Journal of the Brazilian Computer Society, 2025, 31:1, doi: 10.5753/jbcs.2025.5870
This work is licensed under a Creative Commons Attribution 4.0 International License.

Harnessing Foveated Rendering and AI to Tackle VR Cybersickness: A Feature-Centric Perspective

Thiago Malheiros Porcino [**Laboratório Nacional de Computação Científica (LNCC)** | ***thiago@lncc.br***]
Jauvane Cavalcante de Oliveira [**Laboratório Nacional de Computação Científica (LNCC)** | ***jauvane@acm.org***]
Érick Oliveira Rodrigues [**Universidade Tecnológica Federal do Paraná (UTFPR)** | ***erickrodrigues@utfpr.edu.br***]
Horácio Macêdo [**Universidade Federal Fluminense (UFF)** | ***horaciomacedo@id.uff.br***]
Victor Ferrari Sassi [**Universidade Federal Fluminense (UFF)** | ***victorferrari@id.uff.br***]
Daniela Trevisan [**Universidade Federal Fluminense (UFF)** | ***daniela@ic.uff.br***]
Esteban Clua [**Universidade Federal Fluminense (UFF)** | ***esteban@ic.uff.br***]

Laboratório Nacional de Computação Científica (LNCC). Av. Getulio Vargas, 333 - Quitandinha, Petrópolis - RJ, Brazil

Received: 09 April 2025 • **Accepted:** 11 July 2025 • **Published:** 21 August 2025

Abstract As virtual reality becomes increasingly immersive, issues related to cybersickness pose a major challenge. This review investigates how foveated rendering techniques, powered by artificial intelligence, are transforming our response to this topic. We analyze the primary factors that lead to cybersickness, including latency, field of view, vergence-accommodation mismatch, and unnatural locomotion, while demonstrating how adaptive visual strategies can significantly alleviate user discomfort. By considering individual traits like age, previous virtual reality experience, and real-time physiological indicators, including heart rate and skin conductance, modern rendering systems are evolving to be more intelligent and user-specific. We emphasize the role of advanced machine learning models, from interpretable symbolic frameworks to deep neural networks, along with gaze prediction systems that enable real-time adjustments through predictive rendering and user-context-specific optimization. Our findings highlight the promise of closed-loop rendering systems, which aim to preserve visual fidelity while enhancing comfort and engagement, steering to toward safer, more personalized virtual reality experiences.

Keywords: Virtual Reality, Cybersickness, Foveated Rendering, Artificial Intelligence, Adaptive Systems

1 Introduction

Over the last five years (2020–2025), there has been a significant rise in scholarly articles related to cybersickness (CS). A search for "cybersickness" generated around 11,900 results in this period[1], more than double the roughly 4,880 publications from the preceding five years (2015-2020)[2].

Additionally, Market.us [2024] reports that the global Artificial Intelligence (AI) market in Virtual Reality (VR) is rapidly growing, with forecasts estimating a rise from USD 21 billion in 2023 to USD 198 billion by 2033. This growth is driven by the incorporation of AI technologies, including machine learning and natural language processing, into VR systems, resulting in more immersive, adaptive, and personalized virtual experiences. Recent studies focus on optimization strategies [Ye *et al.*, 2024] that boost rendering efficiency and minimize perceptual inconsistencies. Among these, foveated rendering has shown promise. This method prioritizes rendering high-resolution images in the user's focal area while reducing detail in the peripheral vision. As a result, it alleviates the computational load, enhances performance, and lowers latency, potentially helping to mitigate CS [Hussain *et al.*, 2021].

By merging foveated rendering with adaptive techniques, these difficulties can be addressed more successfully, improving performance and user comfort. Additionally, integrating AI enhances the effectiveness of foveated rendering [Liu *et al.*, 2025], enabling real-time, personalized content tailored to individual user data [Valmorisco *et al.*, 2024].

Therefore, this study conducts a literature review of publications from 2015 to 2025 to investigate the research question: "How can AI-driven foveated rendering strategies reduce CS in virtual reality, considering influencing factors, user traits, and adaptive system responses?

2 Research Methodology

A literature review was performed utilizing key academic databases like IEEE Xplore, ACM Digital Library, and Google Scholar to guarantee comprehensive coverage of the topic. The search approach was based on an initial assessment and enhanced through a iterative process of selecting

[1] Search conducted in Google Scholar using the keyword "cybersickness" and filtering results by year (2020–2025). Accessed on March 27, 2025.
[2] Search conducted in Google Scholar using the keyword "cybersickness" and filtering results by year (2015–2020). Accessed on March 27, 2025.

keywords, incorporating Boolean operators. The final search queries included:

```
Query 1: ("cybersickness" OR "motion
sickness" OR "simulator sickness")
AND ("virtual reality" OR "VR") AND
("latency" OR "foveated rendering" OR
"immersion" OR "visual discomfort" OR
"VR displays") OR ("neural rendering")
```

```
Query 2: ("cybersickness" OR "motion
sickness" OR "simulator sickness") AND
("machine learning" OR "deep learning"
OR "Explainable AI") OR ("strategies"
OR "AI-based foveated rendering") OR
("gaussian splatting")
```

This query was applied to the titles, abstracts, and keywords of retrieved publications. The inclusion criteria comprised:

- peer-reviewed journal articles or conference papers;
- focus on foveated rendering, CS, visual fatigue, user experience, or perceptual discomfort in VR environments;
- written in English;
- publication date between 2015 and 2025;
- inclusion of empirical data or systematic reviews.

Exclusion criteria included:

- studies unrelated to VR contexts (e.g., 2D interfaces or desktop simulations);
- lack of focus on CS or perceptual discomfort;
- duplicate or inaccessible full-text articles.

The study selection process is illustrated in the PRISMA flow diagram (Figure 1).

Additionally, its important to note, this work builds on a previous review on cybersickness causes, strategies, and classification methods [Porcino *et al.*, 2021], but it introduces a distinct and novel focus. Although the earlier work provided a broad overview of cybersickness factors and mitigation techniques, the present study advances this discussion by specifically investigating the role of adaptive and AI-driven foveated rendering techniques in mitigating cybersickness within virtual reality environments. Furthermore, this manuscript applies a systematic literature review methodology (PRISMA) to comprehensively analyze recent developments from 2015 to 2025, with particular emphasis on the integration of artificial intelligence, gaze prediction, and real-time adaptive systems in foveated rendering solutions.

2.1 Contributions

This study tackles the research question and offers the following contributions:

- It identifies and examines key factors that contribute to CS, highlighting how foveated rendering techniques can mitigate their impact.
- It assesses AI-powered foveated rendering methods designed to lessen CS in virtual reality (VR).

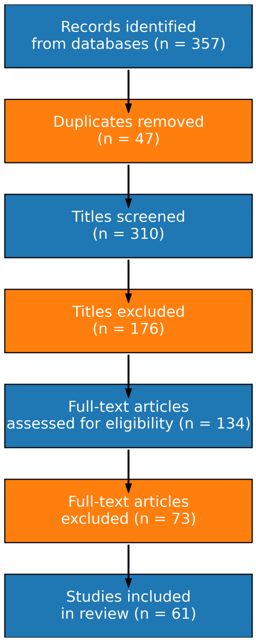

Figure 1. PRISMA flow diagram illustrating the selection process followed in this study. A total of 357 records were initially identified from databases. After removing 47 duplicates, 310 titles were screened, leading to the exclusion of 176 records. Subsequently, 134 full-text articles were assessed for eligibility, with 73 excluded based on predefined criteria. Ultimately, 61 articles were included in the final review (excluding self-citations).

- It identifies essential features for machine learning models, which include subjective factors like age, objective system metrics such as latency and field of view, and physiological signals like electrocardiograms and electrodermal activity.

The structure of this paper is as follows: Section 3 explores the principles and methods of foveated rendering. Section 4 examines the factors contributing to visual discomfort, with an emphasis on subjective experiences, physiological responses, and objective system characteristics. Section 5 presents studies focused on mitigating cybersickness through AI-enhanced foveated rendering techniques. In Section 6, we provide five guidelines and recommendations, along with future directions for improving comfort in virtual reality through adaptive rendering. Finally, Section 7 concludes the study.

3 Background: Foveated Rendering and its Relevance to CS

3.1 Principles of Foveated Rendering

Foveated rendering (FR) is a widely used optimization strategy in computer graphics that seeks to reduce computational demands by adjusting rendering resolutions based on perceptual value. This method capitalizes on the fundamental characteristics of the human visual system, which exhibits the highest visual acuity at the central region known as the fovea and progressively diminishes towards the peripheral vision [Jabbireddy *et al.*, 2022].

The human retina features an uneven arrangement of photoreceptor cells; cones are densely concentrated in the central fovea, which facilitates high-resolution color vision. Conversely, the peripheral areas predominantly contain rods, which are adapted for low-light settings but offer lower spatial resolution [Jabbireddy *et al.*, 2022]. Figure 2 illustrates the difficulty in producing uniformly high-resolution images throughout the visual field, given that the fovea covers merely about 5° of the total field, which is much smaller than the extent of the peripheral visual field [Mohanto *et al.*, 2022].

