ORIGINAL ARTICLE

Perception-based high quality distributed virtual reality

William Naylor[1] · Kurt Debattista[1] · Alan Chalmers[1]

Abstract

Virtual reality has great potential to enable remote collaborative work from anywhere in the world. Developing virtual reality into a platform suitable for natural interaction and immersive collaboration requires the experience to be reliably stable. For a networked collaborative environment, perceived smoothness of motion is limited by the tick rate, that is, the frequency at which information is distributed. As tick rate increases, motion will appear increasingly smooth; however, excessive tick rates may introduce additional load on a network without any perceptible benefit to a user. This paper details two visual psychophysics experiments ($N_1 = 16$, $N_2 = 11$) carried out to evaluate participant sensitivity to tick rate in virtual reality. The influence of three variables, velocity, complexity, and digital medium were investigated. Both velocity and digital medium displayed a significant effect, whilst complexity did not show significance. A model was then built and validated from the results of these experiments. The model predicts for average walking speed within the desktop condition, that 90% of the population will perceive motion to be smooth at 56 Hz, whilst this 90% threshold lies at 113 Hz for the VR condition. This model can predict participant perception of tick rate under given conditions, enabling networks to intelligently optimise participant experience without adding unnecessary further load on the network.

Keywords Virtual reality · Networks · Tick rate · Psychophysics

1 Introduction

Virtual reality (VR) has been proven to be an efficient tool for remote collaboration and education, with experiments exploring its applicability in manufacturing (Dixken et al. 2019; Herder et al. 2019), surgery (Weibel et al. 2020; Chheang et al. 2019), fire safety training (Ha et al. 2016), and school field trips (Zhao et al. 2020). These networked virtual environments inspire immersive and natural face-to-face interaction (Aseeri and Interrante 2021; Roth et al. 2016; Dzardanova et al. 2022); however, this is only attainable with a stable and efficient network (Elbamby et al. 2018).

Tick rate, the frequency at which information is distributed to all participants involved, is an essential component within any network (Parthasarathy et al. 2020). If the tick rate is too low the application will feel unresponsive and jerky. On the other hand, if the tick rate is unnecessarily high, it threatens to overload the network's bandwidth, running the risk of overwhelming the system. This could result in the unfortunate side-effect of introducing jitter or latency spikes, which impairs participant experience and potentially induces cybersickness when using VR (Stauffert et al. 2018). Therefore, a balance needs to be struck that provides the required fluidity for interaction without unnecessary data transmission.

This balance must also consider the high frame rates that VR applications require in order to reduce cybersickness (Kim et al. 2017; Hecht 2016; Brennesholtz 2018). Under optimal circumstances, tick rate should be equal to the frame rate to ensure an update every frame, however this would put a great load upon the network. Therefore, through this research, we may establish whether tick rate can be decoupled from frame rate, and thus operate at sub-frame rate frequencies without affecting a participant's experience.

This paper investigates this balance between network load and satisfaction by evaluating human perception thresholds for fluid tick rates. A visual psychophysics experiment was conducted to establish perceptual thresholds for *smoothness*

✉ William Naylor
W.Naylor@warwick.ac.uk

Kurt Debattista
K.Debattista@warwick.ac.uk

Alan Chalmers
Alan.Chalmers@warwick.ac.uk

1 WMG, University of Warwick, Coventry CV4 7AL, UK

of network tick rate in VR and comparatively to a computer desktop visualisation. An initial exploratory experiment was undertaken to determine the variables that have an effect on perception. A secondary experiment was then conducted to investigate all variables of significance as within-participant variables and their effects to be fully evaluated. The objectives were to establish the psychometric functions defining smoothness perception within VR and to use these to develop a predictive model which could enable networked collaborative VR systems to intelligently optimise a participant's experience. Through dynamically altering network tick rate in accordance with conditions within a scene, the system prevents expending excessive resources on imperceptible benefits or cutting too much that would negatively affect the experience. Figure 1 depicts an example application of such a system, in which the tick rate is adjusted relative to the connected client's device and bandwidth, ensuring quality remains within accepted thresholds of perceptibility.

2 Related work

Although there has been little previous work that investigated the influence of tick rate on human perception, similarities can be drawn between tick rate and frame rate evaluations, which have been studied more often in the past. Tick rate and frame rate both involve still images updating at a fixed frequency, however tick rate enables single objects within a virtual environment to update at sub-frame rate frequencies whilst maintaining rendered performance for the rest of the scene.

Claypool and Claypool (2009) investigated the impact of frame rate and resolution on task performance in a desktop setting. They concluded that as frame rates increased (5–15–30 Hz) the precision of actions improved and one's ability to navigate around a space was also positively affected. Improved resolution similarly improved precision and navigational ability, but to a less significant degree.

