

THE IMPLICATIONS OF INTERPUPILLARY DISTANCE VARIABILITY FOR VIRTUAL REALITY

Paul B. Hibbard, Loes C.J. van Dam*

University of Essex
Departments of Psychology
Wivenhoe Park, Colchester, CO4 3SQ, UK

Peter Scarfe

University of Reading
Department of Psychology and
Clinical Language Sciences
WhiteKnights, Reading, RG6 6AH, UK

ABSTRACT

Creating and presenting binocular images for virtual reality and other 3D displays needs to take account of the interpupillary distance - the distance between the user's eyes. While VR headsets allow for some adjustments of this setting, this does not accommodate the full range found in the population, and may not necessarily be accurately measured and adjusted in practice. A mismatch between the observer's IPD and that assumed in creating and presenting stimuli will tend to cause problems with viewing comfort and accurate depth perception. We identify unnatural eye fixations, visual discomfort and inaccurate depth perception as important considerations for understanding the suitability of VR for use by children. We present a geometrical quantification of each of these factors.

Index Terms— Virtual Reality, Individual differences, 3D vision, interpupillary distance, distance perception

1. INTRODUCTION

The binocular cues of disparity and convergence are an important source of depth information. Through a consideration of the projective geometry of binocular viewing, it is possible to determine the locations of points in 3D space from the images that they form in the two eyes [16]. As a result, binocular information contributes to the quality, magnitude and precision of the perception of metric depth [7]. This is true both for traditional, screen-based displays [10, 11, 12, 13] and immersive virtual reality [8].

In a 3D display, the two dynamic retinal images that would be seen in a naturally viewed scene can be recreated, so that the optic array experienced by the observer can be matched to a simulated or recorded environment. If these are perfectly matched to the simulated scene, the visual experience of the observer is then expected to replicate that of an observer in that real, natural scene. For this to be achieved when rendering a scene for display, the exact locations of the observer's

eyes must be known, so that the scene cameras can be placed at these locations. This requires accurate, rapid, 6 degrees of freedom tracking of the head position, so that the camera locations can be updated in real time [5, 22]. It also requires an accurate estimate of the observer's interpupillary distance (IPD) – the distance between their two eyes. If there is a mismatch between the observer's IPD and that simulated in the rendering of the stimuli, then the binocular disparities in the image will not accurately replicate those that would be experienced by the observer in the actual, natural scene [23].

A mismatch between the effective IPD in a VR headset and the observer's IPD can affect the viewing experience in four important ways. The first effect is the potential to induce divergent orientations of the eyes. When an observer is viewing an object in the far distance, the directions of gaze of the two eyes will be parallel. As the target comes closer, the directions of gaze required to fixate it will become increasingly convergent. Therefore in natural viewing, divergent viewing angles do not occur, at least in people with normal vision and oculomotor control. If in a 3D display the observer's IPD is smaller than that used to render the stimuli, the eyes will need to diverge in order to fixate distant targets. Such divergent viewing angles are a known source of viewing discomfort in 3D displays; our ability to diverge our eyes is very limited and thus this rendering can also easily lead to a break down of binocular fusion [9, 13, 25].

The second effect is that a mismatch between the observer's IPD and that used in capturing or rendering the images for display will lead to distortions in the perception of depth. The binocular disparities in a pair of images result from the lateral separation of the eyes or cameras. To use these disparities to judge the distances to objects and their three-dimensional shape, the observer needs to take their own IPD into account. Any mismatch between the observer's IPD and that used in creating the images, is therefore predicted to lead to biases in 3D perception, relative to the actual structure of the captured or simulated scene.

The third effect is that the range of distances which can be presented comfortably will be altered. Estimates of the mag-

*Thanks to Leverhulme Trust Research Grant RPG-2016-361 and EIRA

nitudes of disparity that can comfortably be presented vary between 1° [13] and 2° [25]. The range of distances that this corresponds to will depend on the IPD used to render the scenes.

The final issue, which we do not discuss in detail here, is the mismatch between the positions of the headset lenses and the observers' eyes. VR headsets contain high-powered lenses to allow the observer to accommodate to a screen which is located just a few centimetres in front of their eyes. Ideally, these lenses should be directly in front of the pupils. If the observer's IPD does not match the separation in the headset, the observer will not be looking directly through the centre of the lens when looking directly ahead. This off-axis viewing will amplify optical artefacts in the display system such as chromatic aberrations [1].