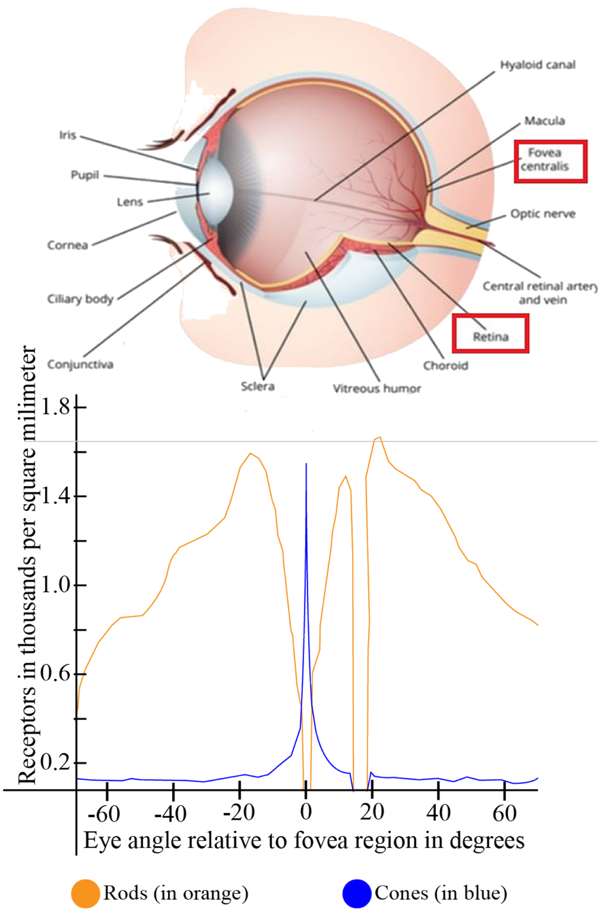

Figure 2. Spatial arrangement of rod and cone photoreceptors within the human retina, highlighting the high density of cones in the central foveal area and the predominance of rods toward the periphery, apart from the blind spot occurring around 20 degrees eccentricity [Jabbireddy *et al.*, 2022; Lisboa *et al.*, 2023].

3.2 Classifying Foveated Rendering

Various researchers have systematically categorized foveated rendering (FR) techniques, considering factors such as the definition of the foveal region, rendering pipelines, hardware, and strategies for reducing computational load [Duchowski *et al.*, 2003; Reingold *et al.*, 2003; Spjut and Boudaoud, 2019; Mohanto *et al.*, 2022; Wang *et al.*, 2023b]. Each taxonomy has its particularities. Spjut and Boudaoud [2019] argue that foveated displays should be classified in relation to user gaze and, therefore, propose a two-axis taxonomy: one for acuity matching and another for gaze contingency. These two dimensions are independent and can be combined in various ways.

The first axis analyses the display regarding how well the foveated rendering angular resolution, measured in cycles per degree, decay matches the human visual system's cones and rods distribution. There are four possible outcomes in this scenario: the resolution is perceptually higher than the cone-rod distribution throughout the retina (acuity matched), the resolution is perceptually higher than the cone-rod distribution only in the fovea and lower in the periphery (fovea matched), the resolution is perceptually higher only in the periphery (peripherally matched), and the resolution does not match the human eye threshold for surpassing our perception in any of the regions (non-acuity matched).

The second axis measures how the technique matches the user's gaze direction, adapting the displayed image to the user's gaze. This is also divided into four classes: the gaze is fully matched by the angular resolution on the display (fully foveated), the gaze is matched by a good amount of angular resolutions on the display (of $\pm 15°$; practically foveated), the gaze is matched by a severely limited amount of angular resolutions (of approximately $\pm 7°$; partially foveated) and the gaze is static on the screen (non-foveated).

Mohanto *et al.* [2022] classifies FR methods into four primary groups: adaptive resolution, geometric simplification, shading simplification with chromatic degradation, and spatiotemporal degradation. These groups are based on the methods used to achieve foveation, and each of them has subdivisions according to the render pipeline and dynamic adjustment of the fovea region. The only rendering pipelines present in this taxonomy are ray tracing and rasterization.

Adaptive resolution techniques adjust the resolution of different regions in an image to enhance computational efficiency while preserving perceptual quality. However, overly reducing resolution may result in visual artifacts like aliasing, necessitating careful management of parameters.

Geometric simplification methods minimize complexity in outer geometries based on factors such as eccentricity, object size, motion, and distance. Although these techniques can efficiently conserve resources, they carry the potential risk of creating visual artifacts like flickering. This issue is frequently mitigated with further visual smoothing techniques.

Techniques for shading simplification and chromatic degradation decrease shading complexity and color accuracy in outer areas, effectively conserving computational resources with little effect on perceived visual quality. These methods are often utilized together to enhance computational efficiency.

Spatio-temporal degradation reduces refresh rates selectively in peripheral vision areas, reallocating resources efficiently. Although effective, this method may induce perceptual artifacts due to variability in human visual sensitivity, such as peripheral flickering or narrowing of the perceived visual field.

Wang *et al.* [2023b] classifies the methods based on three distinct factors: the data type being rendered on screen, how foveated rendering is achieved, and the rendering paradigm employed. This taxonomy covers more variables and tackles more diverse scenarios since it contemplates rendering paradigms and data types not even mentioned by the other two presented works.

When it comes to data types, Wang *et al.* [2023b] considers techniques that aim to render frames (be they single images or videos), volume data, geometric meshes, point clouds, hologram data, or light fields. The foveation principle covers multi-spatial resolutions (conceptually close to Mohanto's adaptive resolution methods), multi-temporal resolutions (close to spatio-temporal degradation), multi-color resolution, multi-luminance resolution and level of detail. Lastly, besides rasterization and ray tracing, Wang [Wang *et al*., 2023b] also considers foveated rendering while making use of ray casting, instant radiosity, shadow mapping, neural rendering, photon mapping, phase retrieval, data transmission, and online/offline simplification.

One of the most recent FR advancements is the Foveated Path Culling, aimed at optimizing ray traced scenes [Henriques *et al*., 2024] with help from radiance fields. This technique replaces the peripheral view of the image with a radiance field reconstruction of the same scene while maintaining the foveal vision ray-traced, blending both the ray-traced view and the radiance field. Its main draw is gaze dependency from radiance fields, which approximates ray-traced effects, allowing for adjusting the peripheral vision in the same way as the foveal vision. This technique originally employed NeRFs [Mildenhall *et al*., 2021; Deng *et al*., 2022] in the periphery, but the advent of 3D Gaussian Splatting [Kerbl *et al*., 2023] proved itself more appropriate to this goal due to its render pipeline that allows for high resolutions with low latency.

These optimizations can play a crucial role in mitigating CS in VR [Hussain *et al*., 2021].

4 CS and Contributing Factors

The human brain relies on the integration of sensory input from multiple systems, including the vestibular apparatus, visual pathways, and proprioceptive feedback, to maintain orientation and balance. When conflicting signals arise among these systems, the brain experiences difficulty in resolving discrepancies, which can cause symptoms such as nausea, dizziness, and general discomfort, commonly called motion sickness (MS) [Kemeny *et al*., 2020]. Such sensory incongruence often manifests itself in real-world scenarios where visual cues suggest motion but corresponding physical movement is absent, as experienced in cars, boats, or airplanes.

A comparable phenomenon emerges in virtual and digital environments, termed visually induced motion sickness (VIMS). coined the term ”gaming sickness” to describe VIMS symptoms observed during video gameplay. Within virtual reality (VR) contexts, this condition is widely referred to as CS [Mazloumi Gavgani *et al*., 2018], encompassing manifestations such as nausea, vertigo, and spatial disorientation, which result from a mismatch between visual perception and vestibular signals. In contrast, when such symptoms occur during the use of flight or driving simulators, they are generally classified as simulator sickness, a term commonly applied in the domain of training simulations and immersive technologies [Porcino *et al*., 2021].

Motion sickness is typically divided into two main categories [Kemeny *et al*., 2020]: transportation sickness, which stems from actual vehicular movement, and simulator sickness, which is induced by artificial or virtual motion. Although both forms exhibit overlapping physiological responses, their underlying causes and effective mitigation techniques differ depending on whether the motion is physically real or perceptually simulated. In particular, the symptoms associated with CS and postural instability closely mirror those found in traditional motion sickness, further supporting the theory that sensory conflict is a central mechanism underlying both conditions. This overlap reinforces the importance of addressing sensory mismatches in the design of immersive virtual environments to reduce the incidence and severity of CS.

Given this theoretical foundation, it is crucial to examine the specific factors that contribute to the onset of CS.

4.1 Key Contributing Factors to CS

The literature identifies multiple factors associated with CS in virtual reality (VR), particularly for users of head-mounted displays (HMDs) [Stanney *et al*., 2020; Biswas *et al*., 2024]. However, in this section, we discuss the key factors that influence CS and, based on our review of the literature, can be specifically addressed using foveated rendering techniques.

4.1.1 Latency

Latency (illustrated in Figure 3), the delay between a user's action and the corresponding visual feedback presented within a head-mounted display (HMD), is a critical factor in virtual reality (VR) systems [Feldstein and Ellis, 2020; Kundu *et al*., 2021]. Elevated latency disrupts the temporal synchronization between physical motion and visual output, exacerbating sensory conflicts and perceptual inconsistencies that often lead to user discomfort, disorientation, and, in certain instances, motion sickness [Wang *et al*., 2023a].

From a perceptual perspective, the human visual system is highly sensitive to temporal discrepancies. Research suggests that humans can detect visual changes occurring at frequencies up to 500 Hz, which implies a latency detection threshold near 2 milliseconds [Davis *et al*., 2015]. However, as noted by Feldstein and Ellis [2020], the mere detection of latency does not necessarily result in CS. The extent to which latency induces discomfort is contingent upon several factors, including the design of the virtual environment, the nature of the task, and individual susceptibility.

Given the strong correlation between latency and user comfort, minimizing delay is crucial for providing seamless and natural interactions, particularly in latency-sensitive applications such as VR gaming [Vlahovic *et al*., 2019], simulation training [Brunnström *et al*., 2020], and collaborative virtual environments [Lei *et al*., 2022].

