# A perceptual model of motion quality for rendering with adaptive refresh-rate and resolution

GYORGY DENES, University of Cambridge
AKSHAY JINDAL, University of Cambridge
ALIAKSEI MIKHAILIUK, University of Cambridge
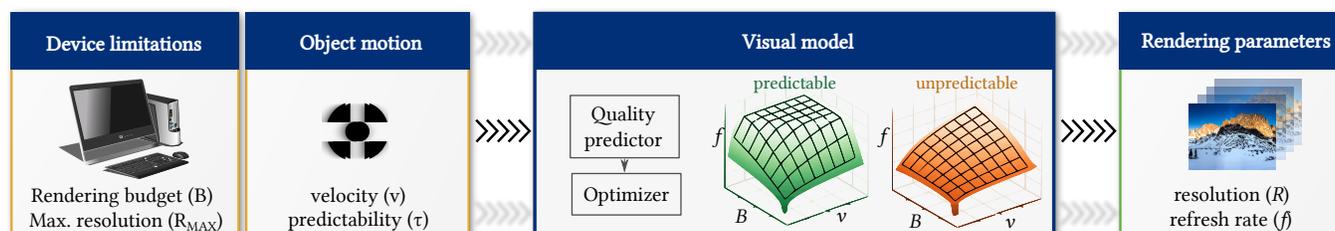RAFAŁ K. MANTIUK, University of Cambridge

Fig. 1. Our proposed perceptual model of motion quality takes object motion, refresh rate and device limitations (such as the rendering budget and the maximum screen resolution) to predict the perceived quality. This model can be then used to find the combination of resolution and refresh rate that produces the highest animation quality under the given conditions. Surface plots visualize model predictions.

Limited GPU performance budgets and transmission bandwidths mean that real-time rendering often has to compromise on the spatial resolution or temporal resolution (refresh rate). A common practice is to keep either the resolution or the refresh rate constant and dynamically control the other variable. But this strategy is non-optimal when the velocity of displayed content varies. To find the best trade-off between the spatial resolution and refresh rate, we propose a perceptual visual model that predicts the quality of motion given an object velocity and predictability of motion. The model considers two motion artifacts to establish an overall quality score: non-smooth (juddery) motion, and blur. Blur is modeled as a combined effect of eye motion, finite refresh rate and display resolution. To fit the free parameters of the proposed visual model, we measured eye movement for predictable and unpredictable motion, and conducted psychophysical experiments to measure the quality of motion from 50 Hz to 165 Hz. We demonstrate the utility of the model with our on-the-fly motion-adaptive rendering algorithm that adjusts the refresh rate of a G-Sync-capable monitor based on a given rendering budget and observed object motion. Our psychophysical validation experiments demonstrate that the proposed algorithm performs better than constant-refresh-rate solutions, showing that motion-adaptive rendering is an attractive technique for driving variable-refresh-rate displays.

CCS Concepts: • **Computing methodologies** → **Perception**; *Rendering*.

Authors' addresses: Gyorgy Denes, University of Cambridge, gyorgy.denes@cl.cam.ac.uk; Akshay Jindal, University of Cambridge; Aliaksei Mikhailiuk, University of Cambridge; Rafał K. Mantiuk, Department of Computer Science and Technology, University of Cambridge, rafal.mantiuk@cl.cam.ac.uk.

Additional Key Words and Phrases: motion quality, adaptive refresh rate

## 1 INTRODUCTION

Modern displays can offer both high spatial resolution (up to 8K) and high refresh rates (above 100 Hz). Such high spatio-temporal resolution is needed to reach the perceptual limits of the visual system and to deliver high-fidelity content. This is particularly important for VR/AR headsets, which still offer resolutions far below the perceptual limits. However, a major obstacle is the limited computational power and bandwidth of modern GPUs: only the most powerful GPUs can render 4K content at 100 Hz or more.

Rendering for modern display technologies often requires a trade-off between spatial and temporal resolution. For example, as VR/AR headsets require constant and sustained refresh rates, the quality control mechanism in rendering engines needs to dynamically adjusts the rendering resolution to fit within the rendering budget. Another approach, employed on G-/Free-Sync capable displays, is to render at a fixed resolution but vary the refresh rate according to the available rendering budget. However, depending on camera and content motion, keeping either the spatial *or* temporal resolution constant may not produce the best visual quality. A better approach is to manipulate the refresh rate and resolution simultaneously, *i.e.* to dynamically adjust the trade-off based on the content of the animation. For example, when a scene is static, the application should maximize spatial resolution, but when movement is fast, the application should optimize for higher refresh rates which result in better perceived quality. Such a mechanism can be introduced

into modern rendering pipelines, but determining the best trade-off requires modeling the perceived quality of motion.

In this paper, we measure the perceived visual quality of motion from 50 Hz to 165 Hz through subjective tests, and propose a novel visual model that predicts this quality taking into account two motion artifacts: non-smooth motion (judder) and blur. We extend existing work by incorporating the velocity and predictability of motion into our model. Our treatment of blur is uniquely broad, fusing hold-type blur, eye motion, and spatial blur arising from resolution reduction. The model isolates individual components contributing to perceived quality such as judder, spatial blur aligned with the direction of motion, and spatial blur orthogonal to the direction of motion. Then, the discrimination of each component is modeled using spatio-temporal contrast sensitivity functions. To find the free parameters of our model, we measured eye movement accuracy and conducted further psychophysical experiments.

The main contributions of this work can be summarized as:

- Measurement of visual quality at refresh rates from 50 Hz to 165 Hz at a granularity of 5 Hz for multiple content and object velocities; we used active sampling to tackle the large dimensionality of the problem (Section 4).
- A velocity-dependent perceptual model for predicting the quality of motion, taking into account judder, eye motion blur, hold-type blur and resolution reduction blur (Section 5).
- Eye tracking measurements of retinal blur for predictable and unpredictable motion at different object velocities and refresh rates between 16 Hz and 165 Hz (Section 6.1).
- Psychophysical measurements on the visibility of non-smooth (juddery) motion when isolated from blur (Section 6.2).
- A real-time rendering algorithm utilizing the proposed visual model, validated in two experiments (Section 7.4).

Additional results, video and the code can be found on the project web page[1].

## 1.1 Limitations

To outline the scope of this paper, we first discuss the assumptions and limitations of the proposed visual model, including hardware limitations that may prevent making full use of the technique.

*1.1.1 Model design limitations.* We derive the model for a single worst-case scenario so that model predictions can be pre-computed for real-time rendering (more in Section 5). However, a less conservative model could provide better control over the resolution and refresh rate. Our approach excludes combination with content-aware methods such as Pellacini et al. [2005]. Furthermore, while our model was designed to generalize across displays, it was calibrated and tested only with high-persistence LCD monitors of standard brightness. More data is required to validate the model with HDR displays and VR headsets. The model assumes that peripheral motion artifacts are not prominent and relies on the velocity in the foveal vision. Lastly, we only address two of the four categories of motion artifacts (Section 2). Rendering artifacts related to ghosting and flicker will not be considered; consequently, temporal aliasing due to resolution changes is not modeled. We demonstrate that even with these limitations, our visual model provides a better fit

to psychophysical data than current state-of-the art, and it allows designing adaptive refresh-rate rendering algorithms.

*1.1.2 Current hardware limitation.* Although the technique could be beneficial for adaptive-refresh-rate rendering in AR/VR, this could not be tested because of the lack of adaptive-refresh-rate headsets. The proposed adaptive-refresh-rate algorithm ideally requires fine control over the display's refresh rate. While G-/Free-Sync gives the capability to synchronize display refresh rate to that of the GPU, this is implemented as a control system. This means that estimating the new refresh rate on the display side and transitioning to it takes time. A direct interface to request from a Free-/G-Sync monitor a specific refresh rate, provided by our adaptive rendering method, could reduce such latency and avoid potential flicker.

## 2 BACKGROUND

In this section, we review the relevant work in resolution and motion perception. Motion artifacts can be divided into four categories: (1) flicker, (2) false edges (ghosting, phantom arrays), (3) motion blur, and (4) non-smooth motion or judder (also known as strobing or stutter) [Daly et al. 2015]. For an illustration, see Figure 2. Motion artifacts in high-persistence and low-persistence (used in VR) displays are significantly different; in this work we restrict our focus to the more common high-persistence LCD displays, and do not consider latency-reduction [Greer et al. 2016] techniques.

## 2.1 Flicker

Displays often switch an image on and off at high frequencies to control brightness or achieve low-persistence. If the frequency of this is below the critical flicker frequency (CFF), the temporal change becomes visible as flicker. The visibility of flicker depends on a number of factors: it increases with the size, the contrast, and the log-luminance of a stimulus (Ferry-Porter law), and varies with spatial frequency and eccentricity [Rutherford 2003]. High-contrast changes are known to cause flicker beyond the typical critical flicker frequency of 60 Hz [Liu et al. 2014]. In this paper, we assume high-persistence LCD displays with photographic and gaming content. In this context, periodic high-contrast changes are uncommon, and hence flicker artifacts are not prominent.

## 2.2 False edges

Temporal change, when combined with object and eye motion, can cause motion artifacts even above the CFF. After the eye integrates the incoming signal in the temporal domain, multiple sharp copies of the target object might be perceived. Such false edges (or ghosting) can occur if a low-persistence displays repeats the same frame (at least) twice. A similar artifact can be observed with DLP projectors, where the color wheel presents the red, green and blue images subsequently. Alternatively, an array of false edges are observable even at surprisingly high refresh rates (500 Hz–1 kHz) during saccadic eye motion. Some authors refer to these *phantom arrays* as flicker [Davis et al. 2015; Roberts and Wilkins 2013], however, we follow Daly's taxonomy [1998] and classify those as false edges. On high-persistence (LCD) displays such false edges are uncommon, as saccadic eye motion results in blur instead.
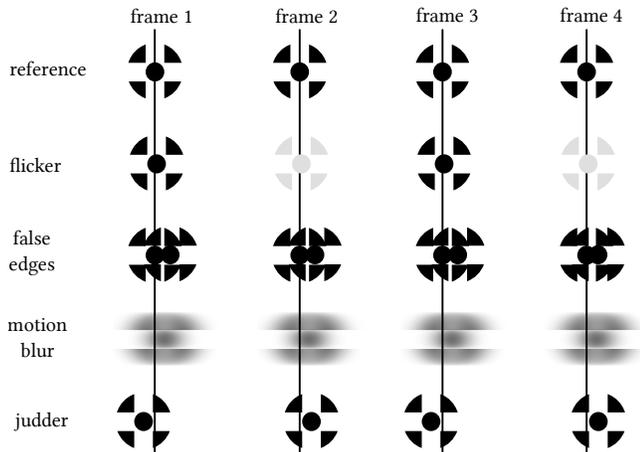
Fig. 2. Motion artifacts (rows 2-5) compared to perfect motion (top row) for consecutive frames (columns). Flicker: luminance change in consecutive frames; false edges: multiple copies of the original object; motion blur: loss of high-frequency detail; judder: object location is inconsistent in consecutive frames (vertical lines indicate reference locations for each frame).
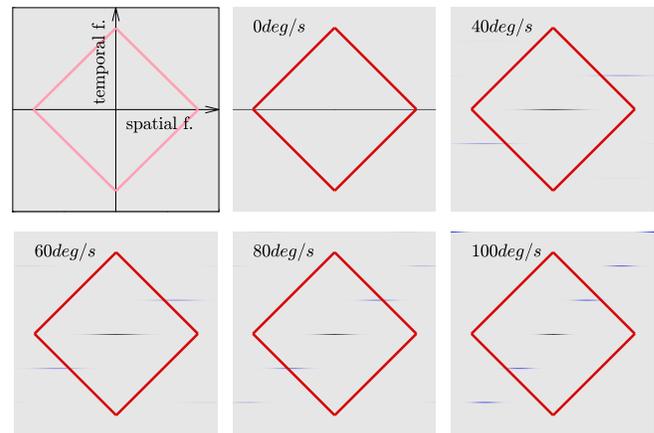


Fig. 3. Illustration of aliasing copies (blue) of the original signal (black) in the frequency domain. The spatio-temporal spectra illustrate a line moving with constant velocity (top-left corners) on a high-persistence display with a fixed refresh rate. The red diamond represents the window of visibility [Watson et al. 2008]. Higher velocities push the copies closer.

## 2.3 Motion blur

Most of motion blur can be attributed to LCD displays holding an image for the full duration of a frame. When the gaze follows a moving object on a screen, it pans over pixels that are stationary. At high refresh rates the visual system integrates the image over time [Tourancheau et al. 2009], resulting in blur. The amount of blur increases proportionately with velocity, and varies inversely with the refresh rate. Because of that, blur is particularly objectionable on VR headsets, where high velocities due to head motion are common, increasing the chance of simulation sickness [Anthes et al. 2016].