These considerations create a challenge when designing VR headsets and applications since there is considerable variation in the IPD across the population. The average IPD for adults has been reported to be 63.4mm, with a standard deviation of 3.8mm [4]. This variability in the population can be accommodated by allowing the IPD of the headset to be varied to match each individual user's IPD. Table 1 shows the IPD range of three current mainstream consumer headsets. Even taken this adjustment into account, a considerable proportion of the population have an IPD outside of this range. In practice, it might also be assumed that not all users will necessarily take the time to vary the IPD, particularly when the headset is shared between multiple people. If the IPD is set to the centre of its range, we can calculate that the average mismatch between observer and headset IPD across the population will tend to be around 5% (table 1). It should also be noted that adjusting the IPD to match the user's will correct the convergence angle when viewing objects at the effective distance of the image plane (around 1.5m), but will not scale the binocular disparities in the images, which are fixed by the image capturing or rendering. Thus, even if the IPD in the headset is correct, systematic errors in the perception of distance for each user are predicted, dependent on the mismatch between their IPD and that assumed when the images are created.

The variation in IPD in the population is a combination of random individual variation, and systematic differences between groups. For example, the mean IPD of 63.4mm represents a combination of a mean of 64.67mm in men and 62.31mm in women [4]. This study also reported some differences in IPD across different ethnic groups. The factor that most influences IPD however is age. The average IPD increases from 50mm at the age of 5, through 53mm at 8, 59mm at 13 and only reaches the final adult value the age of 19 [15]. It is clear from these values that VR headsets are made with adult users in mind. In fact, it is typically recommended that headsets are not used by people under the age of 14. While there is no clear evidence of adverse effects of VR for children, this restriction is typically seen to as a sensible

precautionary approach, in order not to avoid influencing the development of the immature visual system [21]. This does however significantly restrict the potential for the use of VR in many educational, cultural, and entertainment applications. The current inability of headsets to accommodate the IPD of children is a significant limiting factor in this context.

In the current study, we evaluate the implications of the variability of IPD across the population on the quality of 3D experience in virtual reality, and other 3D displays. We investigate specifically the limited distance range across which it is possible to render scenes with plausible (convergent) viewing, predicted biases in the perception of distance, if IPD is not well-matched to the user's IPD, and the comfortable range of distances that can be presented. These calculations show that, while IPD is an important consideration, due to the presence of other depth cues, the implications of a mismatched IPD are much less severe than implied by a consideration of binocular viewing geometry alone.

2. MAXIMUM DISTANCE RANGE

The convergence angle of the two eyes when fixating a target decreases with distance, meaning that vergence can be used as a distance cue [18]. This is illustrated in figure 1a for a simple case of viewing a target directly in front of the observer. In this case, the vergence angle is given by:

$$\tan \frac{\theta}{2} = \frac{I}{2D} \quad (1)$$

Figure 1b also shows how a target can be presented on a 3D display to create this vergence angle, and thus to simulate this target distance. For all distances, the two eyes need to converge in order to fixate the target. If the observer's IPD is smaller than that assumed when creating the stimuli, then far targets would require divergent eye movements. We can calculate the maximum distance that can be rendered without causing divergence. This is illustrated in figure 1c and given by:

$$D_{max} = S \frac{H_R}{H_R - H_O} \quad (2)$$

Where H_R is half the IPD assumed when rendering the disparities, H_O is half the observers actual IPD, and S is the focal distance. Maximum values are plotted for a range of IPDs, for a range of focal distances between 1.25m and 2.5m, in figure 1d, and an assumed IPD of 63mm. Although in a VR the screen is physically just a few centimetres from the eyes, the lenses in the headset mean that the image distance is in this range¹. These calculations are not specific to VR headsets, but are valid for any 3D display.