In this context, foveated rendering techniques can help mitigate latency by selectively reducing the computational load in peripheral regions, thereby enabling higher frame rates and faster visual response times. By prioritizing high-resolution rendering in the foveal region, where visual acuity is highest, and reducing detail in the peripheral vision, foveated rendering alleviates the processing burden on the GPU, resulting in lower latency. Furthermore, gaze-prediction models integrated with foveated rendering can enhance performance by anticipating the user's gaze direction and pre-rendering visual content accordingly, thus minimizing the perception of lag [Patney *et al.*, 2016]. Additionally, adaptive foveated rendering methods that dynamically adjust rendering quality based on system performance and user movement speed can further optimize responsiveness, ensuring a smoother and more immersive VR experience.

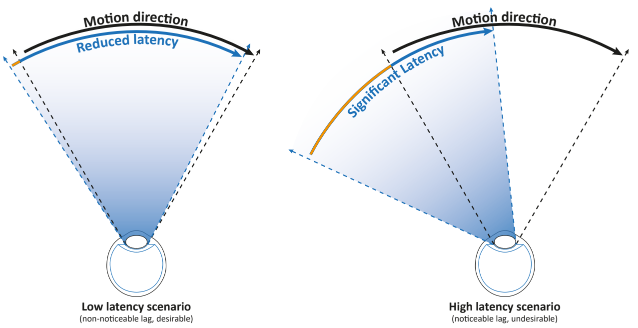

Figure 3. Latency is one of the biggest challenges in virtual reality, as it directly impacts the delay between a user's movement, such as turning their head, and the corresponding update on the display.

4.1.2 Field of View (FoV)

The field of view (FoV) also profoundly affects user experience by balancing immersion and comfort [Kelly *et al.*, 2024]. A wider FoV enhances immersive experiences but simultaneously increases sensory conflicts, thus increasing the risk of cybersickness [Oh and Son, 2022]. In contrast, narrowing the FoV can mitigate discomfort by decreasing peripheral visual stimulation, but negatively impacts the immersive quality of VR experiences. Maintaining an optimal FoV is crucial to preventing CS while ensuring an engaging virtual environment [Bala *et al.*, 2021].

More detailed, a wide FoV increases the amount of peripheral information received by the visual system, which, if not processed correctly, can lead to sensory conflicts between the vestibular and visual systems. This discrepancy is a known cause of CS, as the brain struggles to reconcile perceived motion with bodily cues [Kemeny *et al.*, 2020]. On the other hand, a restricted FoV limits the user's sense of presence and spatial awareness, reducing immersion and potentially diminishing the effectiveness of the VR experience.

Furthermore, Oh and Son [2022] investigated the impact of different Field of View (FOV) sizes on CS using their CYRE application, which included three internal FOV levels: small (30°), middle (45°) and large (90°). Their results did not show statistically significant differences in CS levels between the large and middle FOV conditions, suggesting that marginal reductions in FOV have little effect. However, when FOV was substantially reduced, CS levels decreased significantly by approximately 6% compared to large and medium conditions.

The Figure 4 and Figure 5[3] show the relationship between FoV range, immersion impact, and CS risk.

[3]Figure 5 was adapted from the Interaction Design Foundation, *Field of View (FoV) in Extended Reality*, available at `https://www.interaction-design.org/literature/topics/field-of-view-fov-in-extended-reality`, licensed under CC BY-SA 4.0 (`https://creativecommons.org/licenses/by-sa/4.0/`).

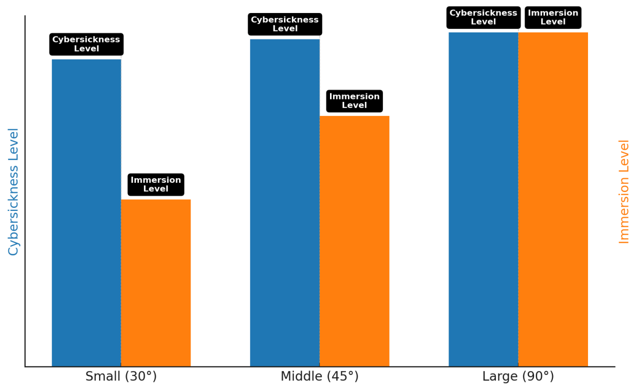

Figure 4. The chart illustrates how increasing the FoV enhances immersive experiences (orange) but also raises the risk of CS (blue). Lower FoV settings, getting the values from [Oh and Son, 2022] as parameters (30° small),(45° middle) result in reduced sensory conflicts, while wider FoV ranges (90° large) provide higher immersion at the cost of increased discomfort for sensitive users.

Foveated rendering offers a balanced solution by dynamically adjusting peripheral visual detail and resolution based on the user's gaze. This technique optimizes rendering performance by preserving high detail in central vision while reducing computational demands in peripheral areas. By leveraging real-time gaze tracking, foveated rendering can adapt the FoV dynamically, ensuring that users experience high-fidelity visuals where they are looking while minimizing unnecessary detail in less critical areas. This adaptive approach not only improves the performance of the system, but also reduces the likelihood of sensory conflicts, thus mitigating symptoms of CS [Wu and Suma Rosenberg, 2022].

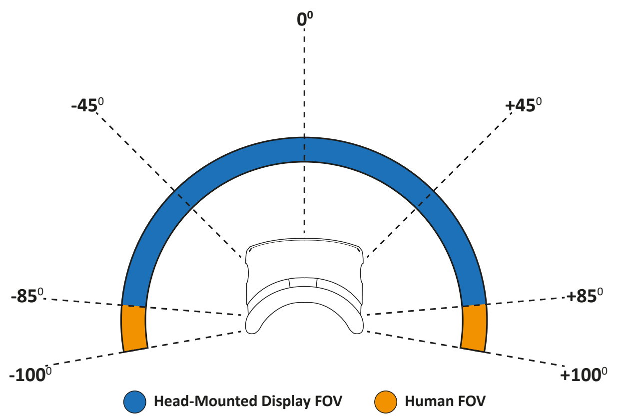

Figure 5. Comparative diagram illustrating the horizontal field of view (FoV) in head-mounted displays (HMDs), with values ranging from 90° (spanning from –45° to +45°) to 170° (spanning from –85° to +85°). These are contrasted with the natural human horizontal FoV, which extends up to approximately 200° (–100° to +100°), highlighting the limitations and variability of peripheral vision coverage across different HMD designs [Xiao and Benko, 2016].

4.1.3 Vergence and Accommodation

Accurate simulation of depth-of-field (DoF) is essential for realistic and comfortable visual experiences in VR. A key challenge is the vergence-accommodation conflict (Figure 6), which arises because virtual reality headsets present images at a fixed focal distance, while our eyes naturally adjust both vergence (eye convergence) and accommodation (lens focus) to perceive depth in the real world. This mismatch can lead to unnatural depth cues, visual discomfort, and fatigue, particularly during prolonged VR use. Studies indicate that users exposed to persistent vergence-accommodation conflicts may experience increased visual strain, headaches, and difficulty focusing after exiting VR environments [Hussain *et al*., 2021].

In HMDs with gaze tracking, improper DoF simulation can exacerbate these issues by producing unrealistic blur effects. Human vision relies on objects outside the focal range appearing blurred as a crucial depth cue, but when this is inaccurately rendered, it disrupts depth perception, making objects appear either unnaturally sharp or incorrectly blurred. This misalignment in visual cues can create perceptual confusion, increasing the likelihood of CS and reducing the overall sense of immersion. Furthermore, inconsistencies in DoF rendering can cause difficulty in distinguishing between near and far objects, impairing spatial awareness in virtual environments.

Foveated rendering techniques can help mitigate these challenges by dynamically adjusting rendering resolution and applying adaptive blurring based on gaze tracking. By ensuring that only objects within the user's focal region appear sharp while others blur naturally, foveated rendering improves visual realism and reduces sensory conflicts that contribute to cybersickness. Advanced implementations of foveated rendering integrate predictive gaze models, which anticipate the user's focal depth and adjust DoF effects accordingly, further refining visual accuracy and enhancing user comfort. Additionally, machine learning-based DoF rendering techniques can personalize blur levels for individual users based on their unique visual processing characteristics, optimizing the VR experience while reducing strain. These adaptive strategies are particularly useful in high-fidelity VR applications such as training simulations, medical imaging, and virtual prototyping, where accurate depth perception is critical to task performance and user well-being.

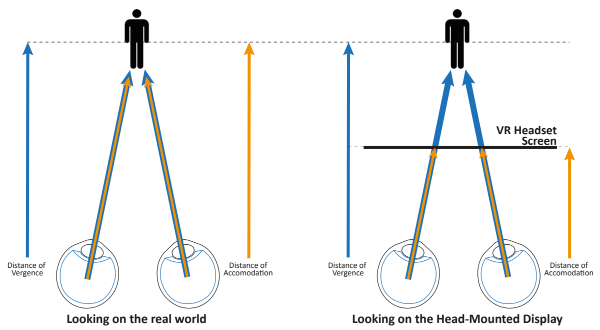

Figure 6. The vergence-accommodation conflict occurs when the eyes converge to focus on different virtual depths (vergence) but still accommodate to a fixed screen distance (accommodation), as seen in HMDs. This mismatch can lead to eye strain and discomfort [Kemeny *et al*., 2020].

4.1.4 Locomotion and Acceleration

Locomotion within VR environments is a key factor influencing CS, with research indicating that techniques closely resembling natural movement and allowing greater user control generally reduce discomfort [Rebenitsch, 2015]. In contrast, vection—visual stimuli that create a sensation of movement without actual physical motion—often induces higher levels of discomfort and limits VR usage duration. The mismatch between visual perception and vestibular cues leads to sensory conflicts, increasing the likelihood of CS symptoms such as nausea and disorientation.