The other source of blur is eye motion. Our gaze is constantly in motion; when fixating, the gaze drifts, performing saccades to shift focus [Robinson 1964]. When observing a moving object with speeds from 0.15 deg/s up to 80 deg/s, the eye follows, keeping the image of the object in the central (foveal) region of the retina. This tracking is known as smooth pursuit eye motion (SPEM) [Robinson 1965]. As SPEM is imperfect, small retinal position deviations are integrated by the visual system, resulting in blur. The exact amount of this blur depends on the nature of the motion. Predictable targets can be followed quite accurately with an average velocity gain of 0.82 [Daly 1998] due to the predictive mechanism of the visual system anticipating the location and speed [Stark et al. 1962]. However, real-world targets often do not follow such predictable patterns. As a fallback, SPEM gets regulated by feedback mechanisms based on the difference between expected and actual retinal location of the target object (also known as retinal slip) [Lisberger 2010; Niehorster et al. 2015]. Numerous factors have been demonstrated to affect the accuracy of this mechanism, including traumatic brain injury [Suh et al. 2006] and manual tracking [Niehorster et al. 2015].

Watson and Ahumada [2011] provides an excellent review of studies investigating the visibility of blur. With a focus on Gaussian blur, the authors unify available psychophysical experimental data, and discuss proposed models for the visibility of blur. In our visual model, we follow similar principles to their visual contrast energy model by computing energy after modulating the blurred signal with the contrast sensitivity function. However, our proposed application of high-refresh-rate rendering requires a more careful consideration of eye blur as a function of motion velocity and predictability.

## 2.4 Judder

At low refresh rates the illusion of motion breaks, and individual frames become visible. This creates juddery, stuttery or strobing motion. Judder is caused by the discrete temporal samples of the display (frames), which produce aliasing copies in the frequency domain (Figure 3). The magnitude and location of these aliasing copies depend on both the refresh rate and motion velocity.

## 2.5 Contrast sensitivity functions

The visual system imposes limits on both the spatial and the temporal resolution we can see. These limits are to some extent observer-dependent, and are known to change with viewing conditions and observer age [Elliott et al. 1995]. A widely-used approach to model whether an artifact is perceivable, is to use the family of contrast sensitivity functions. The sensitivity to luminance contrast at different spatial frequencies is described by the contrast sensitivity function (CSF). The CSF is known to peak on low-to-medium frequencies (1–4 cycles per degrees, or cpd), and fall off exponentially for higher frequencies. The exact shape of the CSF depends on background luminance, orientation, stimulus size, and eccentricity. CSF models exist for standard observers [Barten 2003], but these often need to be fitted to psychophysical data to be of practical use.

When temporal dimension is considered, the sensitivity is explained by the spatio-temporal contrast sensitivity function (stCSF). The spatial and temporal frequency dimensions are not independent, but well-established models capture the joint sensitivity [Daly 1998; Kelly 1979]. If only the sensitivity to higher frequencies is relevant, stCSF can be described with simplified models, such as the window (or pyramid) of visibility [Watson 2015; Watson et al. 2008].
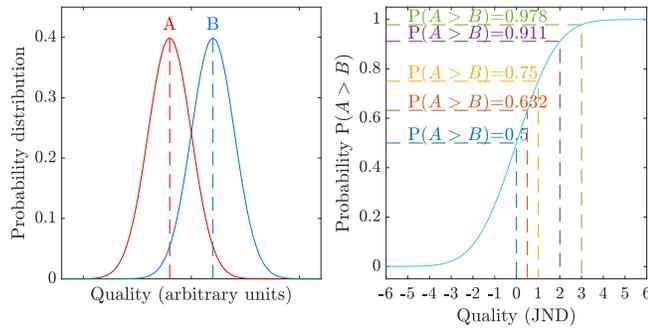
Fig. 4. Left: the quality of two conditions A and B is represented as two normally distributed random variables. Right: a pairwise comparison experiment let us measure $P(A > B)$ or $P((A - B) > 0)$. We use a cumulative normal distribution to recover the quality difference from such probability. The plot shows how JND values correspond to the portion of a population selecting condition A over B. Negative scores indicate that the quality of A is worse than the quality of B.

## 2.6 JND for quality

In this paper, we represent quality in the units of just noticeable differences (JNDs), therefore it is important to explain the interpretation and rationale behind those units. JND units arise from pairwise comparison experiments, in which an observer compares two conditions (*e.g.* two refresh rates) and indicates the one that is of higher quality. When such an experiment is repeated across a number of observers, we can establish a probability that a random observer in a population selects condition A over condition B. A commonly made assumption is that such probabilities arise from unseen quality scores assigned by observers, which can be modeled as a normally distributed random variables with the same variance, as shown in Figure 4-left (Thurstone model case V [Thurstone 1927]). The outcome of a pairwise comparison is explained by sampling both distributions and comparing the resulting scores. To recover unseen quality, we need to map probability $P(A > B)$ into corresponding quality difference. Such a mapping is given by the inverse cumulative normal distribution, shown in Figure 4-right. Moreover, since the scaling of quality units is arbitrary, by convention we select $\sigma$ of the cumulative normal distribution to be 1.4826 so that the quality difference $\Delta Q$ is equal to 1 JND when $P(A > B) = 0.75$, i.e., 75% of the population selects A over B. Unlike Mean-Opinion-Scores (MOS), collected in rating experiments, the quality values in JND units provide a meaningful scale, which can tell us about a practical significance of the difference in quality (effect size).

Pairwise comparison experiments, such as our Experiment 1, often involve comparisons between hundreds of conditions, each helping to estimate a distance between pairs of conditions on the quality scale. To recover a consistent quality scale (and reduce estimation error) across all conditions, the results are scaled, typically by solving for an optimization problem. More information on scaling can be found in [Perez-Ortiz and Mantiuk 2017].

## 3 RELATED WORK

### 3.1 Critical refresh rate

Many authors have attempted to establish a critical monitor refresh rate, beyond which improvements yield no gain in perceived motion quality. Noland et al. [2014] applied traditional sampling theory, combining the CSF with a simple model of eye motion, applying the Nyquist limit to derive the refresh rate that is indistinguishable from perfect motion. Their model for an LCD display predicts that while 140 Hz is sufficient for untracked motion, tracked motion requires at least 700 Hz for the illusion of perfect motion. The authors highlight that the figures should only be considered approximate, partly due to the limitations of Daly's model of SPEM [1998]. Deeper knowledge of the SPEM mechanism suggests that the nature of motion (predictable vs. unpredictable) is also likely to affect these figure. Kuroki et al. [2006; 2007] arrived at a more conservative estimate using psychophysical measurements showing that at least 250 Hz is required to completely remove motion blur and judder. Such refresh rates are unfortunately still beyond the capabilities of most consumer GPUs and monitors, and the threshold numbers provide little intuition as to how the perceived quality of motion increases in the 60-165 Hz range.

### 3.2 Perceived impact of refresh rate

DoVale et al. [2017] measured the just-noticeable-difference (JND) threshold from three anchor frequencies, revealing that a relatively small change from 24 Hz to 27 Hz is perceivable by 75% of the population (1 JND), while from 48 Hz, a more substantial increase to 62 Hz is required. The authors, however, do not attempt to model the JND threshold for arbitrary refresh rates, or relate the findings to models of the visual system. We extend their measurements to the range of frequencies from 50 Hz to 165 Hz and investigate a number of velocities and types of motion.

The perception of motion quality was measured in a number of experiments. Notably Macking et al. [2016] isolated display blur and temporal artifacts such as flicker and judder, collecting mean opinion scores for each as a function of object velocity and monitor refresh rate. The authors concluded that for object velocities below 60 deg/s, about 50% of the critical refresh rate (at which no artifacts are detected) could provide an *acceptable* score, however, they do not provide any guidelines as to how different motion artifacts contribute to the overall perceived quality below such threshold.

### 3.3 Motion quality in films

In the film industry, refresh rate can be considered an artistic tool. While physically higher refresh rates provide more realism, in the context of cinemas, they might be considered less aesthetic — a phenomenon sometimes described as the soap opera effect. Simultaneous management of motion quality and viewer expectations can be achieved by emulating continuously varying frame rates [Templin et al. 2016]. This emulation relies on similar perceptual principles as our paper, such as models of motion blur. However, the authors do not attempt to quantify the visibility or subjective quality of such artifacts, as they merely mimic them for the purpose of producing a "cinematographic look". An acquired taste similar to the soap opera effect has not been reported for computer games,

therefore perceptual video quality metrics calibrated for <60 Hz such as Ou et al. [2010], Kim et al. [2018], Banitalebi-Dehkordi et al. [2015] are not applicable for high-refresh-rate rendering. For this reasons, we also do not discuss papers that introduce simulated camera motion blur to rendered frames [Navarro et al. 2011].

Perceived motion quality of panning (horizontal movement on the screen) was investigated recently by Chapiro et al. [2019]. The authors measured subjective motion artifact scores for refresh rates typical of modern televisions (30, 60, and 120 Hz) across a range of luminance (2.5–40 cd/m$^2$) and camera panning speeds (2–6.6 $deg/s$). A trivariate quadratic empirical model was shown to fit their data well. In contrast to their study, which considered cinematographic content, our work is focused on computer graphics rendering. Therefore, we collect data for refresh rates and velocities more representative of computer games (up to 165 Hz with motion speeds up to 45 $deg/s$) with predictable and unpredictable SPEM. Moreover, by isolating two quality factors (blur and judder), we can find the best trade-off between the spatial and temporal resolution. Such an application is not possible with the empirical model of Chapiro et al. as it does not account for the impact of lower spatial resolution.

## 3.4 Visual quality metrics for games

For computer graphics applications, there are only a few visual quality metrics that take into account both spatial and temporal information. Yee et al. [2001] proposed using the CSF adjusted with object velocity and a model of visual attention to produce a spatio-temporal error tolerance map. Aydin et al. [2010] proposed a full-reference perceptual model based on the spatio-temporal contrast sensitivity function to predict the visibility of distortions in HDR content. While both of these approaches capture some of the visible blur artifacts, the first metric does not generalize to arbitrary refresh rates, and the second metric does not consider SPEM. Unlike previous models, we propose a model that incorporates a more accurate model of eye motion and also considers judder artifacts. Our model is also content-independent, which let us pre-compute the trade-off between resolution and refresh rate and select it adaptively in real-time applications with negligible impact on performance.

## 3.5 Refresh rate vs. resolution

In a system with fixed computational budget, increasing the rendering frame rate entails the reduction of scene complexity, shading quality, bit-depth [McCarthy et al. 2004] or resolution. The computational trade-off between resolution and refresh rate is easy to model and control. In this paper we decided to explore this direction.

Claypool et al. investigated the effect of latency, and trading off resolution for refresh rate in the context of first-person shooters (FPS) [Claypool and Claypool 2007, 2009]. Large-scale user studies revealed that the refresh rate has a significantly larger influence on task performance. On 3-7 Hz users could not target opponents, and there were clear task performance benefits of increasing the refresh rate up to 60 Hz. Perceptual quality and playability gathered with post-experiment questionnaires revealed a similar but less pronounced trend. Unfortunately no measurements were made beyond the capabilities of a standard 60 Hz monitor. Higher refresh rates

result in reduced game latency, another factor known to affect task performance (shooting) in FPS games [Beigbeder et al. 2004].

Debattista et al. [2018] were the first to demonstrate that rendering quality under a constrained budget can be formulated as an optimization problem, where the free parameter is the ratio of refresh rate and resolution Preference data collected in a two-alternative-forced-choice (2AFC) experiment indicated that the optimal ratio for seven-second clips is dependent on the computational budget. We follow their approach in our target application and also optimize rendering quality under a constrained budget using a model. However, our model accounts for the velocity of motion, which was shown to be a major factor affecting motion quality [Chapiro et al. 2019]. Furthermore, our intention is to build an explainable model, accounting for the underlying mechanisms of the visual system, which can generalize across a wide range of input parameters. Such an ability to generalize and explain cannot be achieved with an empirical function fit offered by their study. We compare our model to the model of Debbatista et al. in Section 7.

## 3.6 Perceptual rendering methods

Numerous methods have been proposed to exploit perceptual limitations to overcome hardware constraints. For a comprehensive review, we refer the reader to Masia et al. [2013]. In the temporal domain, such savings are usually achieved by re-projection [Didyk et al. 2010; Niklaus and Liu 2018; Scherzer et al. 2012]. More recently, Denes et al. [2019] demonstrated that removing high spatio-temporal details (rendering every other frame with a reduced resolution) can be a viable alternative. Temporal resolution can be also traded off to increase the apparent display resolution [Didyk et al. 2010; Templin et al. 2011]. While such techniques offer promising results, the motion artifacts they introduce are complex and varied. Since modern graphics pipelines still operate on the assumption that frames are produced at a fixed resolution at a constant refresh rate, we only model this phenomenon and leave it up to future work to extend with features such as models of re-projection accuracy.

