences using the binocular tracking capability of the VR eye-tracker.
- *Robust performance.* Experimental results demonstrate that PipID achieves an average accuracy of 98.65% for user authentication and performs robustly under various impact factors.
- *Public-available dataset.* A dataset containing users' eye movements under 10 lights with different wavelengths will be made available to the public for further research in VR authentication.

Ethical consideration. The dataset collection procedures are under the approval of the Institutional Review Board (IRB) of our institutions. We ensure that all volunteers are fully informed about the purpose, duration, and potential risks of the study. The light sources selected for the experiment are gentle and harmless to human eyes, as the rays from these sources will experience diffuse reflection off the environment and objects before entering the eyes with moderate light intensity. Furthermore, all data collected during the study is handled with the utmost care to protect participant privacy. Identifiable information is removed and data sharing is restricted to research purposes.

Data availability and online appendix. We have open-sourced the datasets and codes of PipID at our project website https://pipid-vr.github.io/ to facilitate further research and development in VR user authentication. Besides, due to the page limit, all appendix are also presented in our project website.

2 Preliminaries

2.1 User Authentication in VR

The rapid development of VR has been evident in the past few years, accompanied by growing popularity [10, 15]. As personal devices, VR systems often store users' sensitive data (e.g., emails, multimedia, browsing history, login accounts) during their usage. For instance, Gucci [22] has launched Gucci Town within the Roblox metaverse recently for online shopping, where users may store their account and credit card information for auto-login and purchase. In this case, leaving the device unguarded to malicious people may expose it to security breach and even financial loss [81]. Therefore, the need for secure and usable ways to authenticate legitimate users becomes even more urgent. The existing user authentication in VR can be categorized into following four categorizes.

Knowledge-based methods. As intuitive authentication solutions, knowledge-based methods, including passwords and PINs entry, are mainly adapted in VR through gesture-based text input. Despite their high deployability, numerous users have voiced dissatisfaction with the unwieldiness of virtual keyboards [62, 73]. We also conducted a user study (in the online appendix) and found that the average time for inputting a digital password consisting of 8 letters, 2 numbers, and 1 symbol in VR exceeds 13.69 seconds. The poor usability discourages users from adopting stronger passwords. What's worse, password is susceptible to side-channel attacks [2, 16, 32, 41], where external observers can easily capture the videos (e.g., Hidden Reality attack in [16]) or sensing signals (e.g., Wi-Fi signal based attack in [2]) of hands typing to infer the entered passwords. The immersive nature of VR often obscures users' awareness of such surrounding attackers in public.

Token-based methods. In token-based approaches, the headset may present a short code for input on an auxiliary device or request scanning of a QR code generated by the auxiliary device. However, this approach is inconvenient for users as it necessitates carrying external devices when using standalone VR devices.

Static biometrics. Static biometrics rely on static images or data of one's biological characteristics which remain relatively stable over time. Renowned for their superior usability and accessibility, static biometrics (primarily iris scan [6]) have emerged favorable for VR user authentication. However, the inherent nature and irrevocability of static biometrics raise privacy concerns. Attackers can exploit replay attacks using fake fingers [67], masks, and high-resolution iris images [54] if such data are leaked. Besides, the sensors required for static biometrics are often specialized and expensive, limiting their widespread adoption in VR.

Dynamic biometrics. Dynamic biometric technologies focus on analyzing the dynamic process of biological characteristics, such as behavioral patterns and voice signatures. The widespread capability of modern VR devices to sense user behaviors makes it a popular research subject for user authentication. Some methods involve performing specific actions like walking [49, 57], typing [49], or grasping objects [47, 49] due to the inherent heterogeneity of human muscle structures. However, these methods are often physically demanding or require substantial space. Unobstructed hand and body movements remain vulnerable to observation-based attacks. In contrast, user authentication leveraging eye-tracking data stands out as a promising approach by presenting specific stimuli to elicit personalized eye movements.

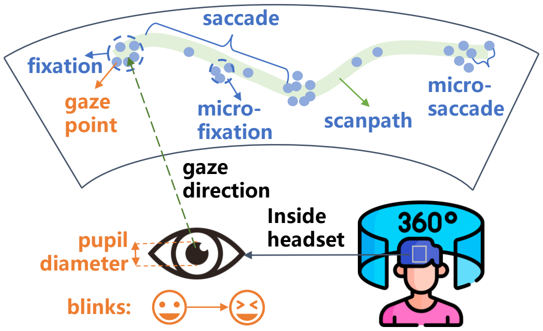


Figure 2: Data type provided by eye-trackers. Blue indicates data exclusive to specialized equipment, while orange marks data obtainable from both specialized and VR devices.

2.2 Eye-tracking Technology

Principle of eye-tracking technique. As illustrated in Figure 2, common eye movement behaviors include *saccades* and *fixations*. Saccades are rapid eye movements that shift the foveal vision[1] from one point to another, while fixations involve maintaining the foveal vision on a target for a certain duration to capture detailed visual information. Fixations are not entirely static, as they involve three types of micro-movements: micro-saccades, nystagmus, and drift.

To acquire rich oculomotor data, eye-tracking technology has emerged prominently. Currently, there are two primary types of eye-tracking: electro-oculography (EOG) and video-based eye-tracking. EOG measures the resting potential of the retina to record eye position and movement [37, 44, 45] via invasive electrodes, which is not suitable for user authentication. Conversely, video-based eye-tracking captures high-resolution eye images including pupil and corneal by cameras. Advanced image processing algorithms calculates the gaze direction vector by tracking the position of the pupil center, while simultaneously computes the pupil diameter and eyelid aperture.

Eye-tracker in specialized device. The specialized device employs a video-based eye-tracker monitoring pupil and corneal reflections. Typically, multiple cameras are leveraged to capture stereo images of both eyes in three-dimensional space to accurate capture gaze points, pupil diameters, and blinks. The professional eye-tracker possesses exceptionally high precision and resolution, with accuracy down to 0.1-0.5 degrees and sampling rates up to 1200 Hz in Tobii for instance [66]. Therefore, eyelid aperture, fixation and

[1] Foveal vision refers to the visual perception originating from the fovea, a small central pit in the retina that is densely packed with cone cells. The fovea is the most sensitive part of the retina and is specialized for high-resolution visual acuity.

saccade patterns, scanpath, and even intricate micro-fixation and micro-saccade can be output or analyzed from the outputted data.

Eye-tracker in commercial VR device. Currently, an increasing number of VR headsets have equipped with eye-tracking technology. The embedded infrared cameras record eye videos, enabling the extraction of gaze point, pupil size, and eye blinks, which provides a better understanding of the user's visual attention in VR.

Figure 2 outlines the types of eye movement data relied upon in existing works, along with their availability in specialized eye-tracking devices and different VR headsets. It is shown that the types of data offered by VR headsets are limited in scope. Even if a VR headset are capable of offering the same types of data as the professional one, they often fall short in terms of data accuracy and frequency [21, 56, 58, 65, 71]. In the online appendix, we present a detailed comparison of several types of specialized and VR devices in terms of their eye-tracking capabilities.

2.3 Eye Motion Based Authentication

Using specialized devices for authentication. The diversity in the optic nerve pathway and muscles controlling eyeballs contributes to the high personal specificity of individuals' eye movements [9, 36]. Several studies [12–14, 19, 20, 53, 61] have achieved user authentication utilizing eye motion data (e.g., fixation point and duration, saccadic speed and direction, micro-fixation and micro-saccade, scanpath and pupil diameter) extracted from specialized devices. The above mechanisms are task-driven and require users to actively engage by gazing at moving targets or performing tasks such as reading text. Meanwhile, several studies have also employed pupil light reflex to achieve passive authentication. [75] utilizes iris patterns and pupil diameter under low or non-visible light conditions for authentication purposes, and [74] proposes using three different light intensities for pupil diameter-based authentication. Nonetheless, all these approaches employ high-precision eye-trackers (e.g., 500 Hz eye-trackers in [61]), which are rarely seen in consumer-grade VR headsets due to cost considerations.

Authentication on commercial VR devices. Recent researches [35, 47, 49, 79, 81, 82] have focused on VR user authentication, and they all necessitate active participation from users, either by performing actions or receiving instructions. In [49] and [47], users are required to execute specific body movements (like grasping, walking, typing, etc.), while their head, hand, and eye motion data are captured to derive unique behavioral patterns. Such methods are space-consuming and effort-intensive, unsuitable for users with impaired motor functions, and the overt body movements make them susceptible to shoulder-surfing attacks. Besides motion patterns, certain interactive stimuli (e.g., moving objects, sounds, videos) can also trigger reflexive eye movements [35, 79, 82]. [35] leverages the public dataset GazeBase and primarily focuses on optimizing authentication models. Though [79] claims to achieve implicit continuous authentication in VR, their implementation depends on smart glasses, and still requires users to focus on a moving cross. [82] employs pupillary responses triggered by audio stimuli, which take a relatively long time for authentication. In addition, the approach proposed in [81] combines knowledge-based and biometric elements by utilizing blinking rhythms as passwords, which also put pressure on user's memory. The blinking inconsistency across different instances contributes to the method's instability. We have summarized the aforementioned representative eye-motion based authentication methods in Table 1 and made comparisons with our proposed method.

Table 1: Summary of Eye Motion Based Authentication.

Study	Stimulus	Feature types	Built-in VR Sensors	Passive Authentication
[61]	moving dots	fixation & saccade	✗	✗
[19]	reading	scan path	✗	✗
[14]	various tasks	fixation & pupil diameter	✗	✗
[75]	non visible light	iris pattern& pupil diameter	✗	✓
[74]	light intensity	pupil diameter	✗	✓
[47]	body motion	gaze position	✓	✗
[79]	moving cross	gaze position	✗	✗
[81]	/	blink rhythm & pupil diameter	✓	✗
PipID	light color	pupil diameter	✓	✓

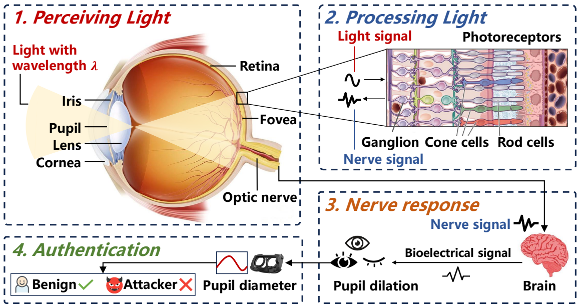


Figure 3: Light perceiving, processing and responding.

3 Threat Model and Motivations

3.1 Threat Model and Goals

In this study, the adversary is an unauthorized individual attempting to gain illicit access to a VR device by impersonating a legitimate user. The ultimate goal is to steal sensitive data from the victim, such as healthcare records, financial information, or commercial secrets. We assume the following two attack scenarios:

- *Basic scenario: physical access.* The adversary may obtain the VR device through theft, loss, or unauthorized borrowing. With physical possession, the adversary attempts to use their own eye motion credentials to bypass the authentication system.
- *Adaptive attacks: access to sensors.* We assume the attacker has the capability to manipulate the eye-tracking sensors after sneaking a user's eye video or historical authentication data from social networks. Thus, the attacker can replace the data transmitted from the eye-tracker to the authentication system by the historical data during authentication.