Foveated rendering can mitigate CS caused by locomotion by dynamically adapting peripheral visual complexity. By reducing the resolution or applying adaptive blurring in peripheral vision while maintaining sharpness in the central gaze region, foveated rendering stabilizes the perceived environment. This approach decreases the impact of vection by ensuring that peripheral movement is perceived as less intense, thereby reducing the sensory conflict responsible for discomfort [Nie *et al*., 2025]. Additionally, adaptive foveated rendering can personalize visual adjustments based on user sensitivity, dynamically modifying rendering intensity to align with individual comfort levels, ultimately extending VR session duration.

Similarly, jerk—defined as the rate of change of acceleration—characterizes how smoothly or abruptly a movement transitions. Higher values of jerk indicate more abrupt changes, while lower values represent smoother motion. Acceleration is a key factor in cybersickness (CS), especially when visual accelerations are not accompanied by matching vestibular feedback; in particular, high-frequency acceleration changes are strongly associated with increased discomfort [Salehi *et al*., 2024].

as illustrated in Figure 7. Importantly, it is not merely the magnitude of acceleration that contributes to CS, but the variability over time; brief, sharp shifts tend to induce fewer symptoms than sustained, fluctuating acceleration and deceleration. Recent work by Salehi *et al*. [2024] supports this view by showing that distinct head movement patterns, especially those involving abrupt changes, can serve as indicators of the onset of CS, offering a promising direction for real-time detection and mitigation strategies.

Foveated rendering strategies help alleviate acceleration-induced sensory conflicts by ensuring smooth visual transitions and limiting abrupt changes in peripheral motion. By reducing peripheral motion sensitivity and controlling the intensity of motion cues outside the central gaze area, foveated rendering prevents excessive visual stimuli from overwhelming the user's sensory processing system. Additionally, predictive gaze-based rendering can anticipate acceleration-related changes in the user's viewpoint and pre-adjust visual elements to minimize latency effects, creating a more stable and comfortable experience [Nie *et al*., 2025]. These adaptations allow for a more natural integration of visual and vestibular cues, enhancing overall VR comfort and usability.

4.2 CS Feature Modeling

Recent research has identified a variety of features that contribute to the onset of CS, along with distinct methodological approaches for analyzing and addressing these issues [Chang *et al*., 2023; Nunes da Silva *et al*., 2024]. For clarity, we categorize these studies according to the type of features used,

Figure 7. In this illustration, a person is using a head-mounted display (HMD) to view immersive content while standing. This creates a conflict because, although the eyes perceive movement through the HMD's screen, the vestibular organs in the inner ear are sending signals to the brain indicating that the person is stationary. This sensory conflict can lead to symptoms of cybersickness.

grouping them into subjective, objective, and physiological categories.

4.2.1 Subjective Features

The study by Dilanchian *et al.* [2021] investigates how age influences the user experience in immersive virtual reality (VR), focusing on aspects such as presence, usability, workload, and CS. These factors are frequently associated with individual profile characteristics, including age and gender. Interestingly, the findings suggest that older adults reported lower levels of CS compared to younger participants, challenging the common assumption that aging increases susceptibility to VR-induced discomfort.

Similarly, Li *et al.* [2024] examined CS experiences in older adults and found no evidence that this group is more prone to symptoms than younger users. This aligns with conclusions from [Drazich *et al.*, 2023], which indicated that the benefits of VR interventions for older populations outweigh the potential risks of CS.

In contrast, earlier findings from Petri *et al.* [2020] suggested that older adults may be more vulnerable to CS, particularly those over 60 years of age or individuals unfamiliar with the content or context of the virtual experience—such as non-karate practitioners in sports-based VR scenarios. These mixed results highlight the importance of considering user familiarity, content type, and interaction style when evaluating CS risk among older adults.

Additionally, in a systematic review, MacArthur *et al.* [2021] emphasize the importance of incorporating gender-sensitive approaches in virtual reality (VR) research, particularly regarding CS susceptibility. By analyzing 71 studies involving both experimental and survey-based methodologies, the authors observed that female participants were more frequently reported to experience heightened levels of CS symptoms compared to their male counterparts. However, they caution against broad generalizations, highlighting the lack of consistency in methodological practices and the need for more nuanced analysis.

Furthermore, other studies have identified additional individual factors contributing to CS, including prior VR experience [Pöhlmann *et al.*, 2024], flicker sensitivity [Almeida *et al.*, 2018], vision impairments [Luu *et al.*, 2021], posture [Kumar *et al.*, 2024], and eye dominance [Sin *et al.*, 2024].

A recent study by Sanaei *et al.* [2024] investigated the interplay among task workload, scene complexity, presence, and CS within a virtual reality (VR) game context. In a between-subjects experiment involving 44 participants, the authors compared user experiences in two conditions: a simple scene (low optic flow and low familiarity) and a complex scene (high optic flow and high familiarity), using the Pendulum Chair VR game for testing purposes. Participants completed a cognitive-motor balancing task while their workload, presence, and CS levels were assessed using the NASA-TLX, immersion questionnaire, and Simulator Sickness Questionnaire (SSQ), respectively. Contrary to expectations, the study found no significant differences in reported CS or workload between the two scenes, and the equivalence tests confirmed statistical similarity across conditions. Interestingly, a moderate negative correlation was observed between workload and CS, suggesting that performing an engaging task may distract participants from discomfort symptoms.

However, in a previous work by Sepich *et al.* [2022], conducted a large-scale study manipulating cognitive demand through a passive condition and two variants of the N-Back task (0-Back and 2-Back), demonstrating that increased cognitive workload significantly exacerbated CS symptoms. Participants in the high-demand condition (2-Back) reported the highest SSQ scores and dropout rates, indicating that excessive cognitive load may intensify sensory conflicts.

In contrast, Sanaei *et al.* [2024] used a cognitive-motor balancing task in visually simple and complex VR scenes and found no significant differences in CS across conditions. Instead, their results suggested a moderate negative correlation between workload and sickness, implying that task engagement might distract users from discomfort. These contrasting outcomes suggest that the effect of workload on CS may be context-dependent—varying with task type, complexity, and visual stimuli, thereby underscoring the importance of considering both cognitive demand and environmental design in mitigating CS in immersive systems. Table 1 summarizes the subjective features identified in this study.

4.2.2 Physiological Features

Several studies have investigated the use of physiological features to detect and predict cybersickness (CS) in virtual reality (VR) environments.

Garcia-Agundez *et al.* [2019] examined the viability of using a two-lead electrocardiogram (ECG) setup to assess CS symptoms. In this experiment, 13 participants were exposed to a 15-minute VR experience while their ECG data were collected alongside responses to the Simulator Sickness Questionnaire (SSQ). The authors observed a notable increase in SSQ scores—averaging 8 points—as well as a reduction in heart rate among individuals reporting higher levels of CS, suggesting that cardiac activity may serve as a physiological marker of discomfort in immersive environments.

In a related study, Islam *et al.* [2020] employed a Unity 3D-based roller coaster simulation to evaluate CS using mul-

Table 1. Summary of subjective features

Feature	Reference
Profile Features	
Age	[Petri *et al.*, 2020; Dilanchian *et al.*, 2021; Drazich *et al.*, 2023; Li *et al.*, 2024]
Gender	[MacArthur *et al.*, 2021; Kelly *et al.*, 2023]
Prior VR experience	[Pöhlmann *et al.*, 2024]
Flicker sensitivity	[Almeida *et al.*, 2018]
Vision impairments	[Luu *et al.*, 2021]
Posture	[Kumar *et al.*, 2024]
Dominant eye	[Sin *et al.*, 2024]
Task Features	
Task workload	[Dilanchian *et al.*, 2021; Sanaei *et al.*, 2024]
Task type	[Sepich *et al.*, 2022; Sanaei *et al.*, 2024]
Complexity	[Sanaei *et al.*, 2024]
Cognitive demand	[Sepich *et al.*, 2022]
Application Features	
Scene complexity	[Sanaei *et al.*, 2024]
Presence	[Sanaei *et al.*, 2024]
Visual stimuli	[Sanaei *et al.*, 2024]
Environmental design	[Sanaei *et al.*, 2024]

tiple physiological features, including heart rate (HR), galvanic skin response (GSR), breathing rate (BR), and heart rate variability (HRV). Participants underwent alternating rest and exposure phases, during which physiological data were recorded. For analysis, the authors applied both a Support Vector Machine (SVM) and a hybrid deep learning model combining a Convolutional Neural Network (CNN) with a Long Short-Term Memory (LSTM) architecture. The CNN-LSTM model achieved higher classification accuracy than the SVM, reinforcing the potential of deep learning approaches for real-time detection of CS based on physiological dynamics.

Expanding the scope to neural responses, Krokos and Varshney [2022] investigated brain activity patterns associated with CS induced by vection—a visual illusion of self-motion. Participants viewed a VR flythrough of a spaceport while their electroencephalogram (EEG) data were collected. The results revealed that increases in Delta, Theta, and Alpha wave activity were positively correlated with self-reported symptoms of CS, captured via a joystick-based assessment interface.

Furthering this line of inquiry, Nunes da Silva *et al.* [2024] proposed a methodology that integrates physiological features from ECG and electrodermal activity (EDA) sensors with motion data from an accelerometer (ACC) to assess CS during VR gaming sessions. Involving 17 participants, their study employed symbolic machine learning to identify correlations between physiological fluctuations and CS symptoms.