In the spatial domain, foveated rendering has received a lot of interest recently. Such techniques exploit the limited spatial resolution of the eye in the periphery [Guenter et al. 2012]. Non-uniform sampling at different eccentricities can be reversed by novel optical design [Akşit et al. 2019], or deep-learning methods [Kaplanyan et al. 2019]. As the per-frame sample count in such schemes is fixed, and we concentrate on motion perception in the fovea, we consider our work orthogonal to foveated rendering.

## 4 EXPERIMENT 1: MEASURING MOTION QUALITY

To understand how motion quality is affected by refresh rate, motion velocity and the predictability of motion, we conducted a quality assessment experiment.Results are later used to train our model.

## 4.1 Experiment description

*4.1.1 Setup.* Observers were shown the same animation at two different refresh rates, each shown on a separate G-sync capable ASUS ROG Swift PG279Q 27" WQHD display. The displays were stacked on top of each other to make the task of comparing horizontal motion easier. The viewing distance was 108 cm (30° field of

view). Every animation was shown at a one of 23 refresh rates from 50 Hz to 165 Hz. The granularity of 5 Hz was chosen to approximate the 1 JND threshold at 50 Hz [DoVale 2017]. For all experiments we ensured that [ITU-R 2016] recommendations were met and that the time for performing one experiment did not exceed 30 minutes to prevent observer fatigue.

*4.1.2 Stimuli.* To cover a range of realistic and synthetic content, we used three animations: checkered circle (*circle*), eye tracker target (*ET*), and a panorama image (*panorama*) (Figure 5). *ET* was a combination of a bull's eye and a cross hair which has been shown to be effective as a fixation target [Thaler et al. 2013]. Animations had only horizontal motion to aid comparison in our vertically-stacked setup. Each of 6 tested conditions involved different content, range of velocities, and type of motion. In conditions (a)–(c) the *circle* underwent predictable sinusoid motion (Figure 6-left) with peak velocities at 15 deg/s, 30 deg/s, and 45 deg/s, respectively. In condition (d) the same *circle* underwent unpredictable motion (Figure 6-right) with mean velocity 23 deg/s. In condition (e) *ET* underwent predictable sinusoid motion (Figure 6-left) with peak 15 deg/s. Finally, in condition (f) *panorama* underwent a predictable motion following a soft staircase function (Figure 6-right): $\theta(t) = 15°\,(\sin(2\pi t)/2\pi t + t)$, peak velocity at 30 deg/s. For unpredictable motion we used the same function as Niehorster et al. [2015] (the sum of non-harmonic sinusoid motions with randomized phases):

$$\theta(t) = 17° \sum_{i=1}^{7} a_i \sin(2\pi\omega_i t + \rho_i), \qquad (1)$$

where $\theta(t)$ is the horizontal object location at time $t$, $a_i = \{2, 2, 2, 2, 2, 0.2, 0.2\}$, and $\rho_i = \{0.1, 0.14, 0.24, 0.41, 0.74, 1.28, 2.19\}$.

*4.1.3 Participants, Task.* Eleven participants aged 20-42, one female ten male, with normal or corrected-to-normal vision took part. Quality was measured using a pairwise comparison protocol because of its relative simplicity and speed [Perez-Ortiz and Mantiuk 2017]. For each trial, participants were asked to select the monitor with higher visual quality. Highr visual quality was defined as *sharp details and smooth motion*. Each participant performed 600 comparisons (6600 across all participants). During training, the researcher highlighted key differences in sharpness and motion smoothness side-by-side at 30 Hz vs. 120 Hz on content that was later not used in the experiment.
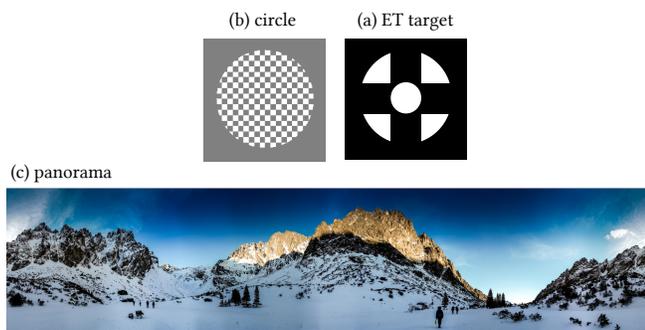


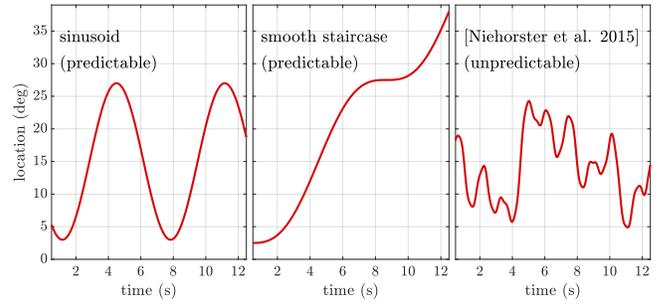Fig. 5. Content used in the motion quality experiment.

Fig. 6. Example object motions used in the motion quality experiment. Sinusoid and smooth staircase motion (left and center) are predictable by the SPEM mechanism, while the sum of non-harmonic sinusoid (right) is unpredictable.

*4.1.4 Sampling.* In order to efficiently utilize observers' time and obtain the most accurate scale possible, we used active sampling [Mikhailiuk et al. 2020]. The next comparison was selected to deliver the most information, *i.e.*, the one that has the highest impact on the posterior distribution of the scores. For that, separately for each animation, we obtained the distribution of quality scores after every performed comparison using Expectation Propagation (EP), assuming Thurstone case V [Murphy 2012], and then estimated the distribution of the scores assuming every possible future comparison. The comparison maximizing the Kulback-Leibler divergence between the current distribution and after every possible comparison was chosen to be performed next.

*4.1.5 Unified velocity scale.* We asked all 11 observers to perform 420 comparisons within each condition (4620 across all participants). To establish reliable quality differences between different velocities, we collected 180 additional comparisons across velocities for conditions (a)–(c) (velocities 15 deg/s, 30 deg/s, 45 deg/s). Both measurements let us obtain a unified quality scale taking into account both the refresh rate and the velocity of the object. The results were scaled using an MLE-based method (Section 2.6). Since JNDs are relative, the quality of the lowest measured refresh rate was set to 0 JND. To show the relative difference between the velocities in the *circle* animation, the quality of 15 deg/s at 50 Hz was set at 0 JND.

## 4.2 Results

Figure 7 shows the result of the cross-velocity scaling for predictable stimuli. The overall shape is consistent with most previous expectations: higher refresh rates imply less-perceivable motion artifacts, but differences above 100 Hz are increasingly more difficult to observe. In this region of refresh rates, the dominant motion artifact is blur — an artifact that is diminishing with refresh rate. There is also a clear preference for lower velocities; the explanation here is two-fold: (1) higher velocities produce more motion blur when displayed at a fixed refresh rate, and (2) content is easier to see at lower velocities (regardless of the refresh rate) so we may have a preference for those. Point (2) can also explain why the velocities differ more in quality at high than at low refresh rates. At high refresh rates, when motion blur is small, the observers are picking slower motion. However, at low refresh rates, when judder is a
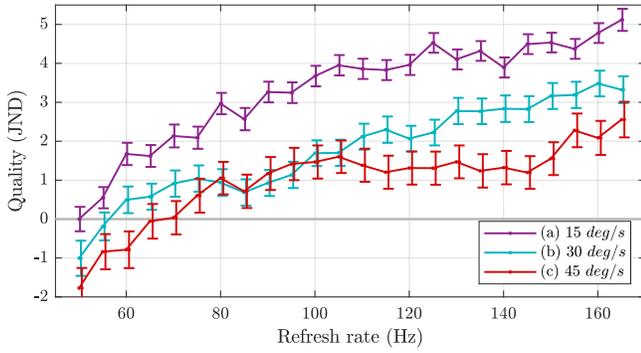
Fig. 7. Quality across different velocities. +1 JND distance indicates prefer-ence by 75% of the population; error bars indicate 75% confidence intervals. Quality increases with refresh rate, but the increase slows down above 100 Hz. Lower velocities are perceived to be of higher quality. This is ex-pected as higher velocities induce more blur and requires higher refresh rates to reproduce.
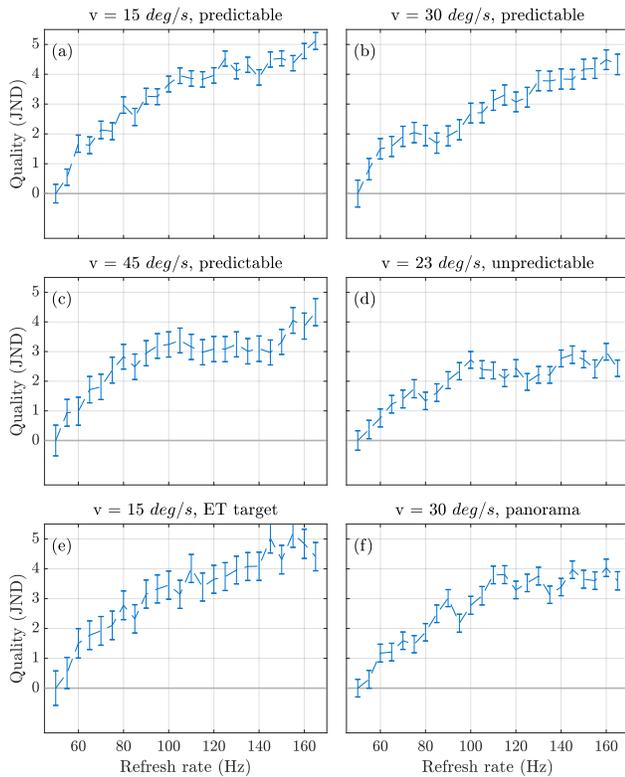


Fig. 8. Results of the motion quality experiment for all six animations. (a-d): circle animation with different velocities; predictable and unpredictable mo-tion. Velocity in title indicates the maximum velocity during the animation. Each curve was anchored such that 50 Hz corresponds to 0 JND.

dominant artifact, the velocities are more difficult to differentiate (all motion looks bad) and the differences in quality are becoming smaller. The predictability of motion influences the shape of the quality curve (Figure 8). On the other hand, image content does

not seem to be a strong factor, as the quality curves for circles and panorama (15 deg/s) look comparable. Similar observations can be made about circles and panorama (30 deg/s). From this, we identify velocity and the predictability of the motion as the key factors of our model. We propose that the quality measured in this experiment can be explained by motion blur and judder.

## 5 A PERCEPTUAL MODEL FOR MOTION QUALITY

In this section we present a perceptual visual model that predicts the effects of refresh rate, resolution, velocity and the type of movement on the perceived quality.

Formally, we define the content-independent quality difference between rendering on display A and display B, each using different spatial resolution and refresh rates:

$$\Delta Q(\ldots) = \Delta Q(f_A, R_A, f_B, R_B, v, \tau), \tag{2}$$

where the quality difference $\Delta Q$ is a function of display refresh rate $f$ (Hz), image resolution $R$ (pixels per degree; ppd), velocity of motion $v$ (deg/s), and predictability of motion $\tau$ (binary input, *predictable* or *unpredictable* SPEM). The unit of the $\Delta Q$ function is JND, as discussed in Section 2.6. When a display is rendering at a reduced resolution, we assume an image is up-sampled to the full screen size using a bilinear filter.

For an overview of our proposed model pipeline, see Figure 9. We approximate $\Delta Q$ as the weighted sum of three components:

$$\begin{aligned} \Delta Q(\ldots) = &w_P \, \Delta Q_P(f_A, R_A, f_B, R_B, v, \tau) + \\ &w_O \, \Delta Q_O(R_A, R_B) + \\ &w_J \, \Delta Q_J(f_A, f_B, v, \tau); \end{aligned} \tag{3}$$

*i.e.*, the amount of blur in the direction of, or parallel to the motion ($\Delta Q_P$), blur orthogonal to the motion, determined by the spatial resolution of the display ($\Delta Q_O$), and the judder or non-smoothness of the motion ($\Delta Q_J$). The following sections describe the steps to derive $\Delta Q_O$ and $\Delta Q_P$, then Section 5.4 describes the model for $\Delta Q_J$.

### 5.1 Blur due to spatio-temporal resolution and eye motion

The first step is to determine the loss of quality caused by the motion blur and the reduction of resolution. We separate the effect of refresh rate and resolution into three blur components: display hold-type blur ($b_D$), eye motion blur ($b_E$), and spatial blur due to the finite screen resolution ($b_R$). We express the amount of each blur as the width of either a box or a triangle filter in visual degrees.

