

# Cybersickness Reduction via Gaze-Contingent Image Deformation

COLIN GROTH, MARCUS MAGNOR, STEVE GROGORICK, and MARTIN EISEMANN, Institute for Computer Graphics, TU Braunschweig, Germany  
PIOTR DIDYK, Università della Svizzera italiana, Switzerland

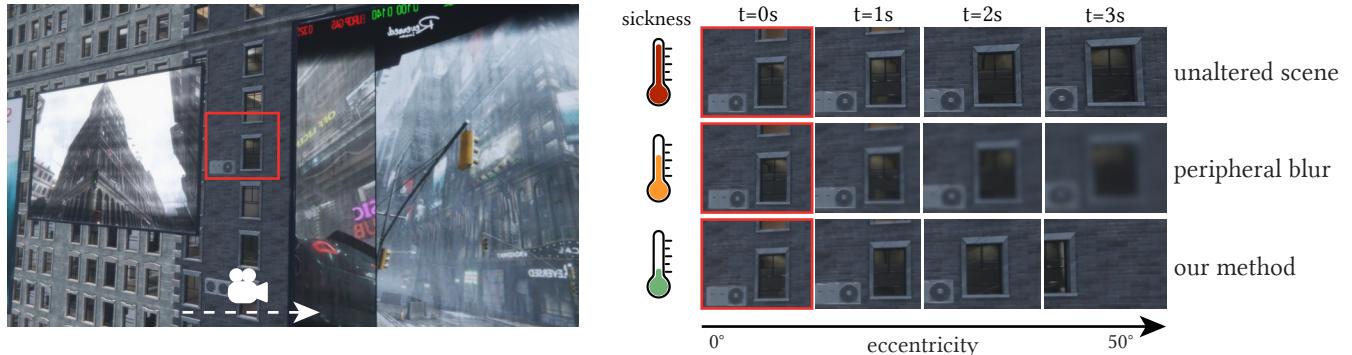


Fig. 1. In this paper, we present a new approach to mitigate cybersickness by applying subtle geometrical distortions to the animation frames. While the modifications significantly reduce cybersickness, the details in the periphery can be preserved better than with the commonly used peripheral blurring.

Virtual reality has ushered in a revolutionary era of immersive content perception. However, a persistent challenge in dynamic environments is the occurrence of cybersickness arising from a conflict between visual and vestibular cues. Prior techniques have demonstrated that limiting illusory self-motion, so-calledvection, by blurring the peripheral part of images, introducing tunnel vision, or altering the camera path can effectively reduce the problem. Unfortunately, these methods often alter the user's experience with visible changes to the content. In this paper, we propose a new technique for reducingvection and combating cybersickness by subtly lowering the screen-space speed of objects in the user's peripheral vision. The method is motivated by our hypothesis that small modifications to the objects' velocity in the periphery and geometrical distortions in the peripheral vision can remain unnoticed yet lead to reducedvection. This paper describes the experiments supporting this hypothesis and derives its limits. Furthermore, we present a method that exploits these findings by introducing subtle, screen-space geometrical distortions to animation frames to counteract the motion contributing tovection. We implement the method as a real-time post-processing step that can be integrated into existing rendering frameworks. The final validation of the technique and comparison to an alternative approach confirms its effectiveness in reducing cybersickness.

CCS Concepts: • Computing methodologies → Virtual reality; Perception.

Additional Key Words and Phrases: Virtual Reality, Cybersickness, VR, Vection, Foveation, Image Distortion

Authors' addresses: Colin Groth, groth@cg.cs.tu-bs.de; Marcus Magnor, magnor@cg.cs.tu-bs.de; Steve Grogorick, grogorick@cg.cs.tu-bs.de; Martin Eisemann, eisemann@cg.cs.tu-bs.de, Institute for Computer Graphics, TU Braunschweig, Germany; Piotr Didyk, piotr.didyk@usi.ch, Università della Svizzera italiana, Switzerland.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2024 Copyright held by the owner/author(s).

0730-0301/2024/7-ART66

<https://doi.org/10.1145/3658138>

## ACM Reference Format:

Colin Groth, Marcus Magnor, Steve Grogorick, Martin Eisemann, and Piotr Didyk. 2024. Cybersickness Reduction via Gaze-Contingent Image Deformation. *ACM Trans. Graph.* 43, 4, Article 66 (July 2024), 14 pages. <https://doi.org/10.1145/3658138>

## 1 INTRODUCTION

Virtual reality (VR) rapidly evolves and captivates users with its highly immersing experiences in virtual worlds. Its potential is vast and varied, from the enhancement of entertainment and education to innovative applications in design and therapy. However, a significant barrier to its universal adoption lies in the onset of cybersickness, a negative symptom similar to general motion sickness. Cybersickness arises from a sensory mismatch between the vestibular system's signals and the visually perceived motion of the virtual experience [Reason and Brand 1975]. A crucial factor for cybersickness isvection, i.e., the sensation of self-motion induced by visual stimuli even when the body is physically motionless [D'Amour et al. 2021]. People often experience such an illusion when sitting on a stationary train and watching another train moving. The visual movement observed through the window leads to a feeling of self-motion. In such cases, human perception attributes a higher significance to the information-dense visual stimuli conflicting with the vestibular signals [Bankieris et al. 2017; ter Horst et al. 2015], often resulting in a negative sensation of sickness or discomfort.

In VR, virtual camera movements that are not accompanied by physical motion are a significant source ofvection, leading to cybersickness. One of the solutions to reduce cybersickness is to limit the visualvection cues. Common approaches modify the visual content by occluding or blurring the peripheral vision or modifying the camera path [Adhanom et al. 2020; Groth et al. 2021b; Hu et al. 2019]. These approaches aim to reduce visible movements, especially in the periphery, which is considered a more motion-sensitive part of the

visual field [Exner 1886; Finlay 1982]. The methods, however, lead to changes in the experience or loss of visual features in the periphery. Other techniques focus on directly controlling the vestibular system. Their active stimulation can reduce the sensory mismatch through alignment with the visual motion [Groth et al. 2022; Sra et al. 2019]. However, such an enhancement requires specialized equipment that is not yet universally accessible.

This paper proposes a novel technique to mitigate cybersickness that is both subtle and effective. We start by studying linear and rotational camera motion to later address arbitrary camera paths. For linear motion we aim to explicitly reduce the magnitude of perceived motion by slowing down objects in the periphery. To maintain the user's perception of the camera velocity, we demonstrate how the reduction of velocity magnitude in the periphery can be compensated in the foveal region without reintroducing cybersickness. The cybersickness caused by visual rotations is even more severe compared to linear camera movements [Groth et al. 2022; Kim et al. 2021]. To address this type of movement, we propose to create an illusion of objects moving along a linear trajectory instead of rotating around a particular axis. We show that simple geometric distortions that modify the screen-space size of the objects are a powerful tool to achieve this goal. By a series of perceptual experiments, we study the effectiveness and visibility of the proposed manipulations (Section 3). Based on the results, we present a perceptual model that describes the manipulations that lead to the maximum reduction ofvection that can be performed without objectionable changes to the content (Section 4). The model provides scene and eccentricity-dependent parameters for the final method. While our initial experiments make use of complete control over 3D scenes, such control is not always feasible in complex scenes. Therefore, we propose a method for applying our manipulations using a simple image-based warping method, which modifies the geometrical information of the scene by introducing subtle image deformations (Section 5). Since the deformations accumulate over time, we exploit saccadic suppression and eye blinks to restore the frame content to the original rendering. The efficiency of the method enables real-time execution and easy integration to any rendering engine. We validate our method in an experiment comparing with subtle blurring of peripheral content (Section 6). While the comparison was not perfectly analogous, the results demonstrate the effectiveness of our approach in reducing cybersickness while the visual fidelity of the scene is preserved, providing a promising direction for future research.

## 2 RELATED WORK

In this section, we describe howvection leads to cybersickness and explore former research on mitigation of cybersickness in virtual reality. Techniques that reduce cybersickness can generally be separated into two categories: visual content manipulations and vestibular perception enhancement. In the following, we will focus on visual methods that do not require further equipment.

**Vection and its Role for Cybersickness.** Unlike traditional displays such as monitors or TVs, VR displays deliver stereoscopic images that isolate the user from the real world, deepening their immersion in the virtual environment. When determining what is real, the human brain tends to prioritize visual cues [Bankieris et al. 2017;

Murovec et al. 2021]. VR delivers strong visual signals that shift the users' feeling of presence to the virtual world. However, the disconnection between visual movements in the virtual environment and actual body motion is a key contributor to cybersickness, a form of motion sickness during VR exposure [Reason and Brand 1975]. When VR users navigate using controllers and continuous movements, they experience motion in the virtual world that does not coincide with their real-world physical state. This discrepancy in perception is evoked byvection – the sensation of self-motion created by visual cues [D'Amour et al. 2021]. In real life, a momentary illusion of self-motion can be evoked, for instance, while sitting in a stationary car and observing another car moving, exemplifying the dominant of our perception in contradictory scenarios. Vection primarily stems from the peripheral vision which is highly sensitive to temporal changes [Guo et al. 2021; Thompson et al. 2007]. Reducing motion cues in the peripheral vision can effectively diminish cybersickness, due to the reduction ofvection [Luu et al. 2021]. However, even when suppressing the subconscious effect ofvection, humans are still able to experience the movements of the virtual camera, due to the conscious perception of the apparent motion in the fovea. This distinction between apparent motion andvection is a key factor in understanding and properly addressing cybersickness.