Taken together, these studies (summarized in Table 2) illustrate the growing sophistication in the use of physiological features for CS analysis. They trace a trajectory from early reliance on single-signal correlations and subjective reporting toward the deployment of real-time monitoring and machine learning-driven prediction. The integration of physiological data into adaptive VR systems represents a promising avenue for enhancing user comfort and reducing CS through individualized system responses.

4.2.3 Objective Features

The classification of objective characteristics is consistent with that described by Kemeny *et al.* [2020], which identifies a range of hardware and software-driven factors that contribute to CS. From a hardware standpoint, elements like latency, refresh rate, resolution, and tracking accuracy are cited as direct contributors to visual-vestibular mismatch and instability. These factors are quantifiable and directly measurable, making them ideal for integration into predictive models designed to enhance VR systems. Their consideration is essential in the development of adaptive VR systems capable of minimizing discomfort in a personalized and context-aware manner.

Expanding on this,Oh and Son [2022] introduced the CS Reference (CYRE) content dataset, which systematically isolates VR content variables, including camera motion and control, FoV, frame reference type, user control level and exposure time to analyze their relationship with CS. Their findings showed that rapid camera movement, wider FoV, and lack of controllability are significantly associated with increased symptom severity, confirming the influence of software-level objective factors.

Similarly, Ramaseri-Chandra *et al.* [2025] proposed a less intrusive strategy by relying on head-tracking data (HTD) as a source of objective features. Captured from consumer-grade VR headsets, the dataset included six degrees of freedom (6DOF) motion data, time stamps, and frame rates. From this, they computed dynamic kinematic features such as linear/angular velocity, acceleration, and rate of change of acceleration, which were successfully used in machine learning models to predict CS.

Table 3 presents a summary of key objective features reported in the literature. High-level design factors such as locomotion design, camera behavior, and scene complexity

Table 2. Summary of physiological features

Feature	Description	Reference
ECG	Electrocardiogram	[Garcia-Agundez *et al*., 2019]
HR	Heart Rate	[Islam *et al*., 2020]
HRV	Heart Rate Variability	[Islam *et al*., 2020]
GSR	Galvanic skin response	[Islam *et al*., 2020]
BR	Breathing Rate	[Islam *et al*., 2020]
EEG	Electroencephalogram	[Krokos and Varshney, 2022]
EDA	Electrodermal Activity	[Plouzeau *et al*., 2018]
ACC	Body motion acceleration	[Salehi *et al*., 2024]

Table 3. Summary of objective features

Feature	Example of Data	Unit	Reference
Camera control	auto/manual switch	categorical	[Oh and Son, 2022]
Camera motion	fast pan	qualitative	[Oh and Son, 2022]
Camera Speed	2.5 m/s	meters/second	[Oh and Son, 2022]
Field of View (FoV)	110°	degrees	[Oh and Son, 2022]
Frame rate	72 fps	frames/second	[Ramaseri-Chandra *et al*., 2025]
Frame reference type	world/local	categorical	[Oh and Son, 2022]
Rate of change of acceleration	0.5 m/s^3	$meters/second^3$	[Salehi *et al*., 2024]
Latency	15 ms	milliseconds	[Kemeny *et al*., 2020]
Motion Data	(x,y,z)+(qx,qy,qz,qw)	meters/quaternions	[Ramaseri-Chandra *et al*., 2025]
Refresh rate	90 Hz	hertz	[Kemeny *et al*., 2020]
Resolution	2160x1200 px	pixels	[Kemeny *et al*., 2020]
Exposure Time	10 min	minutes	[Oh and Son, 2022]
Tracking accuracy	1 mm drift	millimeters	[Kemeny *et al*., 2020]
User control level	free/fixed	categorical	[Oh and Son, 2022]

[Kemeny *et al*., 2020] are excluded, as they are not directly quantifiable for integration into data-driven analyzes.

Recent advancements in AI have led to significant improvements in visual techniques particularly through the use of machine learning to enhance visual fidelity and reduce system latency, both strong contributors to cybersickness.

One promising approach is the use of deep learning models to enhance the resolution of images in the foveal region while minimizing computational demands in the periphery. For instance, Wang *et al*. [2021] introduced Focas, a convolutional neural network (CNN)-based method tailored to optimize spatial video super-resolution within VR. The model emphasizes high-detail reconstruction in the foveal region, thus reducing the computational load in peripheral areas.

Building on this, Ye *et al*. [2024] proposed a neural foveated super-resolution system that integrates a neural accumulator to iteratively gather low-resolution data and a neural super-resolution module to reconstruct high-resolution visuals. This approach achieves real-time performance and significant computational savings while maintaining perceptual accuracy. These methods not only enhance image clarity in the areas where users are focused but also help reduce motion-to-photon latency, thereby improving perceptual stability and lowering the risk of VR-induced discomfort.

Additionally, accurate gaze prediction plays a crucial role in enabling effective foveated rendering, especially in latency-sensitive virtual reality applications. Illahi *et al*. [2022], proposed a model for real-time gaze prediction using only past gaze coordinates and head movement data. The model adopts a dual-branch architecture based on long-short-term memory networks (LSTM/RNN) to forecast future gaze positions within short horizons of up to 150 milliseconds. In particular, it does not rely on image saliency or scene features, which significantly reduces its computational requirements. Evaluated on the OpenNEEDs dataset [Emery *et al*., 2021], the model achieved up to 22 percent improvement in angular error over baseline predictors and demonstrated sub-two millisecond inference times on commodity graphics hardware, making it highly suitable for real-time rendering pipelines, including cloud-based systems.

In this context, Liu *et al*. [2025] introduced FovealNet, a neural model that incorporates event-driven cropping for anticipatory gaze estimation. By leveraging saliency-aware techniques, FovealNet enhances rendering efficiency through better alignment with the user's visual attention. This approach not only minimizes perceptual inconsistencies but also contributes to lower end-to-end latency and improved visual stability—key factors in reducing CS during immersive experiences.

More recently, [Ding *et al*., 2025] proposed FoANet, a multi-task deep learning framework that simultaneously predicts both gaze direction and head movement. FoANet combines convolutional layers, attention mechanisms, and memory units to support predictive pre-rendering in edge-assisted virtual reality systems. The integration of multiple data modalities allows FoANet to outperform single-task baselines in both accuracy and inference time, enabling proactive allocation of rendering resources to regions of future interest and further improving the user experience.

Collectively, these AI-based gaze prediction methods contribute to enhanced immersion and comfort in virtual reality. By dynamically adapting the field of view and depth of field to user focus, these systems reduce peripheral conflicts, mitigate vergence-accommodation mismatches, and improve spatial awareness—ultimately leading to lower levels of CS and a more seamless user experience.

5 AI-Enhanced Foveated Rendering for CS Mitigation

First of all, significant progress in visual computing has been fueled by recent developments in artificial intelligence (AI), especially the use of machine learning techniques to improve visual fidelity and lower system latency, two important factors that are strongly linked to the onset and intensity of cybersickness. Further, with this technology, it is possible to create adaptive rendering algorithms that optimize efficiency and user comfort in immersive virtual worlds by dynamically allocating computational resources according to perceptual importance.

In this context, foveated rendering has emerged as a promising approach, leveraging AI-driven mechanisms to balance visual quality with computational efficiency. The principal AI-enhanced foveated rendering techniques examined and revised in this work are summarized in Table 4, highlighting their key contributions, underlying AI and machine learning components, and their specific roles in mitigating cybersickness across diverse virtual reality applications.

5.1 Super-Resolution

One promising idea is using deep learning models to improve the resolution of images in the foveal region while minimizing computational demands in the periphery. For instance, Wang *et al*. [2021] introduced Focas, a convolutional neural network (CNN)-based method tailored to optimize spatial video super-resolution within VR. The model emphasizes high-detail reconstruction in the foveal region, thus reducing the computational load in peripheral areas.

Building on this, Ye *et al*. [2024] proposed a neural foveated super-resolution system that integrates a neural accumulator to iteratively gather low-resolution data and a neural super-resolution module to reconstruct high-resolution visuals. This approach achieves real-time performance and significant computational savings while maintaining perceptual accuracy. These techniques improve perceptual stability and lessen the possibility of VR-induced discomfort by increasing image clarity in the areas where users are concentrated and reducing motion-to-photon latency.

5.2 Gaze Prediction

Additionally, accurate gaze prediction plays a crucial role in enabling effective foveated rendering, especially in latency-sensitive virtual reality applications. [Illahi *et al*., 2022], proposed a model for real-time gaze prediction using only past gaze coordinates and head movement data. The model adopts a dual-branch architecture based on long-short-term memory networks (LSTM/RNN) to forecast future gaze positions within short horizons of up to 150 milliseconds. In particular, it does not rely on image saliency or scene features, which significantly reduces its computational requirements. Evaluated on the OpenNEEDs dataset [Emery *et al*., 2021], the model achieved up to 22 percent improvement in angular error over baseline predictors and demonstrated sub-two millisecond inference times on commodity graphics hardware, making it highly suitable for real-time rendering pipelines, including cloud-based systems.

In this context, Liu *et al*. [2025] introduced FovealNet, a neural model that incorporates event-driven cropping for anticipatory gaze estimation. By leveraging saliency-aware techniques, FovealNet enhances rendering efficiency through better alignment with the user's visual attention. This approach minimizes perceptual inconsistencies and contributes to lower end-to-end latency and improved visual stability—key factors in reducing CS during immersive experiences.

More recently, Ding *et al*. [2025] proposed FoANet, a multi-task deep learning framework that simultaneously predicts gaze direction and head movement. FoANet combines convolutional layers, attention mechanisms, and memory units to support predictive pre-rendering in edge-assisted virtual reality systems. Integrating multiple data modalities allows FoANet to outperform single-task baselines in both accuracy and inference time, enabling proactive allocation of rendering resources to regions of future interest and further improving the user experience.