*5.1.1 Hold-type motion blur ($b_D$).* When the eye follows a mov-ing object, its motion is continuous, whereas LCD displays can only present a sequence of discrete samples (frames) at a finite re-fresh rate. Current LCD displays do not necessarily emit a constant amount of light throughout a frame; however, as the transition periods have been decreasing in the recent years, and the exact transition profiles are complex, we follow the same practice as Klompenhouwer et al. [2004], and approximate hold-type blur with a box filter of the width (in visual degrees):

$$b_D = \frac{v}{f}, \tag{4}$$

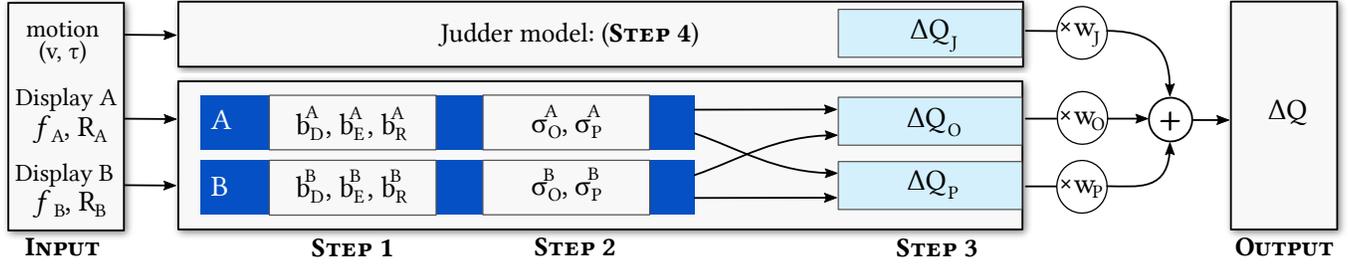where $v$ is the object velocity in degrees per second and $f$ is the refresh rate in Hz.

Fig. 9. Schematic diagram of the model predicting the quality difference between two refresh rates $(f_A, f_B)$ and resolutions $(R_A, R_B)$ assuming the same motion for both refresh rates. Step 1: separately for $A$ and $B$, compute blur factors due to hold-type display, eye motion, and resolution. Step 2: transform to blur kernels orthogonal and parallel to the motion. Step 3: Compute difference between $A$ and $B$, and apply non-linearity (CSF, psychometric function, probability to JND transform) to find $\Delta Q_O$ and $\Delta Q_J$. Step 4: Quality differences due to judder. Each step is described in the corresponding subsections of Section 5.

*5.1.2 Eye motion blur ($b_E$).* When the eye follows an object with SPEM, the tracking is imperfect. As discussed in Section 2, the difference between the object velocity and the gaze velocity is proportionate to the object velocity [Daly 1998]. Hence, such blur can be also modeled as a box filter with width:

$$b_E = p_a\, v + p_b, \tag{5}$$

where $p_a$ and $p_b$ are constant coefficients. We assume this eye motion blur to be independent of the display refresh rate, but expect it to vary with the predictability of motion. Therefore $p_a$ and $p_b$ are different for predictable and unpredictable motions, as demonstrated with experimental data in Section 6.1.

*5.1.3 Spatial resolution blur ($b_R$).* With the general use of bilinear filters for up-sampling images in real-time graphics, the blur due to reduced spatial resolution is well-modeled by a triangle filter with average width $b_R$ (or base width $2b_R$). Given the angular resolution $R$ in pixels per visual degree, the width of the filter is

$$b_R = \frac{1}{R}. \tag{6}$$

## 5.2 Motion-parallel and orthogonal blur

To simplify the combination of different blur types, we approximate each blur component with a Gaussian filter (see Figure 10). A box filter can be approximated with a Gaussian filter of the standard deviation:

$$\sigma = \frac{w}{\pi}, \tag{7}$$

where $w$ is the width of the box filter. This implies:

$$\sigma_D = \frac{v}{\pi f}, \quad \sigma_E = \frac{p_a v + p_b}{\pi}. \tag{8}$$

The triangle filter, used to model the resolution reduction, can be considered as the convolution of two box filters with base width $b_R$. The standard deviation of this combined kernel is then

$$\sigma_R = \sqrt{\left(\frac{b_R}{\pi}\right)^2 + \left(\frac{b_R}{\pi}\right)^2} = \frac{\sqrt{2}b_R}{\pi}. \tag{9}$$

Eye-motion blur ($b_E$), and hold-type blur ($b_D$) will blur the image only in the direction of motion, but lowering spatial resolution ($b_R$) will blur the image equally in all directions. Because of that, we
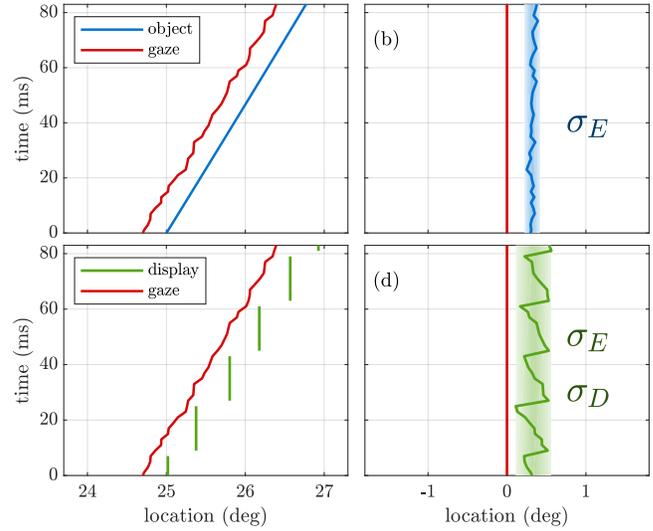
Fig. 10. Combining eye blur ($b_E$) and hold-type display blur ($b_D$). Left column shows object and gaze location in absolute screen co-ordinates; right column shows the same data relative to gaze (i.e. retinal location). Object location (top row) is followed by imperfect eye motion. This introduces eye motion blur, the magnitude of which can be estimated with $\sigma_E$. Displayed object location is also affected by display hold-type behavior ($b_D$). The combined effect of these are shown in the bottom row. Data based on eye tracker measurements (Section 6.1) on 55 Hz monitor.

separately compute the blur that is parallel (P) to the direction of motion and the one that is orthogonal (O) to the direction of motion, as shown in Figure 11.

The blur in the direction parallel to motion ($\sigma_P$) is given by the convolution of individual components:

$$\sigma_P = \sqrt{\sigma_E^2 + \sigma_D^2 + \sigma_R^2}; \tag{10}$$

and the blur that is orthogonal to the direction of motion ($\sigma_O$) is affected only by the resolution reduction:
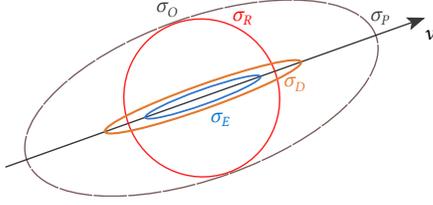
$$\sigma_O = \sigma_R. \tag{11}$$

Fig. 11. Blur is anisotropic. For a given motion in the direction $v$, we distinguish between motion-parallel blur ($\sigma_P$) and motion-orthogonal blur ($\sigma_O$). Motion-parallel blur ($\sigma_P$) consists of resolution reduction blur ($\sigma_R$; red), eye motion blur ($\sigma_E$; blue), and hold-type display blur ($\sigma_D$; orange). Motion-orthogonal ($\sigma_O$) blur consists only of blur due to resolution reduction ($\sigma_R$; red). $\sigma_D$ and $\sigma_E$ ellipses are drawn for visualization only, as we assume these sources of blur to be one-dimensional.

## 5.3 From $\sigma$ to quality

Blur introduced by eye motion, hold-type blur, and spatial resolution will result in the loss of sharpness. To quantify this in terms of loss of perceived quality, we map the physical amount of blur to the perceived quality difference in JND units. Our blur quality function is inspired by the energy models of blur detection [Watson and Ahumada 2011]. Such mapping is applied to the orthogonal ($\sigma_O$) and parallel ($\sigma_P$) components of the anisotropic blur separately, resulting in two independent quality values ($Q_O$ and $Q_P$).

As we are interested in content-independent predictions, we assume the worst-case scenario: an infinitely thin line (Dirac delta function $\delta(x)$), which contains uniform energy across all spatial frequencies. When convolved with a Gaussian blur kernel $\sigma$ in the spatial domain, the resulting image is a Gaussian function with standard deviation $\sigma$. The Fourier transform of this signal is also a Gaussian, given by:

$$m(\omega; \sigma) = \exp\left(-2\pi^2 \omega^2 \sigma^2\right) \tag{12}$$

where $\omega$ is in cpd. To account for the spatial contrast sensitivity of visual system, we modulate the Fourier coefficients with the CSF

$$\tilde{m}(\omega, \sigma) = \text{CSF}(\omega)\, m(\omega; \sigma), \tag{13}$$

where CSF is Barten's CSF model with the recommended standard observer parameters and the background luminance of 100 cd/m$^2$ [Barten 2003].

To compute the overall energy in a distorted signal, we sample a range of frequencies $\omega_i = \{1, 2, \ldots, 64\}$ [cpd], and compute the blur energy as:

$$E_b(\sigma) = \sum_i \left(\frac{\tilde{m}(\omega_i; \sigma)}{\tilde{m}_{t,b}}\right)^{\beta_b}. \tag{14}$$

where $\tilde{m}_{t,b}$ is the threshold parameter and $\beta_b$ is the power parameter of the model. Both of these are fitted to psychophysical data in Section 6.3.

Energy differences can be interpreted as quality differences, yielding:

$$\begin{aligned}\Delta Q_P &= E_b(\sigma_P^A) - E_b(\sigma_P^B),\\ \Delta Q_O &= E_b(\sigma_O^A) - E_b(\sigma_O^B),\end{aligned} \tag{15}$$

substituting in the standard deviations of the blur components for $A$ and $B$, in the directions parallel (P) and orthogonal (O) to SPEM.

We further explain why an energy model is suitable to predict JND differences in the supplemental material (Section S.1).

## 5.4 Judder ($\Delta Q_J$)

On lower refresh rates, finite sampling results in non-smooth, juddery motion. As described in Section 2.4, the visibility of judder can be predicted by transforming the signal to the frequency domain, and examining aliasing copies of the original signal (see Figure 3).

The location of the first aliasing copy, as shown in Figure 12-Left can be determined as follows: the temporal frequency (vertical axis) is equal to the refresh rate; the spatial frequency ($\rho$, horizontal axis) is:
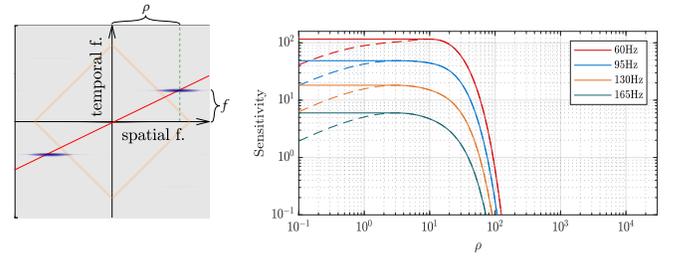
$$\rho = \frac{f}{v}. \tag{16}$$



Fig. 12. Left: the visibility of judder artifacts are determined by the location of the first aliasing copy in the frequency domain (highlighted in blue). The peak of the aliasing copy lies on the red line which in turn is determined by the spatial frequency and the object velocity. Right: spatio-temporal CSF used in the judder model. Kelly's model predicts lower sensitivity values at low spatial frequencies (dashed); in our model, we clamp this conservatively (solid). Colors show different temporal frequencies.

Given two refresh rates $f_A$ and $f_B$, we employ the same energy model architecture as for blur. The unit signal is modulated with the spatio-temporal contrast sensitivity of the eye (stCSF), and normalized by a threshold modulation:

$$E_j(f, v) = \left(\frac{\text{stCSF}(\rho, f)}{\tilde{m}_{t,j}}\right)^{\beta_J}, \tag{17}$$

where $\beta_J$ is the power parameter for judder, and $\tilde{m}_{t,j}$ is the threshold for judder. The threshold is fitted separately for predictable and unpredictable motion. stCSF is Kelly's spatio-temporal CSF [Kelly 1979]; however, to account for the finite width of the alias, we use a truncated low-pass stCSF, as shown in Figure 12-right. As with blur, we express the quality difference due to judder as the difference of energy:

$$\Delta Q_J = E_j(f_A, v) - E_j(f_B, v). \tag{18}$$

## 6 MODEL CALIBRATION

To determine the free parameters of our model, we collected data on eye motion and perceived judder in the following two experiments.
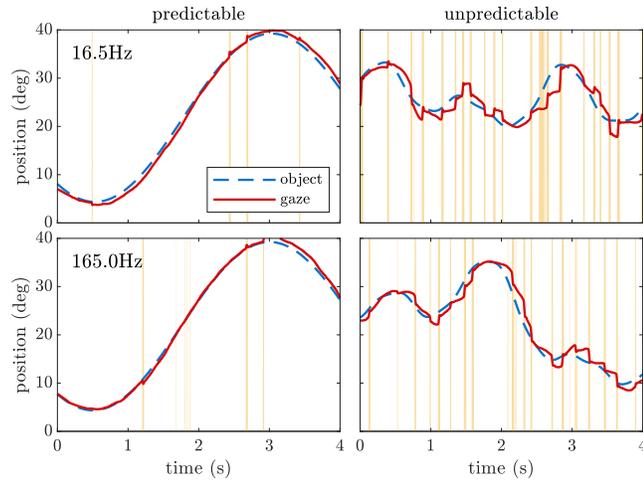
Fig. 13. Traces of gaze location during SPEM of predictable (left) and unpredictable (right) objects at different refresh rates. Vertical yellow lines show interruptions in SPEM (saccades). Unpredictable motion visibly requires more correction saccades, with the gaze lagging behind object motion. Oscillations comparable to the respective display refresh rates are not visible.