¹<https://docs.microsoft.com/en-us/windows/mixed-reality/design/comfort>

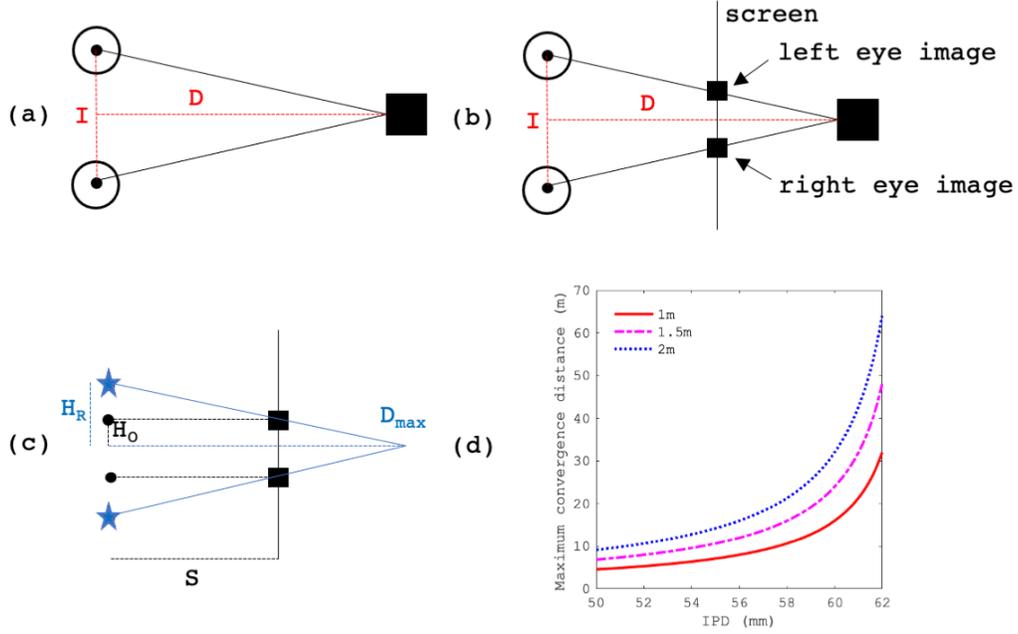


Fig. 1. (a) When fixating an object, the convergence angle depends on the observer’s IPD I and the distance to the object D . As the distance increases, vergence angle decreases. (b) On a binocular display, the apparent distance of an object is simulated by presenting the image for the left and right eyes at different locations, to produce the desired vergence when the object is fixated. (c) If the observer’s IPD (indicated by the black circles) is smaller than the assumed IPD (indicated by the blue stars) then there will be a maximum rendered distance D_{max} beyond which fixation will be divergent rather than convergent. H_O and H_R are half the IPD of the observer and that used in rendering, respectively, and S is the focal distance. (d) D_{max} plotted as a function of the observer’s IPD, for three focal distances, and an assumed IPD of 63mm.

3. MISPERCEPTION OF DISTANCE

When the positions of a target on the screen are such that fixation is convergent, its effective distance can be calculated. When there is a mismatch between the actual IPD and that assumed when rendering the images, the relationship between the rendered and predicted perceived distance, based on binocular geometrical considerations, is given by:

$$D_P = D_R \frac{H_O S}{H_R S + D_R (H_O - H_R)} \quad (3)$$

as illustrated in figure 2a. D_P is the perceived distance, D_R the rendered distance, S the focal distances and H_O and H_R half the observer’s and rendered IPD, respectively. This is shown for an assumed IPD of 63mm when rendering the stimuli, and observer IPDs of 50 to 74mm, and a focal distance of 1.5m in figure 2b. These results shows that a large overestimation of distance is predicted for an observer with a small IPD, and a large underestimation for a user with a large

IPD. The slopes of these functions also differ. For an observer with a small IPD, the 63mm value used in rendering will magnify depth compared with natural viewing, creating a steep slope in the function relating presented to perceived distance, and therefore an expectation that sensitivity to changes in distance will increase. Conversely, for an observer with a large IPD, the slope is reduced, predicting a reduced sensitivity to changes in distance, even after training in this environment.

We can also simulate the effects of this IPD mismatch for populations of observers in different age groups. Measures of the mean and standard deviation of IPD for ages of 5, 8, 13 and 9 years were used [15]. We used Monte Carlo simulations to sample 1000 simulated participants in each age group, and calculated the mean and standard deviation of the predicted perceived distance across this simulated population, for distances between 0.5 and 3m. These are plotted in figure 2c. For younger observers, distance is predicted to be progressively overestimated as the rendered distance increases beyond the focal distance. The variability in perceived dis-

Table 1. The range of IPD adjustment of three example consumer VR headsets. In each case, the minimum, central and maximum range are given. Estimates of the percentage of adult populations with an IPD outside of this range, and the average percentage mismatch between the central value and a user’s actual IPD, are also tabulated.