5.3 Neural Rendering

Building upon advances in gaze prediction and anticipatory rendering, neural rendering offers a complementary mechanism for optimizing VR environments. By integrating neural networks into the rendering pipeline, this approach learns implicit representations of complex scenes, synthesizing photorealistic images from novel viewpoints without relying on explicit geometric models [Garbin *et al*., 2021]. Crucially, neural rendering underpins adaptive scene representation, where rendering precision and resource allocation dynamically adjust to the user's predicted gaze and movement, as facilitated by models like FoANet [Ding *et al*., 2025] and FovealNet [Liu *et al*., 2025].

In the neural rendering context, FoVolNet Bauer *et al*. [2022] illustrates the application of neural rendering for adaptive scene representation in virtual reality by introducing a hybrid two-stage neural network architecture that integrates direct and kernel prediction strategies. As reported by the authors, the method presents certain limitations, particularly in regions characterized by high-frequency intensity variations, which may result in temporal flickering under rapid motion conditions. Nevertheless, FoVolNet achieves up to a 25-fold increase in rendering speed relative to conventional dense ray marching approaches. The rendering efficiency gain is crucial for VR's high frame rates, which results in a reduced latency, which is essential to avoiding CS.

Table 4. Summary of AI-Enhanced Techniques

Technique	Key Contribution	AI/ML Component	Impact on CS Mitigation
		Super-Resolution (SR)	
Focas [Wang *et al.*, 2021]	CNN for foveal super-resolution, reducing peripheral load.	CNN for spatial super-resolution.	Lowers latency and improves stability, reducing VR discomfort.
Neural Foveated SR [Ye *et al.*, 2024]	Neural accumulator + SR for real-time high-res reconstruction.	CNN-based temporal feature accumulation, adaptive mask filtering, and a multi-exit super-resolution network	Real-time efficiency; better perceptual accuracy and reduced discomfort.
		Gaze Prediction (GP)	
Gaze Prediction Model [Illahi *et al.*, 2022]	Predicts gaze via past gaze and head data, no scene features.	Dual-branch LSTM/RNN.	Reduces latency and enhances real-time rendering, lowering CS risk.
FovealNet [Liu *et al.*, 2025]	Saliency-aware event-driven cropping for gaze estimation.	Transformer-based gaze estimation framework.	Cuts latency, improves stability, reducing CS.
FoANet [Ding *et al.*, 2025]	Multi-task model predicting gaze and head movement.	A multi-task CNN-LSTM model with attention-based feature fusion.	Proactive rendering; enhances user experience and reduces CS.
		Neural Rendering (NR)	
FoVolNet [Bauer *et al.*, 2022]	Hybrid network for adaptive VR scene representation.	Two-stage Multi-Layer Perceptron (MLP) neural network.	Up to 25× speedup; supports high frame rates, reducing CS.
		3D Gaussian Splatting (3DGS)	
FPC [Henriques *et al.*, 2024]	Hybrid path tracing + radiance fields for foveated 3DGS.	Neural approximations + Gaussian primitives.	43% speedup; lowers latency and CS risk.
Fov-GS [Fan *et al.*, 2025]	Gaussian forest for foveated 3DGS with multi-level detail.	Hierarchical Gaussian forest guided by vision models.	11× speedup; minimizes frame drops and CS.

5.4 3D Gaussian Splatting

Recent developments in real-time rendering have emphasized perceptually adaptive techniques to meet the computational demands of immersive virtual reality (VR). Among these, 3D Gaussian Splatting (3DGS) offers an efficient solution for photorealistic scene representation [Henriques *et al.*, 2024; Fan *et al.*, 2025]. However, integrating foveated rendering into 3DGS remains challenging due to the mismatch between screen-space foveation and volumetric Gaussian representations. Notable efforts addressing this issue in VR include Foveated Path Culling (FPC) and Fov-GS, each proposing distinct strategies for combining foveated rendering with 3DGS frameworks.

Henriques *et al.* [2024], who proposed Foveated Path Culling (FPC), a hybrid rendering framework that integrates real-time path tracing for the foveal region with precomputed radiance field representations for the periphery. Their method employs recent advances in 3D Gaussian Splatting (3DGS) to render peripheral regions efficiently, significantly reducing computational overhead while maintaining perceptual fidelity. The core innovation in FPC is its culling mechanism: path tracing computations are restricted to the user's gaze-centric region, while peripheral areas rely on less computationally intensive representations, either inferred through neural networks or approximated with splatted Gaussian primitives. This approach yielded up to 43% speedup over conventional full-path tracing in high-resolution VR rendering scenarios.

In contrast, Fan *et al.* [2025] introduced Fov-GS, a comprehensive framework that directly adapts foveated rendering to the 3DGS pipeline for dynamic scene rendering. Unlike FPC, which focuses on hybridizing path tracing and radiance fields, Fov-GS re-engineers the 3D Gaussian representation itself by introducing a hierarchical "Gaussian forest" structure that explicitly supports foveated rendering through multi-level detail representations aligned with human visual system models. The Gaussian forest organizes scene elements into dynamic and static trees, enabling selective deformation of only dynamic components and layer-specific rendering guided by both acuity and contrast sensitivity models.

This architecture addresses the incompatibility between traditional 3DGS and foveated rendering due to the spatial mismatch between screen-space foveation and volumetric Gaussian distributions.

Both FPC and Fov-GS contribute to the mitigation of CS by addressing one of its primary technological determinants, namely, system latency and rendering inefficiency. By optimizing rendering performance through perceptually informed resource allocation that concentrates computational resources on the foveal region while reducing detail in the periphery, both methods facilitate the maintenance of higher frame rates and lower latency.

FPC achieves this optimization by selectively applying path tracing within the foveal region, complemented by using neural radiance fields to represent peripheral areas efficiently. In contrast, Fov-GS improves efficiency by structurally embedding foveated rendering into the 3D Gaussian Splatting pipeline by introducing a hierarchical representation known as a Gaussian forest. In a few words, the Gaussian forest comprises multiple binary trees, each corresponding to a distinct scene segment and organized to support multiple levels of detail (LOD). The selection of appropriate levels within these trees is guided by perceptual models of the human visual system, enabling adaptive rendering that aligns computational resources with visual importance. This structural innovation is particularly advantageous in dynamic scenes, where rendering complexity increases due to object motion and deformation. The real-time capabilities demonstrated by Fov-GS [Fan *et al.*, 2025], achieving up to an eleven-fold improvement in performance, substantially reduce computational demands and the likelihood of frame drops. As a result, both approaches exemplify how artificial intelligence-driven foveated rendering techniques advance visual realism and computational efficiency while aligning with human perceptual constraints to reduce cybersickness and enhance user comfort and immersion in extended reality applications.

Overall, while Henriques *et al.* [2024] introduces a foveated rendering technique that effectively integrates path tracing with precomputed radiance field representations, Fan *et al.* [2025] presents a comprehensive redesign of the 3D Gaussian Splatting framework, enabling native support for foveated rendering in dynamic scenes. Within this context, Foveated Path Culling (FPC) is well-suited for virtual reality applications requiring high-fidelity path tracing confined to specific regions of interest. In contrast, Fov-GS offers a scalable and perceptually optimized approach for real-time rendering of complex and dynamic environments in virtual reality settings.

6 Guidelines and Recommendations

Based on this study, we propose the following guidelines and recommendations for advancing AI-driven foveated rendering (FR) systems that can potentially minimize cybersickness (CS) in virtual reality environments.

- **Prioritize latency** - Latency is among the various factors contributing to CS and is among the most significant. Techniques such as Foveated Path Culling [Henriques *et al.*, 2024] and neural super-resolution models [Ye *et al.*, 2024] offer promising solutions by significantly reducing computational load. Limiting high-resolution rendering only to the foveal region and maintaining perceptual fidelity in peripheral vision makes it possible to improve frame rates and decrease the latency, resulting in a more comfortable VR experience for users.
- **Predictive algorithms** - Integrating predictive algorithms through machine learning has demonstrated substantial potential in enhancing user experience. Gaze prediction models, such as the LSTM-based approach proposed by Illahi *et al.* [2022], along with FovealNet [Liu *et al.*, 2025] and FoANet [Ding *et al.*, 2025], illustrate how anticipatory mechanisms can enable proactive rendering. These systems can pre-render regions of interest with high accuracy and low inference time, resulting in smoother visual transitions and reduced sensory conflict.
- **Physiological monitoring** - Despite recent advances, many features associated with CS remain underexplored in the context of AI-driven foveated rendering. Studies incorporating physiological characteristics such as heart rate variability, electrodermal activity, and EEG signals [Garcia-Agundez *et al.*, 2019; Islam *et al.*, 2020; Krokos and Varshney, 2022; Nunes da Silva *et al.*, 2024] reveal strong correlations between bodily responses and CS symptoms. These studies support the development of closed-loop VR systems that monitor users in real-time and dynamically adjust rendering settings, such as image fidelity and field of view, which contribute to minimizing CS.
- **Interpretable AI** - Works [Nunes da Silva *et al.*, 2024; Islam *et al.*, 2022] have shown promise in producing interpretable, user-adaptive responses. This reinforces the potential of explainable AI in immersive systems, enabling transparent and trustworthy adaptations that respond effectively to user states.
- **Personalized adaptation** - Individual and subjective factors, such as age, sex, prior experience with virtual reality (VR), cognitive workload, and familiarity with the virtual environment, substantially influence susceptibility to cybersickness (CS) [Petri *et al.*, 2020; MacArthur *et al.*, 2021; Sanaei *et al.*, 2024; Pöhlmann *et al.*, 2024]. Although the evidence regarding age-related sensitivity remains inconclusive, a growing consensus exists on the necessity of implementing personalized adaptation mechanisms. These mechanisms should account not only for demographic attributes (e.g., age, sex, and prior VR experience) but also for physiological signals (e.g., heart rate, neural activity) and the cognitive demands imposed by VR interactions.