Zielinski et al. (2015) came to similar conclusions for task performance in VR. Participants were required to undertake a target acquisition task at a low frame rate of 11 Hz and at a higher frame rate of 55 Hz. The results showed that the performance improved with higher frame
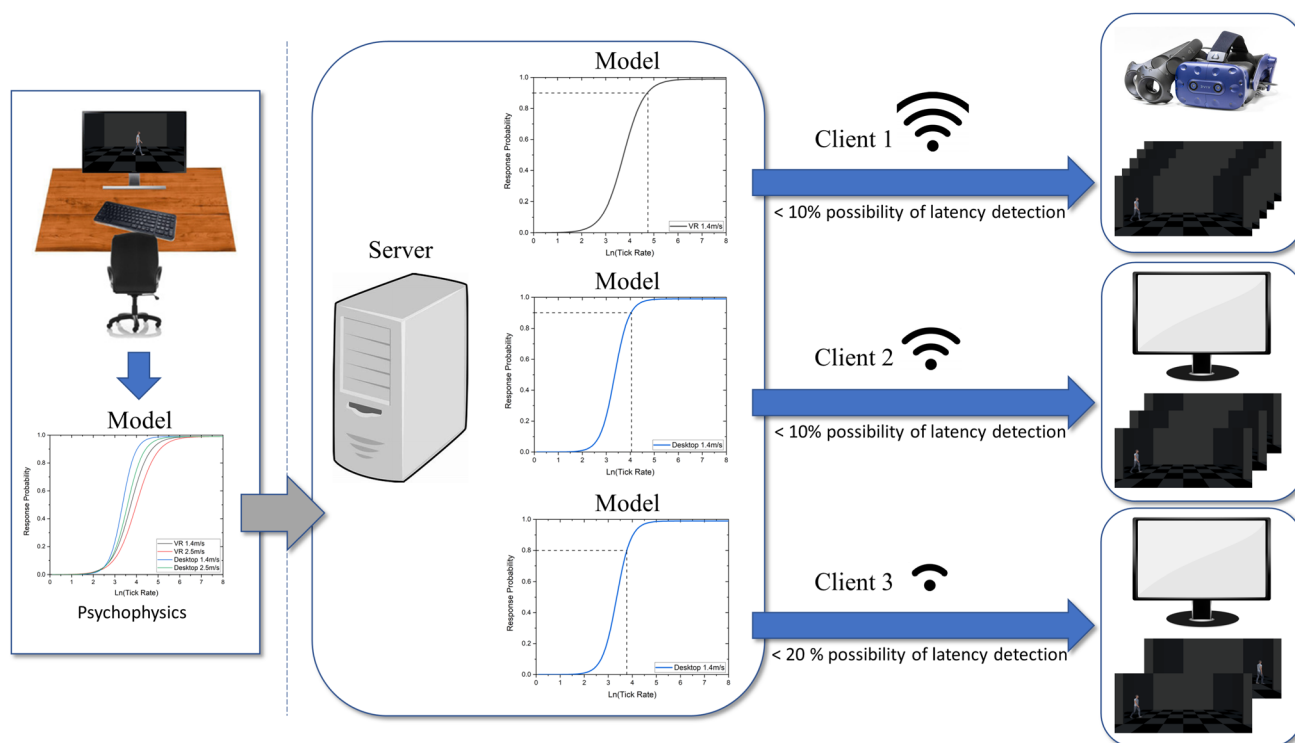

Fig. 1 Framework Diagram depicting the creation and application of the model. A series of psychophysics experiments were conducted to develop a model that describes the influence of three variables upon the psychometric function. The model enables the tick rate of a network to be adjusted based upon the connected client and the conditions within the virtual environment, without any detriment to the experience of either client. The illustrated examples shows how servers can adjust the quality within desired thresholds based on the available bandwidth for available clients

rates. Their study agrees with an earlier target acquisition task experiment by Ware and Balakrishnan (1994). Performance would likely continue to improve for frame rates greater than 55 Hz; however, the magnitude of improvement is likely to diminish, as Denes et al. (2020). found that the perceived improvement in the quality of animations stalled for frame rates above 100 Hz in computer desktop environment. No such experiment at these higher frequencies has been conducted in VR, but potentially a similar trend may occur where the increase in quality of task performance also plateaus.

Contradicting results can be found in the study of stimulus velocities. Sensitivity to motion has been reported to increase for higher velocities (Denes et al. 2020; McDonnell et al. 2007) but also decrease for higher velocities (Hulusić et al. 2011). Denes et al. found sensitivity to motion increased for higher velocities, furthermore they found that sensitivity to motion increased if the trajectory of the moving object was predictable. However, Hulusić et al. (2011) concluded that higher velocities reduced sensitivity to framerate. Hulusić et al. also established that multi-sensory stimuli affected perception of frame rate. Individuals were required to compare the smoothness of the movement and when other sensory factors such as the sound of footsteps were introduced, participants perceived the movement at lower frame rates as more acceptable.

DoVale (2017) studied the just noticeable difference (JND) for three different frame rates, determining the threshold for a 24 Hz frame rate to be between 26 Hz and 28 Hz. JND thresholds for higher frame rates were significantly larger than for the 24 Hz condition, with the JND for 48 Hz at 62 Hz, and 72 Hz producing uncertainty in identification at frame rates as high as 120 Hz.

McDonnell et al. (2007) investigated the thresholds of perceived smoothness for pose update rate in animations on a two-dimensional monitor. Character type and scene complexity had no influence on perceived smoothness, but thresholds increased for higher linear velocities and intensity of movement. When the number of individuals moving increased, sensitivity decreased.

Latoschik et al. (2019) evaluated the performance of a distributed VR environment based upon the number of connected clients. Performance of the networks began to diminish for groups larger than 25, and the server update frequency fell from 120 to 6 Hz for 125 clients. A follow-up subjective experiment affirmed these results, with perceived fluidity and synchrony significantly reducing for crowds larger than 50, increasing dissatisfaction.

The only study dedicated specifically to tick rate was conducted by Lee and Chang (2015), who evaluated the impact of tick rate upon accuracy within an FPS game. They showed that accuracy significantly increased for the higher tick rate of 128 Hz in comparison with 64 Hz.

3 Methodology

The overarching motivation of this work is to improve network efficiency within interactive environments, including VR, without unnecessarily compromising on visual fidelity or the immersive experience. We investigate human perception to various visual conditions, in order to establish thresholds for a given condition so that the bandwidth consumed can be minimised without significant losses in perceived visual fidelity. To help identify these thresholds a psychophysical experiment was conducted.

Psychophysical experiments seek to identify a Psychometric Function (PF) which describes the response from the human visual system when presented with a stimulus (Treutwein 1999). As the stimulus intensity increases, the proportion of trials evoking a positive response will increase in line with the PF. For these experiments, the stimulus intensity will be the tick rate, measured in Hertz (Hz). A logistic function was chosen for the PF, as it can operate with log-transformed variables and the derivative is symmetric around the threshold. This will simplify the construction of a predictive model later on; however, with the data collected, other sigmoidal functions could easily be fitted with the same methodology. Equation 1 gives the generalised form of the PF(Strasburger 2001; Treutwein 1995)

$$\Psi(x \mid \alpha\beta\gamma\lambda) = \gamma + (1 - \gamma - \lambda)\frac{1}{1 + e^{-\beta(x-\alpha)}}. \quad (1)$$

The tick rate is denoted by x. α corresponds to the threshold of the PF, the point at which responses will be positive 50% of the time. β represents the slope of the PF, influencing the gradient of the PF at the threshold. γ is the guess rate, the lower bound of the PF as tick rate reaches its minimum. Finally, λ represents the lapse rate, the result of responses independent of the stimulus, such as misclicks or distraction (Wichmann and Hill 2001).