Headset	Min	Max	Middle	% Outside of Range	Average Error
HTC Cosmos Elite	59	72	65.5	14	5.76
Oculus Quest 2	56	70	63	6.9	4.87
Playstation VR	57	71	64	7.2	5.0

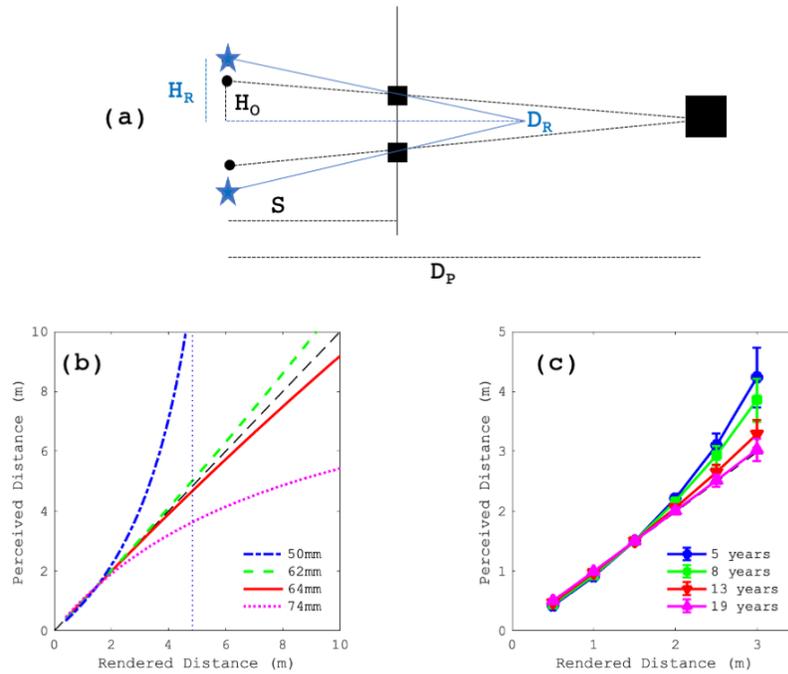


Fig. 2. (a) Errors in distance perception when the observer’s IPD does not match that used to render stimuli. H_O is half the observers IPD. An object at a distance D_R is projected onto a screen at a focal distance S , assuming a half-IPD of H_R . The resulting perceived distance D_P is given in equation 3. The predicted perceived distance as a function of rendered distance for a range of IPDs, with an rendering IPD of 63mm and focal distance of 1.5m. The dotted blue vertical line shows the maximum distance limit for an IPD of 50mm. (c) Predicted perceived distance as a function of rendered distance for four age groups, again with a rendering IPD of 63mm and focal distance of 1.5m. Data points show the mean prediction across 1000 simulated observers, and error bars ± 1 standard deviation.

