Variable Frame Timing Affects Perception of Smoothness in First-Person Gaming

Devi Klein
NVIDIA and UC Santa Barbara
dklein@ucsb.edu

Josef Spjut
NVIDIA
jspjut@nvidia.com

Ben Boudaoud
NVIDIA
bboudaoud@nvidia.com

Joohwan Kim
NVIDIA
sckim@nvidia.com

***Abstract*—With the advent of variable refresh rate (VRR) monitor technologies, gamers experience variable frame timing (VFT) during their gameplay. Combining VRR with low-latency GPU rendering and increased display refresh rates enables smoother variation of frame presentation sequences. Here, we assess how VFT affects self-reported perceived smoothness of game play by introducing frequent but relatively small (4-12 ms) variations in frame time around typical refresh rates (30-240 Hz). Our results demonstrate that VFT degrades the perceived smoothness of game play for large variation in frame time (12 ms) but has a diminished effect on perceived smoothness for small variations in frame time (4 ms).**

***Index Terms*—frame rate, frame time, first-person targeting, first-person games, perception, performance**

I. Introduction

Real-time 3D computer games are designed to present dynamic imagery to our visual system in a sufficiently convincing manner such that we "experience" the game. These games rely on digital simulation to generate a discrete sequence of images that feel continuous and smoothly progressing to us [1]. Despite advances in emerging visualization technologies, the temporal presentation of a sequence of frames can cause undesirable perceptual effects [2], [3]. For example, discrepancies between the render rate of the graphics card and the monitor's refresh rate can cause dropping and doubling of frames (e.g., Fig. 1, bottom render vs. VSync rows, and see Sec. II-B for a more complete description), producing motion artifacts that degrade the perception of smoothness while playing a game.

Prior work has focused on the visibility of motion artifacts for displays with constant frame time (CFT) (Sec. II-C). CFT, however, is a simplified model in gaming due to variations in render time, input sampling, or network delays. Instead, high-end modern display technologies use variable refresh rate (VRR) to present gamers with a sequence of frames that can persist on screen for different periods, a phenomenon we denote as variable frame timing (VFT). That is, one frame could appear on screen for 16 ms (60 Hz) and the next for 20 ms (50 Hz).

Given the prevalence of VFT in modern-day gaming and the relatively little empirical work investigating its influence on the perception of smoothness, our paper makes the following contributions:

1) Introduces VFT. What is it? Why does it occur?
2) Explores the effects of VFT on the perception of smoothness across a broad range of baseline frame rates.
3) Compares perception of VFT and CFT.
4) Provides suggestions for gamers who utilize VFT in both professional and leisure settings and for researchers who want to explore VFT in contexts outside of gaming.

II. Background

The following sections provide a short primer on VFT and frame timing. The discussion points and background information underscore the importance of studying VFT as it becomes more prevalent with modern hardware and software. We follow this discussion with a brief overview of the perceptual quality metric used in this experiment.

A. Frame Rate, Frame Time, and Refresh Rate

Frame rate is the number of frames delivered over some time frame, typically reported in Hertz (Hz) or the number of frames per second. *Frame time* is the time that elapses between a pair of frames during which the first frame is displayed. Assuming frames are completed on a regular schedule, the frame rate is precisely the inverse of the frame time (the frame time represents the period of the rate).

In this work, we use the terms *frame rate* and *refresh rate* interchangeably as the inverse of the frame time. When frame time varies, a single frame rate cannot accurately represent this variability. Thus, we choose to focus primarily on frame times in this work. For those familiar with frame rates, it may be helpful to compute the inverse of a given frame time to gain an intuition for what regular frame rate that frame time corresponds to. We include frame rates for our constant frame time conditions and summarize our variable frame timing conditions using a $\frac{1}{X}$ convention to provide an intuition for the corresponding frame rates.

B. Variable Refresh Rate

Various factors ranging from changes in game geometry or shading workloads, power state and cooling of PC hardware, operating system preemption and scheduling to network communication with a server, and human interaction with the