The PF could hypothetically be affected by a large number of variables, from the resolution and dynamic range of the display to the importance of the stimulus within the scene, and in general, the four PF parameters can be described as a function of variables $p_1, \dots, p_n$ (Debattista et al. 2018)

$$\Psi(x \mid \alpha\beta\gamma\lambda) : \alpha(p_1, \dots, p_n),$$
$$\beta(p_1, \dots, p_n), \gamma(p_1, \dots, p_n), \lambda(p_1, \dots, p_n). \quad (2)$$

Due to the nature of this experiment investigating tick rates, γ and λ can be constrained to reduce the complexity of the problem. As the tick rate reduces to 0, the motion will always appear jittery, so γ can be constrained to 0. Lapse rate is usually fixed at a small nonzero value for psychophysical experiments (e.g. 0.01) and setting equal to zero may introduce significant bias for threshold and slope (Swanson and Birch 1992), so λ will be restricted to 0.01. With these

constraints, Eq. (2) can be simplified to a function depending only upon α and β,

$$\Psi(x \mid \alpha\beta) = (1 - 0.01)\frac{1}{1 + e^{-\beta(x-\alpha)}} : \alpha(p_1, \ldots, p_n), \beta(p_1, \ldots, p_n).$$

(3)

If all variables are known, this PF could be calculated at run-time to dynamically alter the tick rate depending upon the conditions within the scene. In order to reduce the complexity of this equation, this paper will explore three key variables which potentially influence the visual experience. The three variables explored are $p_1 = medium$, $p_2 = velocity$, and $p_3 = complexity$. Once restricted to these three variables, α and β can be described as a function $\alpha(medium, velocity, complexity)$ and $\beta(medium, velocity, complexity)$.

3.1 Motivation for choice of variables

Human perception was investigated in relation to three variables.

The first variable was the digital platform itself. Participants were shown the same stimulus in both VR and on a computer desktop to establish potential differences in perception between the two digital mediums. This variable will hereafter be referred to as *medium*. This is a necessary variable to evaluate whether the increased immersion (Skarbez et al. 2018; Slater 2018) experienced in the VR platform will change the sensitivity to tick rate (Shu et al. 2019; Cao et al. 2021).

The second variable is the visual complexity of the stimulus, referred to here as *complexity*. Complexity of objects significantly affects bandwidth requirements, as more data would be required to update the object on each tick. Therefore, understanding the influence of *complexity* due to increased optical flow (Horn and Schunck 1981) upon human perception is essential to maximising network efficiency without visual detriment.

The third variable we considered is the velocity of the stimulus, denoted throughout as *velocity*. Previous psychophysical experiments have evaluated *velocity* on a computer desktop and have found it to have a significant effect; however, the exact influence of *velocity* has been inconsistent. Hulusić et al. (2011) found perceived smoothness increased as velocities increased, whereas McDonnell et al. (2007) and Denes et al. (2020) found lower velocity stimuli were perceived as smoother. Therefore it is worthwhile to explore *velocity* in VR for it was likely to also display an effect.

These three variables lead to three hypotheses:

H1 Perception thresholds will be higher in VR.

H2 Complex objects will require higher tick rates than simple objects.

H3 Stimulus velocity will affect participant sensitivity.

3.2 Overview of experiments

The predictive models are developed over the course of two experiments. The first experiment explores the influence of all three variables, *medium*, *velocity*, and *complexity*, to establish the effect each may have upon a PF. The second experiment refines this evaluation to explore only the variables that displayed significance in the first experiment. From the results of the second experiment, predictive models are developed and validated. Finally, the results and observations from these experiments are discussed and suggestions for future development are provided.

4 Experiment 1: broad evaluation

Throughout both experiments, a 1 Alternative Forced Choice (1AFC) psychophysical methodology was employed (Kingdom and Prins 2016). This method enabled a full PF to be developed for every participant for each combination of conditions. Therefore, the perception of smooth motion for individuals under each of the three variables at a range of tick rates could be explored, as participant sensitivity influences the shape of the PF.

4.1 Design

The first experiment followed a $2 \times 2 \times 2$ factorial design, with *complexity* and *medium* as within-participant variables, and *velocity* as a between-participant variable. *Velocity* was operated as a between-participant variable to halve the number of trials required for each participant, to reduce the potential error and drift in results from fatigue. For each combination of conditions, the corresponding PF was investigated through three concurrent Weighted Up/Down adaptive procedures targeting the 25%, 50%, and 75% thresholds (Kaernbach 1991). The 50% threshold was targeted through a step size ratio $\Delta^-/\Delta^+ = 1/1$ procedure, whilst the 25% and 75% were targeted through $\Delta^-/\Delta^+ = 3/1$ and $1/3$, respectively. Three thresholds were required to estimate the full shape of the PF, as a single Up/Down method cannot be used to estimate the slope. Each Up/Down procedure was run for 40 trials, with the first 15 trials excluded from the analysis to allow the staircase to find equilibrium. Weighted Up/Down programs with 1/3 step size ratios reach 10 reversals, a common ending condition, after approximately 40 trials (García-Pérez 1998). A 1/1 rule was employed with larger step sizes until the first reversal occurred to accelerate the procedure

to reaching equilibrium (Kingdom and Prins 2016). All Up/Down staircases were run concurrently and presented in a randomised order (Bechara et al. 1997). As visual responses are broadly logarithmic (Varshney and Sun 2013), the Up/Down staircases operated on tick rates transformed by the natural logarithm, and step sizes were performed in increments of 0.05 log-units. All Up/Down staircases began at 4.1 log-units (60.34 Hz). Each participant performed 480 trials and the experiment lasted approximately 40 min.

4.2 Participants

Sixteen individuals volunteered for the first experiment, of which four were female and twelve were male. The participants' age varied from 20 to 83, with an average of 40 ± 18. All involved possessed normal or corrected to normal vision.

4.3 Materials

The experiment was performed in a darkened, empty room to reduce external distractions. For the VR condition, an HTC Vive Pro HMD (1440 × 1600p 90 Hz) was used and Vive Wand Controllers for hand tracking. For the desktop condition, an Acer XB270HU G-Sync enabled monitor (2560 × 1440p 144 Hz) was used. All experiments were run from a PC with an Intel Xeon E5 2690 CPU, 32GB of memory and a NVIDIA GeForce Titan Xp. Temporal Anti-Aliasing was applied throughout all trials, in order to reduce flickering which may have otherwise affected responses(Jimenez et al. 2011; Almeida et al. 2019). All other post-processing techniques were disabled.