## 6.1 Experiment 2: Retinal blur due to motion

Eye motion blur is caused by the differences between object and gaze motion. Daly et al. [1998] suggested that the difference in velocity and therefore also blur amount ($b_E$) can be modeled as a linear function of object velocity within the SPEM-tracking range. There is, however, little data on how this function might change with unpredictable eye motion, and how to incorporate interaction with display refresh rates. To explore this problem and to fit the linear parameters ($p_a, p_b$) of our model described in Section 5.1, we measured the eye's ability to follow predictable and unpredictable objects with an eye-tracker.

*6.1.1 Stimuli.* We used the eye tracker target from Experiment 1 (bull's eye and a cross hair; Figure 5). This object moved left-to-right with predictable or unpredictable motion. For predictable motion the horizontal displacement followed a sinusoidal function with the amplitude of 17° and four different frequencies to give a peak velocity of {12, 18, 24, 36} deg/s. For unpredictable motion we used the same motion as Experiment 1 (Equation 1). The stimuli were rendered at a range of refresh rates $T_i = \{16.5, 27.5, 55, 60, 82.5, 120, 165\}$ Hz.

*6.1.2 Setup.* The fixation target was displayed on an ASUS PG279Q monitor with an Eyelink II eye-tracker sampling the gaze location at 500 Hz (pupil-only mode).

*6.1.3 Procedure.* Participants were asked to follow the fixation target with their gaze. Their head was stabilized on a chinrest 80 cm away from the monitor (field of view of 41°). Each session consisted of 30 trials, each trial lasting 20 s. We performed a binocular 9-point calibration before each trial, selecting the eye that performed better during the 9-point validation. The order of trials was randomized.
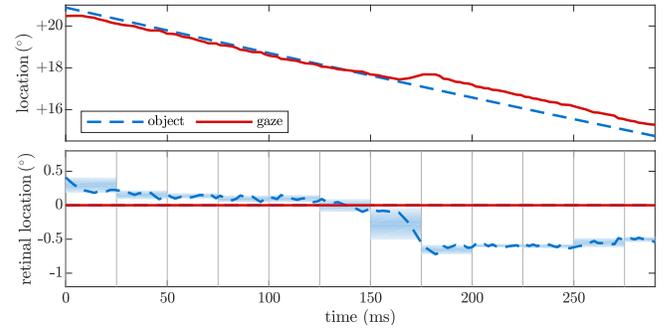
Fig. 14. Eye blur estimation from eye tracking data. Top: gaze does not perfectly follow the target object. Bottom: retinal location is computed as the difference between gaze position and object position. The data is then split into 25-ms windows; for each window, $b_E$ is estimated as the difference between the maximal and minimal retinal location.

*6.1.4 Participants.* Five participants aged 20–27 volunteered to take part in the experiments. Four participants had normal vision, while one participant wore prescription contact lenses.

*6.1.5 Results.* Figure 13 shows examples of measured traces on different refresh rates for predictable and unpredictable motion. The eyetracker reported blinks, and we ignored these blinks in the analysis. Velocities were obtained by discrete differentiation; saccades were then filtered out using a threshold method when either eye velocity exceeded 40 deg/s or acceleration exceeded 9000 deg/s². We verified that the results of the threshold method corresponded to manual labeling. A detailed report of the analysis is provided in the supplemental material (Section S.2). Our results can be summarized as follows: SPEM is not affected by refresh rates above 27.5 Hz, but is affected by the nature of motion (predictable vs. unpredictable). Specifically, saccades are more frequent during unpredictable motion than predictable motion (4.89 vs. 1.56 saccades per second). The delay when tracking unpredictable motion is also significantly higher (0.092 s vs. 0.004 s). We therefore fit the same model parameters ($p_a$, $p_b$) for all refresh rate, but separately for the two motion types.

*6.1.6 Analysis.* To estimate the amount of blur due to eye motion, $b_E$, we split the recorded gaze location traces into segments corresponding in duration to the integration time of the eye. In this analysis, we use 25 ms windows, *i.e.*, the inverse of the approximate foveal flicker fusion frequency [Simonson and Brozek 1952]. Within each integration window, we estimate eye motion blur ($b_e$) as the difference between two extreme retinal positions of an object within the window, effectively measuring the width of the box filter (see Figure 14). To reduce measurement noise, the blur width was averaged for all windows with matching refresh rate and (binned) target velocity.

*6.1.7 Fitting parameters $p_a$ and $p_b$.* Parameters were fitted to minimize the root-mean-squared-error between model predictions from Equation 5 and the average blur values measured in this experiment. Figure 15-top indicates the measured linear relationship between object velocity and $b_E$. The common velocity gain of 0.82 [Daly 1998]
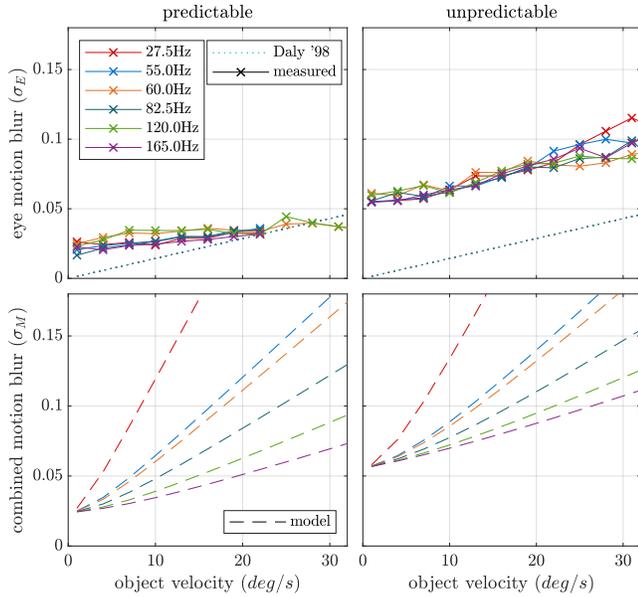
Fig. 15. Top: blur $\sigma$ based on eye tracker data for a range of refresh rates (different colors) and object velocities (x-axis). Blur was computed over 25ms intervals as the distance traveled by the tracked object on the retina. Dotted line: Daly's model. Bottom row: model predictions for blur taking into account both eye motion and display hold-type behavior.

Table 1. Blur model parameters. For details, see text and Figure 15.

|  | $p_a$ | $p_b$ |
|---|---|---|
| Predictable | 0.001648 | 0.079818 |
| Unpredictable | 0.004978 | 0.176820 |

would yield a linear gradient of $p_a = 0.0045$ under the 25ms integration window, which our data for unpredictable SPEM agrees with. However, for predictable motion, our results indicate much more accurate tracking (and hence less blur). For the fitted parameter values (RMSE=0.02), see Table 1.

## 6.2 Experiment 3: Quality loss due to judder

To fit the parameters of the judder model ($\tilde{m}_{t,j}$: energy threshold; $\beta_J$: power parameter), we had to isolate judder from hold-type blur and measure its impact on quality. As one cannot easily remove blur when showing juddery motion at low refresh rates, we instead did the opposite: generated smooth (high refresh rate) and juddery motion (low refresh rate) and artificially introduced blur so that its amount was the same in both conditions.

*6.2.1 Setup.* We used the same setup as Experiment 1 with no eye tracker.

*6.2.2 Stimuli.* Similar to Experiments 1, participants observed predictable or unpredictable horizontal motion, following a fixation target or a checkered circle (Figure 5 right). In the split-screen setup, target 1 was rendered with the refresh rate of $f$ Hz, while target 2 was rendered with twice of that refresh rate ($2f$ Hz) and with simulated hold-type blur so that it was matching the hold-type blur
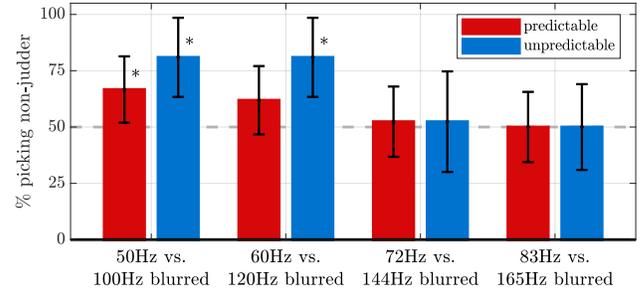


Fig. 16. The probability of selecting the animation with reduced judder (double the refresh rate) but the same amount of motion blur, at a range of refresh rates for predictable (red) and unpredictable motion (blue). Observers were unable to tell the difference between juddery and non-juddery animations above 60 Hz. Error bars denote 95% confidence intervals.

Table 2. Judder model parameters from Section 6.2.

|  | $\tilde{m}_{t,j}$ | $\beta_J$ |
|---|---|---|
| Predictable | 218.712 | 2.5747 |
| Unpredictable | 165.779 | |

seen at $f$ Hz. In practice this was achieved by rendering all content at $2f$ Hz, repeating the frame in the temporal domain for target 1, and overlaying two offset frames for target 2. The spatial offset was computed as $v/(2f)$, where $v$ was the actual velocity of the object.

*6.2.3 Task.* Participants were asked to follow the fixation target with their gaze, then select the animation that provided *smoother* motion. They could view each trail for up to 20 s with the option to replay if needed. Each of the eight voluntary participants completed 108 comparisons.

*6.2.4 Results.* The probability of detecting judder is shown in Figure 16. Judder was detectable for both predictable and unpredictable motion at 50 Hz and 60 Hz. At 72 Hz and 83 Hz the observers could not discriminate between the animations. This indicates that the effect of judder on quality is negligible at 72 Hz and higher refresh rates. Judder was easier to detect for unpredictable motion.

*6.2.5 Model fitting.* As explained in Section 5.4, our measurements can be predicted by the energy difference in the spatio-temporal contrast sensitivity function. The best fit of Equation 17 to the measurement was obtained for the parameters listed in Table 2. The RMSE of the model predictions considering both predictable and unpredictable motion was 0.1074 JND.

## 6.3 Fitting the quality predictions

To find the final weights of the model, we minimized RMSE between the scaled results of Experiment 1 and the predictions of the visual model for the *circle* scene. The data included 3 velocities for predictable motion and one for unpredictable motion (a–d in Figure 17). A power of $\beta_B = 1.83564$ and a threshold value of $\tilde{m}_{t,b} = 383.5854$ provided the best fit with RMSE=0.312. The relative weights of judder and blur indicated that the judder component (when present) is a more significant contributor. Fitting the last parameter of the
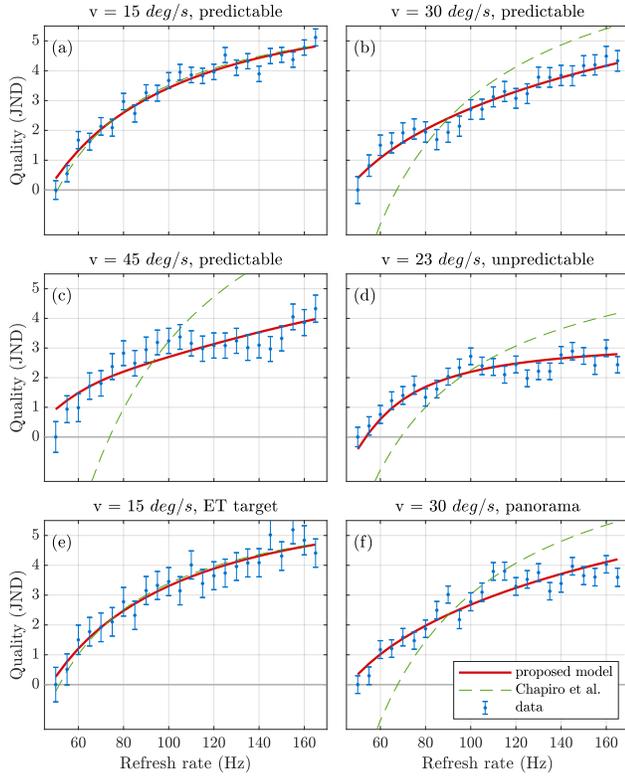
Fig. 17. Fitted model predictions (red lines) against measurement data (blue error bars; 75% confidence). Our model parameters were fitted for the checkered circle scene (a–d). There is a distinct difference in the shape of the quality curves for different velocities (a–c) and predictable vs. unpredictable motion (d). The bottom row shows that predictions are consistent for the eye tracker target and the panorama scenes as well. The empirical model of Chapiro et al. (green dashed) provdes an excellent fit for low object velocity (15 deg/s), but fails for higher velocities.
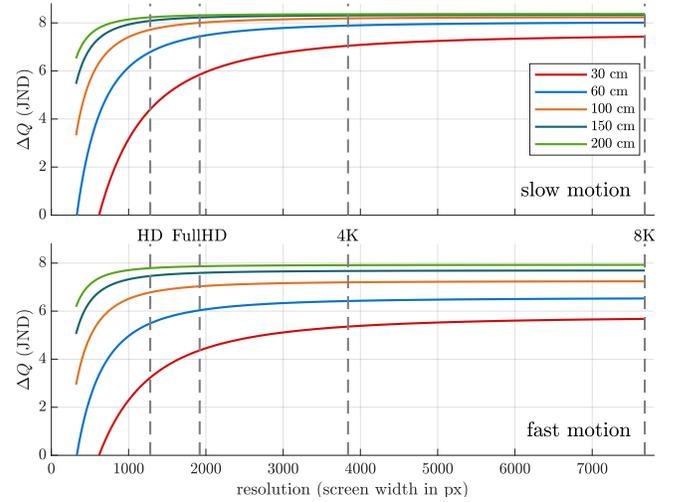


Fig. 18. Predictions for perceived quality on a 15" display with varying resolutions (horizontal axis) and viewing distances (colors). Top: slow panning motion with the content moving horizontally across the entire screen in 6.2 s; Bottom: fast motion with content moving across the entire screen in 1.5 s. Higher resolutions bring diminishing quality gains, especially when viewed from far. Closer viewing distances also result in higher angular velocities with more visible motion artifacts.