**Visual Techniques for Cybersickness Reduction.** Various techniques have been explored that modify the visual stimuli for cybersickness mitigation. Opaque occlusions in the visual field, either in the center region [Bos et al. 2010; Lin et al. 2002; Seay et al. 2001] or based on eye tracking [Adhanom et al. 2020; Groth et al. 2021a,b], are notable for reducingvection and cybersickness. A more common and less intrusive technique includes blurring of the periphery [Groth et al. 2021b; Hillaire et al. 2008; Patney et al. 2016], fixed outer regions [Lin et al. 2020a] or object-dependent areas [Nie et al. 2017, 2019] using gaussian filters. Also learning approaches were proposed for motion reduction [Kaplanyan et al. 2019]. These approaches, acting as a low-pass filter, reduce contrast and information perception, thereby diminishing cybersickness. However, the information reduction in the periphery can be an issue in applications that require fast reactions like VR shooter games. With our method, we preserve the visual details of the scene over the entire field of view (FOV) by subtly reducingvection with content-aware distortions. Recent methods include integrating reverse optical flow visualizations [Kim and Kim 2022; Park et al. 2022]. The idea is based on the larger pooling of the ganglion cells in the peripheral vision [Anderson et al. 1991]. Park et al.'s experiment with reverse optical flow arrows reduced cybersickness but significantly affected participants' experience due to its application in both peripheral and foveal regions [2022]. We calibrate our modulations to stay under the threshold of detectability to keep the virtual experience immersive and enjoyable. The relationship between geometry appearance, motion perception, and cybersickness remains underexplored. However, temporal geometrical modifications can change the basic visual motion perception of objects and are highly interesting for reducingvection. In a simple prototype of Lou et al., the geometry of a building is squeezed towards the viewport's edge during forward movements [2022]. While their project was not generally applied or experimentally validated the authors' approach motivated the use of geometrical

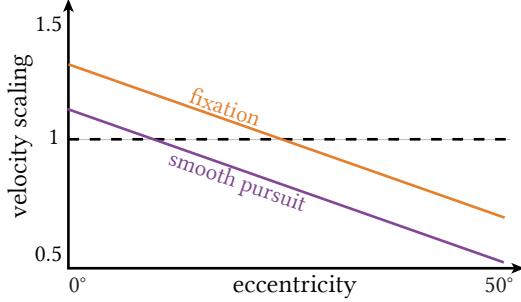


Fig. 2. Visualization of the velocity scaling of an object based on its eccentricity and the type of eye movement.

deformations forvection reduction. Aside from these techniques, the overall quality of the virtual experience as well as the content design play a crucial role. In general, high frame-rate renderings, high quality tracking and reduced latency systems reduce the occurrence of cybersickness symptoms [DiZio and Lackner 1997; LaViola Jr 2000; Sherman 2002].

### 3 PERCEPTION OF FUNDAMENTAL MOVEMENTS

Vection is influenced by visual movements in the FOV. The peripheral area has a much stronger influence on this perception of self-motion than the fovea [Exner 1886; Finlay 1982]. At the same time, modifications in the peripheral region are generally tolerated more by users [Patney et al. 2016]. The following investigations are based on the hypothesis that manipulations can be made to a virtual scene that go unnoticed but reducevection. As a consequence of this hypothesis, we further want to show thatvection – a subconscious effect of perceived self-motion – is decoupled from apparent motion – the conscious estimation of visual movement. This section, describes the psychophysical experiments that we conducted to confirm this hypothesis and derive the detection thresholds that allow for maximizing the reduction ofvection. Our investigations are twofold and examine linear movements and rotational movements in isolation. The results of these psychophysical experiments provide the foundation for the perceptual model derived in the next section. All experiments are conducted in VR with a HTC Vive Pro Eye headset.

#### 3.1 Vection Compensation for Linear Movements

Our first investigations are focused on the reduction ofvection for linear camera movements. In two psychophysical experiments with a simple scene, we directly manipulate the velocity of objects in the VR environment. All modifications are designed to preserve the visually perceived ego-motion in the virtual world. We make two assumptions for our investigations: (1) the human brain derives the apparent motion of a virtual scene as an average of all background motion in the FOV. (2) the velocity of each background object in the visual field can be scaled and the scaling factor can be described by a linear function of eccentricity.

The assumptions are motivated by basic properties of the human visual system (HVS). In the HVS, the effect of information integration of local patterns allows to reach a global consensus even with



Fig. 3. Visual scene of the experiment with linear camera movement. The objects in the left image are scaled to move slower in the periphery. The red dot visualizes the view point. The right image shows the unaltered output.

ambiguities information of local motion (assumption 1). This phenomenon of natural viewing is necessary for humans to characterize ambiguous motion information in local regions that would be raised due to restrictions similar to the aperture problem [Thompson et al. 2011]. The increase in effect strength is motivated with the pooling of peripheral photoreceptor information by the bipolar cells and cortical magnification (assumption 2) [Thompson et al. 2011].

Figure 2 visualizes how the object motion scaling is described by a linear function of eccentricity. The further away an object from the point of view, the more its velocity is reduced. In contrast, objects in the fovea are accelerated. The function has two unknown properties, the slope and the offset, which are calibrated in two separate experiments.

In the first experiment we search for the slope of the linear function that describes the velocity scaling of objects at the threshold of detectability. This threshold describes the intensity of decreased peripheral speed given a constant foveal velocity. However, since we simplified the scene motion to be a weighted average of the objects at different eccentricities, the overall perceived speed of the camera can be altered by the scaling. Such a perceived change in speed would reshape our understanding of the scene with potential implications on the user experience and performance. Therefore, we conduct a second experiment that calibrates the offset of the linear function for the calibrated function slopes (cf. Figure 2).

#### Experiment 1: Slope Calibration

The maximum difference in velocity between the fovea and the outermost point in the periphery is defined by the magnitude of the slope of the scaling function. Here, we use the description of object velocity instead of camera velocity based on the formally made assumption that the linear camera motion derives by averaging the velocity of all background objects in the visual field.

*Experimental Design:* The experiment is conducted with a simple scene that displays an infinite straight wall with multiple lines of windows (see Figure 3). The camera faces the wall while moving to the right. The horizontal spacing between the windows is randomly varied to avoid that the results are influenced by regular patterns. A random color is assigned to each window to increase contrast and recognizability. A small red ball in the forward direction defines the gaze point and participants are asked to always look at this ball. In the experiment the gaze of participants was tracked and the scene was blurred out when the participants' gaze deviates more than 5° from the ball's direction. The experiment is designed as an Up/Down task with the slope of the linear curve being manipulated.

Table 1. Perceptual thresholds of experiment 1 for velocity scaling intensity at two tested camera velocities.

Task	Velocity 1.0		Velocity 1.4	
	Mean	± StDev	Mean	± StDev
Up	0.3344	± 0.1685	0.28	± 0.1582
Down	0.4035	± 0.2657	0.3631	± 0.2721

The linear curve defines how the speed of a window is adjusted based on eccentricity. Thereby, the objects are increased in speed in the foveal area at the same amount as they are slowed down in the periphery. For clarification, a slope value of 0.3 describes a scene where the windows are 30% increased in speed at the focal point and 30% slowed down in speed at the outermost point in the periphery. For the up trials the slope start value is 0 which gets continuously increased by 0.05 per second. The down trials start with a 0.95 slope value which is decreased by 0.05 per second. For every trial participants are asked to press a button when they *notice a difference in speed or distance of the windows between the inner and outer area of the FOV*. With the button press the result is logged and the next trial begins. Each trial starts with a random idle period between 2 and 5 seconds to avoid that participants recognize regular patterns. We investigated the slope for two different speeds, 1 m/s and 1.4 m/s, which are counterbalanced. The higher speed, thereby, corresponds to half of the maximum speed for smooth movement perception of the eye [Daly 2001], allowing us to stay within this limit even with a maximum scaling of 200%. In our within-subject experiment we had 5 repetitions for each, the Up and Down tasks, and both speeds, yielding a total of 20 trials per participant. A total of 11 participants took part in the experiment (2 females, age =  $25.5 \pm 4.31$ ).

**Results:** Table 1 shows the results of the experiment. In general, the results of the psychophysical experiment support the initial assumptions. Even in the worst case scenario, the speed of objects is reduced by 28% at an eccentricity of 50°. Stronger modifications go unnoticed for slower camera movements. To ensure subtlety, we interpret the results in a conservative manner and apply the lower values of the Up task in the following.

## Experiment 2: Speed Adjustment

The offset property of the scaling function derives the default scaling factor at zero eccentricity. Therefore, this parameter weights the influence of objects in the inner and outer half of the FOV on the overall perceived scene motion. Consider the two curves in Figure 2 that describe the same perceptual camera velocity: since people rely more on the foveal content during smooth pursuit eye movements, the objects' speed in the periphery can be reduced more.

**Experimental Design:** The scene of the second experiment remains unchanged, with the camera moving along a wall with multiple lines of windows. While we consider a constant slope based on the results of Experiment 1, the overall speed of the camera, i.e. the function offset, can be adjusted by the participants. We show two versions of the scene to the participants, which can be switched as often as

Table 2. Summary of the results for the second experiment assessing participants' perception of movement speeds during fixations and smooth pursuit eye movements. The table presents the matched speeds for two different camera velocities (1 m/s and 1.4 m/s) across two conditions: with slope (as per the findings of Experiment 1) and control (uniformly moving windows). Note: when the values get lower, the perceived scene speed is higher than the actual camera velocity and, consequently, scene objects can be decelerated more.

Scenario	Condition	Speed: Mean ± Std
Fixation	With Slope	1.0: $0.9906 \pm 0.0905$ 1.4: $1.3750 \pm 0.1132$
	Control	1.0: $0.9922 \pm 0.0663$ 1.4: $1.4078 \pm 0.1133$
	With Slope	1.0: $0.8047 \pm 0.1018$ 1.4: $1.2141 \pm 0.1501$
	Control	1.0: $1.0266 \pm 0.0848$ 1.4: $1.4219 \pm 0.1075$

required. The first version provides the unaltered baseline with slope 0 as a reference. The second version is constructed similarly, but participants are able to change the speed of the camera. We ask participants to match the speed of the adjustable scene to the baseline. They are allowed to take as much time and scene switches as needed. When the scenes are toggled, a gray screen of 1.3 seconds is displayed to avoid direct comparison from the change and rather rely on the participants' memory. Similar to the first experiment we display a focus point and occlude the scene when participants deviate from it. However, additionally to the fixed focus point we also investigate the perception for moving gaze. In this moving gaze scenario the focus point is moved at the speed of the windows and displaced to the right when it gets close to the edge of the screen. With this scenario, we induce smooth pursuit eye movements, while the static focal point investigates eye fixations. These two types of fundamental eye movements are highly relevant for real-world scenarios, but can significantly diverge in the perception of movement speeds. As in the former experiment, we investigate both scenarios for the baseline camera velocities of 1 m/s (slope: 0.334) and 1.4 m/s (slope: 0.28). To verify the validity of the results, we further introduce a control condition where the adjustable version of the scene is without manipulation and shows uniformly moving windows just like the baseline. The purpose of this control condition is to assess participants' general ability to match the speed of objects during fixations and smooth pursuit to a baseline. We conducted the within-subject experiment with three repetitions for both velocities, conditions and scenarios. The order of the trials within each scenario was randomly distributed and participants did not know about the different conditions. A total of 17 participants took part in the experiment (8 females, Age =  $24.2 \pm 3.68$ ).