All trials were performed using a plain darkened textureless scene, constructed in Unity version 2021.2.8f1, to reduce external stimuli. The stimulus was positioned 6.25 m in front of the participant and would travel a straight path perpendicular to the facing direction. Other paths were explored in the pilot study and a straight path was found to be most natural. Similar to a theatre stage, the displayed stimulus would travel from one side of the scene to the other, appearing from behind one wall before disappearing

behind the other, repeating the same path until a response was provided by the participant. After receiving an answer, an inter-trial interval of 2 s was shown in which no stimulus was present. The direction of motion was randomised for all trials to negate any potential directional biases. Likewise, the controls for responding true or false were randomised, though kept constant for each participant to avoid confusion and to reduce the quantity of lapse results due to misclicks.

Stimuli would travel past the participant at two distinct velocities, the low *velocity* was selected to be 1.4 m/s as this is equal to average human walking speed (Mohler et al. 2007) and a higher *velocity* of 2.5 m/s was chosen as the pilot study suggested it appeared significantly faster than the alternative stimulus without appearing unnatural and negatively influencing immersion throughout the experiment. Both of these velocities lie within the band of greatest sensitivity found by Orban et al. (1984) from their experiment exploring velocity differential detection.

Two different stimuli were presented to participants, in order to investigate the influence of visual noise from a more complex moving stimulus. The simple stimulus was a dark grey, 0.5m radius sphere without textures, meanwhile the complex stimulus was a photogrammetry of a human, purchased from the RenderPeople asset store. The two stimuli are depicted in Fig. 2. The human model was rigged to use a walking animation for the low velocity and a jogging animation for the high velocity.

4.4 Procedure

Participants were informed about the proceedings of the experiment, after which they gave informed consent. After providing their age, gender, and previous VR experience at their own discretion, they were randomly assigned to perform the VR or desktop task first. Before the experiment began, two examples were shown, firstly a stimulus with a low tick rate (12 Hz) to demonstrate the noticeable jitter within the movement. Secondly, a high tick rate example (200 Hz) was presented to explain how motion will appear smooth at sufficiently high tick rates. Following the two

Fig. 2 Two different stimuli were presented in Experiment 1. Left: The human stimulus Right: The sphere stimulus

Table 1 Descriptive statistics for Experiment 1

Variable	α Mean	α SD	β Mean	β SD
Desktop	3.756	0.113	4.781	1.248
VR	4.157	0.131	4.072	1.578
Sphere	3.968	0.098	4.426	1.474
Human	3.945	0.146	4.426	1.352
Velocity low	4.176	0.140	4.325	1.482
Velocity high	3.736	0.104	4.528	1.344

examples, participants were given an opportunity to ask questions before the main experiment began, after which no interaction with the participant occurred until the experiment concluded. After the completion of the first half of the experiment, participants took a 10 min break before continuing with the alternate VR or desktop task.

4.5 Results

Each participant's 25%, 50%, and 75% measurements were fitted through a Levenberg-Marquardt fitting algorithm (Levenberg 1944) to determine their PF for each of the eight combinations of *medium × complexity × velocity*. The PFs are reported as threshold α and slope β, and descriptive statistics are provided in Table 1.

From the calculated set of thresholds and slopes, a Shapiro-Wilk normality test (Shapiro and Wilk 1965) was performed on all the separate combinations of variables and normality was found for all cases.

A three-way mixed ANOVA was performed on the thresholds and slopes for all participants. For the main effect of *medium*, sensitivity to tick rate was significantly higher within VR than on a desktop monitor and the effect size calculated by partial omega squared analysis displayed a large effect ($p < 0.01$, $\omega_p^2 = 0.407$). Reverting the log-transformation, the desktop mean threshold translates to 42.9 Hz, whilst the VR mean threshold is 64.1 Hz. For the main effect of *complexity*, no statistical significance was found ($p = 0.873$, $\omega_p^2 = -0.063$), with mean thresholds at approximately 52 Hz. Negative effect sizes were obtained (Okada 2017) and the mean difference between the two conditions was 0.024 whilst the 95% confidence interval was (-0.153, 0.201), therefore the null hypothesis cannot be rejected. For the main effect of *velocity*, significance was displayed ($p = 0.020$, $\omega_p^2 = 0.223$), with sensitivity decreasing for increased velocities, from 65.4 Hz at 1.4 m/s to 42.1 Hz at 2.5 m/s. No statistical significance was found for slopes under any conditions. Similarly, no significance was found for any cross-variable interaction for thresholds or slopes. Analysis on the participants' data found there to be no significant difference ($p > 0.05$) for age, gender, or previous VR experience.

Tukey post-hoc tests for means comparison agreed with the single variable ANOVA results, suggesting that for thresholds, a significant difference was found for *medium* ($p < 0.01$) and *velocity* ($p = 0.020$), whilst no statistical significance was found for *complexity* ($p = 0.785$) or slopes under any conditions.

5 Experiment 2: refinement

Since *complexity* did not display any significance in the first experiment, in the second experiment it was disregarded and a spherical stimulus was used for all trials. Therefore, a 2×2 factorial design was employed, where *velocity* and *medium* were both evaluated as within-participant variables. As such the total number of trials remained at 480, but now with both *velocities* presented to every participant.

The same up/down methodology targeting the 25%, 50%, and 75% thresholds as Experiment 1 was utilised. The experimental procedure for the participant was also identical to Experiment 1, with a single stimulus presented per trial and participants were given the same task of identifying whether the movement of the stimulus appeared smooth.

Eleven participants volunteered for the second experiment, none of whom took part in the first experiment. Two were female and nine were male. The average age of the participants was 20.1 years. All had normal or corrected to normal vision.

5.1 Results

For each of the four *medium × velocity* combinations, the same Levenberg-Marquardt fitting procedure was applied as in Experiment 1. Descriptive statistics are provided in Table 2.

To compare the results, two-way repeated measures ANOVA was performed and effect sizes were calculated from partial omega squared. Both *medium* ($p = 0.040$, $\omega_p^2 = 0.406$) and *velocity* ($p < 0.001$, $\omega_p^2 = 0.737$) displayed a significant effect upon the threshold, though no significant interaction effect was found for *medium × velocity* ($p = 0.650$, $\omega_p^2 = -0.094$). Sensitivity increased in VR in comparison with desktop, and similarly higher tick rates

Table 2 Descriptive statistics for Experiment 2

Variable	α Mean	α SD	β Mean	β SD
Desktop	3.463	0.325	7.339	5.565
VR	3.856	0.344	2.670	0.882
Velocity low	3.539	0.326	5.510	5.123
Velocity high	3.780	0.411	4.499	4.049

were required for motion to appear smooth at increasing velocities.