In this study, PipID should defend against adversaries with any of the above-mentioned attack capabilities. Note that we do not consider adaptive attacks that rely on any of the following three stronger assumptions: (1) the adversary can obtain the challenge stimuli, which consist of a randomized sequence of colored lights displayed on the virtual screen; (2) launch attacks at the system level of VR devices; and (3) manipulate their own eye movements to imitate the victim's pupillary response. We discuss the feasibility of implementing the above three attacks on PipID in Section 6.

3.2 Motivation of PipID

3.2.1 Basic Insights and Research Questions. The human pupil is a small circular aperture at the center of the iris, serving as the gateway for incoming light. Pupil diameter is influenced by various factors, including individual age, ethnicity, as well as visual stimuli such as light intensity, color, and the distance of target. Existing researches have demonstrated that human pupils exhibit varying degrees of constriction under different colored light [5, 11, 18, 23, 26, 33, 68, 77]. As illustrated in Figure 3, when light traverses the pupil, it is received and processed by distinct photoreceptor cells located on the retina. These photoreceptors facilitate visual phototransduction, converting the energy of photons into a neuronal signal. The major types of photoreceptors encompass cone cells, rod cells, and intrinsically photosensitive retinal ganglion cells, among which cone cells primarily govern the perception of colored light.

Concentrated in the fovea centralis or parafovea[2] of the retina, cone cells' sensitivity and response speed to light vary with color due to their possession of three distinct visual pigments, each capable of absorbing red, green, and blue light. As light changes, cone cells detect these alterations and transmit different signals to the nervous system (i.e., the brain). In turn, the brain modulates the tension of iris muscles via neural pathways such as the oculomotor nerve, thereby altering pupil size. The iris muscles consist of multi-unit smooth muscles. The circular muscle layer (pupillary sphincter) leads to pupil miosis, while the radial muscle fibers (pupillary dilator muscle) results in pupil dilation. Though the superficial process appears to be the impact of different colored light on pupil diameter, our study actually adopts the terminology of using light sources of different wavelengths given that the mapping relationship between RGB colors and wavelengths.

In this section, we aim to validate our motivation through a series of preliminary case studies which answer the following critical questions.

- Does people perform unique patterns of pupil reactions compared to others when seeing light of same wavelength?
- Is the pupil response of the same user hold consistency and stability under the same wavelength?
- Do visual stimuli with different wavelengths cause diverse changes in pupil diameter for one person?

3.2.2 Observations Derived from Case Studies. To address the above questions, we conduct experiments in which a total of 52 volunteers are recruited. During the experiments, volunteers are instructed to wear the HTC VIVE Pro Eye [70] headset and observe the virtual scenarios presented on the headset's screen. The light source within the scenario varied sequentially across 10 different wavelengths: 400 nm, 450 nm, 475 nm, 500 nm, 525 nm, 550 nm, 570 nm, 600 nm, 630 nm and 700 nm. There is no precise mapping for converting a wavelength (λ) directly to an RGB color; instead, the correlation between them is more accurately described as an approximate conversion, as illustrated in Table 2. Each wavelength is maintained for a duration of 5 seconds, while the Pupil Lab eye tracker [29] integrated into the headset records the participants' eye movement data.[3] The sampling process is repeated 20 times per user, resulting in a total of 1040 samples with each comprising a time-varying sequence of pupil diameter.

Table 2: Light color and approximate wavelength (λ), the mapping adopted is based on [60].

Color	RGB value	λ	Color	RGB value	λ
🟪	(131, 0, 181)	400 nm	🟩	(163, 255, 0)	550 nm
🟦	(0, 70, 255)	450 nm	🟨	(225, 255, 0)	570 nm
🟦	(0, 192, 255)	475 nm	🟧	(255, 190, 0)	600 nm
🟩	(0, 255, 146)	500 nm	🟧	(255, 79, 0)	630 nm
🟩	(74, 255, 0)	525 nm	🟥	(255, 0, 0)	700 nm

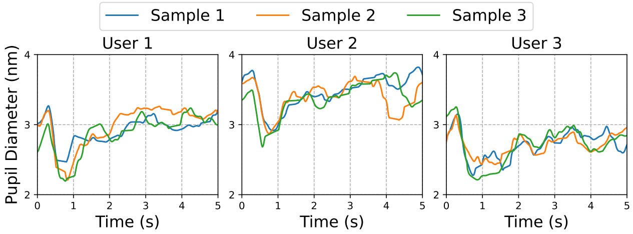


Figure 4: Pupillary response of same users and between different users when λ = 550 nm.

Observation 1: *inter-subject pupillary responses show distinctive patterns at the same stimulus.*

Figure 4 presents pupillary response of three users under the wavelength λ = 550 nm, with three samples per user. The graph illustrates significant distinctions in the pupil response patterns among users, which are sufficient for authentication purposes.

Observation 2: *the pupillary response is consistent for the same user at the same wavelength.*

Three pupillary response trials from each user under identical stimulus are depicted in Figure 4, exhibiting a high degree of similarity. Furthermore, based on the original data, we evaluate the inter-user variability and intra-user consistency by conducting the Kolmogorov-Smirnov (KS) test for each user. The test returns a value ranging from 0 to 1, representing the maximum vertical deviation between two sample sets. A higher KS value indicates greater divergence between the two sample sets. For each user, we randomly select five samples and calculate the average KS value for all pairwise comparisons to measure the consistency among the user's own samples. Meanwhile, we compute the average KS value between user's own samples and those from other users to evaluate their differences.

In Figure 5, the KS values of intra-user groups (with an average of 0.39) are significantly smaller than those of inter-user groups (with an average of 0.76). This observation underscores the intra-user homogeneity and the inter-user divergence under the same visual stimulus, further validating the feasibility and effectiveness of light-pupillary response in user authentication.

[2] Fovea centralis is surrounded by the parafovea belt.

[3] Pupil Lab is an optional accessory (i.e., a commercial VR eye tracker) for HTC VIVE Pro Eye.

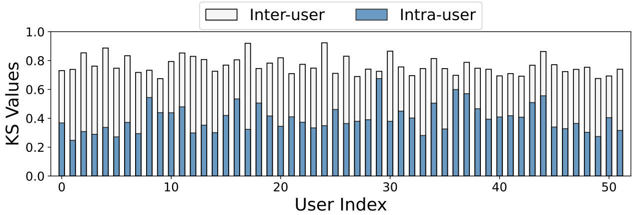


Figure 5: KS values of inter-user and intra-user groups.

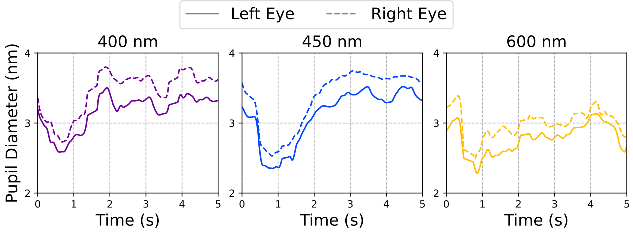


Figure 6: Left and right pupil size of same user under varying wavelengths.

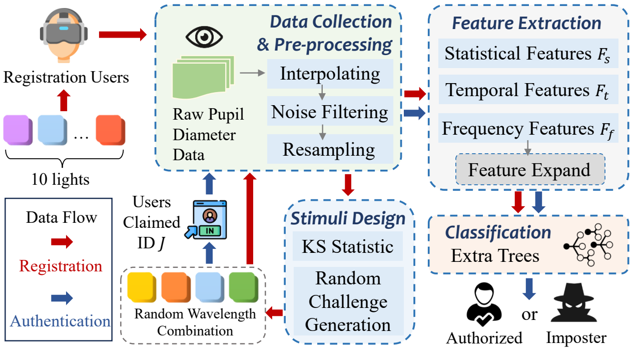


Figure 7: The system overview of PIPID.

Observation 3: *different wavelengths of ambient light trigger diverse binocular pupillary responses for the same user.*

Figure 6 plots the left and right pupil size of a given volunteer under three distinct wavelengths of light. It highlights variations in pupillary responses not only across wavelengths but also in binoculus under the same stimulus. This represents a fine-grained biometric with the potential for random challenge-response mechanisms. By altering the wavelength combinations within the stimuli, substantial response credentials can be generated. This reduces the risk of impersonation attacks through credentials theft.

4 System Design of PIPID

Figure 7 shows an overview of PIPID, which consists of four modules. Initially, the *Data Collection and Pre-processing* module acquires the raw pupil size sequence during both the registration and authentication phases and implements signal smoothing and de-noising to mitigate the negative effects of spontaneous blinks and undesirable noise. Next, the *Stimuli Design* module calculates the light wavelength combination that generates the most distinctive pupillary response for each registered user and instructs the Data Collection and Pre-processing module to collect data for any user requesting authentication. The de-noised data are then fed into the *Feature Extraction* module, where statistical, temporal, and frequency features are extracted. Finally, the *Classification* module trains the classifier based on the registered datasets and performs authentication. In this section, we will elaborate on the implementation details of these four modules.

4.1 Data Collection and Pre-processing

As illustrated in Figure 7, during both registration (collecting users' pupil data with real identity labels) and authentication (when the user claims an identity whose authenticity is unclear), PIPID adopts standard data collection and pre-processing methods as described below.

4.1.1 Data Collection. We utilize the HTC VIVE Pro Eye to gather users' eye-tracking data. According to the official manufacturer, eye-tracker of this device has a sampling frequency of 120 Hz, with an accuracy ranging from 0.5° to 1.1°, and a field of view (FOV) of 110°. We design a virtual scenario with dynamically changing colored light source on Unity[4] for the participants, with data recorded simultaneously through the official SDK.

During the registration phase, this study involves 52 users, with each user being sampled 20 times. Each round includes 10 distinct wavelengths of light, and each wavelength lasts for 5 seconds. Consequently, we represent pupil diameter recordings $R = [R^u_{k,i}]$, where $k \in [1, N_k]$ represents the k-th wavelength as described in Table 2 with $N_k = 10$, $i \in [1, N_s]$ denotes the sample index with $N_s = 20$, and $u \in [1, N_u]$ is the user index with $N_u = 52$. For the authentication phase, where a user claims identity J, to achieve lightweight authentication, we display only N_w colors with different wavelengths $w^J = [w_1, w_2, \ldots, w_{N_w}]$ provided by the Stimuli Design module (details can be found in Section 4.2), and the collected data is represented as $A = [A^J_{w_i}]$, where $i \in [1, N_w]$.

4.1.2 Data Pre-processing. To optimize the quality of data and mitigate the impact of noise, we stick to the following steps to handle the collected raw data:

(1) Interpolating: Due to the inevitable spontaneous blinks, outliers (values < 0)[5] exist in the pupil size readings from eye-tracking. PIPID employs linear interpolation to address this issue.