**Results:** Table 2 illustrates the perceived velocities for a given baseline. The results for the two scenarios differ considerably. For fixations our assumption that the influence of the inner and the outer part of the vision contribute evenly to the perceived camera velocity holds true, letting the measured values match the baseline

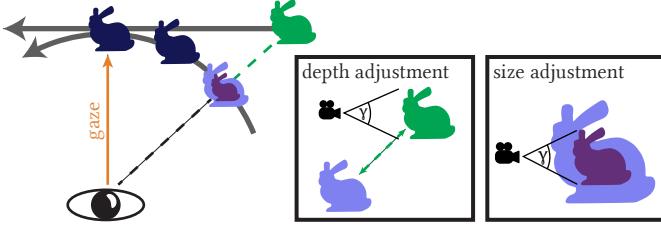


Fig. 4. Visualization of the depth adjustment and scaling method forvection reduction of rotating objects. Blue bunnies represent the unmanipulated state, while green and violet are the adjusted state.

almost perfectly. For smooth pursuit eye movements the influence of the foveal region for the overall perceived camera velocity is significantly higher. The decelerated periphery contributes around three times less to the scene speed perception highlighting a significant shift compared with the fixation scenario (see Figure 2). The results of the control condition confirm the validity of participants' speed assessment ability. Accordingly, the deviating results of different eye movement scenarios are not based on a change in accuracy of speed estimations but rather on the shift in the importance of different image areas for motion perception during active eye movement. In conclusion, for both types of eye movement, thevection in the motion-sensitive periphery can be greatly reduced without changing the apparent motion of the virtual camera, which we later exploit to effectively mitigate cybersickness.

### 3.2 Compensation of Visual Rotation

Next, we take a closer look at rotational camera movements inside virtual environments. In a psychophysical experiment with a simple scene, we transform objects in the periphery to create the illusion of the objects moving along a linear trajectory. The manipulations are designed to preserve the conscious scene understanding while reducing cybersickness. The study is motivated by a key observation of former research that rotational movements in virtual environments induce considerably more cybersickness than linear camera movements [Groth et al. 2021b, 2022; Kim et al. 2021]. We make an assumption based on preliminary findings: the subconscious interpretation of rotational scene movements can be tricked to perceive linear motion, while the active scene understanding still results with the actual angular movement. The assumption, therefore, suggests a decoupling of the illusory self-movements (vection) and the apparent motion of the scene. If this assumption holds true, the applied scene adjustments will change the perceivedvection and have a positive effect on cybersickness even when they are below the threshold of detectability.

#### Experiment 3: Rotation Reduction

Humans derive visual rotations of a static scene by the change of object sizes over time as well as from depth clues e.g. disparity [Leigh and Zee 2015]. This experiment investigates how much both factors contribute to the visual perception of angular movement.

*Scaling methods:* Figure 4 illustrates how the scene objects are transformed in our investigations to create the illusion of movement

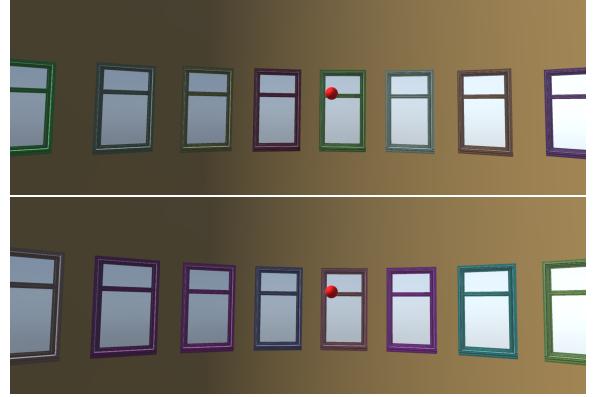


Fig. 5. Visual scene of the experiment with rotating camera. In the upper image the objects are scaled in size to create the visual impression of a linear movement trajectory. The lower image is the unaltered output.

on a linear path. For this visual linear motion, we scale the size of the 3D objects by a function of eccentricity. While the position of the modified objects in the 3D space remains on the circular path, their scaled size in relation to the other objects provide the visual cues of a position further away (see Figure 2). The distance  $d$  of the object on the virtual linear path, orthogonal to the cameras' forward vector  $f$  can be derived by  $d = \frac{r}{\cos \phi}$  with  $r$  representing the radius and  $\phi$  the angle between the forward vector and the vector from the camera to the object. Consequently, the adjusted size  $s = \frac{w}{d}$  for an object is calculated using distance  $d$  and the objects width  $w$ . However, visual size is not the only cue to provide information about the 3D position. In our experiment, we introduce another condition in which the actual position of the scene objects is modified (depth condition). Therefore, in this condition, both the object's size on the screen and its disparity are altered. For the depth condition, the depth of the objects is modulated by displacing the position of the objects in accordance with  $d$  in the direction of the vector from the camera to the object. The depth condition is meant as a baseline and provides further insights about the importance of different cues for the perception of visual movements.

*Experimental design:* The scene in the experiment displays a circular wall with lines of windows and the camera rotating in the center (see Figure 5). The angular velocity of the camera is modulated by a sine function. We investigate two different approaches, firstly the scaling of the objects in size and secondly the adjustment of the objects' positions in depth. In the experiment we want to find the intensity of the methods which describe the threshold of detectability. Intensity reflects the amount to which the methods manipulate the perceptual trajectory of the objects in the visual field at maximum angular velocity. At an intensity of 1 the methods manipulate the perceived object trajectory into a straight line and for 0 they are on the circular path. We use a 1 up/1 down procedure as psychophysical estimation method to find the conservative detection threshold (CDT) [Zenner et al. 2021]. In this experimental method the intensity of the manipulation technique is increased (+0.1) when participants do not notice any distortions in a trial and decreased

Table 3. Results of the calibration experiment for rotational movement compensation. The values indicate the CDT of the visual manipulation methods applied to the scene for different rotations. A value of 1 would correspond to a perceptually linear movement of the objects (full manipulation), while at 0 the initial circular path is displayed.

Condition	Yaw	Pitch
Scale	$0.624 \pm 0.2052$	$0.674 \pm 0.2314$
Depth	$0.716 \pm 0.2126$	$0.782 \pm 0.184$

(-0.1) in the opposite case. Furthermore, we apply a staircase design with interleaved ascending (start with intensity = 0) and descending (start with intensity = 1) sequences. The sequences are terminated after five reversals and the average of the last four reversals yield the sequence threshold estimate [Zenner et al. 2021]. In the experiment, the participants are asked to indicate by a button press when noticing any changes of the scene compared to the unaltered reference. The reference movement (normal circular path) is given in the first trial of each block which is properly communicated with the participants. We decided to show the reference condition only at the beginning of a block because we are interested in the participants' comparison with conscious expectations rather than a side-by-side comparison that does not compare to the final application. Each sequence of acceleration and deceleration of the camera to 0 is considered as one trial without pauses in-between. Also, each method is tested for yaw and pitch rotations. These 4 blocks (2 methods \* 2 directions) are counterbalanced by a 4x4 Williams design Latin square. In the experiment, participants are asked to keep their head straight. A total of 15 participants took part in the experiment (7 females, Age =  $28.8 \pm 2.66$ ).

**Results:** Table 3 shows the CDTs for both methods and rotation directions. The results of the third experiment are in line with our assumption that subtlevection reduction is feasible. Substantial modulations can be made to the visual trajectory of the scene objects before participants report any difference. Overall, we were able to compensate for 62% of the rotational movement in the horizontal FOV of the VR glasses. Pitch rotations, that are due to hardware limitations displayed with smaller vertical angles can be compensated more. In the control condition, where the 3D objects are displaced in depth, compensations are around 10% higher compared to the modification in size. Overall, the results suggest that the majority of the movement information is derived by the contents' temporal change in size rather than from the disparity which motivates an important part of our method for cybersickness reduction to rely on transformations in geometrical size. As a limitation, the standard deviation is relatively high, which could reflect uncertainty among the participants. On the other hand, the thresholds are conservative and manipulations can be expected to be less detected in practice when users do not pay attention to possible deformations. Also, more complex scenes could mask the deformations to a certain degree.

#### 4 PERCEPTUAL MODEL

In this section, we derive a perceptual model that estimates a scaling factor for each objects' optical flow between frames. The model is

designed to maximize the reduction ofvection while maintaining the observers' perception of the scene. For linear and angular camera movements, we provide two separate components  $F_{lin}$  and  $F_{rot}$  which both are functions of eccentricity. For the definition of  $F_{lin}$ , we assume a linear function that defines the scaling factor of the objects' motion based on eccentricity. To keep the perceptual scene speed intact, we compensate the reduced speed in the periphery with faster movements around the gaze point (cf. Section 3). For angular movements the function  $F_{rot}$  computes a scaling factor that modulates the objects' geometrical size to be smaller in the periphery. In the image domain, the modulations of rotational movements with  $F_{rot}$  give the impression of the scene content moving on a linear trajectory that is anchored at the gaze point. Our model is based on the observations of the psychophysical experiments of Section 3. The supplementary material provides an additional list of all parameters defined here.