979-8-3503-5067-8/24/$31.00 ©2024 IEEE

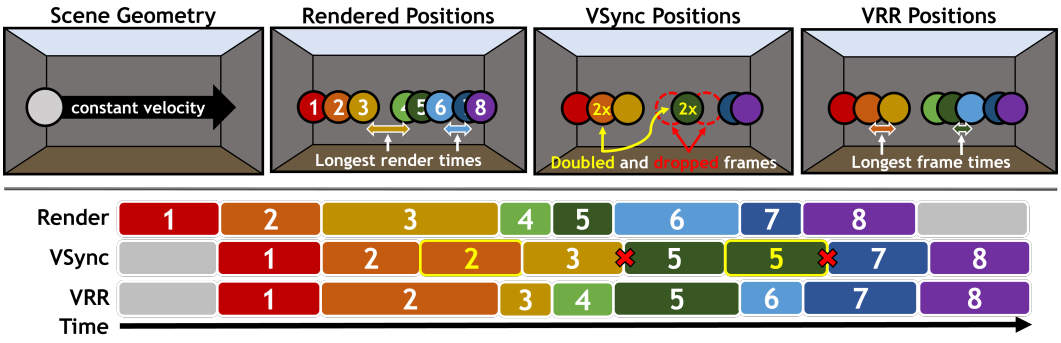


Fig. 1. (Top) A demonstration of the visual effects of variable frame timing. (Bottom) A demonstration of how variable render/simulation workload timing impacts VSync and variable refresh rate (VRR) differently. In the VSync time series red crosses and corresponding red dashed regions in the VSync position figure represent frames dropped due to timing of render relative to scanout (frame 4 and 6). Frame 5, highlighted in yellow, is duplicated as a result of frame 6 not being completed in time for scanout. When using VRR all rendered frames are displayed, but not at their rendered durations, the result is mismatch of the time a frame is displayed relative to how long it took to draw and small but apparent changes in the resulting speed of the target.

game can interfere with rendering and cause frame time to vary. Frame time variability itself can vary across games, from small compared to the total frame time to significant enough to be easily noticed.

VRR implementations, like G-Sync [4] and VESA Adaptive Sync [5], adjust monitor frame time dynamically as frames are completed. In other words, VRR-enabled monitors initiate the presentation of a frame whenever the frame arrives from the GPU. VRR avoids the potential large magnitude stutter caused by minor frame time variations when VSync is on while avoiding tearing[1] caused by disabling VSync. The primary functionality that VRR enables is presenting the variation in frame completion time directly to the display, emitting light to the eyes of the viewer. Fig. 1 provides a graphical depiction contrasting how VSync displays frames vs how VRR-enabled displays provide a one-to-one mapping between rendered and displayed frames (not necessarily preserving frame time).

Since the introduction and widespread availability of VRR-enabled displays and computing systems, researchers have begun to conduct studies using the paradigm. For instance, Poth et al. [6] used G-Sync to limit the presentation duration of a stimulus to below 10 ms. They varied the time in small step sizes to characterize visual processing time. Others have promoted VRR for potential energy savings by reducing power when display updates are unnecessary [7], [8]. Another study measured (constant) refresh rates for *just noticeable differences* of varying target motion, then leveraged VRR to adjust refresh rate dynamically based on object speed. However, they did not study the effects of varying refresh rate itself [2], [9]. Jindal et al. [10] added adaptive rate shading to the optimizations for motion, spatial shading rate, and temporal rate but once again neglected to study the effects of varying frame rate in isolation. This work hints that not all spatiotemporal content has the same presentation time requirements for the equal perception of quality but does not answer whether a single target's motion is the primary driver of the perception of smoothness or stutter in VRR contexts.

More relevant work studying VRR's effect on task performance and perception has indicated positive outcomes [11], [12]. While subjects struggled to identify when VRR was enabled, results show significant effects of VRR on both performance and engagement. Liu et al. [3] studied the effects of varied frame timing on the quality of experience in 3 real games, as measured by mean observer score. Instead, they relied on a (set of) CPU-based infinite Fibonacci sequence generators running in the background, to artificially create uncontrolled variations in frame times. They find that a 95% frame rate floor predicts the quality of experience across three games. None of these studies carefully control the sequence nor the distribution of frame times experienced by the gamer. Additionally, they neglected to study any refresh rates above 120 Hz. Our study aims to fill in some gaps in the current understanding of VRR's effects on human's perception of smoothness during gameplay.

C. Motion Artifacts Define Perceptual Smoothness

To this point, we have alluded to the psychological construct of *perceptual smoothness* and tangentially related it to the motion artifacts. Prior work has done a fine job of operationally defining perceptual smoothness regarding object motion by binning it into various distinct visual percepts: false edges, motion blur, and judder [2], [13], [14], [15]. The perception of smoothness comes down to how well people detect (i.e., the visibility of) motion artifacts on displays with different frame rates [16], [17], [13].