(2) Noise Filtering: The collected recordings $R = [R^u_{k,i}]$ usually contains noise due to environmental light interference, electromagnetic emissions from the device, as well as the inherent complexity of ocular movements and device accuracy limitations. Through Fourier transformation, we identify that the noise is concentrated in the high-frequency region. Therefore, we use a low-pass filter with a cutoff frequency of 50 Hz to remove the noise.

(3) Resampling: As data transmission and hardware synchronization are not completely ideal, frame skipping or duplication may exhibit in the obtained data, resulting in inconsistent frame counts for each 5-seconds sample $R^u_{k,i}$. To facilitate subsequent processing, we leverage interpolation to resample $R^u_{k,i}$ at a rate of 60 Hz , ensuring a uniform count of 300 frames per sample[6].

[4] Unity is a development platform for VR apps. For details, see Section 5.1 and Figure 12.
[5] The eye-tracker is able to detect blinking and outputs a pupil diameter value of −1.
[6] Despite the official sampling rate being 120 Hz, we have observed issues such as frame skipping and frame repetition in the timestamps of the output eye-tracking sequences. The average sampling rate is around 60 Hz in practice.

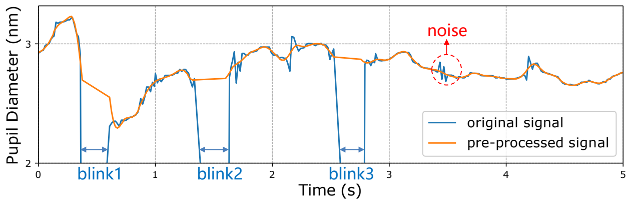


Figure 8: Pre-processing of original data.

Figure 8 illustrates the original and the pre-processed pupillary response signals. Ultimately, during the registration phase, we obtain the pre-processed response traces $I^u_{k,i}[n]$ from each $R^u_{k,i}$, which represents a pupil diameter sequence of length $L = f_s \times T$ under the k-th wavelength, where the resampling rate f_s is 60 Hz and the duration T is 5 seconds. Similarly, we have pre-process trace $K^J_{w_i}$ from $A^J_{w_i}$ during the authentication phase.

4.2 Stimuli Design

In this subsection, we design customized visual stimuli tailored to users' pupillary responses across various wavelengths. Our objective is to find the best combinations of wavelengths, which is assumed to elicit the most distinctive eye movement signatures (i.e., maximal inter-individual heterogeneity while preserving intra-personal consistency). In addition, we will introduce the generation of random challenges to defend against replay attack.

4.2.1 User-Specific Light Color Selection. We employ the KS test described in Section 3.2.2 to quantify both variability and consistency. For pre-processed pupil data from registration users, we compute the KS statistic between their pupillary signals and those of other users across each wavelength. We average all samples from user u_i at a given wavelength:

$$I^{u_i}_{k,avg}[n] = avg(I^{u_i}_{k,1}[n], I^{u_i}_{k,2}[n], ...I^{u_i}_{k,N_s}[n]), \tag{1}$$

and similarly, average the samples from another user $u_j \neq u_i$ at the same wavelength:

$$I^{u_j}_{k,avg}[n] = avg(I^{u_j}_{k,1}[n], I^{u_j}_{k,2}[n], ...I^{u_j}_{k,N_s}[n]). \tag{2}$$

The KS test (denoted as $KS(\cdot,\cdot)$) is calculated between $I^{u_i}_{k,avg}[n]$ and $I^{u_j}_{k,avg}[n]$. This process is repeated for every other user, and the average of these KS values represents the final KS score for user u_i as $\overline{KS}(u_i, k) = \frac{1}{N_u - 1} \sum_{i \neq j} KS(I^{u_i}_{k,avg}[n], I^{u_j}_{k,avg}[n])$, signifying the inter-user difference at the specified k-th wavelength. PipID could establish a prior database to store a certain amount of pupil data. This will mitigate the cold start problem for the initial users of the system. Whenever a new user joins, PipID collects the user's eye-tracking data and performs a KS test on it.

To assess personal consistency, we perform pairwise KS tests on all samples for each user at each wavelength and average the results. Finally, we obtain measures of both variability and consistency for each user across different wavelengths. We select the top N_w wavelengths with the largest difference between variability and consistency as the user-specific stimuli (i.e., wavelength combination). In this study, unless otherwise specified, N_w is set to 4. During the registration phase, we preserve only the user's pre-processed data corresponding to the top N_w wavelengths for subsequent feature extraction and classifier training. During each login phase, the user first claims their identity J, after which the corresponding N_w wavelengths are presented as visual stimuli for authentication.

4.2.2 Random Challenge Generation. To defend against replay attacks, we further introduce random challenges. After customizing the most suitable N_w wavelength for each user, PipID alters the sequence of light presented to the user. For instance, during one login attempt, the system presented light wavelengths as [400 nm, 450 nm, 500 nm, 600 nm]; in a subsequent login, the stimuli may be rearranged to [450 nm, 600 nm, 400 nm, 500 nm]. In the following sections, for a given authentication attempt, we denote the random challenge w^J as $[w_1, w_2, \ldots, w_{N_w}]$. We assume that the adversary in replay attacks possesses samples from the N_w optimal wavelengths but does not know their randomized order during a given login attempt. The performance of this defense approach is evaluated in Section 5.4.1.

4.3 Feature Extraction

After obtaining the pre-processed data $I^u_{k,i}[n]$ or $K^J_{w_i}$, we extract three types of features: statistical features, temporal morphological features, and frequency-domain features. Taking a sample $\hat{i}$ at a specific wavelength $\hat{k}$ of user $\hat{u}$ as an example (abbreviated as $\hat{I}[n]$), we will introduce the extraction process below in detail.

4.3.1 Statistical and Temporal Features. Statistical features refer to the descriptions of signal statistical analysis. For $\hat{I}[n]$, we calculate the average, variance, median, skewness, and kurtosis of the pupil diameter to constitute its statistical features F_s. Besides, PipID extracts temporal features F_t from $\hat{I}[n]$ to outline the morphology of light-pupillary response curves, as shown in Figure 9. By introducing short intervals of darkness between each wavelength, we allow the subject's pupil size to approximately recover to baseline before the next wavelength. As observed in Figure 9, upon exposure to a specific wavelength, the pupil undergoes an evident constriction after a response lag, with notable fluctuation from the 1st to 3rd seconds. Therefore, we select the signal segment $C[n], n \in [1, f_s \times 2]$, from $\hat{I}[n]$ spanning 1st to 3rd seconds. Normalization mechanism is conducted on it to obtain $C^{'}$. The $k-$th element of $C^{'}$ is:

$$C^{'}[k] = \frac{C[k] - min}{max - min}. \tag{3}$$

The normalization value is set with $min = 0$ and $max = 8$ based on the general pupil diameter's fluctuation range under light influence. To further improve PipID's ability, we also propose following feasible features as a component of F_t:

- Area under the curve (AUC). This represents the area beneath the pupillary response curve, calculated by integrating over time. A larger AUC tends to indicate a higher sensitivity for user to light stimulus.
- Peak/Valley magnitudes. These refer to the local maxima and minima in the pupil fluctuations. More acute fluctuations result in higher peaks and lower valleys. PipID employs a classical peak detection algorithm to identify peaks and valleys in the response waveform. To unify the feature vector length, we select the first N peaks and valleys from $\hat{I}[n]$. If the number of

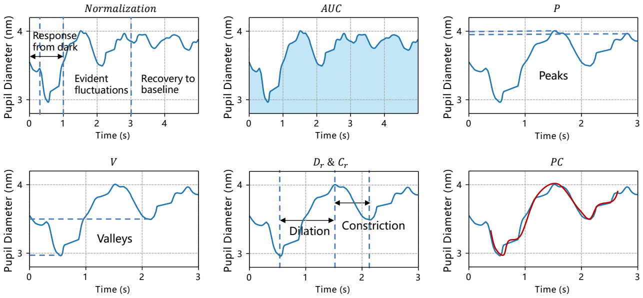


Figure 9: Illustration of pupillary response under stimuli and its temporal features (i.e., normalization, AUC, peak magnitudes, valley magnitudes, dilation rate, miosis rate, and polynomial coefficients).

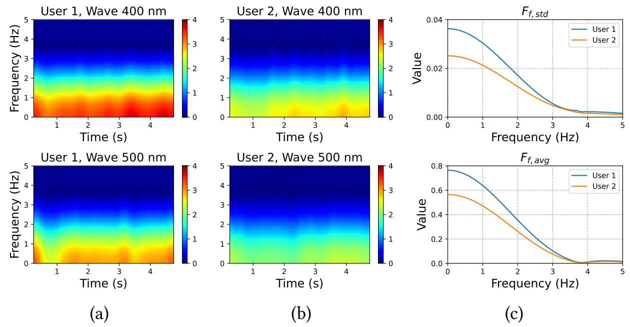
peaks or valleys is less than N, we pad with the minimum peak or maximum valley values. PipID also calculates the average, median, and variance of the peaks and valleys as supplements of the Peak/Valley magnitude features (P, V).

- Dilation/Miosis rates. Dilation (Miosis) rate is defined as the ratio of pupil dilation (miosis) amplitude to the corresponding time interval. We determine pupil dilation and miosis by leveraging the adjacent peaks and valleys identified in the previous step. Similarly, we select M dilation/miosis rates to standardize the feature vector length, denoted as Dr and Mr.
- Polynomial coefficients. The k-degree polynomial can be used to approximate the shape of the pupillary response curve, with its coefficients PC serving as features. We set k to 15, as it provides a good fit to the ground truth.

In summary, we can obtain the temporal feature of PipID as $F_t = [C', AUC, P, V, Dr, Mr, PC]$.

4.3.2 Frequency Features. To further observe the distribution of the pupillary response signal in the frequency domain, we perform the Short-Time Fourier Transform (STFT) on the $\hat{I}[n]$ with length of $L = f_s \times T$. We adopt a sampling rate f_{fft} of 60 Hz with a Hanning window of size 32 to separate the signal into small chunks across time domain. The overlap step of moving windows is set to 16 and the FFT points number N_{fft} is 512. Then, a spectrogram S is generated. Thereafter, PipID truncates the spectrogram below a cutoff frequency f_{AF}, which is set to 5 Hz based on the concentration of amplitude. Subsequently to the cutoff, we get $Spec = S(: M_{spec}, :)$ with $M_{spec} = \frac{f_{AF} \times N_{fft}}{f_{fft}} = 43$.

Figures 10(a) and 10(b) demonstrate the spectrograms under two different light stimuli for two users (the overlap step is set to 30 in the figures). The differences exhibited in the spectrograms indicate their effectiveness as frequency domain features for user authentication. We take the average of $Spec$ across the time domain and generates the frequency domain features $F_f = \overline{Spec}$ with a length of M_{spec}. Figure 10(c) has validates the differences of the average and standard deviation of F_f across wavelengths between two users. Ultimately, for a sequence of pupillary response signals $\hat{I}[n]$, the extracted features $F = [F_s, F_t, F_f]$.