**Linear Movement Component.** First, we define the linear movement component  $F_{lin}$  of our perceptual model. The results of the psychophysical experiments for linear movement compensation confirmed our initial assumption that objects can be scaled by a linear function of eccentricity (see Section 3.1). In these experiments, the function was calibrated to the threshold that allows for the strongest modulations that stay undetected. Based on these findings, the linear component  $F_{lin}$  is defined as a linear function with slope  $a$  and offset  $b$ . The function derives a scaling factor of the object motion per input point with the current camera speed  $v_{cam}$ , the gaze velocity  $v_{eye}$  and a set of geometry points  $p \in P$  as an input.

$$F_{lin}(P, v_{cam}, v_{eye}) = 1 - (a(v_{cam}) \cdot \min(\theta(p) \cdot n_1, 1) + b(v_{eye})) \quad (1)$$

Here, the eccentricity function  $\theta$  is normalized by  $n_1 = \frac{1}{\pi} \cdot \frac{9}{5}$  to be within  $[0, 1]$ . In the calibration experiments we placed the gaze point in the middle of the screen and, therefore, normalize with half the FOV ( $50^\circ$ ). In the experiments, we found that for higher movement speeds the scaling slope has to be slightly less intense (see Section 3.1). We assume a linear dependency of the velocity for the change of the slope function and interpolate the slope for velocities that are between the calibrated ones. For fast movements above  $1.4 \text{ m/s}$  the slope is extrapolated following this trend making it more flat. This means that an increase in scene speed decreases the intensity of the modifications. Velocities below  $1 \text{ m/s}$  are handled conservatively and a constant slope of  $s_{max} = 0.33$  (33% velocity reduction at the highest eccentricity) is assumed to avoid that the modifications become overly intensive and are recognized by the user. Specifically, we derive the following slope function with the former calibration values:

$$a(v_{cam}) = \min(s_{max}, m \cdot v_{cam} + s_{max} - m) \quad (2)$$

The value  $m$  models the decrease of the slope with increasing speed,  $m = (s_f - s_{max}) / (v_2 - V_1) = -0.125$ , defined by the calibrated slopes  $s_{max} = 0.33$  and  $s_f = 0.28$  for the speeds  $v_1 = 1$  and  $v_2 = 1.4$ .

The function offset parameter  $b$  of  $F_{lin}$  defines the perceptual difference in velocity between the actual linear speed of the scene camera and the perceived movement speed. The formulation of  $b$  is based on the experimental results for velocity adjustment for different eye movements, described in Section 3.1. We found that the

perceived velocity of the modified scene highly varies with the type of eye movement the user performs. For object following, the HVS relies intensively on the foveal region. For fixations, on the other hand, different parts of the visual field contribute more equally to the velocity perception. We define offset  $b$  by a fixation parameter  $b_{fix}$  and an additional term that adjusts the offset of the scaling function for smooth pursuit eye movements.

$$b(v_{eye}) = b_{fix} + o_{sp} \cdot \min(v_{eye} \cdot n_2, 1) \quad (3)$$

With  $b_{fix} = -0.25 \cdot a$  the default offset for full fixations allows for slowing down the object velocity in the outer half of the FOV proportionally to the speed up in the inner circle. The constant  $o_{sp} = 0.2$  is derived from the former calibration and defines the offset adjustment that is necessary to address the increased importance of the fovea for smooth pursuit eye movements. The formula considers a normalization of the gaze velocity  $v_{eye}$  with  $n_2 = \frac{\delta t}{\pi} \cdot \frac{9}{3}$  for gaze movements below 30°/s, where smooth pursuit eye movements fully take place, based on former findings [Leigh and Zee 2015]. We interpolate gaze velocities below 30°/s, since the studies of Leigh *et al.* show that no hard threshold can be defined between fixations and smooth pursuit [2015]. Note that the exact velocity where a movement is categorized as smooth pursuit is debatable and other works also suggest lower values [Komogortsev and Karpov 2013]. By relying on the gaze velocity, we can implicitly model both fundamental types of eye movements in  $F_{lin}$ .

*Rotational Movement Component.* Next, we define the rotational component of the perceptual model by the function  $F_{rot}$  using the discoveries of the psychophysical experiments in Section 3.2. The function output is a scaling factor that moderates the objects' geometrical size to visually align to a linear movement path. For this illusion, the size decreases stronger with higher eccentricity in the direction of motion. We define:

$$F_{rot}(P, p_{eye}) = (1 - \cos(\min(|d_{eye} - d_p|, |d_p|) \cdot \frac{\text{FOV} \cdot \pi}{180})) \cdot \alpha \quad (4)$$

The scalars  $d_{eye}$  and  $d_p$  are the respective distances of the projected object point and gaze point onto the axis of movement. The scaling modifications by  $F_{rot}$  are anchored on the gaze point as well as the forward vector of the head to increase temporal stability, represented by  $\min(|d_{eye} - d_p|, |d_p|)$ . The free parameter  $\alpha$  determines the strength of the effect and was calibrated in the psychophysical experiments to  $\alpha = 0.624$ . A value of 1 would correspond to a visually perfect linear trajectory. The scalars  $d_{eye} := \langle p_{eye}, \hat{\mathbf{u}}_{cam} \rangle$  and  $d_p := \langle p, \hat{\mathbf{u}}_{cam} \rangle$  are the distances of the gaze point  $p_{eye}$  and the geometrical point  $p$  to the origin after projection onto the axis of movement  $\mathbf{u}_{cam}$ . We assume the origin to be at the forward vector position in the image plane.

## 5 APPLICATION

In this section, we describe the implementation of our method for cybersickness mitigation based on our perceptual model. In the psychophysical experiments of Section 3, we directly transformed the objects' 3D geometry in the scene. However, such a full control is not feasible in complex scenes. For general applicability, our implementation therefore operates as an effective and lightweight post-process on the rendering output of a VR application. Unlike

before, the scaling modifications are applied using image warping in the 2D screen-space. Figure 6 provides a general overview of the shader operations that are performed on the RGB image to generate the distorted output. Since our framework operates in the image domain of the rendering, it can be applied to arbitrary scenes without further adjustment. The implementation runs in Unity 2022.3 using renderer features and GPU-based shader operations in combination with OpenXR for VR support. The motion of the pixels between frames is derived by motion flow.

*Warping.* In our implementation, we apply geometrical distortions to reducevection. Without knowledge of the 3D scene, the modifications are applied to the rendered RGB frame. Since the topology of objects is unknown in screen-space, it cannot be adjusted directly. Therefore, the implementation relies on the optical flow to adjust the content movement. Rotational movements are compensated by foveated displacements that transform the content towards the axis of movement. This operation results in a reduction of object sizes by distortions comparable to a concave lens laying over the image. For an efficient implementation that leverages the standard capabilities of graphics hardware, deformations are realized with a sparse vertex grid and UV warping in the vertex shader. The final image is rendered in the fragment shader by simple UV sampling.

The overall displacement  $D$  of the vertices  $p \in P$  in the deformation grid are derived by combining the displacement vectors for rotations  $D_{rot}$  and the vectors of the linear movement modulation  $M_{opt}$ . The vector field of  $M_{opt} = M_{raw} + F_{lin}$  combines the scenes' motion flow  $M_{raw}$  with the linear movement scaling  $F_{lin}$  of the perceptual model (see paragraph 'Optimization').

$$D(P) = D_{rot} + M_{opt} \cdot \max(1 - \frac{\omega_{cam}}{\omega_{max}}, 0) \quad (5)$$

The camera's angular velocity  $\omega_{cam}$  is normalized with a maximum speed constant  $\omega_{max}$  which can be adjusted per scene. In our experiments, we calibrated  $\omega_{max}$  to 40°/s. The predominance of rotational movements makes the compensation of linear movements negligible during strong rotational movements [Groth *et al.* 2022; Kim *et al.* 2021]. Accordingly, the deformations by the linear component, described by  $M_{opt}$ , are scaled by the angular velocity of the camera  $\omega_{cam}$ . The displacement vector field  $D_{rot}$  is derived by applying the scaling factor  $F_{rot}$  of the perceptual model to a vector orthogonal to the camera movement direction  $\mathbf{u}_{cam}$ .

$$D_{rot}(P, \omega_{cam}, \mathbf{u}_{cam}) = F_{rot} \cdot (\mathbf{u}_{cam} \cdot d_p - p) \cdot \frac{\omega_{cam}}{\omega_{max}} \cdot 2 \quad (6)$$

The length of the displacement  $D_{rot}$  depends on the eccentricity of the vertex point  $p$ . Here,  $\mathbf{u}_{cam} \cdot d_p - p$  gives us the vector that points from  $p$  to the axis of movement.

*Optimization.* The movement of the scene content between frames is derived by motion flow. To slow-down the velocity of peripheral content, the movement can be transformed to the inverse of the motion vectors. This counter-movement is modulated by the scaling parameter of the linear component of the perceptual model  $F_{lin}$  to derive the initial vector representation  $M_{init}$ . However, image overlaps can arise from opposing motion vectors. To avoid overlaps and find a globally optimal solution, we optimize  $M_{init}$  to find the

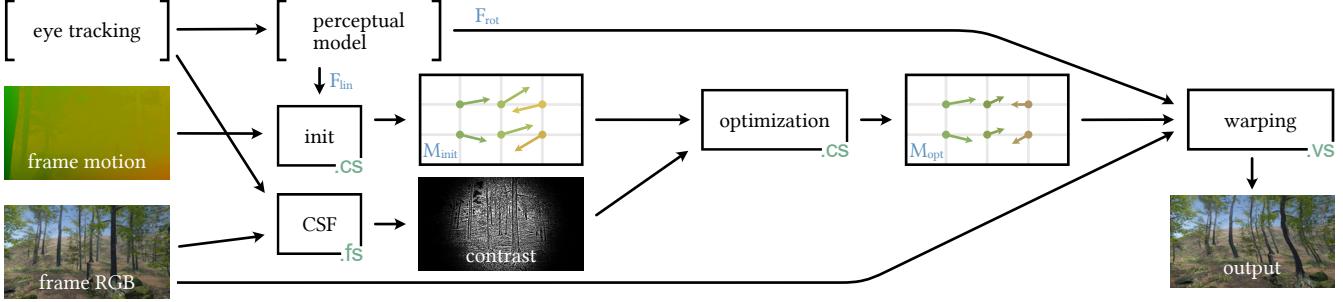


Fig. 6. With a series of shader operations (shader identifier in box corners), we apply screen-space geometrical modifications that reduce vection.

final vector field  $M_{opt}$  which is used for the image warping. The goal is to determine a final state, ensuring each vertex closely aligns with its desired target position including considerations of the importance of different image areas. The optimization is required to maintain the non-penetration constraint.