We study how VRR influences first-person gamer's perception of smoothness. First-person gamers constantly perform visuomotor tasks that demand extremely high spatial precision within a very tight timeline. In FPS games, the gamer typically uses the mouse to control egocentric view rotation and the keyboard to translate the camera within a scene. We choose to study the FPS aiming context as it represents an area where high-skill human-computer interaction intersects

[1]Tearing is when 2 different frames are displayed simultaneously with a horizontal "tear" between the top and bottom frame.

TABLE I
HARDWARE SPECIFICATIONS FOR THE SYSTEMS USED IN OUR EXPERIMENT.

Parameter	Value
Display	360 Hz Alienware AW2521H
Resolution	1920 × 1080
G-Sync (VRR)	Enabled
VSync	Disabled
Viewing Distance	∼70 cm
CPU	Intel i7-9700k @ 3.60 GHz
GPU	NVIDIA RTX 2080Ti
RAM	32GB DDR4
Mouse	Logitech G Pro Wireless
Mouse DPI	800
Keyboard	Logitech G Pro YU0039
Click-to-Photon Latency	20 ms @ 120 Hz

high-performance rendering and corresponding variable frame timing. Specifically, we study the camera rotation aspect of FPS aiming in this work, disallowing translation within the scene to avoid additional confounding factors. Because of the highly interactive nature of the task and the accumulated experience, the conventional characterization of motion artifacts may not capture the visual experience of first-person gamers. For this reason, our work does not attempt to define perceptual smoothness in one way or another (i.e., by the detection of judder or motion blur). Instead, we leave it up to the gamer to define perceptual smoothness while interacting with the gaming environment.

III. METHODS

This experiment used an open-source FPS experimentation platform, First Person Science [18], [19] (release v22.05.01), to run and collect data from participants. First Person Science provides a feature whereby an experimenter can specify a target frame time to be used on a per-frame basis. We use this virtual sandbox to study the effects of variable frame timing on the gamer. The hardware specifications for this experiment can be found in Table I.

A. Realization of Variable Frame Timing

Across the conditions in this experiment, subjects experienced a range of constant and variable frame time sequences over which they ranked smoothness while completing a targeting task.

In deciding which type of VFT scheme to use, we had many different candidates. Initially, we considered a "1:N stutter" scheme where a series of N frames occur at constant frame timing, followed by a single (or small set of) frames(s) that is (are) longer/shorter. This model creates a rapid change in frame timing that emulates a single performance hitch or unpredictable rendering which are common in real games. However, this approach is not desirable as it emulates only one of the factors that cause VFT in the wild and varies rapidly, potentially maximizing the objectionability of timing changes.

Instead, we elected to use a sine wave sampling scheme of frame time over time to produce our VFT sequences. The signal had a frequency of 3 Hz, and the waveform spanned

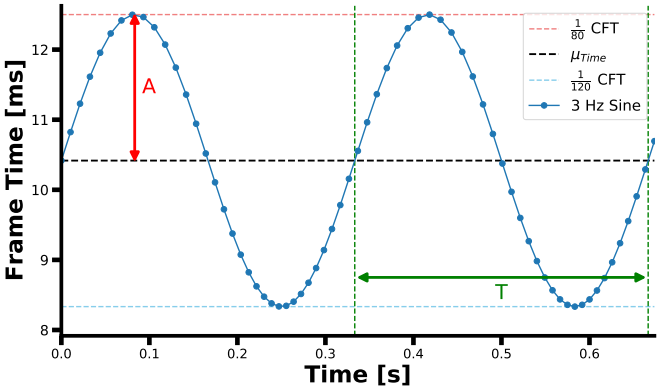


Fig. 2. Two periods of an example 3 Hz sine wave of varying frame time over time with 2 ms amplitude (A), 333.33 ms period (T), and time-weighted mean frame time depicted as a dashed black line (μ_{Time} = 10.42 ms). Note that the scatter points (and their relative order over time) overlaid on the waveform are the frame times presented to participants in the task. The dashed red/blue lines indicate the minimum and maximum frame duration denoted $(\frac{1}{120}, \frac{1}{80})$ (see the supplement for more details).

between pairs of CFTs also tested. This choice of VFT pattern is not selected to be representative of an in the wild frame time pattern, but tests smoothly varying frame time changes and avoids overemphasizing the noticeability of frame timing changes. The 3 Hz frequency was chosen because it is far (90%) below the lowest tested frame rate (30 Hz), ensuring that the Nyquist limit is met and that resampling artifacts such as aliasing or lower frequency "pulsing" were minimized. The VFT amplitudes of 2, 4, and 6 ms were kept small to (1) align better with variations in timing common in modern rendering and simulation pipelines and (2) avoid duplicating results from previous VSync studies.