Table 3. Model parameters. For details, see text.

| $w_J$ | $w_P$ | $w_O$ | $\tilde{m}_{t,b}$ | $\beta_B$ |
|---|---|---|---|---|
| 2.218677 | 1.472819 | 1.472819 | 383.5854 | 1.83564 |

model, the weight of orthogonal blur relative to parallel blur and judder ($w_O$), requires careful observation and comparison of spatial blur and motion artifacts. High $w_O$ values bias the model to reject resolution reductions, while low $w_O$ values result in insensitivity to orthogonal blur in slowly-moving images. The stimuli in Experiment 1 did not contain spatial blur orthogonal to the direction of the motion; however, the fitted value of $w_P$ provides a reasonable starting point, as both $\Delta Q_P$ and $\Delta Q_O$ consider artifacts due to spatial blur. An expert observer verified on the content of Experiment 1 with $0 - 80$ deg/s that $w_O = w_P$ provided consistent quality. Further experiments on such high velocities are challenging for an average observer, but could further refine this estimate. In the next section we consider the predictions of the model, then propose and validate an application showing that the collection of parameters together can predict a good trade-off between resolution and refresh rate.

### 6.4 Comparison with the model of Chapiro et al.

Figure 17 shows the empirical model of Chapiro et al. in green. It must be noted that the maximum velocity measured in their study

was 6.6 deg/s while our minimum velocity was 15 deg/s, therefore, measurements are not directly comparable. For a better illustration, we aligned their model with our measurements at low velocities by linear rescaling of quality predictions. Their model almost perfectly matches our data for the velocity of 15 deg/s. However, it is also clear that their model cannot extrapolate predictions for higher velocities, nor can it distinguish between predictable or unpredictable motion. For a fair comparison, we refitted their model to our data by linearly rescaling the quality and reported results in Table 4. Their functional model does not seem to improve the fit over a fitted logarithmic function of refresh rate: $p_1 \log(p_2 f)$.

### 6.5 Ablation study

To justify the importance of each component of our visual model, we perform an ablation study. We isolate three key features: parallel quality ($\Delta Q_P$), judder ($\Delta Q_J$), and the isolation of unpredictable vs. predictable motion (pred). We refit the model to the checkered circle scene for each combination of features, minimizing the RMSE error in linear JND space. We report the goodness of fit (RMSE) for the training set (*circle* scenes), and the RMSE for the eye tracker target and panorama scenes as a restricted test set. Orthogonal blur cannot be separated in this study, as Experiment 1 did not manipulate orthogonal resolution.

Our results as presented in Table 4 indicate that the judder model ($\Delta Q_J$) on its own provides a poor fit to the quality curves (RMSE>0.51).

Table 4. RMSE for different combination of model components. Stimuli are labeled as in Figure 17. *pr.* indicates whether the model distinguishes between predictable and unpredictable motion.

| features | train | test | (a) | (b) | (c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|---|---|
| Chapiro et al. | 0.54 | 0.64 | 0.82 | 0.39 | 0.54 | **0.27** | 0.83 | 0.37 |
| log | 0.41 | 0.42 | 0.44 | 0.29 | 0.44 | 0.46 | 0.44 | 0.40 |
| $\Delta Q_P$ | 0.42 | 0.43 | 0.45 | 0.28 | 0.46 | 0.48 | 0.46 | 0.40 |
| $\Delta Q_P$, pr. | 0.36 | 0.38 | 0.30 | 0.27 | 0.50 | 0.30 | 0.33 | 0.43 |
| $\Delta Q_J$ | 0.51 | 0.58 | 0.57 | 0.66 | 0.36 | 0.39 | 0.63 | 0.51 |
| $\Delta Q_J$, pr. | 0.67 | 0.73 | 0.78 | 0.74 | 0.46 | 0.67 | 0.81 | 0.65 |
| $\Delta Q_P$, $\Delta Q_J$ | 0.36 | 0.35 | 0.28 | 0.36 | **0.33** | 0.45 | 0.34 | **0.35** |
| $\Delta Q_P$, $\Delta Q_J$, pr. | **0.31** | **0.34** | **0.26** | **0.26** | 0.42 | 0.29 | **0.30** | 0.38 |

The parallel quality factor ($\Delta Q_P$) captures some trends, but cannot correctly distinguish between varying object velocities. Best predictions are provided when all model features are enabled (RMSE=0.31). Quantifying the significance of each component is non-trivial; further ablation is available in the supplemental material (S.3).

## 6.6  The effect of resolution

One advantage of our model is that we can extrapolate our findings to different screen resolutions and viewing distances. In Figure 18 we show how the perceived quality of slow (top) and fast (bottom) panning motion changes with the screen resolution (x-axis) and viewing distance. As expected, an increased screen resolution brings diminishing returns when viewed from far, and the motion looks worse from a close distance because of higher retinal velocity.

## 7  APPLICATION

G-Sync capable monitors offer the freedom to refresh the monitor at arbitrary rates and without introducing tearing artifacts. However, under limited rendering budget, this may result in images that are sharp but juddery if the resolution is too high, or blurry but smooth animation if the resolution is too low. We propose a motion-adaptive resolution and refresh rate (MARRR) rendering algorithm, where the quality predictions of the visual model are used in real-time to establish the relative quality of different configurations of refresh rate and resolution for a fixed rendering budget bandwidth. We follow [Debattista et al. 2018] and express this as an optimization problem:

$$\underset{R,f}{\operatorname{argmax}} \ \Delta Q\left(R, f, R_\kappa, f_\kappa, v, \tau\right) \quad \text{s.t.} \quad R\,f\,\Phi\,\Theta \leq B \ \wedge \ f \geq 50\,Hz$$

$$(19)$$

where $B$ is the rendering budget in pixels per second, $\Phi$ and $\Theta$ are the horizontal and vertical viewing angles of the monitor respectively. Quality is a relative value, so we anchor to resolution $R_\kappa$ and refresh rate $f_\kappa$. The optimal refresh rate will be dependent on the current object velocity, and hence, does not necessarily remain constant throughout the animation sequence. We found that the choice of the anchor point did not have a significant impact on predictions, with $f_\kappa = 150\,Hz$ and $R_\kappa = B/(150\,\Phi\,\Theta)$ yielding stable results.

Figure 19 shows the complex shape of model predictions for the ASUS PG279Q display at fixed viewing distance (108 cm). For high
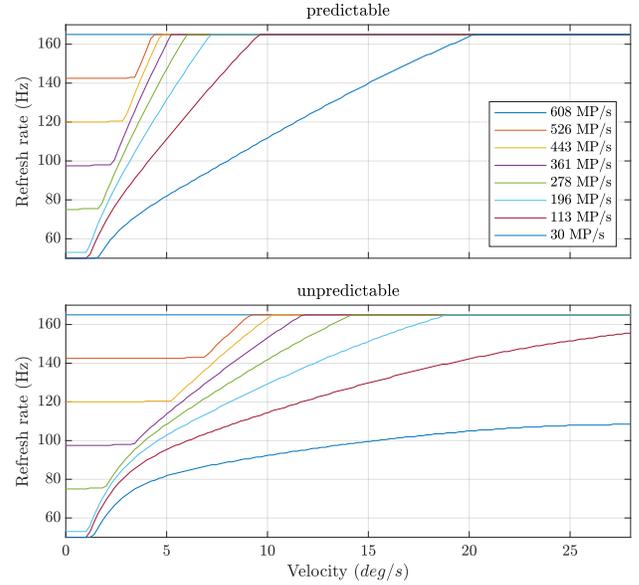


Fig. 19. Model predictions for different rendering bandwidths (colors; measured in Megapixels per second) for predictable (top) and unpredictable motion (bottom). The plots show only the refresh rate as the resolution is determined by the fixed rendering budget.

budgets ($> 443$ megapixels-per-second; MP/s), the model recommends keeping the refresh rate and the resolution constant up to a certain velocity and then to gradually increase the refresh rate at the cost of the resolution. The transition is more gradual for smaller rendering budgets and unpredictable motion, with slower increase in refresh rate.

### 7.1  Real-time implementation

To avoid solving an optimization problem (Equation 19) for each frame, the relation between the pixel budget ($B$), velocity ($v$) and the optimum refresh-rate/resolution ($R$, $f$) can be precomputed as a look-up table (LUT). Two such LUTs, one for predictable and another unpredictable motion, are shown in Figure 19. In our experiments, we set the anchor frame rate to $f_\kappa = 150\,Hz$, and sample velocity once per deg/s.

### 7.2  Low-persistence displays:

Although our model was fitted and validated only on a high-persistence LCD display, we can easily extrapolate predictions to low-persistence displays, such as the ones used in VR/AR headsets. For this, we assume shorter integration time when estimating the amount of hold-type blur; we replace ($v/f$) with ($vp/f$) in Equation 4, where $p$ is the fraction of the frame when the display is on. The resolution/refresh rate plot in Figure 20 suggests that high-persistence (high percentage) demands higher refresh rates even at low velocities, whereas low persistence can keep the resolution higher under the same budget ($278\,MP/s$). Such a model could be potentially used to dynamically control the persistence of a display to avoid visible flicker.
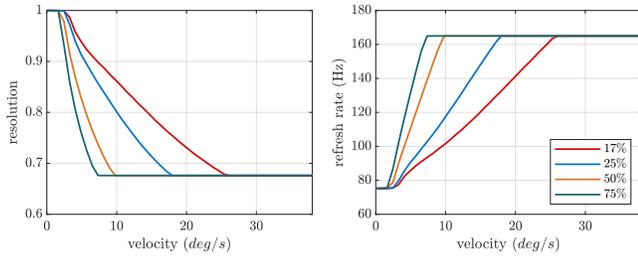
Fig. 20. Model predictions for different display persistence values. Colors denote the percentage of the frame duration when the display is on.
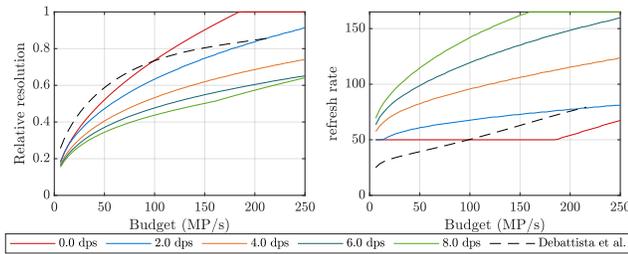


Fig. 21. Model predictions for predictable motion for different velocities (colors) in deg/s (dps) plotted against Debattista et al. [2018] (dashed line). Assuming viewing distance of 108 cm and a field of view of 30°. Resolution is relative to linear image size; Refresh rate predictions are for QHD resolution.

### 7.3 Comparison with Debattista et al.:

As discussed in Section 3.5, we build on this work by formulating the application as an optimization problem, but introduce an explainable visual model, which considers object velocity. In Figure 21, we show how the optimum resolution and refresh rate vary with the computational budget (x-axis). Although our model shows the same trends as current state-of-the-art, there are notable differences. (1) Our model, intended for real-time graphics, recommends overall higher refresh-rates and lower resolutions, especially when the velocity of motion is high. (2) For high budgets (>100 MP/s) and low velocities, our model recommends higher resolutions. (3) The results of Debattista et al. [2018] agree with the 0–2 deg/s curves the most, and differ significantly from the higher-velocity curves.