Like for the warping, we opt for an efficient sparse grid (32x32p) for the optimization rather than pixel-level resolution. The optimization uses a semi-implicit Euler approach (step constant = 0.02). The forces acting on the vertices are the accumulated sum of the displacement forces to each of the four neighboring vertices. The rest length between vertices is the sum of the grid cell width (32p) and the temporally accumulated deformation of former frames. This dynamic definition of the spring rest length allows for more temporal stability than a static definition. In the optimization, we want to preserve content with high visibility, e.g. high contrast edges, while content under the threshold of sensibility can be deformed arbitrarily. The stiffness parameter of each spring between vertices in the grid is defined by the luminance contrast of the surrounding content given by the eccentricity-based contrast sensitivity function (CSF) [Tursun et al. 2019]. After the contrast sensitivity is masked by a transducer model [Zeng et al. 2000], the final visibility value is derived by summing up the individual frequency layers. Higher visibility corresponds to greater stiffness which preserves the content. The damping constant was chosen with  $d = 0.9$  based on experimental validation. For each frame, we optimize the grid for 50 iterations on the GPU, taking approximately 3 ms.

For each vertex computation, the positions are evaluated in a local coordinate system, centered at the vertex's default grid position. This local approach simplifies displacement calculations, independent of global positioning and actual grid dimensions. Although the vertices do not have a global understanding, the global effects still have an influence as they spread over the multiple iterations of one timestep. Lastly, to preserve image content, the line of vertices along the image border remains fixed in position.

*Recovery of the Original Rendering.* The geometrical distortions accumulate over time since the modulations operate as a temporal effect. During blinks and saccades, the phenomenon of change blindness occurs, which masks changes made to the visual content. In our implementation, we leverage this natural effect to recover to the original rendering. Human eye blinks occur around every

3 seconds and allow for an instantaneous reset of the entire image [Nakano et al. 2013]. The recovery with saccadic suppression, on the other hand, is more challenging, because a full reset would alter the eye's target position and disrupt the inherent expectation of edge consistency. Therefore, we only reset content  $c$  that based on the gaze position  $g$  is located against the direction of movement  $\mathbf{m}$ . It applies  $\langle(c - g), \mathbf{m}\rangle < 0$ . To avoid a sharp edge between the recovered and deformed content, the reset intensity is scaled with a gradient. Formally, the reset intensity  $I$  is quantified by:

$$I(b, p_v, p_{eye}) = \|\langle \hat{\mathbf{m}}, p_v \rangle - \langle \hat{\mathbf{m}}, p_{eye} \rangle\| \quad (7)$$

Here,  $p_v$  represents the position of the respective vertex in UV coordinates, and  $p_{eye}$  denotes the gaze point of the eye. Sun et al. found that scene displacements during saccades are undetectable for movements up to 12.6°/s even in the target region [Sun et al. 2018]. During saccades, we use this threshold to subtly reduce content deformations located in the direction of movement. In the case of insufficient resets, the distortions are kept below a maximum offset of 20° from the original content. The threshold is chosen based on our empirical analysis and ensures that the warping effect is both effective and visually pleasing. For the detection of eye blinks, we rely on the capabilities of the SRanipal library and the HTC Vive Pro Eye HMD. The detection of saccades is based on the gaze velocity using the algorithm of Imaoka et al. [2020].

## 6 VALIDATION

We validate the effectiveness of our implementation in a comprehensive experiment using a realistic VR scenario with multiple scenes. We run a within-subjects experiment with three sessions per participant to explore the effect of the proposed deformation technique. While the virtual environment was the same for all sessions, the visual post-processing was altered, comparing our method against blurring of the peripheral vision and an unaltered control condition. We counterbalanced the order in which the three conditions were shown to the participants. Furthermore, a recovery time of at least 48 hours was maintained between sessions to avoid carry-over effects. The experiment is approved by the university's ethics committee.

### 6.1 Experiment Design

#### 6.1.1 Stimuli

*Conditions.* We investigate three different conditions: Our technique that does content-specific adjustments to the scene content



Fig. 7. In the experiment, we used two scenes, both displayed twice per session with a smooth transition.

while preserving visual fidelity, the commonly used method of peripheral blurring and a control condition as baseline. Peripheral blurring is one of the most common techniques to subtly prevent occurrences of cybersickness. Based on former work, we implemented a foveated blurring where the quality degradation from the foveal point is scaled by the camera's linear and angular velocity [Groth et al. 2021a,b]. At maximum speed, the foveal region is kept the smallest with a diameter of  $10^\circ$  [Groth et al. 2021b]. The strength of the blur increases linearly with eccentricity. Following the results of Lin *et al.* [Lin et al. 2020b], the linear increase is set with a maximum kernel size of 13 which was found to be at a 50% detection probability. Comparably, the underlying principles of our warping are likewise calibrated to this commonly used detection threshold (see Section 3) and adjusted based on the methodology described in the implementation section. We perform the manipulations in image space for a more dynamic response to scene content and broad applicability of the technique to a variety of applications.

*Virtual Environment.* In the experiment we present a photorealistic virtual environment with two different scenes, urban and nature. While the urban scene mainly contains geometrical structures and straight lines, the nature scene explores the perception of a more heterogeneous shape composition. The camera path in the experiment is predefined and cannot be altered by the participants. While the path completes its circular trajectory, both environments are shown for the same amount of time, resulting in four scene changes per session in the same order. We carefully design the camera path to feature a balanced mixture of different camera movements. For linear movements, the camera constantly accelerates and decelerates in a sinusoidal manner to investigate movements that are most relevant for cybersickness [Groth et al. 2022; O'Hanlon et al. 1974]. Furthermore, we include sections with unpredictable movements with a linear and angular component and full  $360^\circ$  rotations.

**6.1.2 Apparatus.** For the experiment we used a HTC Vive Pro Eye head-mounted display (HMD) with a visible FOV of  $100^\circ$  and a frame rate of 90 Hz. The resolution of that HMD is 2880 x 1600 pixel. The rendering is performed on a commodity computer with a NVIDIA RTX 4090 graphics card.

**6.1.3 Participants.** A total of 25 participants completed all three sessions (12 females, 1 non-binary, Age range = 19 - 35, Avg age = 24.9, SD = 4.29). The participation was compensated with 40€. Due to the within-subjects design, every participant experienced all of the conditions. The order of the conditions are counterbalanced. For the analysis, we separate the participants in two disjoint groups based on the occurrence of sickness symptoms in the baseline condition. In

line with former research, a simulator sickness questionnaire (SSQ) total score of 20 is chosen as clustering factor [Groth et al. 2022]. The group with individuals that got negatively affected consists of 17 participants (11 females).

**6.1.4 Measurement.** We used the SSQ for participant feedback on cybersickness [Kennedy et al. 1993]. Following common procedure, we let participants fill in the SSQ twice, before and after each session of the experiment, to counteract different daily conditions. The total sickness score as well as the corresponding subscores of the SSQ are calculated according to the original procedure of Kennedy *et al.* [Kennedy et al. 1993]. During the experiment we also asked the participants to press different buttons every time their feeling in comfort got worse or better. This discomfort includes all symptoms of the SSQ and can be in intervals as coarse or fine as participants choose. The responses allow for the calculation of the participants' individual level of discomfort over time [Groth et al. 2022]. After each session of the experiment we hold a semi-structured interview with the participants. The questions were: (1) Have you noticed anything unusual, and if so, what? (2) *Optionally:* How often did the effect occur? (3) *Optionally:* How disturbing did you find the effect? (4) How was your overall feeling during this session? (5) How immersed were you in the VR environment?

We chose an interview over subjective rankings of subtleness and immersion since the nature of the two manipulations is very different and comparing them on one numerical scale is prone to misinterpretation. The interview gave us the chance to dig deeper into the effects that were actually perceived by the participants and gain substantial insights about the impact of the investigated methods.

**6.1.5 Procedure.** The experiment for every participant was conducted in three sessions with one condition in each session (control, blur, or warp). The procedure of each session followed the same structure. In the beginning, every participant gave written consent and was informed about the experimental procedure and the possibility that negative symptoms may arise. Demographics have already been provided for participant registration. In each session, participants first filled the pre-experimental SSQ to capture their initial state, adjusted the VR glasses and performed a calibration routine for the eye tracker. In the experiments, participants were seated and ask to keep their head straight to increase comparability between sessions. The in-game camera was moved along the predefined path and participants could constantly indicate their well-being over the respective buttons. We ask participants to indicate when they experience severe negative feelings and the session was ended immediately in that case. Otherwise, the total time of the VR experience was 13 minutes. After every session of the experiment, participants first filled in the second SSQ, before we then conducted a semi-structured interview.

## 6.2 Analysis and Results

For the analysis of the experimental results, we performed pairwise two-sided dependent t-tests for repeated measures comparing both manipulation techniques to the control condition and to each other. Qualitative results of the interviews were determined by thematic

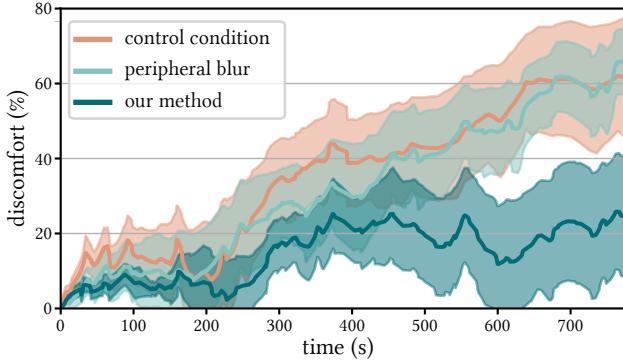


Fig. 8. Progression of the relative discomfort of all participants over the sessions, with shaded areas for the standard error of the mean (SEM).

analysis [Braun and Clarke 2006]. The further results focus only on the group of participants for whom the virtual simulation evoked significant cybersickness, since mitigation methods are less interesting for participants that do not get sick in the first place. As expected, the statistical results of the group of participants that were unaffected by the simulation show no significant change in the SSQ scores or times participants spent in the VR environment. Data from two participants had to be discarded from the time-based discomfort analysis due to improper task execution.