Fig. 2 plots two periods of an example waveform and demonstrates how a thoughtful sampling of the intended signal is required to avoid pitfalls. Note that frame times are sampled more frequently at the trough of the sine wave than at the peak due to the interaction between the variables represented on the x (time) and y (frame time) axes.

While we believe the sine pattern is a reasonable starting point for study, it is neither a complete nor representative model of VFT in the wild. However, one could use our approach of sampling frame time over time to create more complex VFT patterns representative of particular and realistic use cases of VFT. In Sec. II-B, we enumerated sources of VFT in gaming scenarios, including user action(s), game state and background workloads. These factors interact in complex and non-trivial ways creating a distribution of frame times with a particular temporal spectrum that is not well-characterized or generalizable. Therefore, there is no one-size-fits-all pattern of VFT we could have selected for testing.

B. Hypotheses

The experiment is concerned with the visual perception of smoothness of motion for sphere-shaped targets in a first-person targeting task. Based on our understanding of the

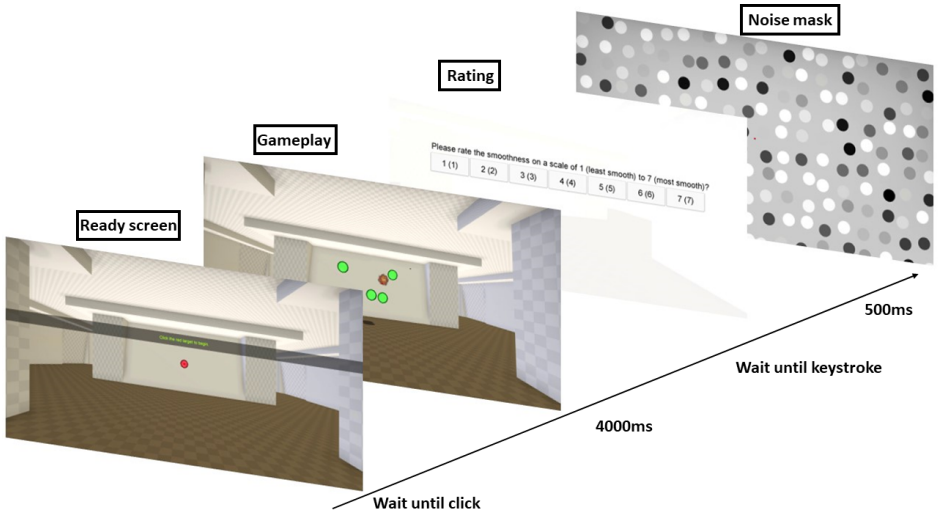


Fig. 3. Example trial flow for the experiment. From left to right: the user is first prompted with a ready screen. Clicking on the red sphere initiates the trial. Next 4 green targets moved in random direction for 4 seconds. The user destroyed as many targets as possible (infinite target respawn). After experiencing the gameplay, the user is prompted with a rating scale to indicate their perceived smoothness of the experience. Lastly, a visual noise mask is presented at the end of the trial.

effects of frame time, and its variation, from prior art and pilot studies in FPS aiming, we expect that gamers can perceive 4-12 ms variations in frame time, across a range of base frame rates (30-240 Hz).

C. Task Procedure

1) Trial flow: A graphical depiction of the first-person targeting task procedure is provided in Fig. 3. Each trial began with a red reference target centered on the screen, with the reticle's default position set to the center of this target. Participants initiated the trial by clicking the reference target. Afterward, four green targets appeared on the screen. The targets moved randomly within a predefined region in world-space coordinates for four seconds. We instructed participants to try their best to destroy the targets in each trial and warned them that their data would be discarded if the number of targets destroyed did not meet an unspecified threshold. We kept this threshold vague to ensure that participants focused on the judgment of smoothness while simultaneously completing the task of destroying the targets. We decided to have the targets respawn infinitely to ensure that participants experienced the complete sequence of frame times before making a perceptual judgment.

After four seconds of gameplay elapsed, a rating screen appeared. Participants indicated their perception of smoothness on a Likert scale ranging from 1-7, with 1 mapping to the least smooth perceptual experience and 7 mapping to the smoothest perceptual experience. Participants marked their responses by pressing the appropriate number key on the keyboard. To end the trial, a dynamic visual noise mask (run as a 2D shader) [20] would appear on the screen for half a second. We added the mask in an attempt to remove any carryover effects or motion adaptation from the previous trial.