There are also other crucial directions to consider. We envision a time when responsive and intelligent systems will incorporate foveated rendering. Real-time physiological feedback, predictive gaze modeling, and visual adaptation will work together to create highly customized virtual reality experiences that maximize comfort and immersion based on the individual characteristics of each user. Nevertheless, existing strategies are still dispersed, with several projects concentrating solely on particular facets. Because of this, our goal for a future project is to create a complete, cross-modal system that can dynamically adjust in response to real-time physiological and visual data.

7 Conclusion

This study examines the convergence of foveated rendering techniques, artificial intelligence models, and reducing cybersickness (CS) in virtual reality environments. While the primary focus is on software-driven approaches and experimental frameworks related to foveated rendering and CS mitigation, it also highlights the increasing need for personalization in immersive settings. In particular, adapting rendering methods to accommodate individual user characteristics, such as eye movements, physiological reactions, and cognitive load, which emerges as a crucial approach for lessening sensory mismatches that frequently lead to CS in VR. However, current research remains fragmented, with many strategies addressing only certain aspects of the problem. Future efforts should prioritize real-time, multimodal, and explainable AI technologies that boost system efficiency while ensuring user comfort, safety, and inclusivity. We contend that this research helps establish a foundation for scientists striving to develop multimodal and adaptive systems to enhance CS reduction in virtual reality.

Declarations

Authors' Contributions

All authors contributed to the writing of this article, read andapproved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The authors confirm that the data supporting the findings of this study (i.e., the selected papers) are available within the article.

References

Almeida, A., Rebelo, F., Noriega, P., and Vilar, E. (2018). Virtual reality self induced cybersickness: an exploratory study. In *Advances in Ergonomics in Design: Proceedings of the AHFE 2017 International Conference on Ergonomics in Design, July 17- 21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8*, pages 26–33. Springer. DOI: 10.1007/978-3-319-60582-1_3.

Bala, P., Oakley, I., Nisi, V., and Nunes, N. J. (2021). Dynamic field of view restriction in 360 video: Aligning optical flow and visual slam to mitigate vims. In *Proceedings of the 2021 CHI conference on human factors in computing systems*, pages 1–18. DOI: 10.1145/3411764.3445499.

Bauer, D., Wu, Q., and Ma, K.-L. (2022). Fovolnet: Fast volume rendering using foveated deep neural networks. *IEEE transactions on visualization and computer graphics*, 29(1):515–525. DOI: 10.1109/tvcg.2022.3209498.

Biswas, N., Mukherjee, A., and Bhattacharya, S. (2024). "are you feeling sick?"–a systematic literature review of cybersickness in virtual reality. *ACM Computing Surveys*, 56(11):1–38. DOI: 10.1145/3670008.

Brunnström, K., Dima, E., Qureshi, T., Johanson, M., Andersson, M., and Sjöström, M. (2020). Latency impact on quality of experience in a virtual reality simulator for remote control of machines. *Signal Processing: Image Communication*, 89:116005. DOI: 10.1016/j.image.2020.116005.

Chang, E., Billinghurst, M., and Yoo, B. (2023). Brain activity during cybersickness: a scoping review. *Virtual reality*, 27(3):2073–2097. DOI: 10.1007/s10055-023-00795-y.

Davis, J., Hsieh, Y.-H., and Lee, H.-C. (2015). Humans perceive flicker artifacts at 500 hz. *Scientific reports*, 5(1):7861. DOI: 10.1038/srep07861.

Deng, N., He, Z., Ye, J., Duinkharjav, B., Chakravarthula, P., Yang, X., and Sun, Q. (2022). Fov-nerf: Foveated neural radiance fields for virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3854–3864. DOI: 10.1109/tvcg.2022.3203102.

Dilanchian, A. T., Andringa, R., and Boot, W. R. (2021). A pilot study exploring age differences in presence, workload, and cybersickness in the experience of immersive virtual reality environments. *Frontiers in Virtual Reality*, 2:736793. DOI: 10.3389/frvir.2021.736793.

Ding, D., Cao, Z., Gu, Z., Chen, H., Qi, C., and Dong, F. (2025). Foanet: Focus of attention prediction for foveated pre-rendering to enable high-quality edge vr. *ACM Transactions on Sensor Networks*. DOI: 10.1145/3722222.

Drazich, B. F., McPherson, R., Gorman, E. F., Chan, T., Teleb, J., Galik, E., and Resnick, B. (2023). In too deep? a systematic literature review of fully-immersive virtual reality and cybersickness among older adults. *Journal of the American Geriatrics Society*, 71(12):3906–3915. DOI: 10.1111/jgs.18553.

Duchowski, A. T., Cournia, N., and Murphy, H. (2003). Gaze-contingent displays: Review and current trends. Available at:`http://andrewd.ces.clemson.edu/gcd/adc04.pdf`.

Emery, K. J., Zannoli, M., Warren, J., Xiao, L., and Talathi, S. S. (2021). Openneeds: A dataset of gaze, head, hand, and scene signals during exploration in open-ended vr environments. In *ACM Symposium on Eye Tracking Research and Applications*, pages 1–7. DOI: 10.1145/3448018.3457996.

Fan, R., Wu, J., Shi, X., Zhao, L., Ma, Q., and Wang, L. (2025). Fov-gs: Foveated 3d gaussian splatting for dynamic scenes. *IEEE Transactions on Visualization and Computer Graphics*. DOI: 10.1109/tvcg.2025.3549576.

Feldstein, I. T. and Ellis, S. R. (2020). A simple video-based technique for measuring latency in virtual reality or teleoperation. *IEEE Transactions on Visualiza-*

tion and Computer Graphics, 27(9):3611–3625. DOI: 10.1109/tvcg.2020.2980527.

Garbin, S. J., Kowalski, M., Johnson, M., Shotton, J., and Valentin, J. (2021). Fastnerf: High-fidelity neural rendering at 200fps. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 14346–14355. DOI: 10.1109/iccv48922.2021.01408.

Garcia-Agundez, A., Reuter, C., Caserman, P., Konrad, R., and Göbel, S. (2019). Identifying cybersickness through heart rate variability alterations. *International Journal of Virtual Reality*, 19(1):1–10. DOI: 10.20870/ijvr.2019.19.1.2907.

Henriques, H., de Oliveira, A., Oliveira, E., Trevisan, D., and Clua, E. (2024). Foveated path culling: A mixed path tracing and radiance field approach for optimizing rendering in xr displays. *Journal on Interactive Systems*, 15(1):576–590. DOI: 10.5753/jis.2024.4352.

Hussain, R., Chessa, M., and Solari, F. (2021). Mitigating cybersickness in virtual reality systems through foveated depth-of-field blur. *Sensors*, 21(12):4006. DOI: 10.3390/s21124006.

Illahi, G. K., Siekkinen, M., Kämäräinen, T., and Ylä-Jääski, A. (2022). Real-time gaze prediction in virtual reality. In *Proceedings of the 14th international workshop on immersive mixed and virtual environment systems*, pages 12–18. DOI: 10.1145/3534086.3534331.

Islam, M. S., Hussain, I., Rahman, M. M., Park, S. J., and Hossain, M. A. (2022). Explainable artificial intelligence model for stroke prediction using eeg signal. *Sensors*, 22(24):9859. DOI: 10.3390/s22249859.

Islam, R., Lee, Y., Jaloli, M., Muhammad, I., Zhu, D., Rad, P., Huang, Y., and Quarles, J. (2020). Automatic detection and prediction of cybersickness severity using deep neural networks from user's physiological signals. In *2020 IEEE international symposium on mixed and augmented reality (ISMAR)*, pages 400–411. IEEE. DOI: 10.1109/ismar50242.2020.00066.

Jabbireddy, S., Sun, X., Meng, X., and Varshney, A. (2022). Foveated rendering: Motivation, taxonomy, and research directions. *arXiv preprint arXiv:2205.04529*. DOI: 10.48550/arXiv.2205.04529.

Kelly, J. W., Doty, T. A., Gilbert, S. B., and Dorneich, M. C. (2024). Field of view restriction and snap turning as cybersickness mitigation tools. *IEEE Transactions on Visualization and Computer Graphics*. DOI: 10.31234/osf.io/ksw63.

Kelly, J. W., Gilbert, S. B., Dorneich, M. C., and Costabile, K. A. (2023). Gender differences in cybersickness: Clarifying confusion and identifying paths forward. In *2023 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, pages 283–288. IEEE. DOI: 10.31234/osf.io/qrkdx.

Kemeny, A., Chardonnet, J.-R., and Colombet, F. (2020). Getting rid of cybersickness. *Virtual reality, augmented reality, and simulators*, 12:15. DOI: 10.1007/978-3-030-59342-1.

Kerbl, B., Kopanas, G., Leimkühler, T., and Drettakis, G. (2023). 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4):139–1. Available at: `https://sgvr.kaist.ac.kr/~sungeui/ICG_F23/Students/[CS482]%203D%20Gaussian%20Splatting%20for%20Real-Time%20Radiance%20Field%20Rendering.pdf`.

Krokos, E. and Varshney, A. (2022). Quantifying vr cybersickness using eeg. *Virtual Reality*, 26(1):77–89. DOI: 10.1007/s10055-021-00517-2.