(a) (b) (c)

Figure 10: (a) Cutoff spectrogram of user 1 in 400 nm and 500 nm. (b) Cutoff spectrogram of user 2 in 400 nm and 500 nm. (c) $F_{f,std}$ and $F_{f,avg}$ of user 1 and user 2.

4.3.3 Feature Expanding Based on Left and Right Eye Difference. The eye tracker embedded in HTC VIVE headset is able to acquire individual eye movement data separately for the left and right eyes. Thus, features (F_{left}, F_{right}) can be extracted from the two eyes respectively. By taking the average of pupil sizes from both eyes, we are able to derive a sequence representing the change in cross-eye pupil diameter, which allows PipID to extract the feature F_{cross}. Furthermore, we observe subtle yet significant differences in pupillary response variations between the left and right eyes of each person, as stated in Section 3. We generate a sequence representing the bilateral difference by subtracting the right value from the left one. Statistical and temporal features from this sequence are designated as F_{diff}. We will not extract frequency-domain features from the sequence of differences.

Overall, for a given traces on each wavelength from one user (i.e., $I_{k,i}[n]$), PipID is able to obtain four features: F_{left}, F_{right}, F_{cross}, and F_{diff}. We denote $F_{I_{k,i}} = [F_{left}, F_{right}, F_{cross}, F_{diff}]$. Accordingly, the total samples for one user (N_s recordings for each of N_k wavelengths), i.e., $I[n] = [I_{k,i}[n]]_{N_k \times N_s}$, yields feature samples F_I:

$$F_I = [F_{I_{k,i}}]_{N_k \times N_s}, k \in [1, N_k], i \in [1, N_s]. \quad (4)$$

4.3.4 Features Expanding Based on Wavelength Subgroups. For each user, we have devised a unique combination consisting $N_w = 4$ wavelengths as visual stimuli in Section 4.2, and only the features derived from these wavelengths are utilized for the user’s classification training. To achieve high-accuracy authentication, we only consider a fixed-order wavelength combination for usability. Additionally, with a heightened focus on security measures, especially in mitigating the risk of replay attacks, we will integrate randomized factors into the design of our features.

Authentication Features. To enhance the performance of the classification model, for the dataset collected during user registration, we randomly combine the feature samples used for training. Assuming $w^J = [w_1, w_2, w_3, w_4]$ are the $N_w = 4$ (i.e., four) selected wavelengths for a given user, each time PipID randomly selects one sample from each wavelength, which can be combined to generate a new sample $F_c = [F_{I_{w_1,p}}, F_{I_{w_2,q}}, F_{I_{w_3,r}}, F_{I_{w_4,s}}]$, where p, q, r, s are time indexes. For the authentication phase, to comply with the design goal of lightweight authentication, we do not apply random combination processing to them. Therefore, the feature combination that servers as test sample is $[F_{K_{w_1}}, F_{K_{w_2}}, F_{K_{w_3}}, F_{K_{w_4}}]$.

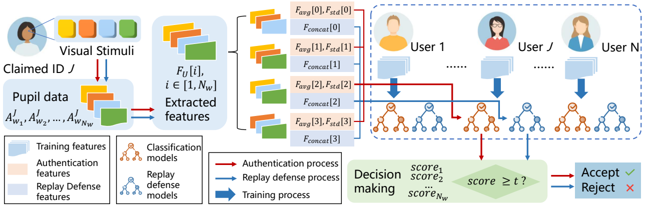


Figure 11: PipID's training and authentication processes.

Further, we apply the same process manner for both combination feature during registration and authentication. As shown in Figure 11, we take F_c as an example, each feature sample is divided into $\binom{4}{3} = 4$ subgroups, the i-th subgroup represents the feature samples from wavelengths $w^J[i \bmod 4]$, $w^J[i+1 \bmod 4]$ and $w^J[i{+}2 \bmod 4]$, denoted as $[F_c(w^J[i \bmod 4]), F_c(w^J[i{+}1 \bmod 4]), F_c(w^J[i+2 \bmod 4])]$, and we calculate the average and standard deviation (F_{avg}, F_{std}) across the wavelengths within each subgroup:

$$\begin{aligned} F_{avg}[i] =& avg([F_c(w^J[i \bmod 4]), F_c(w^J[i+1 \bmod 4]), \\ & F_c(w^J[i+2 \bmod 4])]), \end{aligned} \tag{5}$$

$$\begin{aligned} F_{std}[i] =& std([F_c(w^J[i \bmod 4]), F_c(w^J[i+1 \bmod 4]), \\ & F_c(w^J[i+2 \bmod 4])]). \end{aligned} \tag{6}$$

We can also following equations 5 and 6 to address subgroups during authentication. Thus, for the user's authentication request, our ultimate feature $F_U = [F_U[i]], i \in [1, 4]$ comprises four subgroups, with each subgroup $F_U[i] = [F_U[i][n]], n \in [1, N_f]$ containing N_f feature samples, with a length of $L_f = 1600$ for each sample.

Features to thwart replay attacks. We consider incorporating randomized factors into the features, specifically, the sequence of various lights employed in the stimuli. Taking F_c as an example, for the i-th subgroup, we concatenate the samples corresponding to different wavelengths:

$$\begin{aligned} F_{concat}[i] =& [F_c(w^J[i \bmod 4]), F_c(w^J[i+1 \bmod 4]), \\ & F_c(w^J[i+2 \bmod 4])]. \end{aligned} \tag{7}$$

4.4 Classification

4.4.1 Overview of Classification Module. The classification module leverages the extracted features to distinguish between legitimate and illegal users. The training and authentication process of PipID are illustrated in Figure 11. Given the relatively small dataset scale and subtle differences in pupillary response, using Deep Neural Networks or one-class classifiers tends to result in over-fitting. Therefore, we employ a binary classification method to train our network for each user. The training and test sets are independently divided. The model is fed with two types of explicitly labeled data: legitimate user's data and impersonator's data. Both are the feature samples under the wavelength combination of current legitimate user, which is chosen through KS test stated in Section 4.2. The former are the features extracted from the user's own pupillary response, following the steps outlined in Section 4.3. The latter are randomly selected from the features sets of other users, with an equal sample number as the legitimate one.

We leverage the Extremely Randomized Trees (Extra-Trees) to train our classifier. The Extra-Trees builds multiple decision trees and outputs the class that is the mode of the classes of the individual trees. It differs from traditional Random Forests in its randomized attribute selection and split criteria, making it more robust to noise.

4.4.2 Defending Against Basic Impersonation. During the authentication, we utilize the method described in Section 4.3.4 to generate $N_w = 4$ groups from a single trial. For i-th sub-sample, the trained model assigns a probability of being classified as legal as its score (denoted as score$_i$). PipID then calculates the product of the scores across all groups as the final score for the test sample, i.e., score $= \prod_{i=1}^{N_w}$ score$_i$. When score $\geq t$, where t is a given threshold, the model considers the sample from the legitimate user.

4.4.3 Defending Against Replay Attack. To further defend against replay attack, we focus on samples classified as legitimate user. Features calculated from equation 7 will be utilized to train a new Extra Trees model. For a specific challenge, the positive samples are features derived from the current challenge, while the negative sample set we employ consists of samples corresponding to wavelength combinations that an attacker might generate (using the same wavelengths but with different arrangements). For instance, if a random challenge is wavelengths of [400 nm, 450 nm, 500 nm, 600 nm], the corresponding samples are labeled as positive, while other samples (e.g., those from [450 nm, 600 nm, 400 nm, 500 nm]) are labeled as negative. Since these samples can be obtained during the enrollment phase, for each possible wavelength combination, we are able to pre-train its classification model offline; in total, there are 24 possible arrangements, corresponding to 24 mini-models for replay attack detection. Similar to the process of authentication, for a test sample, we calculate the final multi-score based on the scores of each of its subgroups, and only classify it as a positive sample if the multi-score exceeds a predefined threshold.

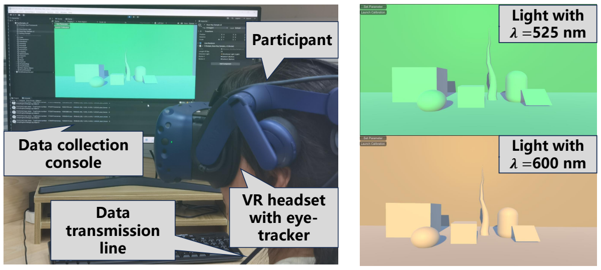
5 Evaluation of PipID

In this section, we evaluate the performance of PipID. Due to the page limit, the evaluations of impact factors and the user study are presented in the online appendix.

5.1 Experiment Settings

Hardware and software setup. In this study, we adopt the HTC VIVE Pro Eye headset integrated with Pupil Labs' eye tracker to collect users' eye motion data. The experimental virtual scenes are implemented on Unity, a platform for VR application development. We harness relevant APIs from the *Eye and Facial Tracking SDK* and the *SRanipal SDK* to record users' eye movements scripted in C#. The VR headset is connected to a computer system equipped with an Intel Core 14th Generation i7 processor and an NVIDIA GeForce RTX 4060 Ti graphics card, via a streamer and USB interface. Subsequent data processing and training are performed on this computer using Python.

Data collection. In the data collection phase, we recruit 52 volunteers comprising of 21 males and 31 females, with ages ranging from 16 to 29. Each participant is instructed to wear an HTC VIVE Pro Eye headset and seat comfortably at a designated desk in a laboratory with tranquil atmosphere and mild indoor lighting. Before



(a) Data collection settings.

(b) Virtual scenario.

Figure 12: Details of data collection procedure.

data acquisition, an eye-tracking calibration process inherent to the headset is implemented for each one, ensuring the reliability of the subsequent eye movement recordings. Then, the participants are immersed in a virtual scenario with the wavelength of its light source varies, as shown in Figure 12. They are instructed to relax and naturally observe the scene, allowing their gaze to wander freely across the virtual landscape or fixate on any point of interest within the environment.

The light source within the scenario sequentially transitions through a predetermined series of wavelengths: 400 nm, 450 nm, 475 nm, 500 nm, 525 nm, 550 nm, 570 nm, 600 nm, 630 nm and 700 nm, each maintains for a duration of 5 seconds. To facilitate pupil recovery and adaptation to the next wavelength, a brief 0.5-second interval of complete darkness is interposed between each wavelength. The variations of light across 10 wavelengths constitutes a sampling round, generating a sequence of user's pupil diameter over time (i.e., sample fragment). A total of 20 sampling rounds are launched for each participant. To conduct the longitudinal experiment, 10 participants are re-invited to repeat the sampling process on the 1st, 2nd, 3rd, 7th, and 14th days following their initial sampling. After data collection, we extract features corresponding to the wavelengths within the user-specific stimuli as shown in Section 4.2 and 4.3. Then a total of 200 feature samples are obtained after feature expanding for each person.