**Results for Effectiveness.** Figure 9 shows the SSQs scores for total sickness and the three subscales nausea, disorientation, and oculomotor effects. The results confirm our technique’s effectiveness in mitigating cybersickness. Our method significantly reduced the sickness scores (-31.8% over control) across all SSQ subscales (total:  $T = 4.87, p = 0.0002$ ; nausea:  $T = 3.81, p = 0.0015$ ; disorientation:  $T = 4.37, p = 0.0005$ ; oculomotor:  $T = 2.68, p = 0.0166$ ). While the peripheral blur technique also reduces the average cybersickness (-20.2%), it does not achieve a significant effect in our experiment. Simultaneously, the time participants are willing to spend in the VR environment (600.3s in control) was significantly increased when post-processing the output (Blur: +54.8s, Warp: +92.9s). However, only our warp condition had a significant influence on the increase of time in the VR environment ( $T = -2.97, p = 0.009$ ). By introducing content distortions, the drop-out rate was reduced by 50% over the control session (control: 41.7%, blur: 29.2%, warp: 20.8%). From the qualitative results, a majority of 84% of participants stated that the session with our technique was the most enjoyable without being aware of the differences.

The positive impact of content-aware image warping on the comfort of users is further supported by the real-time data of the discomfort analysis (see Figure 8). When applying our modulations, the general level of relative discomfort over time was significantly lower. While the relative level of discomfort is around 25% at the end, the temporal progression of the warp session only shows a rising trend between 240 and 300 seconds. The much larger remaining intervals, on the other hand, tend to oscillate around their constant mean. Participants were even able to recover from some of their

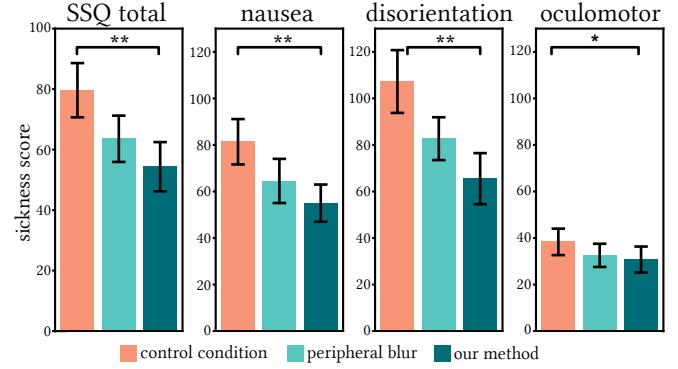


Fig. 9. SSQ results for all conditions. Error bars represent the SEM. Significance is denoted by \*\*\* ( $p \leq 0.01$ ) and \*\* ( $p \leq 0.05$ ).

discomfort in certain parts of the simulation, e.g., between 450 to 520 seconds. Such a recovery was not evident in the other conditions.

With the pre-defined camera trajectory and continuous discomfort metrics, we can further assess the relative efficacy of individual motion types (linear and angular). The mean discomfort increment (as percentage per minute) for rotational motions registered at 15.97 in the control scenario, 15.88 with peripheral blur applied, and 11.37 when warping was used. Linear motion was less severe and increased discomfort rates by 0.53% per minute without post-processing and 0.70% per minute with peripheral blur. With our warp implementation the data suggests even a decrease in discomfort during linear motion scenarios with -2.29% per minute.

**Results for Subtlety.** A decisive factor for the widespread use of methods for cybersickness reduction is, in addition to efficiency, the unobtrusiveness of the methods. The perceptual model of our method is based on novel findings from our studies of the HVS (Section 3). The free parameters of the model are calibrated to the conservative detection threshold in dedicated experiments. Further constraints in the implementation prevent excessive distortion of the scene. In the validation experiment, we investigated the subtleness of the applied image manipulations and the extent of distracting participants from their virtual experience. From the results of the semi-structured interviews, for 36% of the participants it was impossible to detect any changes made to the scene over the whole 13 minutes of the experiment. Surprisingly, the peripheral blur was less noticed in our experiment (undetected for 64%) than in former research with equivalent parameters. On the other hand, it was noted to be more distracting when perceived by participants. A detailed investigation of the eye tracking results leads to the assumption that most of the detected artifacts of the warping were not caused by the visibility of the deformations themselves, but rather by insufficient resets during the saccades. Better hardware and saccade detection techniques have the potential to improve the subtleness of the resetting of scene deformations.

**Further Findings.** Overall, the results show the same trend for men and women, with the control condition perceived as the most sickness inducing and the warping condition rated as most pleasant (see Figure 11). In line with former findings [Groth et al. 2022;

Narciso et al. 2019], women experienced higher sickness symptoms, have a higher average sickness score, and end the experiment earlier. However, with our scene deformation technique applied, cybersickness was mitigated significantly for both, men ( $T = 3.01, p = 0.0299$ ) and women ( $T = 3.86, p = 0.0032$ ). The result give a positive indication for the use of warping methods for effective reduction of the general gender bias of VR experiences. This results, however, should be considered with care due to the small group size of male participants that actually got sick in the experiment.

Based on our qualitative inquiry, the influences on cybersickness are grouped into two themes: Influences of the scene and the type of movement. Most of the participants reported the forest scene to be worse than the city presentation. This was further reported to be influenced by the closeness of the surrounding objects, e.g. trees, which were further away in the urban area. In line with former research, rotational movements caused more severe cybersickness [Groth et al. 2022; Kim et al. 2021]. Participants reported that pitch rotations had an especially strong influence. One participant related the severity of pitch rotations to the high buildings of the city environment [P17]. Movements to the side were further claimed to be critical for sickness symptoms, while forward movements were not seen as a problem. The majority of participant experienced a high level of presence in the VR experience (15 out of 25 participants; 60%). Another 8 individuals reported the presence level to be at a decent amount, describing the experience as “medium immersing” or “somewhat real”. Participants’ preferences among the session comparing immersion was less clear. Most of the participants reported that all sessions were equally immersing (10 participants). Followed by a preference for the control session and warping with 7 and 6 votes, respectively. The blurring was less accepted for an immersive experience (2 votes), which likely be attributed to the attenuation of details in the periphery.

## 7 LIMITATIONS AND FUTURE WORK

The experiment reveals the peripheral blur to be more subtle than in prior studies using comparable intensities. Therefore, our method’s predominance in effectiveness, though indicative, might differ when stronger blurring is applied. Notably, more pronounced blurring would further diminish peripheral visual details, while our method preserves the visual fidelity of the scene.

The current implementation of our method uses the motion flow of the rendering pipeline to retrieve content motion between frames. For content not relying on a 3D scene, like 360° videos, the computation would require a different input, e.g. vision-based optical flow. Hardware-accelerated optical flow algorithms promise real-time computation in milliseconds [Fast Optical Flow 2023].

Another limitation of our method is the prediction of complex camera movements in short time intervals. Unpredictable changes in direction can disrupt the motion energy accumulation, leading to artifacts in the rendered scene. The full scene resets, triggered by the user’s blinking, recover the original rendering every three seconds on average.

The distortions of the warped frames may also become visible in scenarios with near-camera presentations and scenes that are extensively regular, e.g. very regular low-poly scenes. Especially in the

forest scene, multiple objects of the virtual environment got close to the camera and some participants reported recognizing artifacts of the distortions in these scenarios. Strong regularity, on the other hand, has the potential to disrupt the subtlety of the method because the optimization often results in a curvature effect of straight lines. In a scene consisting mostly of high contrast lines and regular patterns, the manipulations can become visible. However, while the presented city scene already contains considerable regularity, our method remained subtle in this scenario.

Against our expectations, the peripheral blur method was more subtle and less effective than in comparable implementations that use the same gradient [Groth et al. 2021a; Lin et al. 2020b]. Compared with the experiments of Lin et al. [2020b], our participants could not control the camera movements. Also, the scenes in our experiment were more complex, which may had an influence on the detectability of the blurring. In contrast, our warping method was adjusted to a 50% detection rate which is confirmed by the experimental results. Using a more aggressive blurring of the peripheral content can result in a more effective mitigation of cybersickness. However, increasing the blur would result in even stronger suppression of peripheral details. The key motivation of our work is to allow effective cybersickness mitigation without the suppression of any visual details.

Our implementation does not yet separate between background and dynamic foreground objects. Dynamic objects do not inducevection, and therefore no cybersickness [Seno et al. 2009]. For scenes with extensively moving objects, a separation between dynamic and static objects would avoid unnecessary scene deformations. All visual movements that originate to the users’ body movements, e.g. turning the head, should not be compensated because the corresponding vestibular signals are triggered accordingly. Our implementation considers this fact and only counteracts passive camera movements.

Potential future work that exploits the HVS are modulations of disparity to counteract forward and backward movements. These modulations involve adjusting the virtual environment’s disparity to negate motion effects, potentially reducing sensory mismatches.

## 8 CONCLUSION

This research introduces a new technique for mitigating cybersickness in virtual reality environments without compromising the visual details of the scene. The perceptual model of our technique is based on two key findings: Firstly, reducing the visual motion of objects in the motion-sensitive peripheral field does not alter the camera velocity perception as long as the movement speed in the fovea is adjusted accordingly. Secondly, angular camera movements can be adjusted to visually move on a linear trajectory by scaling the image content with eccentricity. By means of calibration experiments, perceptual thresholds were found that keep the effects unnoticeable while maximizing the reduction ofvection. We presented an effective implementation of our method that runs in real-time as a post-process on the rendered image. Experimental validation confirms the effectiveness of our approach to mitigate cybersickness, while the modifications often remained completely undetected.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge funding by the German Science Foundation (DFG projects MA2555/15-1 “Immersive Digital Reality” and 491805996) the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No. 804226 PERDY).