6 Building the model

From the results of the second experiment, a model was developed to describe the variation within the entire sampling group. Because *complexity* was disregarded after the first experiment, the model was defined for *medium* and *velocity* only.

Logistic regression via maximum likelihood estimation is the most common technique for estimating a group PF from a collection of results (Akaike 1998; Prins 2019). However, it did not accurately describe the collective sampling group from these experiments. This is because the method consistently underestimated the slope of participants, resulting in maximum likelihood estimates with slope parameters significantly lower than the measured data, such as $\alpha_{\mathrm{VR\ 2.5\,m/s}} = 3.93$ and $\beta_{\mathrm{VR\ 2.5\,m/s}} = 0.91$, and with $R^2 = 0.037$.

We present an alternative method for combining the collective sampling group based upon the decomposition of means and standard deviations (Altman et al. 2000). The logistic function can be described as a cumulative distribution function and the derivative of this logistic function produces the probability distribution function of the threshold, with a defined standard deviation (Warren et al. 2022). The derivative of Eq. 3 is the logistic distribution scaled by 0.99 due to the lapse rate assumption

$$\frac{\mathrm{d}\Psi}{\mathrm{d}x} = 0.99 \frac{\beta e^{-\beta(x-\alpha)}}{(1 + e^{-\beta(x-\alpha)})^2} \tag{4}$$

and as such the slope β can be converted into a standard deviation σ through the variance of the logistic distribution

$$\sigma^2 = \frac{1}{0.99^2} \frac{\pi^2}{3\beta^2}. \tag{5}$$

By converting the slope into a standard deviation of the threshold, it enables the results of individual participants to be combined through the decomposition of means and standard deviations. The combined threshold and standard deviation can be transformed back into a PF through Eq. 5, thus producing a PF for the whole sampling group.

The merged data for the conditions of *medium* × *velocity* are listed in Table 3 and depicted in Fig. 3. The models for α and β clearly display the results from Table 2, with VR requiring higher thresholds over a desktop and threshold increasing for higher *velocities*.

The Bayesian Information Criterion (BIC) is a popular method for evaluating the suitability of various models attempting to describe a set of psychophysical results (Schwarz 1978). The BIC can be transformed into a Schwarz weight for probability-based optimal model selection (Wagenmakers and Farrell 2004). Four different models were developed and evaluated through the BIC analysis. Model 1 is the null model in which no variables hold any influence. Model 2 and 3 are models where significance is displayed for only *medium* or *velocity*, respectively. Finally, Model 4 is for the *medium* × *velocity* condition where *medium* and *velocity* are both significant.

The Schwarz weights for each model are displayed in Table 4. The model with the highest probability is Model 4 with the *medium* × *velocity* condition ($p = 1.000$). Because Model 4 was selected with the highest Schwarz weight, it can be concluded that both *velocity* and *medium* have a significant effect on sensitivity with respect to tick rate, which is in line with the ANOVA results. The BIC for Model 2 is lower than for Model 3, suggesting *medium* has a stronger
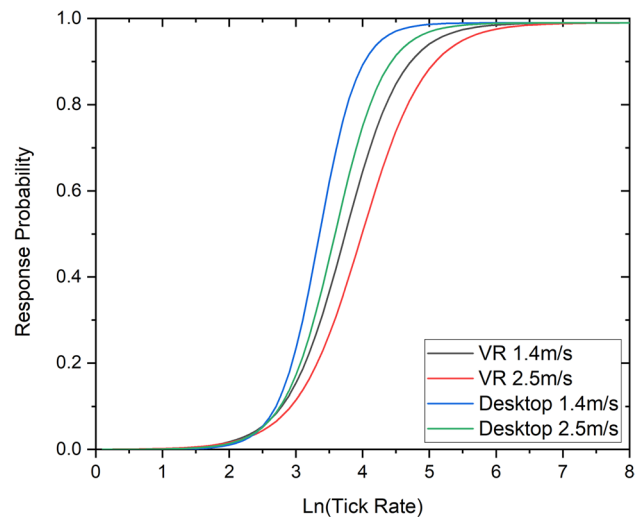

Fig. 3 Predictive models for *medium* × *velocity* conditions

Table 3 Models for *medium* × *velocity* conditions

Variables (m/s)	α	β
Desktop 1.4	3.347	3.388
Desktop 2.5 m/s	3.578	2.697
VR 1.4 m/s	3.731	2.320
VR 2.5	3.981	2.074

Table 4 BIC results from Experiment 2

Model	BIC	Schwarz weight
1. Null	7397	0.000
2. *Medium*	7226	0.000
3. *Velocity*	7320	0.000
4. *Medium* × *velocity*	7176	1.000

influence on sensitivity than for the different velocities. This can also be observed through the α values in the predictive model, where changing the display *medium* evokes a change in the threshold of $\Delta\alpha \approx 0.4$ log-units, whilst changing the *velocity* only induces a shift of $\Delta\alpha \approx 0.25$ log-units.

A simpler, *velocity*-independent model may be easier to implement for most networks, as it removes the requirement to calculate the velocity relative to the viewer, thus the Model 2 parameters are provided here. For the desktop condition, $\alpha_{\text{desktop}} = 3.463$ and $\beta_{\text{desktop}} = 2.932$, whilst for the VR condition, $\alpha_{\text{VR}} = 3.856$ and $\beta_{\text{VR}} = 2.162$.

6.1 Application

Threshold predictions for a defined probability can be obtained from the models through the inverse of Eq. (3),

$$x = \frac{-\ln\left(\frac{0.99}{p} - 1\right)}{\beta} + \alpha, \tag{6}$$

where p is the desired response probability, α and β are the model parameters from Table 3, and x is the log-transformed tick rate. Due to the nature of this function as well as the guess rate and lapse rate assumptions, p is restricted to the range $0 < p < 0.99$. Threshold predictions for a range of common probabilities are stated in Table 5.