Kumar, N., Lim, C. H., Sardar, S. K., Park, S. H., and Lee, S. C. (2024). Effects of posture and locomotion methods on postural stability, cybersickness, and presence in a virtual environment. *International Journal of Human–Computer Interaction*, 40(20):6314–6326. DOI: 10.1080/10447318.2023.2250611.

Kundu, R. K., Rahman, A., and Paul, S. (2021). A study on sensor system latency in vr motion sickness. *Journal of Sensor and Actuator Networks*, 10(3):53. DOI: 10.3390/jsan10030053.

Lei, Z., Ren, S., Hu, Y., Zhang, W., and Chen, S. (2022). Latency-aware collaborative perception. In *European Conference on Computer Vision*, pages 316–332. Springer. DOI: 10.1007/978-3-031-19824-$3_1$9.

Li, X., Ren, X., Suzuki, X., Yamaji, N., Fung, K. W., and Gondo, Y. (2024). Designing a multisensory vr game prototype for older adults-the acceptability and design implications. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–18. DOI: 10.1145/3613904.3642948.

Lisboa, T., Macêdo, H., Porcino, T., Oliveira, E., Trevisan, D., and Clua, E. (2023). Is foveated rendering perception affected by users' motion? In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 1104–1112. IEEE. DOI: 10.1109/ismar59233.2023.00127.

Liu, W., Duinkharjav, B., Sun, Q., and Zhang, S. Q. (2025). Fovealnet: Advancing ai-driven gaze tracking solutions for efficient foveated rendering in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*. DOI: 10.1109/tvcg.2025.3549577.

Luu, W., Zangerl, B., Kalloniatis, M., Palmisano, S., and Kim, J. (2021). Vision impairment provides new insight into self-motion perception. *Investigative Ophthalmology & Visual Science*, 62(2):4–4. DOI: 10.1167/iovs.62.2.4.

MacArthur, C., Grinberg, A., Harley, D., and Hancock, M. (2021). You're making me sick: A systematic review of how virtual reality research considers gender & cybersickness. In *Proceedings of the 2021 CHI conference on human factors in computing systems*, pages 1–15. DOI: 10.1145/3411764.3445701.

Market.us (2024). Ai in virtual reality market size, growth, trends, and forecast 2024–2033. Available at:`https://market.us/report/ai-in-virtual-reality-market/`.

Mazloumi Gavgani, A., Walker, F. R., Hodgson, D. M., and Nalivaiko, E. (2018). A comparative study of cybersickness during exposure to virtual reality and "classic" motion sickness: are they different? *Journal of Applied Physiology*, 125(6):1670–1680. DOI: 10.1152/japplphysiol.00338.2018.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. (2021). Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106. DOI: h10.1007/978-3-030-58452-8$_2$4.

Mohanto, B., Islam, A. T., Gobbetti, E., and Staadt, O. (2022). An integrative view of foveated rendering. *Computers & Graphics*, 102:474–501. DOI: 10.1016/j.cag.2021.10.010.

Nie, T., Pospick, C. H., Cantory, V., Zhang, D., DeGuzman, J. J., Interrante, V., Adhanom, I. B., and Rosenberg, E. S. (2025). Peripheral teleportation: A rest frame design to mitigate cybersickness during virtual locomotion. *IEEE Transactions on Visualization and Computer Graphics*. DOI: 10.1109/tvcg.2025.3549568.

Nunes da Silva, W., Porcino, T. M., Castanho, C. D., and Jacobi, R. P. (2024). Analysis of cybersickness through biosignals: an approach with symbolic machine learning. In *Proceedings of the 26th Symposium on Virtual and Augmented Reality*, pages 11–20. DOI: 10.1145/3691573.3691582.

Oh, H. and Son, W. (2022). Cybersickness and its severity arising from virtual reality content: A comprehensive study. *Sensors*, 22(4):1314. DOI: 10.3390/s22041314.

Patney, A., Salvi, M., Kim, J., Kaplanyan, A., Wyman, C., Benty, N., Luebke, D., and Lefohn, A. (2016). Towards foveated rendering for gaze-tracked virtual reality. *ACM Transactions on Graphics (TOG)*, 35(6):1–12. DOI: 10.1145/2980179.2980246.

Petri, K., Feuerstein, K., Folster, S., Bariszlovich, F., and Witte, K. (2020). Effects of age, gender, familiarity with the content, and exposure time on cybersickness in immersive head-mounted display based virtual reality. *American Journal of Biomedical Sciences*, 12(2). Available at: `https://www.nwpii.com/ajbms/papers/AJ20200203.pdf`.

Plouzeau, J., Chardonnet, J.-R., and Merienne, F. (2018). Using cybersickness indicators to adapt navigation in virtual reality: A pre-study. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 661–662. IEEE. DOI: 10.1109/vr.2018.8446192.

Pöhlmann, K. M., Li, G., Wilson, G., McGill, M., Pollick, F., and Brewster, S. (2024). Is video gaming a cure for cybersickness? gamers experience less cybersickness than non-gamers in a vr self-motion task. *IEEE Transactions on Visualization and Computer Graphics*. DOI: 10.1109/tvcg.2024.3456176.

Porcino, T., Trevisan, D., and Clua, E. (2021). A cybersickness review: causes, strategies, and classification methods. *Journal on Interactive Systems*, 12(1):269–282. DOI: 10.5753/jis.2021.2058.

Ramaseri-Chandra, A. N., Reza, H., and Pothana, P. (2025). Exploring the feasibility of head-tracking data for cybersickness prediction in virtual reality. *Electronics*, 14(3):502. DOI: 10.3390/electronics14030502.

Rebenitsch, L. R. (2015). Cybersickness prioritization and modeling. Michigan State University. DOI: doi:10.25335/nw4q-va87.

Reingold, E. M., Loschky, L. C., McConkie, G. W., and Stampe, D. M. (2003). Gaze-contingent multiresolutional displays: An integrative review. *Human factors*, 45(2):307–328. DOI: 10.1518/hfes.45.2.307.27235.

Salehi, M., Javadpour, N., Beisner, B., Sanaei, M., and Gilbert, S. B. (2024). Cybersickness detection through head movement patterns: A promising approach. In *International Conference on Human-Computer Interaction*, pages 239–254. Springer. DOI: 10.1007/978-3-031-60611-3$_1$8.

Sanaei, M., Gilbert, S. B., Javadpour, N., Sabouni, H., Dorneich, M. C., and Kelly, J. W. (2024). The correlations of scene complexity, workload, presence, and cybersickness in a task-based vr game. In *International Conference on Human-Computer Interaction*, pages 277–289. Springer. DOI: 10.1007/978-3-031-61041-7$_1$8.

Sepich, N. C., Jasper, A., Fieffer, S., Gilbert, S. B., Dorneich, M. C., and Kelly, J. W. (2022). The impact of task workload on cybersickness. *Frontiers in Virtual Reality*, 3:943409. DOI: 10.3389/frvir.2022.943409.

Sin, Z. P., Jia, Y., Li, R. C., Leong, H. V., Li, Q., and Ng, P. H. (2024). illumotion: An optical-illusion-based vr locomotion technique for long-distance 3d movement. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pages 924–934. IEEE. DOI: 10.1109/vr58804.2024.00111.

Spjut, J. and Boudaoud, B. (2019). Foveated displays: toward classification of the emerging field. In *ACM SIGGRAPH 2019 Talks*, pages 1–2, Los Angeles California. ACM. DOI: 10.1145/3306307.3328145.

Stanney, K., Lawson, B. D., Rokers, B., Dennison, M., Fidopiastis, C., Stoffregen, T., Weech, S., and Fulvio, J. M. (2020). Identifying causes of and solutions for cybersickness in immersive technology: reformulation of a research and development agenda. *International Journal of Human–Computer Interaction*, 36(19):1783–1803. DOI: 10.1080/10447318.2020.1828535.

Valmorisco, S., Raya, L., and Sanchez, A. (2024). Enabling personalized vr experiences: a framework for real-time adaptation and recommendations in vr environments. *Virtual Reality*, 28(3):128. DOI: 10.1007/s10055-024-01020-0.

Vlahovic, S., Suznjevic, M., and Skorin-Kapov, L. (2019). The impact of network latency on gaming qoe for an fps vr game. In *2019 Eleventh international conference on quality of multimedia experience (QoMEX)*, pages 1–3. IEEE. DOI: 10.1109/qomex.2019.8743193.

Wang, J., Shi, R., Zheng, W., Xie, W., Kao, D., and Liang, H.-N. (2023a). Effect of frame rate on user experience, performance, and simulator sickness in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2478–2488. DOI: 10.1109/tvcg.2023.3247057.

Wang, L., Hajiesmaili, M., and Sitaraman, R. K. (2021). Focas: Practical video super resolution using foveated rendering. In *Proceedings of the 29th ACM International Conference on Multimedia*, pages 5454–5462. DOI: 10.1145/3474085.3475673.

Wang, L., Shi, X., and Liu, Y. (2023b). Foveated rendering: A state-of-the-art survey. *Computational Visual Media*, 9(2):195–228. DOI: 10.1007/s41095-022-0306-4.

Wu, F. and Suma Rosenberg, E. (2022). Adaptive field-of-view restriction: Limiting optical flow to mitigate cybersickness in virtual reality. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*, pages 1–11. DOI: 10.1145/3562939.3565611.

Xiao, R. and Benko, H. (2016). Augmenting the field-of-view of head-mounted displays with sparse peripheral displays. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 1221–1232. DOI: 10.1145/2858036.2858212.

Ye, J., Meng, X., Guo, D., Shang, C., Mao, H., and Yang, X. (2024). Neural foveated super-resolution for real-time vr rendering. *Computer Animation and Virtual Worlds*, 35(4):e2287. DOI: 10.22541/au.169955755.52710616/v1.