## REFERENCES

- Isayas Berhe Adhanom, Nathan Navarro Griffin, Paul MacNeilage, and Eelke Folmer. 2020. The Effect of a Foveated Field-of-view Restrictor on VR Sickness. In *IEEE Conf. Virtual Reality and 3D User Interfaces*. IEEE, 645–652.
- Stephen J Anderson, Kathy T Mullen, and Robert F Hess. 1991. Human peripheral spatial resolution for achromatic and chromatic stimuli: limits imposed by optical and retinal factors. *The Journal of physiology* 442, 1 (1991), 47–64.
- Kaitlyn R Bankieris, Vikrant Rao Bejjanki, and Richard N Aslin. 2017. Sensory cue-combination in the context of newly learned categories. *Scientific Reports* 7, 1 (2017).
- Jelte E Bos, Sjoerd C de Vries, Martijn L van Emmerik, and Eric L Groen. 2010. The effect of internal and external fields of view on visually induced motion sickness. *Applied ergonomics* 41, 4 (2010), 516–521.
- Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101.
- Scott Daly. 2001. Engineering observations from spatiotemporal and spatiotemporal visual models. In *Vision Models and Applications to Image and Video Processing*.
- Paul DiZio and James Lackner. 1997. Circumventing side effects of immersive virtual environments. *Advances in human factors/ergonomics* 21 (1997), 893–896.
- Sarah D’Amour, Laurence R Harris, Stefan Berti, and Behrang Keshavarz. 2021. The role of cognitive factors and personality traits in the perception of illusory self-motion (vection). *Attention, Perception, & Psychophysics* 83 (2021), 1804–1817.
- Sigm Exner. 1886. Ein Versuch über die Netzhautperipherie als Organ zur Wahrnehmung von Bewegungen. *Archiv für die gesamte Physiologie des Menschen und der Tiere* 38, 1 (1886), 217–218.
- Fast Optical Flow 2023. NVIDIA Optical Flow SDK. <https://developer.nvidia.com/optical-flow-sdk>. Accessed: 2024-01-17.
- David Finlay. 1982. Motion perception in the peripheral visual field. *Perception* 11, 4 (1982), 457–462.
- Colin Groth, Jan-Philipp Tauscher, Nikkel Heesen, Susana Castillo, and Marcus Magnor. 2021a. Visual Techniques to Reduce Cybersickness in Virtual Reality. In *IEEE Conf. Virtual Reality and 3D User Interfaces*. 486–487.
- Colin Groth, Jan-Philipp Tauscher, Nikkel Heesen, Steve Grogorick, Susana Castillo, and Marcus Magnor. 2021b. Mitigation of Cybersickness in Immersive 360° Videos. In *IEEE VR Workshop on Immersive Sickness Prevention*. 169–177.
- Colin Groth, Jan-Philipp Tauscher, Nikkel Heesen, Max Hattenbach, Susana Castillo, and Marcus Magnor. 2022. Omnidirectional Galvanic Vestibular Stimulation in Virtual Reality. *IEEE Trans. Vis. Comput. Graph.* 28, 5 (2022), 2234–2244.
- Xuanru Guo, Shinji Nakamura, Yoshitaka Fujii, Takeharu Seno, and Stephen Palmisano. 2021. Effects of luminance contrast, averaged luminance and spatial frequency onvection. *Experimental Brain Research* 239 (2021), 3507–3525.
- Sébastien Hillaire, Anatole Lécuyer, Rémi Cozot, and Géry Casiez. 2008. Depth-of-field blur effects for first-person navigation in virtual environments. *IEEE computer graphics and applications* 28, 6 (2008), 47–55.
- Ping Hu, Qi Sun, Piotr Didyk, Li-Yi Wei, and Arie E. Kaufman. 2019. Reducing Simulator Sickness with Perceptual Camera Control. *ACM Trans. Graph.* 38, 6, Article 210 (2019), 12 pages.
- Yu Imaoka, Andri Flury, and Eling D De Bruin. 2020. Assessing saccadic eye movements with head-mounted display virtual reality technology. *Frontiers in Psychiatry* 11 (2020), 1–19.
- Anton Kaplanyan, Anton Sochenov, Thomas Leimkuhler, Mikhail Okuney, Todd Goodall, and Gizem Rufo. 2019. DeepFovea: Neural Reconstruction for Foveated Rendering and Video Compression using Learned Statistics of Natural Videos. *ACM Trans. Graph. (Proc. SIGGRAPH Asia)* 38, 4 (2019), 212:1–212:13.
- Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The Int. Journal of Aviation Psychology* 3, 3 (1993), 203–220.
- Juno Kim, Stephen Palmisano, Wilson Luu, and Shinichi Iwasaki. 2021. Effects of linear visual-vestibular conflict on presence, perceived scene stability and cybersickness in the Oculus Go and Oculus Quest. *Front. Virtual Real.* 2 (2021), 42.
- Song Min Kim and Gerard J. Kim. 2022. Sparse Peripheral Countervection Flow Visualization with Reverse Optical Flow for VR Sickness Reduction. In *International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 798–799.
- Oleg V Komogortsev and Alex Karpov. 2013. Automated classification and scoring of smooth pursuit eye movements in the presence of fixations and saccades. *Behavior research methods* 45 (2013), 203–215.
- Joseph J LaViola Jr. 2000. A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin* 32, 1 (2000), 47–56.
- John Leigh and David Zee. 2015. *The Neurology of Eye Movements*. Oxford University Press.
- JJ-W Lin, Henry Been-Lirn Duh, Donald E Parker, Habib Abi-Rached, and Thomas A Furness. 2002. Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Proc. IEEE Virtual Reality*. IEEE, 164–171.
- Yun-Xuan Lin, Rohith Venkatakrishnan, Roshan Venkatakrishnan, Elham Ebrahimi, Wen-Chieh Lin, and Sabarish V. Babu. 2020a. How the Presence and Size of Static Peripheral Blur Affects Cybersickness in Virtual Reality. *ACM Trans. Appl. Percept.* 17, 4, Article 16 (Nov. 2020), 18 pages. <https://doi.org/10.1145/3419984>
- Ruding Lou, Richard So, and Dominique Bechmann. 2022. Geometric Deformation for Reducing Optic Flow and Cybersickness Dose Value in VR. In *Eurographics 2022 - Posters*.
- Wilson Luu, Barbara Zangerl, Michael Kalloniatis, and Juno Kim. 2021. Effects of stereopsis onvection, presence and cybersickness in head-mounted display (HMD) virtual reality. *Scientific reports* 11, 1 (2021), 12373.
- Brandy Murovec, Julia Spaniol, Jennifer L Campos, and Behrang Keshavarz. 2021. Multisensory effects on illusory self-motion (vection): The role of visual, auditory, and tactile cues. *Multisensory Research* 34, 8 (2021), 869–890.
- Tamami Nakano, Makoto Kato, Yusuke Morito, Seishi Itoi, and Shigeru Kitazawa. 2013. Blink-related momentary activation of the default mode network while viewing videos. *Proc. Natl. Acad. Sci. USA* 110, 2 (2013), 702–706.
- David G. Narciso, Maximino Bessa, Miguel C. Melo, António Coelho, and José Vasconcelos-Raposo. 2019. Immersive 360° video user experience: impact of different variables in the sense of presence and cybersickness. *Univers. Access Inf. Soc.* (2019).
- Guangyu Nie, Yue Liu, and Yongtian Wang. 2017. Prevention of Visually Induced Motion Sickness Based on Dynamic Real-Time Content-Aware Non-salient Area Blurring. In *IEEE Int. Symp. Mixed and Augmented Reality Adjunct*. 75–78.
- Guang-Yu Nie, Henry Been-Lirn Duh, Yue Liu, and Yongtian Wang. 2019. Analysis on Mitigation of Visually Induced Motion Sickness by Applying Dynamical Blurring on a User’s Retina. *TCVG* (2019).
- James F O’Hanlon, Michael E McCauley, et al. 1974. Motion sickness incidence as a function of the frequency and acceleration of vertical sinusoidal motion. *Aerospace medicine* 45, 4 (1974), 366–369.
- Su Han Park, Bin Han, and Gerard Jounghyun Kim. 2022. Mixing in reverse optical flow to mitigatevection and simulation sickness in virtual reality. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. Article 189, 11 pages.
- Anjul Patney, Marco Salvi, Joohwan Kim, Anton Kaplanyan, Chris Wyman, Nir Bentz, David Luebke, and Aaron Lefohn. 2016. Towards Foveated Rendering for Gaze-Tracked Virtual Reality. *ACM Trans. Graph.* 35, 6, Article 179 (Nov. 2016), 12 pages. <https://doi.org/10.1145/2980179.2980246>
- James T Reason and Joseph John Brand. 1975. *Motion sickness*. Academic press.
- A Fleming Seay, David M Krum, Larry Hodges, and William Ribarsky. 2001. Simulator sickness and presence in a high FOV virtual environment. In *Proc. IEEE Virtual Reality*. IEEE, 299–300.
- Takeharu Seno, Hiroyuki Ito, and Shoji Sunaga. 2009. The object and background hypothesis forvection. *Vision research* 49, 24 (2009), 2973–2982.
- Craig Sherman. 2002. Motion sickness: review of causes and preventive strategies. *Journal of travel medicine* 9, 5 (2002), 251–256.
- Misha Sra, Abhinandan Jain, and Pattie Maes. 2019. Adding Proprioceptive Feedback to Virtual Reality Experiences Using Galvanic Vestibular Stimulation. In *Conf. Human Factors in Computing Systems*. 1–14.
- Qi Sun, Anjul Patney, Li-Yi Wei, Omer Shapira, Jingwan Lu, Paul Asente, Suwen Zhu, Morgan McGuire, David Luebke, and Arie Kaufman. 2018. Towards virtual reality infinite walking: dynamic saccadic redirection. *ACM Trans. Graph.* 37, 4 (2018).
- Arjan C ter Horst, Mathieu Koppen, Luc P. J. Selen, and W. Pieter Medendorp. 2015. Reliability-based weighting of visual and vestibular cues in displacement estimation. *PloS one* 10, 12 (2015), 1–15.
- Benjamin Thompson, Bruce C Hansen, Robert F Hess, and Nikolaus F Troje. 2007. Peripheral vision: Good for biological motion, bad for signal noise segregation? *Journal of Vision* 7, 10 (2007), 1–12.
- William Thompson, Roland Fleming, Sarah Creem-Regehr, and Jeanine Kelly Stefanucci. 2011. *Visual Perception from a Computer Graphics Perspective*. CRC Press.
- Okan Tarhan Tursun, Elena Arabadzhyska-Koleva, Marek Wernikowski, Radoslaw Mantiuk, Hans-Peter Seidel, Karol Myszkowski, and Piotr Didyk. 2019. Luminance-contrast-aware foveated rendering. *ACM Trans. Graph.* 38, 4 (2019), 1–14.
- Wenjun Zeng, Scott Daly, and Shawmin Lei. 2000. Point-wise extended visual masking for JPEG-2000 image compression. In *International Conference on Image Processing*, Vol. 1. 657–660.
- André Zenner, Kora Persephone Regitz, and Antonio Krüger. 2021. Blink-suppressed hand redirection. In *IEEE Conf. Virtual Reality and 3D User Interfaces*. 75–84.