Training procedure. As mentioned in Section 4.4, to achieve user authentication, we train a binary classifier for each user, utilizing data with two labels: legitimate one and impersonators. Among all the datasets, 80% is allocated for training purposes, while the remaining 20% is reserved for testing.

Evaluation metrics. In this study, we use accuracy, false acceptance rate (FAR), false rejection rate (FRR) and F1-score as metrics to evaluate the performance of PipID. The accuracy is defined as the proportion of correct predictions for both genuine and impostor attempts. FAR measures the proportion of impostor access attempts that are incorrectly accepted by the system, while FRR measures the proportion of the genuine one being incorrectly rejected. The F1-score is a harmonic mean of precision and recall.

5.2 Overall Performance of PipID

Overall accuracy. PipID leverages Extra Trees to train classifiers for all 52 users. The testing results achieve an average accuracy of 98.65% and F1-score of 98.61%. Notably, 41 models (78.85% of the total) reaches the peak accuracy of 100%. Specifically, for the entire 10400 samples, the overall FAR and FRR is 0.77% (i.e., 40 samples being falsely accepted out of 5200) and 1.92% (i.e., 100 samples are falsely rejected out of 5200) respectively.

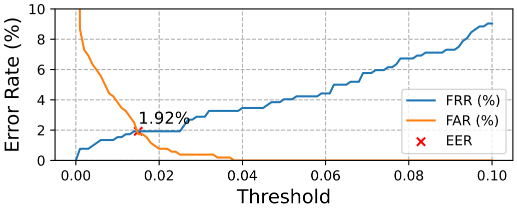


Figure 13: FAR, FRR and EER of PipID.

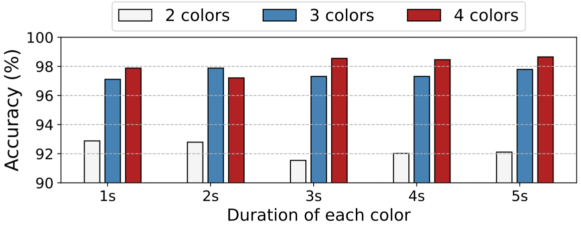


Figure 14: Impact of stimuli types and duration on PipID.

To visualize how different thresholds affect the classifier's performance, the changes of FAR and FRR under various thresholds are illustrated in Figure 13. The results of FAR 0.77% and FRR 1.92% are obtained when the threshold is set to 0.02. The EER (Equal Error Rate) is represented by the intersection point of the two curves, which achieves the lowest (i.e., 1.92%) at a threshold of 0.015.

Time overhead. We present the time overhead of PipID, including the authentication time (i.e., the time spent collecting user data during authentication), the enrollment time (i.e., the time spent on training), and the authentication response time (i.e., the time PipID takes to process the data and make a decision).

(1) Authentication time. Figure 14 shows the classification accuracy of PipID when the stimuli consist of varying numbers of colored light, each with different sample durations. Specifically, we test scenarios with 4, 3, and 2 colors in stimuli, coupled with sample duration of 5s, 4s, 3s, 2s, and 1s. The time reduction is achieved by extracting the first few seconds from the original sequences. Generally, an increase in both the number of colors and the sample duration leads to an enhancement in model accuracy.

Fortunately, shortening duration time has not make a considerable difference to PipID's performance. For instance, with 4 colors and a 3-second duration, PipID still achieves an accuracy of 98.55%. Besides, a duration time of 2 seconds yields the highest accuracy rate of 97.88% among scenarios with 3 colors in stimuli. This may be attributed to the fact that the pupillary response is most significant within the first 2 to 3 seconds when exposed to a specific wavelength of light, during which time sufficient identity-related information can be obtained. On the other hand, when decreasing the number of colors in stimuli from 4 to 3, the accuracy remains relatively stable as well, which generally hovers above 97%. These observations underscore the potential to compress the stimuli duration of PipID, thereby facilitating a shorter authentication time for better usability.

We compare PipID with other prevalent biometric user authentication methods in terms of authentication time and system performance. We consider three-color visual stimuli, each lasting for 2

Table 3: Comparison of Authentication Time and Performance Between PipID and Existing Biometrics User Authentication Methods.

Methods	FAR(%)	FRR(%)	Auth Time(s)	Biometrics
Blinkey [81]	4.00	4.00	7.3 - 11.7	Blink, Pupil Size
BioMove [47]	0.0032	1.30	≈20	Body, Eye Motion
Reflexive [61]	6.30	6.30	≤10	Saccade
OcuLock(EOG) [37]	3.55	3.55	≤10	EOG
Iris scan [6]	6.35	6.35	≤1	Iris Image
SkullConduct [55]	6.90	6.90	≤23	Bone Sound
GaitLock [57]	2.90	2.90	≤5	Gait Pattern
PipID (ours)	2.69	1.54	≤7.5	Pupil Size

Table 4: Impact of numbers of original training samples on PipID.

N_t	Accuracy (%)	FAR (%)	FRR (%)	F1-score (%)
10	98.65	0.77	1.92	98.61
9	97.81	0.87	3.49	97.68
8	97.68	0.80	3.85	97.58
7	97.71	1.03	3.55	97.59
6	97.66	1.92	2.75	97.65
5	97.63	2.56	2.18	97.69

seconds, as a good balance between security and usability. With this setup, the FAR and FRR are 2.69% and 1.54%, respectively, with an authentication time of 7.5 seconds (including a 0.5 s black interval before each light). As shown in Table 3, PipID outperforms most of the methods in both authentication time and accuracy. While some works offer faster authentication times, such as Iris scan [6] and GaitLock [57], the former is susceptible to replay attacks while the latter is vulnerable to shoulder-surfing attacks and physically demanding. Therefore, when considering both security and usability comprehensively, PipID demonstrates significant advantages.

(2) Enrollment time. To collect sufficient training and testing samples, we sampled from volunteers for 20 rounds (approximately 18 minutes) in our data collection process, which is not practical considering enrollment effort. To reduce the enrollment time, we conduct tests to evaluate the model performance with different numbers of original training samples (denoted as N_t). The results are illustrated in Table 4. We keep the training-testing dataset proportions consistent (training set proportion $P_t = 0.8$), as corresponding numbers of training samples can be expanded through random combination from original training set as stated in Section 4.3.3. The optimal accuracy (98.65%) and F1-score (98.61%) are achieved when $N_t = 10$.

As the number of training samples decreases, there is a subtle but not significant decline in authentication accuracy. Even with only 5 original training samples, the accuracy remains at 97.64%, suggesting that the number of sampling rounds during the enrollment phase can be reduced to approximately 5 rounds. Additionally, considering the analysis of authentication time, the duration of each wavelength segment can also be shortened. This implies that the enrollment time can be reduced to $10 \times (2 + 0.5) \times 5 = 125$ seconds (i.e., approximately 2 minutes).

(3) Authentication response. We also evaluate the authentication response speed of PipID. For each authentication attempt, the average time overhead $T_{overall}$ of PipID covers the data pre-processing time $T_p = 0.005$ s, feature extraction time $T_f = 0.226$ s, and classification time $T_c = 0.289$ s. Thus, the overall time $T_{overall} = 0.52$ s, which is accessible.

Table 5: Performance of PipID on different groups.

Groups (headcount)	Accuracy (%)	FAR (%)	FRR (%)
Male (21)	97.86	0.48	3.81
Female (31)	98.55	0.64	2.26
Non-myopia (13)	98.46	2.31	0.77
Uncorrected Myopia (12)	99.17	0.83	0.83
Wearing Spectacles (27)	97.22	2.96	2.59

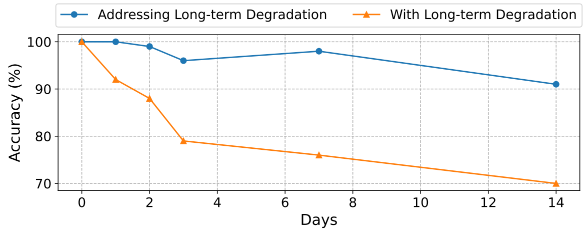


Figure 15: Long-term Performance of PipID.

Performance on different groups. We collect pupillary data from 52 volunteers, comprising 21 males and 31 females. Based on their vision status, 13 individuals are non-myopic, and among the 39 myopic subjects, 27 wear spectacles under headsets, while the remaining 12 not. We test the performance of PipID across these subgroups, as shown in Table 5. The average accuracy achieved for the male is 97.86%, and 98.55% for female. For the individuals of non-myopic, myopic without spectacles, and myopic with spectacles, the average accuracy are 98.46%, 99.17%, and 97.22%, respectively. The accuracy of the group wearing spectacles may be affected by lens refraction.

5.3 Longitudinal Study

To investigate the long-term performance of PipID, we invite 10 volunteers back for a multi-session sampling protocol, conducted on the 1st, 2nd, 3rd, 7th, and 14th days following their initial experiment, with 10 rounds of sampling per session for each volunteer. As a biometric, the response of pupillary diameter under different lights will exhibits variability over time, leading to a decline in authentication performance [27, 34]. To mitigate the long-term effects of biometric degradation, we implement a method where new samples are periodically added to the training datasets. Then, the classification model is retrained to incorporate these new data.

Figure 15 showcases the results of the longitudinal experiment. For comparison, we also test the authentication performance without addressing long-term degradation over the same time span, where the accuracy continuously decline over time. In contrast, with periodic updates of the training dataset, PipID maintains a relatively stable authentication accuracy. Specifically, on the 7th and 14th day, the accuracies are 98% and 91% respectively. This demonstrates that PipID is effective in counteracting the temporal variations in pupillary response. We acknowledge the limitations of the current approaches and will discuss the countermeasures in Section 6.

Table 6: The impact of the number of correctly guessed positions by the attacker on the success rate of replay attack. We have excluded scenarios where the attacker can exactly guess the correct challenge.

Number of matching positions	2	1	0
Expected occurrences (out of 23)	6	8	9
Attack success rate (%)	3.01	2.82	1.61

5.4 Robustness Against Adaptive Attacks

5.4.1 Replay Attacks. We conduct experiments to evaluate the resilience of PipID against replay attacks, where adversaries may attempt to access the VR device by exploiting pupillary responses generated from previously used visual stimuli of the user. For each user, we select $N_w = 4$ wavelengths as stated in Section 4.2. We assume that the attacker has obtained the user's pupillary responses from these 4 wavelengths but is unaware of light order in each random challenge. For a login attempt, the attacker randomly guesses the current wavelength permutation and sends the corresponding response to the model (10 samples are selected as test sets for each guessed permutation). We evaluate the system's effectiveness in defending against replay attacks under scenarios where 0, 1 or 2 positions of wavelength in the attacker's guessed stimuli match the actual stimuli.