Note that the proposed model has several further advantages, such as adapting to any viewing distance, maximal display resolution and refresh rate. In the next section, we demonstrate that the inclusion of motion velocity results in subjective preference.

### 7.4 Experiment 4: Validation on controlled motion

To compare MARRR with the current state-of-the-art approach, we pick three computation budgets and their corresponding refresh rates from Debattista et al. [2018], and demonstrate that MARRR produces subjectively preferred results. The selected bandwidths were $B_i = \{28, 55, 221\}$ MP/s. To account for potential viewing condition differences between the experiment setups, we tested three fixed refresh rates on and around the reference values reported in [Debattista et al. 2018].



Fig. 22. Content of Experiment 4. © Marcell Szmandray

*7.4.1 Setup.* The experiment used a 2AFC design with the same setup as in Section 4 — two G-sync capable monitors stacked on top of each other. We implemented a custom C++ OpenGL application that allowed the users to scroll across a panorama image using either a mouse or predetermined motion. On one monitor, the renderer used a single refresh rate throughout the entire animation; on the other monitor, the renderer established the optimal refresh rate (from 50 Hz to 165 Hz) frame-by-frame according to our model. For this, we used a pre-computed look-up table. The application reduced rendering resolution to meet the budget requirements. The two monitors displayed the same content but at different resolutions and refresh rates. The mouse movement was synchronized over the network.

*7.4.2 Stimuli.* For content, we picked four high-quality panorama images (Figure 22). For predictable motion, the observer could pan the panoramas by moving the mouse. Such user-controlled motion is predictable and is similar to *e.g.* camera rotation in a first-person game or camera panning in real-time strategy or simulator games. For unpredictable motion, we used the same formula as previously (Equation 1). The experiment was more difficult to implement for unpredictable motion. The rapid changes in refresh rates combined with mis-predictions in the control system of G-Sync caused occasional skipped frames. Such motion flaw is not a limitation of the visual model or the algorithm, but the lack of ability to aid the G-Sync control system from an application side when picking the current refresh rate. To reduce these artifacts, we discretized the predicted refresh rates to integer divisors of 165 Hz: {55, 82.5, 165} Hz.
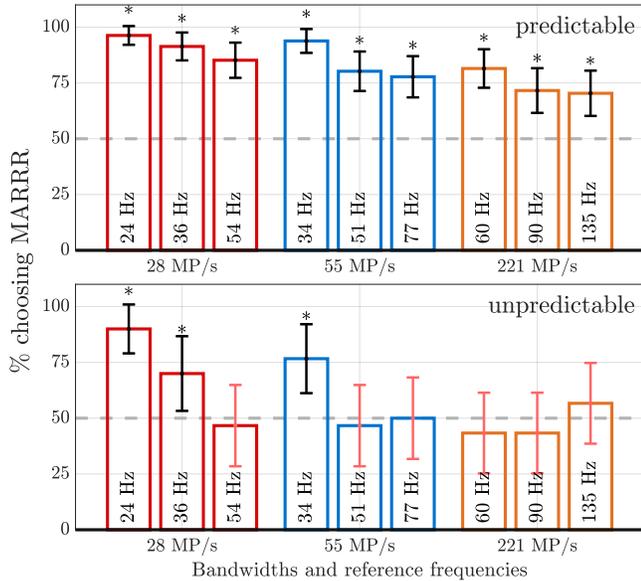
Fig. 23. Results of validation experiment showing the percentage of participants picking the proposed adaptive MARRR algorithm over standard *constant*-resolution-and-*constant*-refresh-rate rendering, viewing predictable (top) and unpredictable (bottom) motion. Colors denote different rendering budgets. The refresh rates were selected around the predictions of [Debattista et al. 2018]. Error-bars denote 95% confidence intervals.

*7.4.3 Participants and procedure.* Nine voluntary participants took part in the experiment (aged 20–40) with normal or corrected-to-normal vision. Participants were asked to imagine a scenario where they were purchasing a new monitor for playing computer games, and for each trial to pick either the top or the bottom monitor based on overall visual quality (including motion and sharpness). The order of the comparisons and the presentation on the monitors were randomized. Each observer completed 81 and 27 comparisons for predictable and unpredictable motion, respectively. All participants reported only casual gaming experience (playing only a few times a month) with little-to-no exposure to high-refresh-rate monitors.

*7.4.4 Results.* The results of the validation experiment, shown in Figure 23, indicate that inclusion of motion velocity resulted in an overall preference for MARRR as compared to the fixed rates from [Debattista et al. 2018]. The difference is consistently significant for user-induced (predictable) motion: MARRR was picked with over a 70% probability. The trend indicates that the impact of activating adaptive rendering was lower for higher bandwidths, which is consistent with expectations. For unpredictable motion, MARRR provided better overall results, but for high-refresh-rate conditions the experiment was inconclusive. Better synchronization capabilities with the monitor should allow for a less noisy comparison in the future.

## 7.5 Experiment 5: Validation in a video game

To validate the applicability of our model in more complex motion scenarios, we built a sandbox video game running our proposed
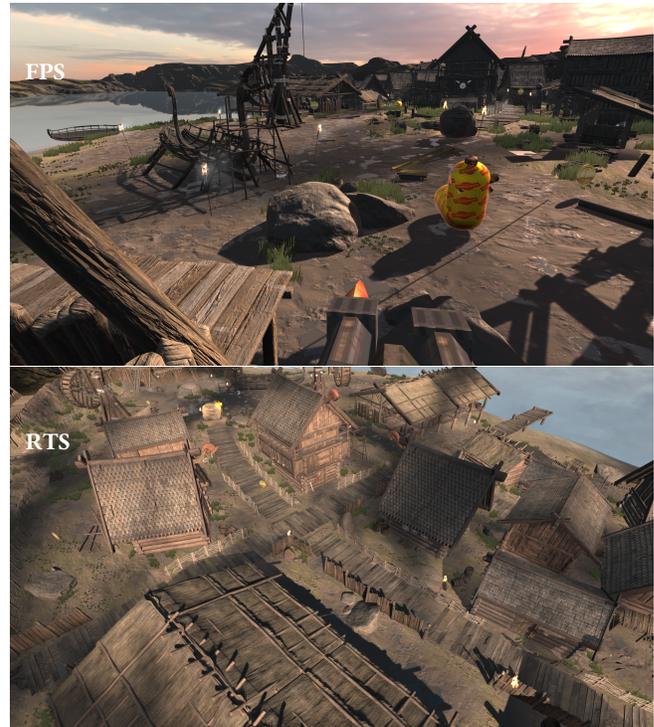


Fig. 24. Viking village stimulus used in Experiment 5 with FPS camera (top) and RTS camera (bottom). © Unity Technologies

MARRR algorithm and ran a preference experiment to compare with a fixed refresh-rate rendering.

*7.5.1 Setup.* This experiment used a pairwise comparison design where the conditions were played on a single G-sync capable ASUS display. Users had the freedom to toggle between the two conditions but did not know which condition was which. We built the video game in Unity3D and integrated the same eye-tracking setup as in Experiment 2 (Section 6.2). In each trial, the two conditions involved in the same 3D scene and the same fixed bandwidth (1/3rd of maximum available bandwidth: 203 MP/s) but one with MARRR and other with a constant resolution and refresh rate. MARRR assumed predictable motion and established the optimal refresh rate (from 50 Hz to 165 Hz) frame-by-frame by querying a pre-computed look-up table. Refresh rates chosen for constant resolution-refresh rate were 30 Hz, 60 Hz and 120 Hz to mimic conventional gaming hardware.

*7.5.2 Stimuli.* For video game content, we used the Viking Village scene from Unity. We added two camera modes to this 3D environment: (1) FPS mode where players could explore the scene and interact with objects such as collecting coins and shooting enemies similar to a typical first-person shooter-style video game (Figure 24-top). (2) RTS mode where the players could explore the scene from a bird's eye view similar to a simulator and real-time strategy style video games (Figure 24-bottom). The trials were randomized with an equal number of FPS and RTS conditions.
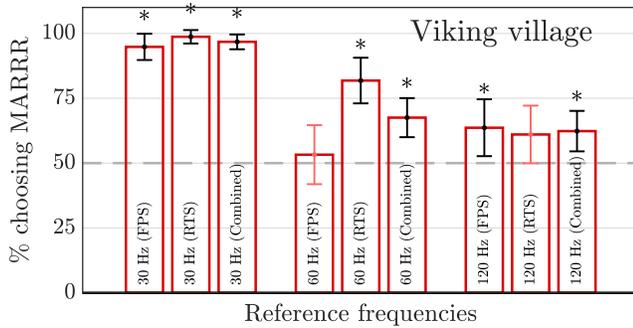
Fig. 25. Results of validation experiment in a video game setting showing the percentage of participants picking the proposed adaptive MARRR algorithm over standard *constant*-resolution-and-*constant*-refresh-rate rendering. The notation is the same as in Figure 23. Our method consistently performed better across all conditions except at FPS-60Hz, where the results are not statistically significant. We believe the latter could be due to sample size and participants' habituation to FPS gaming at 60Hz.

*7.5.3 Estimating velocity.* To estimate object target velocity for our model in real-time, we used the gaze velocity averaged over 17 samples as reported by the EyeLink SDK. When there was no camera movement, this velocity was prone to be noisy. Hence, when there was no mouse or keyboard interaction, we instead reverted to screen-space object velocities at the gaze location as reported by the game engine.

*7.5.4 Participants and procedure.* 11 people (aged 20-40) volunteered to take part in the experiment. All had normal or corrected-to-normal vision. In each trial, participants were asked to explore the 3D scene in both rendering conditions and choose the one with overall better visual quality. The order of the comparisons was randomized. Each observer completed 42 comparisons. All participants reported only casual gaming experience (playing only a few times a month) with little-to-no exposure to high-refresh-rate monitors.

*7.5.5 Results.* Similar to Experiment 4, the results of this experiment indicate an overall preference for MARRR as compared to the fixed-refresh-rate rendering (Figure 25). At the chosen bandwidth, the difference in preference is particularly large for 30 Hz. The difference is significant, but less pronounced for comparisons with 60 and 120 Hz. We believe this could have been caused by the complexity of the scene with many factors such as aliasing and eye-tracking noise affecting the results.

*7.5.6 Additional results.* Experiment 5 compared MARRR to conventional refresh rates ({30, 60, 120} Hz). Results in Debattsita et al. [2018] suggest rendering the scene at a fixed refresh-rate between 60 Hz and 90 Hz, which we were unable to test due to the COVID-19 lockdown. Those additional results, comparing directly to the Debattsita et al. [2018] method with the game setup for both FPS and RTS, will be made available on the project web page[1] and in a revised version of the paper, once such experiments can be safely conducted.

## 8 CONCLUSIONS AND FUTURE WORK

As display capabilities are exceeding the available rendering and transmission bandwidth, it is now necessary to adaptively control the quality of rendered images. To that end, we measured how quality changes as a function of refresh rate, and argued that the main factors influencing the quality of motion are spatial resolution, refresh rate, predictability and velocity of motion. We proposed a new visual model that explains the relationship between these factors, the blur and judder artifacts they introduce, and their impact on perceived quality. The parameters of the model are fitted to the results of eye-tracking and quality assessment experiments. The dataset and the model are made available[1]. Such a model can be used to drive adaptive rendering algorithms, which can deliver better perceived quality of animation than the non-adaptive approach.

The collected psychophysical data and the proposed visual model are steps in the direction of developing more comprehensive models in the future. A few major omissions include the data collection for low-persistence displays, and the inclusion of ghosting and flicker artifacts. A higher dynamic range and extending the model to be content-aware would be also desirable. Furthermore, we only demonstrated the utility of such models with a single adaptive rendering algorithm (MARRR). However, image resolution could be traded for lower complexity of shading, for variable shading rate, reduced AA sampling, reprojected samples, reduced LODs or others. Some of those techniques could be analyzed in the framework of our model, others will require extensions. The introduction of multiple velocities and the estimation of these from eye tracker data or saliency maps is also an interesting challenge. We hope that our work will accelerate further work on such motion-adaptive rendering methods.

## REFERENCES

Kaan Akşit, Praneeth Chakravarthula, Kishore Rathinavel, Youngmo Jeong, Rachel Albert, Henry Fuchs, and David Luebke. 2019. Manufacturing Application-Driven Foveated Near-Eye Displays. *IEEE transactions on visualization and computer graphics* 25, 5 (2019), 1928–1939.