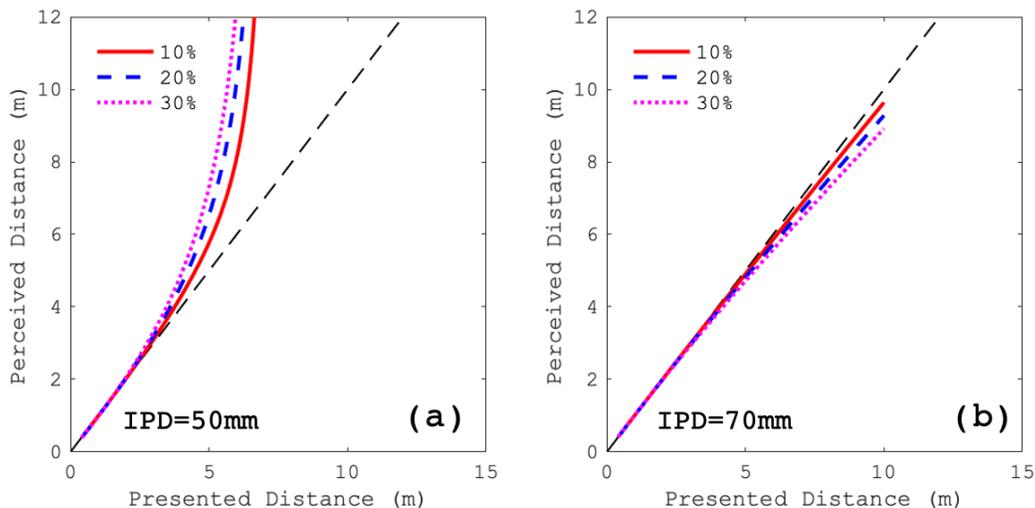


Fig. 3. Predicted perceived distance for three different weights of binocular contribution to estimated depth, with an assumed IPD of 63mm and focal distance of 1.5m. Results are plotted for observer IPDs of (a) 50mm and (b) 70mm.

tance is also greater for younger observers, and with greater simulated distance.

4. PURELY BINOCULAR CONSIDERATIONS WILL OVERESTIMATE THE IMPACT OF MISMATCHED IPD

The above calculations assume that apparent distance is determined purely by binocular cues. In practice, a broad range of cues are available to inform our distance judgements, including motion parallax, perspective, texture, height in the visual field and familiar size [3]. Since these cues are not affected by IPD, the estimated impact of a mismatch in this parameter are likely to be significantly overestimated. This can be quantified if we assume that perceived distance is a weighted average of the available cues [14]. If the distance estimated from binocular cues (D_B) is combined with the distance estimated from all other cues (D_O) then the predicted perceived distance, in the case of independent cues, is given by:

$$D_P = w_B D_B + (1 - w_B) D_O \quad (4)$$

where w_B is the weight assigned to binocular cues. The effect of this is to downweight the biases caused by an incorrect IPD. It should be noted that these calculations assume that the cues are weighted regardless of that fact that they provide strongly conflicting information [19, 24]. It has been suggested that the weight assigned to the discrepant cue may be reduced in this case [14], and bistable depth perception has been found when disparity is in conflict with perspective information [27]. Figure 3 plots the expected errors in perceived distance for a small (50mm) and large (75mm) IPD, with binocular cues contributing 10%, 20% or 30% to the

estimated distance. As this weight is reduced, these biases decrease.

5. ZONE OF COMFORTABLE DISPARITIES

Conflict between accommodation and vergence is a source of discomfort in 3D displays [9, 13, 25]. The range of disparities that can comfortably be presented has been estimated to be between 1° [13] and 2° [25]. For a given focal distance, we can calculate the nearest and furthest comfortable distance than can be presented, and from this estimate a depth budget, or the total range of distances at which objects can be presented. To do this, we calculated the convergence angle at the fixation distance, and the nearest and furthest comfortable convergence angles as $\pm 2^\circ$ from this. We then calculated the rendered distance, with a potentially different IPD, that would create the screen disparities corresponding to these angles. The depth budget was then calculated as the difference between the far and near cut-off distances for this zone of comfort. Figure 4 shows this depth budget as a function of IPD, assuming a comfortable range of disparities of $\pm 2^\circ$ and a focal distance of 1.5m. The black line plots the budget for the case that the rendered IPD is matched to that of the observer, and shows that this range will increase with decreasing IPD. The calculations can also be performed assuming an IPD that is not matched to the observer's. The magenta line shows the budget when the IPD used in rendering is fixed at the average of 63mm. This will tend to decrease the depth budget for observers with a small IPD, and increase it for observers with a large IPD. This is because in this case the depth budget is determined by the IPD used to render the stimuli, not by the observer's own IPD.

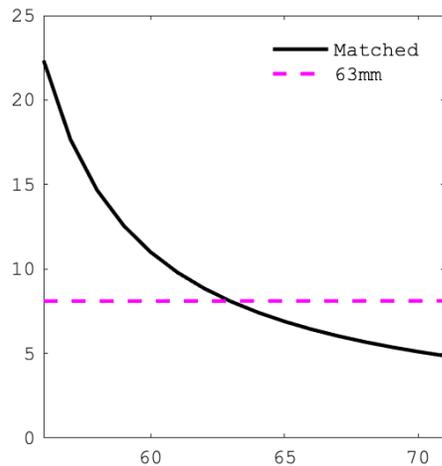


Fig. 4. Available depth budget, for a focal distance of 1.5m and a comfortable disparity range of $\pm 2^\circ$, as a function of IPD. Data are plotted for the case in which the IPD used for rendering is (1) matched to that of the observers (black line) and (2) fixed at the average of the population (magenta line). The range of IPDs plotted covers 95% of the adult population.