The results shown in Table 6 indicate an average attack success rate (i.e., FAR) of 3.01%, 2.82%, and 1.61%, respectively, when the guessed challenge by the attacker differs from the current challenge in 2, 3, and 4 positions. There are a total of 4! = 24 permutations for 4 wavelengths and 23 permutations that differ from the current challenge. Among these, the number of permutations that differ from the current challenge in 2, 3, and 4 positions are 6, 8, and 9, respectively. Therefore, excluding scenarios where the attacker perfectly guesses the challenge, we can calculate the weighted average replay attack success rate (ASR) as $\frac{6}{23} \times 3.01\% + \frac{8}{23} \times 2.82\% + \frac{9}{23} \times 1.61\% = 2.40\%$. When considering the correct guesses by the attacker, we regard the true accept rate (TAR) of PipID in Section 5.2 as the ASR, which is 98.08%. Based on this, the overall ASR is recalculated as $\frac{23}{24} \times 2.40\% + \frac{1}{24} \times 98.08\% = 6.39\%$, indicating that PipID still achieves a protection rate of 93.61%. We discuss the limitations of PipID when defending against replay attacks and propose potential countermeasures in Section 6.

5.4.2 Open-World Attacks. In above-mentioned experiments, all pupillary information about impersonators is recorded in the known dataset and potentially used for model training. However, in an open-world scenario, there may be external attackers endeavoring to masquerade as legitimate users. To evaluate the model's performance in an open-world scenario, we divide 52 users into 35 known users and 17 foreign users at a ratio of 2:1. For each of the known users, an Extra Trees model is trained using our proposed method. We then collect test samples from both known and foreign users. Each test sample is run through all models of the known users. We consider a test sample to be originating from a specific user only if it yields substantially superior results on its own model in comparison to other models. Conversely, if the results are relatively indistinguishable across all models, the sample is presumed to originate from an impersonator who is not part of the known dataset.

We denote score_m as the final classification score of a sample obtained by the model of user m, as detailed in Section 4.4. In a close-world scenario, if score_m exceeds a given threshold, the sample is recognized as originated from user m. In the open-world scenario, if a test sample claims to belong to user i, we calculate the confidence parameter:

$$confidence = \frac{\text{score}_i - \frac{1}{M}\sum_{m=1}^{M}\text{score}_m}{\text{score}_i}. \tag{8}$$

For test samples with $\text{score}_m \geq t$ and $confidence > 0.5$, we consider them as belonging to user i; otherwise, they are deemed as illegal users. The final results show that in the open-world scenario, the model achieve an average accuracy of 97.22%, with a FAR of 3.12% and a FRR of 2.08%. When only considering test samples from foreign users, the probability of them being misclassified as a known legitimate user is 2.09%.

5.4.3 Shoulder-surfing Attacks. Since VR headsets fully cover the eyes, capturing eye movement is challenging for external observers. Besides, the visual stimuli used for authentication are only displayed on the virtual screen of the VR device for the user. No valid eye-tracking information is accessible to external attackers.

6 Limitations and Discussion

In this section, we discuss the limitations of PipID and identify potential countermeasures for further improving its performance. Due to the page limit, several discussions are presented in the online appendix.

Advanced adaptive attacks bypassing PipID. Regarding the three adaptive attacks with stronger assumptions discussed in Section 3, we analyze their feasibility to compromise PipID. For the first adaptive attack where the adversary is aware of the randomized stimuli, PipID would not be effective. However, implementing such an attack is highly complex. Given that VR headsets provide full coverage, it is extremely challenging for attackers to obtain the light stimuli through external observation. On the other hand, accessing the information displayed on virtual screens requires special permissions. Consequently, PipID can maintain highly robust performance in practice. Then, considering the second type of attack, in which the adversary gains system-level operational permissions, it would effectively neutralize all challenge-response authentication mechanisms. Nevertheless, VR platforms typically have controlled environments with restricted access to system-level resources. Therefore, launching an attack at the system level is considered a highly improbable assumption. Finally, regarding the third type of attack, where the attacker attempts to mimic another person's pupillary response patterns, studies [43] have shown that the pupillary light reflex is extremely difficult to replicate, even with conscious effort and extensive training. Therefore, such imitation is highly impractical, especially without deep biometric knowledge or precise control.

Replay attack defense limitations. Our defense strategy may prove ineffective if an attacker manages to determine the correct sequence of lights. To address this, a promising future optimization would involve expanding the security space of the challenge, for instance, by randomly varying the duration of each light, thus making the attacker's brute-force search approach futile. A more resilient defensive measure would be to integrate liveness detection directly into the system, ensuring that it can verify whether the current data is sourced from a genuine individual rather than outdated response data.

Longitudinal performance degradation. To mitigate long-term degradation in biometric features, we periodically add new samples to retrain the classification model (i.e., Extra Trees) in response to concept drift (e.g., dynamic changes in pupillary responses). However, due to the model's fixed parameters, its accuracy may still degrade if the distribution of new data diverges significantly from the original training set. To better accommodate evolving pupillary patterns, future work may explore adaptive learning (AL) and incremental learning (IL). AL supports real-time updates via online optimization, while IL preserves historical knowledge through continual learning strategies. We leave the exploration of a hybrid AL-IL framework that balances new and historical data as future work to enhance the system's robustness to long-term variability.

The impact of physiological impairments. According to medical literature [1, 8, 17, 25, 28, 38, 46], physiological impairments such as alcohol, prescription drugs or fatigue do have an impact on pupil diameter. Alcohol-induced central nervous system paralysis relaxes the pupillary sphincter muscle, resulting in abnormally enlarged pupil diameter [8, 25, 46], especially under low-light conditions. Moreover, larger quantities of alcohol consumption tend to result in more pronounced pupil dilation. Similarly, certain medications, particularly stimulants and antidepressants that affect the parasympathetic or sympathetic nervous systems, can cause dilated pupils [1, 17, 28]. Besides, [38] has shown that during periods of fatigue, pupil diameter exhibits a tendency to decrease as the mental load on the nervous system intensifies. Thus, we consider PipID applicable to users in a normal state both mentally and physically. Since VR devices are not as easily accessible for use as mobile phones, most users are unlikely in situations including heavy alcohol consumption or excessive fatigue.

7 Related Work

Wavelength of light affects pupil size. Many previous works [11, 18, 26] have shown that the blue light stimulation produces a larger pupil constriction amplitude compared with red light stimulation at all intensities. Young and Kimura [77] further indicate that high-intensity blue light leads to more sustained pupillary contraction than photopically equivalent red light. Ishikawa et al. [23] proves that the initial rate of pupil constriction induced by blue light is notably higher. Kawasaki and Kardon [11, 26] suggest that this consistent difference in amplitude is likely due to the added input from intrinsic activation that occurs when only blue light is used. Besides, [33] systematically investigates pupil response amplitude, latency, and constriction speed under single light stimuli of four distinct wavelengths, reporting the greatest amplitude and shortest latency with white and green light. Furthermore, [5] offers a detailed assessment of the pupillary light reflex elicited by purple, blue, and red monochromatic stimuli (administered individually or in combination for 5 minutes), demonstrating notable variations in constriction speed and magnitude across different spectral ranges.

Eye motion based authentication. Current research on eye motion-based authentication such as [49] and [47] instruct users to perform a series of controlled body movements (e.g., grasping, typing) while monitoring their unique head, hand, and eye movement patterns in VR environments. Besides, certain interactive stimuli (e.g., images, sounds, videos) can also trigger reflexive eye movements [13, 53, 61, 82]. Among them, [82] utilizes users' pupillary responses upon hearing several audio clips to achieve user authentication in VR devices. The approaches in [13, 53, 61] authenticate users by recording their gaze patterns on moving targets by specialized eye-tracker. Notably, Sluganovic et al. [61] employ a challenge-response paradigm, where the stimulus is a red dot changing positions on a dark screen, and the response is the reflexive saccade. These task-driven authentication methods require active user engagement through the act of focusing on dynamic targets. In contrast, Yano et al. [75] propose an authentication method based on the pupillary light reflex, which combines iris patterns and pupil features captured by a professional eye tracker under low or non-visible light conditions. Similarly, [74] chooses the light intensity as visual stimulus, and concludes that better authentication performance could only be achieved under moderate intensity.

Additionally, Sitzmann et al. [59] demonstrate a universal finding that 70% of people's attention is focused on the 20% "hotspot" regions of an image, making the distribution of areas of interest a potential feature. Several studies have integrated knowledge-based methods with biometrics. For instance, [81] enables users to select a blinking rhythm as a password and combines the pupil diameter changes during blinking intervals for authentication. What's more, as proposed by [64], different task-driven scenarios lead to distinct eye movement patterns, which paves the way for continuous authentication by employing eye-movement behaviors during daily tasks. Continuous authentication can constantly detect users' identities while they are using the system, without interfering with their normal work. [79] once attempted to implement continuous authentication in VR environments, but the implicit stimuli remain a moving target that users need to be concerned about. In order to select biometric features that are as unrelated to specific triggers as possible, Holland et al. [19, 20] investigate users' eye movement behavior during reading tasks and Eberz et al. [14] test the impact of task familiarity on eye movement and feature stability. Eberz et al. [12] address three challenges in practical eye movement authentication, including the dependency of eye movements on tasks, indicating that achieving the cross-task eye movement authentication requires significant effort.

8 Conclusion

In this study, we propose PipID, a seamless and robust VR user authentication system by employing the unique pupillary diameter responses under various light stimuli. We theoretically analyze and experimentally validate the potential of light-pupillary response as a novel biometric for authentication. Moreover, the incorporation of inter-ocular differences as a novel feature enriches our feature set. PipID is tested on dataset including 52 subjects and achieve an average accuracy of 98.65%, demonstrating its effectiveness in differentiating legitimate users from impersonators.

Acknowledgements

The authors from Shanghai Jiao Tong University were supported by the National Natural Science Foundation of China (No. 62302298, 62325207, 62132013), Young Elite Scientists Sponsorship Program by CAST (YESS20230589), and Startup Fund for Young Faculty at SJTU (23X010502192).

References

[1] Lakeside Academy. [n. d.]. What Kind of Drugs Cause Dilated Pupils? https://boysteenchallenge.mntc.org/what-kind-of-drugs-cause-dilated-pupils/.

[2] Abdullah Al Arafat, Zhishan Guo, and Amro Awad. 2021. VR-Spy: A Side-Channel Attack on Virtual Key-Logging in VR Headsets. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 564–572. doi:10.1109/VR50410.2021.00081

[3] Oliver Bergamin, Andreas Schoetzau, Keiko Sugimoto, and Mario Zulauf. 1998. The influence of iris color on the pupillary light reflex. *Graefe's Archive for Clinical and Experimental Ophthalmology* 236, 8 (1998), 567–570. doi:10.1007/s004170050122

[4] Marco I Bonelli and Nicolas Hamelin. 2022. Virtual Reality in Financial Trading and Investing: A Review of Literature and Applications. In *2022 8th International Conference on Virtual Reality (ICVR)*. IEEE, 97–101.