C. Anthes, R. Garcia-Hernandez, M. Wiedemann, and D. Kranzlmuller. 2016. State of the art of virtual reality technology. In *2016 IEEE Aerospace Conference.* IEEE, 1–19. https://doi.org/10.1109/AERO.2016.7500674

Tunç Ozan Aydin, Martin Čadík, Karol Myszkowski, and Hans-Peter Seidel. 2010. Video Quality Assessment for Computer Graphics Applications. *ACM Trans. Graph.* 29, 6, Article 161 (Dec. 2010), 12 pages. https://doi.org/10.1145/1882261.1866187

Amin Banitalebi-Dehkordi, Mahsa T Pourazad, and Panos Nasiopoulos. 2015. The effect of frame rate on 3D video quality and bitrate. *3D Research* 6, 1 (2015), 1.

---

[1]Project web page: https://www.cl.cam.ac.uk/research/rainbow/projects/motion_quality_model

P. G. J. Barten. 2003. Formula for the contrast sensitivity of the human eye, Yoichi Miyake and D. Rene Rasmussen (Eds.). 231–238. https://doi.org/10.1117/12.537476

Tom Beigbeder, Rory Coughlan, Corey Lusher, John Plunkett, Emmanuel Agu, and Mark Claypool. 2004. The effects of loss and latency on user performance in unreal tournament 2003®. In *Proceedings of 3rd ACM SIGCOMM workshop on Network and system support for games*. ACM, 144–151.

Alexandre Chapiro, Robin Atkins, and Scott Daly. 2019. A Luminance-aware Model of Judder Perception. *ACM Transactions on Graphics (TOG)* 38, 5 (2019), 1–10.

K. T. Claypool and M. Claypool. 2007. On frame rate and player performance in first person shooter games. *Multimedia Systems* 13, 1 (jul 2007), 3–17. https://doi.org/10.1007/s00530-007-0081-1

M. Claypool and K. Claypool. 2009. Perspectives, frame rates and resolutions. In *Proceedings of the 4th International Conference on Foundations of Digital Games - FDG '09*. ACM Press, New York, New York, USA, 42. https://doi.org/10.1145/1536513.1536530

S. Daly, N. Xu, J. Crenshaw, and V. J. Zunjarrao. 2015. A Psychophysical Study Exploring Judder Using Fundamental Signals and Complex Imagery. *SMPTE Motion Imaging Journal* 124, 7 (October 2015), 62–70. https://doi.org/10.5594/j18616

S. J. Daly. 1998. Engineering observations from spatiovelocity and spatiotemporal visual models, Bernice E. Rogowitz and Thrasyvoulos N. Pappas (Eds.). 180–191. https://doi.org/10.1117/12.320110

James Davis, Yi-Hsuan Hsieh, and Hung-Chi Lee. 2015. Humans perceive flicker artifacts at 500 Hz. *Scientific reports* 5 (2015), 7861.

K. Debattista, K. Bugeja, S. Spina, T. Bashford-Rogers, and V. Hulusic. 2018. Frame Rate vs Resolution: A Subjective Evaluation of Spatiotemporal Perceived Quality Under Varying Computational Budgets. *Computer Graphics Forum* 37, 1 (feb 2018), 363–374. https://doi.org/10.1111/cgf.13302

G. Denes, K. Maruszczyk, G. Ash, and R. K. Mantiuk. 2019. Temporal Resolution Multiplexing: Exploiting the limitations of spatio-temporal vision for more efficient VR rendering. *IEEE Transactions on Visualization and Computer Graphics* 25, 5 (May 2019), 2072–2082. https://doi.org/10.1109/TVCG.2019.2898741

Piotr Didyk, Elmar Eisemann, Tobias Ritschel, Karol Myszkowski, and Hans-Peter Seidel. 2010. Apparent display resolution enhancement for moving images. In *ACM Transactions on Graphics (TOG)*, Vol. 29. ACM, 113.

E. DoVale. 2017. High Frame Rate Psychophysics: Experimentation to Determine a JND for Frame Rate. *SMPTE Motion Imaging Journal* 126, 9 (nov 2017), 41–47. https://doi.org/10.5594/JMI.2017.2749919

David B Elliott, KC Yang, and David Whitaker. 1995. Visual acuity changes throughout adulthood in normal, healthy eyes: seeing beyond 6/6. *Optometry and vision science: official publication of the American Academy of Optometry* 72, 3 (1995), 186–191.

Trey Greer, Josef Spjut, David Luebke, and Turner Whitted. 2016. 8-3: Hybrid Modulation for Near Zero Display Latency. In *SID Symposium Digest of Technical Papers*, Vol. 47. Wiley Online Library, 76–78.

Brian Guenter, Mark Finch, Steven Drucker, Desney Tan, and John Snyder. 2012. Foveated 3D graphics. *ACM Transactions on Graphics (TOG)* 31, 6 (2012), 164.

ITU-R. 2016. Subjective assessment methods for 3D video quality. ITU-R Recommendation P.915.

Anton Kaplanyan, Anton Sochenov, Thomas Leimkuehler, Mikhail Okunev, 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)* (2019).

D. H. Kelly. 1979. Motion and vision II Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America* 69, 10 (oct 1979), 1340. https://doi.org/10.1364/JOSA.69.001340

Woojae Kim, Jongyoo Kim, Sewoong Ahn, Jinwoo Kim, and Sanghoon Lee. 2018. Deep video quality assessor: From spatio-temporal visual sensitivity to a convolutional neural aggregation network. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 219–234.

Michiel A Klompenhouwer and Leo Jan Velthoven. 2004. Motion blur reduction for liquid crystal displays: motion-compensated inverse filtering. In *Visual Communications and Image Processing 2004*, Vol. 5308. International Society for Optics and Photonics, 690–699.

Y. Kuroki, T. Nishi, S. Kobayashi, H. Oyaizu, and S. Yoshimura. 2006. 3.4: Improvement of Motion Image Quality by High Frame Rate. *SID Symposium Digest of Technical Papers* 37, 1 (2006), 14. https://doi.org/10.1889/1.2433276

Y. Kuroki, T. Nishi, S. Kobayashi, H. Oyaizu, and S. Yoshimura. 2007. A psychophysical study of improvements in motion-image quality by using high frame rates. *Journal of the Society for Information Display* 15, 1 (2007), 61. https://doi.org/10.1889/1.2451560

S. G. Lisberger. 2010. Visual Guidance of Smooth-Pursuit Eye Movements: Sensation, Action, and What Happens in Between. *Neuron* 66, 4 (may 2010), 477–491. https://doi.org/10.1016/j.neuron.2010.03.027

Jing Liu, Soja-Marie Morgens, Robert C Sumner, Luke Buschmann, Yu Zhang, and James Davis. 2014. When does the hidden butterfly not flicker?. In *SIGGRAPH Asia 2014 Technical Briefs*. ACM, 3.

A. Mackin, K. C. Noland, and D. R. Bull. 2016. The visibility of motion artifacts and their effect on motion quality. In *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2435–2439. https://doi.org/10.1109/ICIP.2016.7532796

Belen Masia, Gordon Wetzstein, Piotr Didyk, and Diego Gutierrez. 2013. A survey on computational displays: Pushing the boundaries of optics, computation, and perception. *Computers & Graphics* 37, 8 (2013), 1012–1038.

J. D. McCarthy, M. A. Sasse, and D. Miras. 2004. Sharp or smooth?. In *Proceedings of the 2004 conference on Human factors in computing systems - CHI '04*. ACM Press, New York, New York, USA, 535–542. https://doi.org/10.1145/985692.985760

Aliaksei Mikhailiuk, Clifford Wilmot, Maria Perez-Ortiz, Dingcheng Yue, and Rafal Mantiuk. 2020. Active Sampling for Pairwise Comparisons via Approximate Message Passing and Information Gain Maximization. arXiv:cs.LG/2004.05691

K. P. Murphy. 2012. *Machine Learning: A Probabilistic Perspective* (1 ed.). MIT Press.

Fernando Navarro, Susana Castillo, Francisco J Serón, and Diego Gutierrez. 2011. Perceptual considerations for motion blur rendering. *ACM Transactions on Applied Perception (TAP)* 8, 3 (2011), 1–15.

D. C Niehorster, W. F. Siu, and L. Li. 2015. Manual tracking enhances smooth pursuit eye movements. *Journal of vision* (2015). https://doi.org/10.1167/15.15.11

S. Niklaus and F. Liu. 2018. Context-Aware Synthesis for Video Frame Interpolation. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 1701–1710. https://doi.org/10.1109/CVPR.2018.00183

K Noland. 2014. The application of sampling theory to television frame rate requirements. *BBC Research & Development White Paper* 282 (2014).

Yen-Fu Ou, Zhan Ma, Tao Liu, and Yao Wang. 2010. Perceptual quality assessment of video considering both frame rate and quantization artifacts. *IEEE Transactions on Circuits and Systems for Video Technology* 21, 3 (2010), 286–298.

Fabio Pellacini. 2005. User-configurable automatic shader simplification. *ACM Transactions on Graphics (TOG)* 24, 3 (2005), 445–452.

M. Perez-Ortiz and R. K. Mantiuk. 2017. A practical guide and software for analysing pairwise comparison experiments. *arXiv:1712.03686v2* (2017).

JE Roberts and AJ Wilkins. 2013. Flicker can be perceived during saccades at frequencies in excess of 1 kHz. *Lighting Research & Technology* 45, 1 (2013), 124–132.

D. A. Robinson. 1964. The mechanics of human saccadic eye movement. *The Journal of Physiology* (1964). https://doi.org/10.1113/jphysiol.1964.sp007485

D A Robinson. 1965. The mechanics of human smooth pursuit eye movement. *The Journal of Physiology* 180, 3 (oct 1965), 569–591. https://doi.org/10.1113/jphysiol.1965.sp007718

Andrew Rutherford. 2003. Handbook of perception and human performance. Vol 1: Sensory processes and perception. Vol 2: Cognitive processes and performance. In *Applied Ergonomics*. Vol. 18. Chapter 6, 340. https://doi.org/10.1016/0003-6870(87)90144-x

D. Scherzer, L. Yang, O. Mattausch, D. Nehab, P. V. Sander, M. Wimmer, and E. Eisemann. 2012. Temporal coherence methods in real-time rendering. *Computer Graphics Forum* 31, 8 (2012), 2378–2408. https://doi.org/10.1111/j.1467-8659.2012.03075.x

E. Simonson and J. Brozek. 1952. Flicker Fusion Frequency: Background and Applications. *Physiological Reviews* 32, 3 (jul 1952), 349–378. https://doi.org/10.1152/physrev.1952.32.3.349

L. Stark, G. Vossius, and L. R. Young. 1962. Predictive Control of Eye Tracking Movements. *IRE Transactions on Human Factors in Electronics* HFE-3, 2 (sep 1962), 52–57. https://doi.org/10.1109/THFE2.1962.4503342

M. Suh, R. Kolster, R. Sarkar, B. McCandliss, and J. Ghajar. 2006. Deficits in predictive smooth pursuit after mild traumatic brain injury. *Neuroscience Letters* 401, 1-2 (jun 2006), 108–113. https://doi.org/10.1016/j.neulet.2006.02.074

K. Templin, P. Didyk, K. Myszkowski, and H. Seidel. 2016. Emulating displays with continuously varying frame rates. *ACM Transactions on Graphics* 35, 4 (jul 2016), 1–11. https://doi.org/10.1145/2897824.2925879

Krzysztof Templin, Piotr Didyk, Tobias Ritschel, Elmar Eisemann, Karol Myszkowski, and Hans-Peter Seidel. 2011. Apparent resolution enhancement for animations. In *Proceedings of the 27th Spring Conference on Computer Graphics*. ACM, 57–64.

L. Thaler, A.C. Schütz, M.A. Goodale, and K.R. Gegenfurtner. 2013. What is the best fixation target? The effect of target shape on stability of fixational eye movements. *Vision Research* 76 (jan 2013), 31–42. https://doi.org/10.1016/j.visres.2012.10.012

L. L. Thurstone. 1927. A law of comparative judgment. *Psychological Review* 34, 4 (1927), 273–286. https://doi.org/10.1037/h0070288

S. Tourancheau, P. Le Callet, K. Brunnström, and B. Andrén. 2009. Psychophysical study of LCD motion-blur perception, Bernice E. Rogowitz and Thrasyvoulos N. Pappas (Eds.). 724015. https://doi.org/10.1117/12.811757

A. B. Watson. 2015. High Frame Rates and Human Vision: A View through the Window of Visibility. *SMPTE Motion Imaging Journal* (2015). https://doi.org/10.5594/j18266xy

Andrew B Watson and Albert J Ahumada. 2011. Blur clarified: A review and synthesis of blur discrimination. *Journal of Vision* 11, 5 (2011), 10–10.

Andrew B. Watson, Albert J. Ahumada, and Joyce E. Farrell. 2008. Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. *Journal of the Optical Society of America A* (2008). https://doi.org/10.1364/josaa.3.000300

Hector Yee, Sumanita Pattanaik, and Donald P Greenberg. 2001. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. *ACM Transactions on Graphics (TOG)* 20, 1 (2001), 39–65.