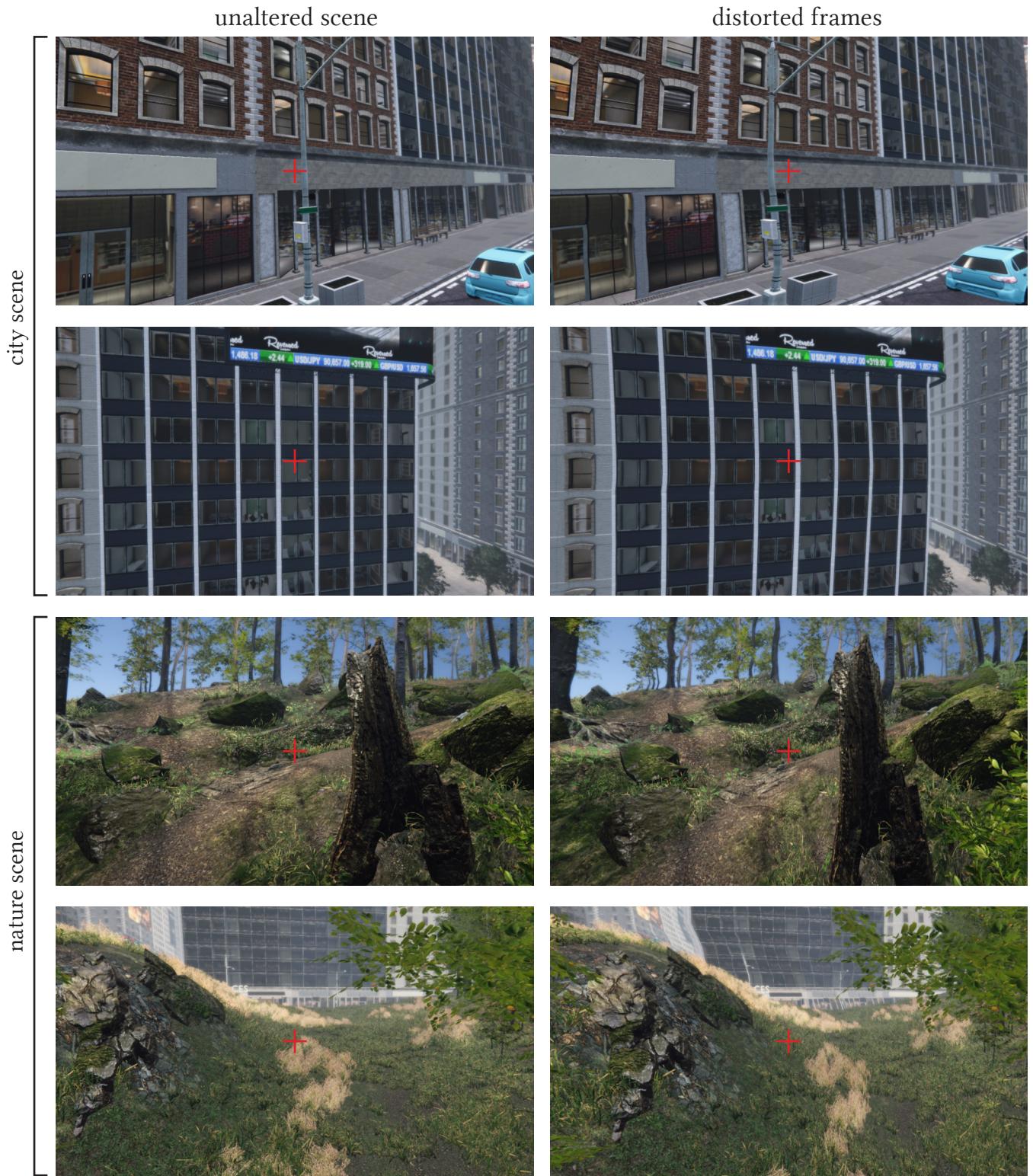


Fig. 10. Exemplary frames of the validation experiment. In the left column we present the unmanipulated frames of the control condition. After applying our method, the frames are distorted to reduce the visual motion (right column). The gaze point is marked with the red cross. In the VR glasses, the effect is gaze-contingent and mostly unnoticeable to users.

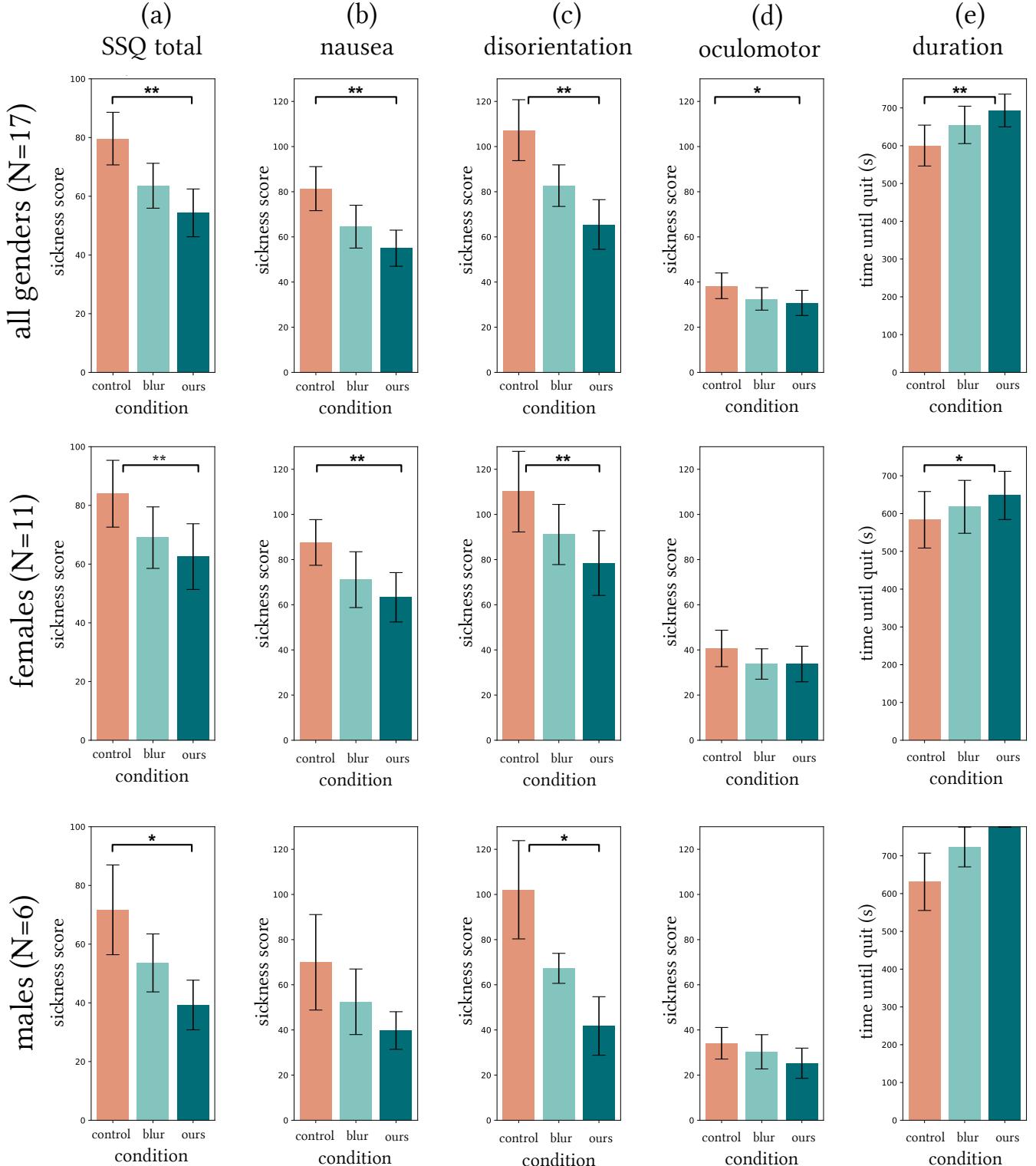


Fig. 11. Results of the SSQ and durations for the control condition, the peripheral blur and our warping technique. This analysis only includes participants that were negatively affected (SSQ total score > 20). Error bars represent the SEM. (a) SSQ results for the total score. (b-d) results for each of the SSQ subscales. (e) duration people were willing to spend in the virtual environment. Significant results are denoted by \*\*\* ( $p \leq 0.01$ ) and \*\* ( $p \leq 0.05$ ).