[5] Maria A. Bonmati-Carrion, Konstanze Hild, Cheryl M. Isherwood, Stephen J. Sweeney, Victoria L. Revell, Juan A. Madrid, Maria A. Rol, and Debra J. Skene. 2018. Effect of Single and Combined Monochromatic Light on the Human Pupillary Light Response. *Frontiers in Neurology* 9 (2018). doi:10.3389/fneur.2018.01019

[6] Fadi Boutros, Naser Damer, Kiran Raja, Raghavendra Ramachandra, Florian Kirchbuchner, and Arjan Kuijper. 2020. Iris and periocular biometrics for head mounted displays: Segmentation, recognition, and synthetic data generation. *Image and Vision Computing* 104 (2020), 104007. doi:10.1016/j.imavis.2020.104007

[7] Abraham G. Campbell, Thomas Holz, Jonny Cosgrove, Mike Harlick, and Tadhg O'Sullivan. 2020. Uses of Virtual Reality for Communication in Financial Services: A Case Study on Comparing Different Telepresence Interfaces: Virtual Reality Compared to Video Conferencing. In *Advances in Information and Communication*, Kohei Arai and Rahul Bhatia (Eds.). Springer International Publishing, Cham, 463–481.

[8] Swift River Addiction Campuses. [n. d.]. Alcohol and Your Eyes: Exploring the Effects on Pupil Dilation. https://swiftriver.com/blog/alcohol-and-your-eyes-exploring-the-effects-on-pupil-dilation/.

[9] M. S. Castelhano and J. M. Henderson. 2008. Stable individual differences across images in human saccadic eye movements. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale* 62, 1 (2008), 1–14. doi:10.1037/1196-1961.62.1.1

[10] Ciklum. 2024. AR/VR Trends and Predictions For 2025 & Beyond. https://www.ciklum.com/resources/blog/ar/vr-trends-and-predictions-for-2025-beyond.

[11] Michael Tri Hoang Do and King-Wai Yau. 2010. Intrinsically photosensitive retinal ganglion cells. *Physiological reviews* (2010).

[12] Simon Eberz, Giulio Lovisotto, Kasper B. Rasmussen, Vincent Lenders, and Ivan Martinovic. 2019. 28 Blinks Later: Tackling Practical Challenges of Eye Movement Biometrics. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security* (London, United Kingdom) *(CCS '19)*. Association for Computing Machinery, New York, NY, USA, 1187–1199. doi:10.1145/3319535.3354233

[13] Simon Eberz, Kasper Bonne Rasmussen, Vincent Lenders, and Ivan Martinovic. 2015. Preventing Lunchtime Attacks: Fighting Insider Threats With Eye Movement Biometrics. In *Network and Distributed System Security Symposium*. https://api.semanticscholar.org/CorpusID:750589

[14] Simon Eberz, Kasper B. Rasmussen, Vincent Lenders, and Ivan Martinovic. 2016. Looks Like Eve: Exposing Insider Threats Using Eye Movement Biometrics. *ACM Trans. Priv. Secur.* 19, 1, Article 1 (jun 2016), 31 pages. doi:10.1145/2904018

[15] eMarkerter. 2024. Tech giants rethink XR strategies as MR adoption plateaus and mobile AR accelerates. https://www.emarketer.com/content/tech-giants-rethink-xr-strategies-mr-adoption-plateaus-mobile-ar-accelerates.

[16] Sindhu Reddy Kalathur Gopal, Diksha Shukla, James David Wheelock, and Nitesh Saxena. 2023. Hidden Reality: Caution, Your Hand Gesture Inputs in the Immersive Virtual World are Visible to All!. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 859–876. https://www.usenix.org/conference/usenixsecurity23/presentation/gopal

[17] Healthline. [n. d.]. What Prescribed and Nonprescribed Drugs Cause Pupils to Dilate (and Why). https://www.healthline.com/health/dilated-pupils-drugs.

[18] Kristina Herbst, Birgit Sander, Dan Milea, Henrik Lund-Andersen, and Aki Kawasaki. 2011. Test–retest repeatability of the pupil light response to blue and red light stimuli in normal human eyes using a novel pupillometer. *Frontiers in neurology* 2 (2011), 9196.

[19] Corey Holland and Oleg V. Komogortsev. 2011. Biometric identification via eye movement scanpaths in reading. In *Proceedings of the 2011 International Joint Conference on Biometrics (IJCB '11)*. IEEE Computer Society, USA, 1–8. doi:10.1109/IJCB.2011.6117536

[20] Corey D. Holland and Oleg V. Komogortsev. 2013. Complex eye movement pattern biometrics: Analyzing fixations and saccades. In *2013 International Conference on Biometrics (ICB)*. 1–8. doi:10.1109/ICB.2013.6612953

[21] Zehao Huang, Gancheng Zhu, Xiaoting Duan, Rong Wang, Shuai Zhang, Yongkai Li, and Zhiguo Wang. 2024. Measuring eye-tracking accuracy and its impact on usability in Apple Vision Pro. doi:10.48550/arXiv.2406.00255v2

[22] Inverse. 2022. Gucci partners with Roblox to launch 'Gucci Town' metaverse world. https://www.inverse.com/input/style/gucci-roblox-metaverse-world-gucci-town.

[23] Hitoshi Ishikawa, Asami Onodera, Ken Asakawa, Satoshi Nakadomari, and Kimiya Shimizu. 2012. Effects of selective-wavelength block filters on pupillary light reflex under red and blue light stimuli. *Japanese journal of ophthalmology* 56 (2012), 181–186.

[24] John M. Jones, Reyhan Duezguen, Peter Mayer, Melanie Volkamer, and Sanchari Das. 2021. A Literature Review on Virtual Reality Authentication. In *Human Aspects of Information Security and Assurance*, Steven Furnell and Nathan Clarke (Eds.). Springer International Publishing, Cham, 189–198.

[25] A Kaifie, M Reugels, T Kraus, and M Kursawe. 2021. The pupillary light reflex (PLR) as a marker for the ability to work or drive - a feasibility study. *J Occup Med Toxicol* 16, 1 (2021), 39. doi:10.1186/s12995-021-00330-2

[26] Randy Kardon, Sara C. Anderson, Theresa G. Damarjian, Elizabeth M. Grace, Eric Stone, and Akio Kawasaki. 2009. Chromatic pupil responses: preferential activation of the melanopsin-mediated versus outer photoreceptor-mediated pupil light reflex. *Ophthalmology* 116, 8 (August 2009), 1564–1573. arXiv:19501408 doi:10.1016/j.ophtha.2009.02.007 Epub 2009 Jun 5.

[27] Sanjeev Kasthurirangan and Adrian Glasser. 2006. Age related changes in the characteristics of the near pupil response. *Vision Research* 46, 8 (2006), 1393–1403. doi:10.1016/j.visres.2005.07.004

[28] D Koller, M Saiz-Rodríguez, P Zubiaur, D Ochoa, S Almenara, M Román, D Romero-Palacián, A de Miguel-Cáceres, S Martín, M Navares-Gómez, G Mejía, A Wojnicz, and F Abad-Santos. 2020. The effects of aripiprazole and olanzapine on pupillary light reflex and its relationship with pharmacogenetics in a randomized multiple-dose trial. *Br J Clin Pharmacol* 86, 10 (Oct 2020), 2051–2062. doi:10.1111/bcp.14300 Epub 2020 Apr 27.

[29] Pupil Labs. [n. d.]. VR & AR-Eye tracking. https://pupil-labs.com/products/vr-ar.

[30] Ajey Lele. 2013. Virtual reality and its military utility. *Journal of Ambient Intelligence and Humanized Computing* 4 (2013), 17–26.

[31] Lixiang Li, Xiaohui Mu, Siying Li, and Haipeng Peng. 2020. A Review of Face Recognition Technology. *IEEE Access* 8 (2020), 139110–139120. doi:10.1109/ACCESS.2020.3011028

[32] Zhen Ling, Zupei Li, Chen Chen, Junzhou Luo, Wei Yu, and Xinwen Fu. 2019. I Know What You Enter on Gear VR. *2019 IEEE Conference on Communications and Network Security (CNS)* (2019), 241–249. https://api.semanticscholar.org/CorpusID:84834207

[33] L L Lobato-Rincón, Mdel C Cabanillas-Campos, C Bonnin-Arias, E Chamorro-Gutiérrez, A Murciano-Cespedosa, and C Sánchez-Ramos Roda. 2014. Pupillary behavior in relation to wavelength and age. *Frontiers in Human Neuroscience* 8 (2014), 221. doi:10.3389/fnhum.2014.00221

[34] L.L. Lobato-Rincón, M.del C. Cabanillas-Campos, C. Bonnin-Arias, E. Chamorro-Gutiérrez, A. Murciano-Cespedosa, and C. Sánchez-Ramos Roda. 2014. Pupillary behavior in relation to wavelength and age. *Frontiers in Human Neuroscience* 8 (2014), 221. doi:10.3389/fnhum.2014.00221

[35] Dillon Lohr and Oleg V. Komogortsev. 2022. Eye Know You Too: Toward Viable End-to-End Eye Movement Biometrics for User Authentication. *Trans. Info. For. Sec.* 17 (Jan. 2022), 3151–3164. doi:10.1109/TIFS.2022.3201369

[36] Kaidi Lõo, Jacolien van Rij, Juhani Järvikivi, and Harald Baayen. 2016. Individual Differences in Pupil Dilation during Naming Task. *Cognitive Science* (2016). https://api.semanticscholar.org/CorpusID:30144980

[37] Shiqing Luo, Anh Nguyen, Chen Song, Feng Lin, Wenyao Xu, and Zhisheng Yan. 2020. OcuLock: Exploring Human Visual System for Authentication in Virtual Reality Head-mounted Display. *Proceedings 2020 Network and Distributed System Security Symposium* (2020). https://api.semanticscholar.org/CorpusID:211266673

[38] Takashi Matsui, Shion Takahashi, Genta Ochi, Seiji Yoshitake, Daisuke Funabashi, Hiroki Matsuoka, Junichi Kagesawa, Shohei Dobashi, Hisashi Yoshimoto, Yosuke Sakairi, and Hideki Takagi. 2024. Cognitive decline with pupil constriction independent of subjective fatigue during prolonged esports across player expertise levels. *Computers in Human Behavior* 156 (2024), 108219. doi:10.1016/j.chb.2024.108219

[39] David H. McDougal and Paul D. Gamlin. 2010. The influence of intrinsically-photosensitive retinal ganglion cells on the spectral sensitivity and response dynamics of the human pupillary light reflex. *Vision Research* 50, 1 (2010), 72–87. doi:10.1016/j.visres.2009.10.012

[40] Joshua McVeigh-Schultz and Katherine Isbister. 2022. A "beyond being there" for VR meetings: envisioning the future of remote work. *Human–Computer Interaction* 37, 5 (2022), 433–453.