6.2 Validation

A k-fold validation was performed to test the accuracy of model predictions and to compare the proposed model to the traditional maximum likelihood estimation algorithm. One-fold or leave-one-out analysis evaluates the validity of a model by removing a single data point and comparing it to a predictive model developed without the removed data point. The method applied here will compare the validity of tick rate predictions through the Wasserstein distance (Wasserstein 1969). The Wasserstein distance compares two curves through the absolute difference integrated over all space and for cumulative distribution functions, it can be defined as $\int_{\mathfrak{R}} | PF_{\text{model}} - PF_{\text{true}} | \, dx$.

This metric was calculated for both the proposed model and the maximum likelihood method and out of the 44 total

Table 5 Threshold predictions based upon a probability of perceived smoothness. All values have the log-transformation reverted and have units of Hz

Prob (%)	Desk. 1.4 m/s	Desk. 2.5 m/s	VR 1.4 m/s	VR 2.5 m/s
50	28.60	36.08	42.08	54.14
75	39.79	54.64	68.17	92.88
90	56.09	81.10	112.55	162.74

data points, 43 were more closely modelled by the proposed model. The proposed model had average Wasserstein distance of 0.353 with a standard deviation of 0.157, whilst the maximum likelihood model had an average Wasserstein distance of 0.694 and a standard deviation of 0.279. This is due to the maximum likelihood model consistently underestimating the slopes of individuals, resulting in large errors at the higher response probabilities. The individual results of various conditions were equally compared, and the largest differences were found in the desktop 2.5 m/s condition with an average of 0.406 log-units, whilst the best predicted was the VR 1.4 m/s condition with an average difference of 0.296 log-units. From the results of this validation, it can be concluded that the proposed model more closely predicts the individual results of participants when compared to the traditional maximum likelihood estimation technique.

7 Discussion

The results of these two experiments support our hypotheses that *medium* and object *velocity* display a significant effect upon perceived smoothness of tick rate; however, no significance was found for the visual *complexity* of a stimulus.

VR consistently required higher tick rates for motion to appear smooth in comparison with the desktop equivalent, thus we accept H1. VR has been found to elicit stronger sensations of presence within virtual environments when compared to computer desktops (Shu et al. 2019), and human perception has been shown to be more sensitive in VR, with Niu et al. observing participants were more sensitive to surface roughness and specularity in VR (Niu and Lo 2022). Table 5 quantifies the requirement for higher tick rates in VR. The results suggest that the minimum required tick rate lies at 29 Hz for a desktop application and at 42 Hz for a VR application. However, for the majority of the population to consistently perceive the motion as smooth, the model predicts 56 Hz would be required for a desktop whilst VR would need 113 Hz. The 113 Hz prediction exceeds the maximum frame rate of the HMD (90 Hz), thus a tick rate to match the frame rate of the HMD may be applied without detriment to the experience.

H2 predicted that participants would be more sensitive to complex objects, however, *Complexity* displayed no significant effect upon tick rate thresholds, therefore H2 cannot be accepted. Two different stimuli types were presented in Experiment 1. The first was a simple sphere, whilst the second was a human avatar displaying increased *complexity* due to the motion of the hands and feet in the walking animation. ANOVA results found no significance for the *complexity* variable, therefore our findings suggest that tick rate may be controlled independently of the visual intricacy of an object. For this study, the visual *complexity* was

investigated in relation to update frequency; however, further research should be performed into the *complexity* of stimuli. For objects with complex geometry, more data must be sent each update and therefore will require greater bandwidth to maintain the same tick rate. In bandwidth constrained scenarios, complex objects would be forced to shift down the presented models to remain within the limitations of the bandwidth; therefore, a predictive model could be built based upon the size of the data transmitted and bandwidth availability, to find the balance between reducing the *complexity* of the object and reducing the tick rate to minimise the impact on the participant's experience.

The results from this study support H3, because *Velocity* displayed a significant effect in both experiments; however, the direction of the effect changed between Experiment 1 and Experiment 2. In the first experiment, *velocity* was a between-participant variable and results suggested that sensitivity decreased for increasing velocities, whilst in the second experiment *velocity* changed to a within-participant variable, and sensitivity changed to increase for increasing velocities. The difference observed in our experiments is likely to be due to the difference in methodology between experiments. In the first experiment, only one velocity was presented to each participant, whereas in the second experiment both velocities were shown and as such judgements of smoothness could be made relative to how smooth the other stimulus class appeared. It is this change in relative perception that may have caused the shift in results. This contradiction of results is not unique to this study. McDonnell et al. (2007) found for animations at higher velocities, a higher frequency was required, whilst Hulusić et al. (2011) found increased velocities were perceived as smoother in comparison with lower velocities, however, only for trials with no multimodal stimuli. As participants in a virtual environment are likely to be exposed to multiple velocities in a short duration, and thus relative perception can affect their experience, the results from the within-participant evaluation are more likely to accurately describe participant sensitivity to tick rate in a practical application. The predictions from Table 5 suggest that a faster moving stimulus required a 26% to 30% higher tick rate for both mediums.

8 Conclusions and future work

The influence of tick rate upon human perception of smoothness was investigated and predictive models were developed. Over the course of two experiments, three key variables were studied through a 1 alternative forced choice psychophysics methodology, the display *medium*, the *velocity*, and the *complexity* of the object.

The results suggest that VR requires significantly higher tick rates before motion appears smooth, and similarly sensitivity increases for higher velocities. The complexity of the object had no significant effect on tick rate sensitivity. The predictive models developed here could be used to predict participant tick rate thresholds at any desired probability and they establish a set of guidelines for appropriate tick rates in the future applications based upon the conditions in the scene. An important observation is that the majority of the predictions from the models lie significantly below the 90 Hz frame rate of the HMD, therefore tick rate may be decoupled from frame rate and thus operate at sub-frame rate frequencies without negatively affecting participant experience.

There are likely to exist other variables not investigated in this study, which may also exert an influence upon an individual's PF. VR enables binocular vision unlike a monitor, and in this work stimuli were only presented at one distance from the observer, so further research could be performed to establish whether a changing distance from the observer will affect perception. Throughout these experiments, the stimulus was always the focal point of the participant's attention and attention has previously been shown to significantly affect temporal sensitivity (Carver and Brown 1997). Additional research could evaluate whether objects within a virtual environment that are not the centre of attention could update at sub-threshold frequencies without any perceptible difference to a user.