6. DISCUSSION

We identified four potential problems that arise in virtual reality when the headset IPD is not well matched with the observer’s IPD: divergent fixation, biases in the perception of depth, visual discomfort and off-centre viewing. We outline how divergent viewing and discomfort are a particular concern for observers with a small IPD, particularly children. A small IPD will also tend to result in an overestimation of distances, whereas a larger IPD will result in underestimation. However, because binocular vergence and disparity are not the only cues to distance in typical scenes, the biases predicted from simple geometrical calculations will tend to overestimate the effect on an unmatched IPD.

Nevertheless, these problems will have a differential effect on different participant groups. For example, IPD is on average smaller for women than men, contributing to the notion that current headset design might be sexist [20, 6]. As IPD increases with age, the small IPD of children will be a significant barrier to the usability of VR with children. The potential for divergent fixation, and large conflicts between binocular and other cues, are important considerations for their possible effects on the developing visual system.

In aiming to accurately replicate the user’s IPD, there is an assumption that the goal is to precisely recreate the visual cues in the rendered scenes. In the case of 3D displays, it has been argued that there may be situations where this should not be the desired outcome. For example, in an approach known as ‘microstereopsis’, a very small effective IPD is deliberately

used [26, 17]. The intention here is to introduce just enough binocular disparity to create a good stereoscopic effect, while minimising the conflict between disparity and accommodative cues [25]. Alternatively, the use of a larger effective IPD can be used as an effective way of magnifying depth information to increase sensitivity [2]. This increase in sensitivity would however create a bias in the way that disparity is scaled and combined with other cues, which would require adaptation if accurate scaled depth estimation or visuomotor control are required [28], as is typical in virtual reality.

In summary, we have identified viewing comfort, accurate depth perception and optical distortions as factors affected by considerations of IPD in virtual reality. These are an important focus for empirical research for understanding the accessibility of virtual reality, particularly for children.

7. REFERENCES

- [1] Ryan Beams, Andrea S Kim, and Aldo Badano. Transverse chromatic aberration in virtual reality head-mounted displays. *Optics express*, 27(18):24877–24884, 2019.
- [2] Mark F Bradshaw, Paul B Hibbard, Rob van der Willigen, Simon J Watt, Ian RL Davies, Neil S Stringer, Nick Beagley, and Andrew R Willis. Can observers exploit enhanced-disparity information to control reaching movements within telepresence environments? In *Stereoscopic Displays and Virtual Reality Systems VII*, volume 3957, pages 253–261. International Society for Optics and Photonics, 2000.
- [3] James E Cutting and Peter M Vishton. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *Perception of space and motion*, pages 69–117. Elsevier, 1995.
- [4] Neil A Dodgson. Variation and extrema of human interpupillary distance. In *Stereoscopic Displays and Virtual Reality Systems XI*, volume 5291, pages 36–46. International Society for Optics and Photonics, 2004.
- [5] Stuart J Gilson, Andrew W Fitzgibbon, and Andrew Glennerster. Quantitative analysis of accuracy of an inertial/acoustic 6dof tracking system in motion. *Journal of neuroscience methods*, 154(1-2):175–182, 2006.
- [6] Simone Grassini and Karin Laumann. Are modern head-mounted displays sexist? a systematic review on gender differences in hmd-mediated virtual reality. *Frontiers in Psychology*, 11, 2020.
- [7] Paul B Hibbard, Alice E Haines, and Rebecca L Hornsey. Magnitude, precision, and realism of depth perception in stereoscopic vision. *Cognitive Research: Principles and Implications*, 2(1):1–11, 2017.