[41] Ülkü Meteriz-Yıldıran, Necip Fazıl Yıldıran, Amro Awad, and David Mohaisen. 2022. A Keylogging Inference Attack on Air-Tapping Keyboards in Virtual Environments. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 765–774. doi:10.1109/VR51125.2022.00098

[42] Tahrima Mustafa, Richard Matovu, Abdul Serwadda, and Nicholas Muirhead. 2018. Unsure How to Authenticate on Your VR Headset? Come on, Use Your Head!. In *Proceedings of the Fourth ACM International Workshop on Security and Privacy Analytics* (Tempe, AZ, USA) *(IWSPA '18)*. Association for Computing Machinery, New York, NY, USA, 23–30. doi:10.1145/3180445.3180450

[43] Daniel C Niehorster, Simon Bengtsson, Nicklas Brodin, and Anders Rasmussen. 2022. No evidence of conditioning of pupillary constriction despite overtraining. *PeerJ* 10 (15 Feb 2022), e12948. doi:10.7717/peerj.12948

[44] Nweon. 2023. Meta Patent |Eye-tracking using embedded electrodes in a wearable device. https://patent.nweon.com/27217.

[45] Nweon. 2024. Apple Patent |Eye detection methods and devices. https://patent.nweon.com/35321.

[46] American Academy of Ophthalmology. [n. d.]. Does alcohol affect pupil size in a typical person? https://www.aao.org/eye-health/ask-ophthalmologist-q/does-alcohol-affect-pupil-size-in-typical-person.

[47] Ilesanmi Olade, Charles Fleming, and Hai-Ning Liang. 2020. BioMove: Biometric User Identification from Human Kinesiological Movements for Virtual Reality Systems. *Sensors* 20, 10 (2020). doi:10.3390/s20102944

[48] Adedoyin Tolulope Oyewole, Omotayo Bukola Adeoye, Wilhelmina Afua Addy, Chinwe Chinazo Okoye, Onyeka Chrisanctus Ofodile, and Chinonye Esther Ugochukwu. 2024. Augmented and virtual reality in financial services: A review of emerging applications. *World Journal of Advanced Research and Reviews* 21, 3 (2024), 551–567.

[49] Ken Pfeuffer, Matthias J. Geiger, Sarah Prange, Lukas Mecke, Daniel Buschek, and Florian Alt. 2019. Behavioural Biometrics in VR: Identifying People from Body Motion and Relations in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3290605.3300340

[50] PICO. 2022. PICO 4 PRO. https://www.picoxr.com/cn/products/pico4-pro.

[51] Mihai Pop, Yves Payette, and Emma Santoriello. 2002. Comparison of the pupil card and pupillometer in measuring pupil size. *Journal of Cataract & Refractive Surgery* 28, 2 (2002), 283–288. doi:10.1016/S0886-3350(01)01222-6

[52] WGRM PS Rathnayake. 2018. Usage of mixed reality for military simulations. In *2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)*. IEEE, 1–5.

[53] Ioannis Rigas and Oleg V. Komogortsev. 2014. Biometric Recognition via Probabilistic Spatial Projection of Eye Movement Trajectories in Dynamic Visual Environments. *IEEE Transactions on Information Forensics and Security* 9, 10 (2014), 1743–1754. doi:10.1109/TIFS.2014.2350960

[54] Chris Roberts. 2007. Biometric attack vectors and defences. *Computers & Security* 26, 1 (2007), 14–25. doi:10.1016/j.cose.2006.12.008

[55] Stefan Schneegass, Youssef Oualil, and Andreas Bulling. 2016. SkullConduct: Biometric User Identification on Eyewear Computers Using Bone Conduction Through the Skull. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1379–1384. doi:10.1145/2858036.2858152

[56] Immo Schütz and Katja Fiehler. 2022. Eye Tracking in Virtual Reality: Vive Pro Eye Spatial Accuracy, Precision, and Calibration Reliability. *Journal of Eye Movement Research* 15 (09 2022), 3. doi:10.16910/jemr.15.3.3

[57] Yiran Shen, Hongkai Wen, Chengwen Luo, Weitao Xu, Tao Zhang, Wen Hu, and Daniela Rus. 2019. GaitLock: Protect Virtual and Augmented Reality Headsets Using Gait. *IEEE Transactions on Dependable and Secure Computing* 16, 3 (2019), 484–497. doi:10.1109/TDSC.2018.2800048

[58] A Sipatchin, S Wahl, and K Rifai. 2021. Eye-Tracking for Clinical Ophthalmology with Virtual Reality (VR): A Case Study of the HTC Vive Pro Eye's Usability. *Healthcare (Basel)* 9, 2 (2021), 180. doi:10.3390/healthcare9020180

[59] Vincent Sitzmann, Ana Serrano, Amy Pavel, Maneesh Agrawala, Diego Gutierrez, Belen Masia, and Gordon Wetzstein. 2018. Saliency in VR: How Do People Explore Virtual Environments? *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (2018), 1633–1642. doi:10.1109/TVCG.2018.2793599

[60] Vaclav Skala, Tristan Claude Louis Bellot, and Xavier Berault. 2023. Wavelength Computation from RGB. In *Computational Science and Its Applications – ICCSA 2023*, Osvaldo Gervasi, Beniamino Murgante, David Taniar, Bernady O. Apduhan, Ana Cristina Braga, Chiara Garau, and Anastasia Stratigea (Eds.). Springer Nature Switzerland, Cham, 423–430.

[61] Ivo Sluganovic, Marc Roeschlin, Kasper B. Rasmussen, and Ivan Martinovic. 2016. Using Reflexive Eye Movements for Fast Challenge-Response Authentication. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (Vienna, Austria) *(CCS '16)*. Association for Computing Machinery, New York, NY, USA, 1056–1067. doi:10.1145/2976749.2978311

[62] Sophie Stephenson, Bijeeta Pal, Stephen Fan, Earlence Fernandes, Yuhang Zhao, and Rahul Chatterjee. 2022. SoK: Authentication in Augmented and Virtual Reality. In *2022 IEEE Symposium on Security and Privacy (SP)*. 267–284. doi:10.1109/SP46214.2022.9833742

[63] Meta Store. 2022. Meta Quest Pro: Premium Mixed Reality. https://www.meta.com/quest/quest-pro/.

[64] Benjamin Tatler, Nicholas Wade, Hoi Kwan, John Findlay, and Boris Velichkovsky. 2010. Yarbus, Eye Movements, and Vision. *i-Perception* 1 (07 2010), 7–27. doi:10.1068/i0382

[65] Tobii. 2023. Eye tracking performance assessment — for VR/AR headsets and wearables. https://www.tobii.com/resource-center/reports-and-papers/eye-tracking-performance-assessment#cta-section.

[66] Tobii. 2024. Most advanced eye tracking system: Tobii Pro Spectrom. https://www.tobii.cn/products/eye-trackers/screen-based/tobii-pro-spectrum.

[67] Ruben Tolosana, Marta Gomez-Barrero, Christoph Busch, and Javier Ortega-Garcia. 2020. Biometric Presentation Attack Detection: Beyond the Visible Spectrum. *IEEE Transactions on Information Forensics and Security* 15 (2020), 1261–1275. doi:10.1109/TIFS.2019.2934867

[68] Felleman DJ Van Essen DC, Anderson CH. 1992. Information processing in the primate visual system: an integrated systems perspective. *science* 255, 5043 (1992), 419–23. doi:10.1126/science.1734518

[69] Varjo. 2023. Varjo XR-3, the first true mixed reality headset. https://varjo.com/products/varjo-xr-3/.

[70] VIVE. 2020. Next-level immersion with precision eye tracking. https://www.vive.com/sea/product/vive-pro-eye/overview/.

[71] Shu Wei, Desmond Bloemers, and Aitor Rovira. 2023. A Preliminary Study of the Eye Tracker in the Meta Quest Pro. In *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences* (Nantes, France) *(IMX '23)*. Association for Computing Machinery, New York, NY, USA, 216–221. doi:10.1145/3573381.3596467

[72] B Winn, D Whitaker, D. B. Elliott, and N. J. Phillips. 1994. Factors affecting light-adapted pupil size in normal human subjects. *Investigative Ophthalmology and Visual Science* 35, 3 (March 1994), 1132–1137.

[73] Chien-Min Wu, Chih-Wen Hsu, Tzu-Kuei Lee, and Shana Smith. 2017. A virtual reality keyboard with realistic haptic feedback in a fully immersive virtual environment. *Virtual Reality* 21 (2017), 19–29.

[74] Shilin Yan, Shan Chang, Jiacheng Wang, and Shanila Azhar. 2020. Using Pupil Light Reflex for Fast Biometric Authentication. In *Proceedings of the ACM Turing Celebration Conference - China* (Hefei, China) *(ACM TURC '20)*. Association for Computing Machinery, New York, NY, USA, 139–143. doi:10.1145/3393527.3393551

[75] Vitor Yano, Alessandro Zimmer, and Lee Luan Ling. 2015. Extraction and application of dynamic pupillometry features for biometric authentication. *Measurement* 63 (2015), 41–48. doi:10.1016/j.measurement.2014.12.001

[76] A Yasukouchi, T Hazama, and T Kozaki. 2007. Variations in the light-induced suppression of nocturnal melatonin with special reference to variations in the pupillary light reflex in humans. *Journal of Physiological Anthropology* 26, 2 (March 2007), 113–21. doi:10.2114/jpa2.26.113

[77] Rockefeller S.L. Young and Eiji Kimura. 2008. Pupillary correlates of light-evoked melanopsin activity in humans. *Vision Research* 48, 7 (2008), 862–871. doi:10.1016/j.visres.2007.12.016

[78] Zhen Yu, Hai-Ning Liang, Charles Fleming, and Ka Lok Man. 2016. An exploration of usable authentication mechanisms for virtual reality systems. In *2016 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*. 458–460. doi:10.1109/APCCAS.2016.7804002

[79] Yongtuo Zhang, Wen Hu, Weitao Xu, Chun Tung Chou, and Jiankun Hu. 2018. Continuous Authentication Using Eye Movement Response of Implicit Visual Stimuli. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 177 (Jan. 2018), 22 pages. doi:10.1145/3161410

[80] Zhihui Zhang, Josep M. Fort, and Lluis Giménez Mateu. 2023. Facial expression recognition in virtual reality environments: challenges and opportunities. *Frontiers in Psychology* 14 (2023). doi:10.3389/fpsyg.2023.1280136

[81] Huadi Zhu, Wenqiang Jin, Mingyan Xiao, Srinivasan Murali, and Ming Li. 2020. Blinkey: A two-factor user authentication method for virtual reality devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–29.

[82] Huadi Zhu, Mingyan Xiao, Demoria Sherman, and Ming Li. 2023. SoundLock: A Novel User Authentication Scheme for VR Devices Using Auditory-Pupillary Response. (2023). https://api.semanticscholar.org/CorpusID:257501425