Alongside exploring new variables, the present variables could be expanded to explore greater extremities of conditions. Participants were only tested on two different media, VR and desktop, so further research could be performed to expand these results to different levels of immersive media, from a small mobile phone screen to Mixed Reality headsets such as the Microsoft HoloLens. Additionally, only two velocities were presented, selected as velocities that are likely to be encountered in a virtual environment. The testing range could be expanded to explore more extreme values and such results could be compared to the results from Orban et al. (1984) to establish whether tick rate smoothness follows a similar shape.

In this experiment, the PFs for each participant were evaluated from the 25%, 50%, and 75% thresholds; however, the shape of the curve at more extreme thresholds such as 10% and 90% were not explored. As a result, in a practical implementation of this model, there is greater uncertainty in the true position of the high probability thresholds. Therefore, further research needs to be performed to establish whether the logistic function is the true shape of the PF or an alternative sigmoidal function, such as the Weibull or Cumulative Normal distribution functions would more closely describe the data for probabilities closer to certainty.

Data availability The datasets generated during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

Ethical approval Ethical approval was granted by BSREC (BSREC 143/20-21) and informed consent was received from every participant.

References

Akaike H (1998) Information theory and an extension of the maximum likelihood principle. Sel Pap Hirotugu Akaike. https://doi.org/10.1007/978-1-4612-1694-0_15

Almeida T, Pereira JM, Jorge J (2019) Evaluation of antialiasing techniques on mobile devices. In: 2019 international conference on graphics and interaction (ICGI). IEEE, pp 64–71, https://doi.org/10.1109/ICGI47575.2019.8955084

Altman D, Machin D, Bryant T et al (2000) Statistics with confidence: confidence intervals and statistical guidelines, 2nd edn. Wiley, Incorporated, Hoboken

Aseeri S, Interrante V (2021) The influence of avatar representation on interpersonal communication in virtual social environments. IEEE Trans Visual Comput Graphics 27(5):2608–2617. https://doi.org/10.1109/TVCG.2021.3067783

Bechara A, Damasio H, Tranel D et al (1997) Deciding advantageously before knowing the advantageous strategy. Science 275:1293–1295. https://doi.org/10.1126/science.275.5304.1293

Brennesholtz MS (2018) 3–1: Invited paper?: VR standards and guidelines. In: SID symposium digest of technical papers, vol 49, no 1. pp 1–4. https://doi.org/10.1002/sdtp.12476

Cao R, Zou-Williams L, Cunningham A, et al (2021) Comparing the neuro-physiological effects of cinematic virtual reality with 2d monitors. In: Proceedings - 2021 IEEE conference on virtual reality and 3D user interfaces, VR. pp 729–738. https://doi.org/10.1109/VR50410.2021.00100

Carver RA, Brown V (1997) Effects of amount of attention allocated to the location of visual stimulus pairs on perception of simultaneity. Percept. Psychophys. 59:534–542. https://doi.org/10.3758/BF03211862

Chheang V, Saalfeld P, Huber T, et al (2019) Collaborative virtual reality for laparoscopic liver surgery training. In: 2019 IEEE international conference on artificial intelligence and virtual reality (AIVR). IEEE, pp 1–17, https://doi.org/10.1109/AIVR46125.2019.00011

Claypool M, Claypool K (2009) Perspectives, frame rates and resolutions: It's all in the game. In: Proceedings of the 4th international conference on foundations of digital games - FDG '09. ACM Press, New York, p 42, https://doi.org/10.1145/1536513.1536530

Debattista K, Bugeja K, Spina S et al (2018) Frame rate vs resolution: a subjective evaluation of spatiotemporal perceived quality under varying computational budgets. Comput. Graphics Forum 37:363–374. https://doi.org/10.1111/cgf.13302

Denes G, Jindal A, Mikhailiuk A, et al (2020) A perceptual model of motion quality for rendering with adaptive refresh-rate and resolution. ACM Transactions on Graphics (Proc of SIGGRAPH 2020)

Dixken M, Diers D, Wingert B, et al (2019) Distributed, collaborative virtual reality application for product development with simple avatar calibration method. In: 2019 IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 1299–1300, https://doi.org/10.1109/VR.2019.8797884

DoVale E (2017) High frame rate psychophysics: experimentation to determine a JND for frame rate. SMPTE Motion Imaging J 126(9):41–47. https://doi.org/10.5594/JMI.2017.2749919

Dzardanova E, Kasapakis V, Gavalas D et al (2022) Virtual reality as a communication medium: a comparative study of forced compliance in virtual reality versus physical world. Virtual Real. 26(2):737–757. https://doi.org/10.1007/s10055-021-00564-9

Elbamby MS, Perfecto C, Bennis M et al (2018) Toward low-latency and ultra-reliable virtual reality. IEEE Netw. 32(2):78–84. https://doi.org/10.1109/MNET.2018.1700268

García-Pérez MA (1998) Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties. Vision Res. 38(12):1861–1881. https://doi.org/10.1016/S0042-6989(97)00340-4

Ha G, Lee H, Lee S, et al (2016) A vr serious game for fire evacuation drill with synchronized tele-collaboration among users. In: Proceedings of the 22nd ACM conference on virtual reality software and technology. ACM, pp 301–302, https://doi.org/10.1145/2993369.2996306

Hecht J (2016) Optical dreams, virtual reality. Opt Photonics News 27(6):24. https://doi.org/10.1364/OPN.27.6.000024

Herder J, Brettschneider N, Mooij JD, et al (2019) Avatars for co-located collaborations in hmd-based virtual environments. In: 2019 IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 968–969. https://doi.org/10.1109/VR.2019.8798132

Horn BK, Schunck BG (1981) Determining optical flow. Artif Intell 17:185–203. https://doi.org/10.1016/0004-3702(81)90024-2

Hulusić V, Debattista K, Aggarwal V et al (2011) Maintaining frame rate perception in interactive environments by exploiting audio-visual cross-modal interaction. Vis Comput 27:57–66. https://doi.org/10.1007/s00371-010-0514-2

Jimenez J, Lottes T, Malan H, et al (2011) Filtering approaches for real-time anti-aliasing. In: ACM SIGGRAPH 2011 courses on - SIGGRAPH '11. ACM Press, New York, pp 1–329, https://doi.org/10.1145/2037636.2037642

Kaernbach C (1991) Simple adaptive testing with the weighted up-down method. Percept. Psychophys. 49:227–229. https://doi.org/10.3758/BF03214307