- [8] Paul B Hibbard and Rebecca L Hornsey. Contributions of pictorial and binocular cues to the perception of distance in virtual reality. *Submitted*.
- [9] David M Hoffman, Ahna R Girshick, Kurt Akeley, and Martin S Banks. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, 8(3):33–33, 2008.
- [10] Wijnand A IJsselsteijn, Huib de Ridder, and Roelof Hamberg. Perceptual factors in stereoscopic displays: the effect of stereoscopic filming parameters on perceived quality and reported eyestrain. In *Human Vision and Electronic Imaging III*, volume 3299, pages 282–291. International Society for Optics and Photonics, 1998.
- [11] Wijnand A IJsselsteijn, Huib de Ridder, and Joyce Vliegen. Subjective evaluation of stereoscopic images: effects of camera parameters and display duration. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(2):225–233, 2000.
- [12] Marc Lambooi, Wijnand IJsselsteijn, Don G Bouwhuis, and Ingrid Heynderickx. Evaluation of stereoscopic images: Beyond 2d quality. *IEEE transactions on broadcasting*, 57(2):432–444, 2011.
- [13] Marc TM Lambooi, Wijnand A IJsselsteijn, and Ingrid Heynderickx. Visual discomfort in stereoscopic displays: a review. In *Stereoscopic Displays and Virtual Reality Systems XIV*, volume 6490, page 64900I. International Society for Optics and Photonics, 2007.
- [14] Michael S Landy, Laurence T Maloney, Elizabeth B Johnston, and Mark Young. Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision research*, 35(3):389–412, 1995.
- [15] Colleen MacLachlan and Howard C Howland. Normal values and standard deviations for pupil diameter and interpupillary distance in subjects aged 1 month to 19 years. *Ophthalmic and Physiological Optics*, 22(3):175–182, 2002.
- [16] JEW Mayhew and HC Longuet-Higgins. A computational model of binocular depth perception. *Nature*, 297(5865):376–378, 1982.
- [17] John P McIntire, Paul R Havig, Lawrence K Harrington, Steve T Wright, Scott NJ Watamaniuk, and Eric Heft. Microstereopsis is good, but orthostereopsis is better: Precision alignment task performance and viewer discomfort with a stereoscopic 3d display. In *Three-Dimensional Imaging, Visualization, and Display 2018*, volume 10666, page 106660Q. International Society for Optics and Photonics, 2018.
- [18] Mark Mon-Williams, James R Tresilian, and Andrew Roberts. Vergence provides veridical depth perception from horizontal retinal image disparities. *Experimental brain research*, 133(3):407–413, 2000.
- [19] Chris MP Muller, Eli Brenner, and Jeroen BJ Smeets. Testing a counter-intuitive prediction of optimal cue combination. *Vision Research*, 49(1):134–139, 2009.
- [20] Justin Munafo, Meg Diedrick, and Thomas A Stoffregen. The virtual reality head-mounted display oculus rift induces motion sickness and is sexist in its effects. *Experimental brain research*, 235(3):889–901, 2017.
- [21] Simon K Rushton and Patricia M Riddell. Developing visual systems and exposure to virtual reality and stereo displays: some concerns and speculations about the demands on accommodation and vergence. *Applied Ergonomics*, 30(1):69–78, 1999.
- [22] Peter Scarfe and Andrew Glennerster. Using high-fidelity virtual reality to study perception in freely moving observers. *Journal of vision*, 15(9):3–3, 2015.
- [23] Peter Scarfe and Andrew Glennerster. The science behind virtual reality displays. *Annual review of vision science*, 5:529–547, 2019.
- [24] Peter Scarfe and Paul B Hibbard. Statistically optimal integration of biased sensory estimates. *Journal of vision*, 11(7):12–12, 2011.
- [25] Takashi Shibata, Joohwan Kim, David M Hoffman, and Martin S Banks. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of vision*, 11(8):11–11, 2011.
- [26] Mel Siegel and Shojiro Nagata. Just enough reality: comfortable 3-d viewing via microstereopsis. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(3):387–396, 2000.
- [27] Raymond Van Ee, Loes CJ Van Dam, and Casper J Erkelens. Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slants. *Journal of Vision*, 2(9):2–2, 2002.
- [28] Robert Volcic, Carlo Fantoni, Corrado Caudek, John A Assad, and Fulvio Domini. Visuomotor adaptation changes stereoscopic depth perception and tactile discrimination. *Journal of Neuroscience*, 33(43):17081–17088, 2013.