Kim JW, Choe WJ, Hwang KH et al (2017) 78–2: the optimum display for virtual reality. In: SID symposium digest of technical papers, vol 48, no 1. pp 1146–1149. https://doi.org/10.1002/sdtp.11845

Kingdom FA, Prins N (2016) Psychophysics: a practical introduction, 2nd edn. Elsevier, Amsterdam. https://doi.org/10.1016/C2012-0-01278-1

Latoschik ME, Kern F, Stauffert JP et al (2019) Not alone here?! Scalability and user experience of embodied ambient crowds in distributed social virtual reality. IEEE Trans Visual Comput Graphics 25(5):2134–2144. https://doi.org/10.1109/TVCG.2019.2899250

Lee WK, Chang RKC (2015) Evaluation of lag-related configurations in first-person shooter games. In: 2015 international workshop on network and systems support for games (NetGames), vol 1. IEEE, pp 1–3, https://doi.org/10.1109/NetGames.2015.7382997

Levenberg K (1944) A method for the solution of certain non-linear problems in least squares. Q Appl Math 2(2):164–168. https://doi.org/10.1090/qam/10666

McDonnell R, Newell F, O'Sullivan C (2007) Smooth movers: perceptually guided human motion simulation. In: 2007 ACM SIG-GRAPH/Eurographics symposium on Computer animation (SCA '07). Association for Computing Machinery Inc, pp 259–270, https://doi.org/10.5555/1272690.1272726

Mohler BJ, Thompson WB, Creem-Regehr SH et al (2007) Visual flow influences gait transition speed and preferred walking speed. Exp Brain Res 181(2):221–228. https://doi.org/10.1007/s00221-007-0917-0

Niu M, Lo CH (2022) Do we see rendered surface materials differently in virtual reality? A psychophysics-based investigation. Virtual Real 2021:1–15. https://doi.org/10.1007/S10055-021-00613-3

Okada K (2017) Negative estimate of variance-accounted-for effect size: how often it is obtained, and what happens if it is treated as zero. Behav Res Methods 49(3):979–987. https://doi.org/10.3758/s13428-016-0760-y

Orban GA, de Wolf J, Maes H (1984) Factors influencing velocity coding in the human visual system. Vision Res 24:33–39. https://doi.org/10.1016/0042-6989(84)90141-X

Parthasarathy V, Simiscuka AA, O'Connor N, et al (2020) Performance evaluation of a multi-user virtual reality platform. In: 2020 international wireless communications and mobile computing (IWCMC). IEEE, pp 934–939, https://doi.org/10.1109/IWCMC48107.2020.9148390

Prins N (2019) Too much model, too little data: How a maximum-likelihood fit of a psychometric function may fail, and how to detect and avoid this. Attent. Percept. Psychophys. 81:1725–1739. https://doi.org/10.3758/s13414-019-01706-7

Roth D, Lugrin JL, Galakhov D, et al (2016) Avatar realism and social interaction quality in virtual reality. In: 2016 IEEE virtual reality (VR). IEEE, pp 277–278, https://doi.org/10.1109/VR.2016.7504761

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464. https://doi.org/10.2307/2958889

Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52(3–4):591–611. https://doi.org/10.1093/biomet/52.3-4.591

Shu Y, Huang YZ, Chang SH et al (2019) Do virtual reality head-mounted displays make a difference? A comparison of presence and self-efficacy between head-mounted displays and desktop computer-facilitated virtual environments. Virtual Real 23:437–446. https://doi.org/10.1007/s10055-018-0376-x

Skarbez R, Brooks FP Jr, Whitton MC (2018) A survey of presence and related concepts. ACM Comput Surv 50(6):1–39. https://doi.org/10.1145/3134301

Slater M (2018) Immersion and the illusion of presence in virtual reality. Br J Psychol 109(3):431–433. https://doi.org/10.1111/bjop.12305

Stauffert JP, Niebling F, Latoschik ME (2018) Effects of latency jitter on simulator sickness in a search task. In: 25th IEEE conference on virtual reality and 3D user interfaces, VR 2018 - proceedings. Institute of Electrical and Electronics Engineers Inc., pp 121–127, https://doi.org/10.1109/VR.2018.8446195

Strasburger H (2001) Converting between measures of slope of the psychometric function. Percept Psychophys 63:1348–1355. https://doi.org/10.3758/BF03194547

Swanson WH, Birch EE (1992) Extracting thresholds from noisy psychophysical data. Percept Psychophys 51:409–422. https://doi.org/10.3758/BF03211637

Treutwein B (1995) Adaptive psychophysical procedures. Vision Res 35:2503–2522. https://doi.org/10.1016/0042-6989(95)00016-X

Treutwein B (1999) Fitting the psychometric function. Percept Psychophys 61:87–106

Varshney LR, Sun JZ (2013) Why do we perceive logarithmically? Significance 10(1):28–31. https://doi.org/10.1111/j.1740-9713.2013.00636.x

Wagenmakers EJ, Farrell S (2004) AIC model selection using Akaike weights. Psychon Bull Rev 11(1):192–196

Ware C, Balakrishnan R (1994) Reaching for objects in VR displays. ACM Trans Comput-Hum Interact (TOCHI) 1(4):331–356. https://doi.org/10.1145/198425.198426

Warren PA, Bell G, Li Y (2022) Investigating distortions in perceptual stability during different self-movements using virtual reality. Perception 51(10):681–697. https://doi.org/10.1177/03010066221116480

Wasserstein LN (1969) Markov processes with countable state space describing large systems of automata. Probl Inf Transm 5(3):47–52

Weibel N, Gasques D, Johnson J, et al (2020) Artemis: mixed-reality environment for immersive surgical telementoring. In: Extended abstracts of the 2020 CHI conference on human factors in computing systems. ACM, pp 1–4, https://doi.org/10.1145/3334480.3383169,

Wichmann FA, Hill NJ (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys 63:1293–1313. https://doi.org/10.3758/BF03194544

Zhao J, LaFemina P, Carr J, et al (2020) Learning in the field: comparison of desktop, immersive virtual reality, and actual field trips for place-based stem education. In: 2020 IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 893–902, https://doi.org/10.1109/VR46266.2020.00012

Zielinski DJ, Rao HM, Sommer MA et al (2015) Exploring the effects of image persistence in low frame rate virtual environments. In: 2015 IEEE virtual reality conference, VR 2015 - proceedings, pp 19–26. https://doi.org/10.1109/VR.2015.7223319