Before the experiment started, participants were allowed to adjust their mouse sensitivity to their desired setting (in °/mm), and the mouse DPI was set to 800. Uniform mouse DPI ensures equal view rotation quantization, while individual sensitivity selection provides ecological validity and ensures that we do not confound a participant's preference for a particular sensitivity with their ability to discern effects of variable frame timing. Next, they completed a single trial, except it ended once all 4 targets were destroyed with no time limit. This design allowed them to continually update the DPI setting until they were ready to begin the experiment.

Afterward, participants were provided task instructions and block of practice trials (same trial flow as shown in Fig. 3) to get comfortable with the experiment design. The instructions consisted of two separate trials. The first trial lasted ten seconds instead of four, with the game rendered at 15 Hz. Participants were instructed to pay attention to the smoothness of the gameplay, as this visual experience represented a smoothness rating of 1. Next, they experienced a trial at 360 Hz for ten seconds and were told that this visual experience mapped to a smoothness rating of 7. As a note, we did not include 15 Hz or 360 Hz frame rates in the actual experiment. After being introduced to the rating scale, participants experienced five warm-up trials lasting four seconds each. The block of warm-up trials consisted of variable frame time sequences and constant frame time sequences and were designed to introduce the participant to the primary task of the experiment.

2) Stimuli: The target was green and had an in-game world-space diameter of 0.5 m (in-game field of view of the target was 3.82 °). Note that we did not calculate degrees of visual angle of the target as we did not use a chin rest to control for viewing distance. Additionally, the apparent size of the target changed slightly on-screen as it moved

TABLE II

EXPERIMENT 1 CONDITIONS ARE ORDERED BY THE AMPLITUDE OF FRAME TIME VARIATION. CONSTANT (0 MS) CONDITIONS, WITH EQUAL MINIMUM AND MAXIMUM FRAME DURATION, ARE REPRESENTED BY A SINGLE NUMBER INSTEAD OF A PAIR. ALL FRAME DURATION IS PRESENTED AS FRACTIONS IN SECONDS. THE COUNT COLUMN REPRESENTS THE NUMBER OF LEVELS TESTED IN EACH OF THE FOUR CONDITIONS.

Amplitude	**Frame Duration Levels (min – max) [s]**							
0 ms (constant)	$1/240 - 1/240$	$1/120 - 1/120$	$1/80 - 1/80$	$1/60 - 1/60$	$1/48 - 1/48$	$1/40 - 1/40$	$1/34 - 1/34$	$1/30 - 1/30$
2 ms	$1/240 - 1/120$	$1/120 - 1/80$	$1/80 - 1/60$	$1/60 - 1/48$	$1/48 - 1/40$	$1/40 - 1/34$	$1/34 - 1/30$	
4 ms	$1/240 - 1/80$	$1/120 - 1/60$	$1/80 - 1/48$	$1/60 - 1/40$	$1/47 - 1/34$	$1/40 - 1/30$		
6 ms	$1/240 - 1/60$	$1/120 - 1/48$	$1/80 - 1/40$	$1/60 - 1/34$	$1/48 - 1/30$			

due to perspective projection. The target moved at a constant world-space speed of 3 m/sec (22.92 °/s). At some point, the target changed direction every 700-1000 ms (uniformly sampled), with the new direction also sampled uniformly from all possible directions. The targets spawned in a 0.5 x 2 x 2 m cube centered 7.5 m away from the participant spawn position. Those targets moved within a larger 1 x 2.3 x 6 m cube centered on the same position as the spawn bounding box and bounced off the edges when they moved to that location.

3) Conditions: This experiment employed four conditions, one constant frame time and three variable frame time conditions that differed in amplitude (see Section III-A above). The constant frame time condition was comprised of eight levels. From the longest frame duration to the shortest, we included: 1/30, 1/34, 1/40, 1/48, 1/60, 1/80, 1/120, and 1/240 s. We selected these frame duration levels based on two factors. First, we wanted to adequately sample the domain of monitor refresh rates available in the wild. Older displays and TV content update at 30 Hz refresh rates, while new monitors can run up to 500 Hz or higher, but we only tested up to 240 Hz as this is what is widely available in the gaming market. Second, the VFT conditions, using a sinusoidal variation, were composed of frame sequences that oscillated periodically between two frame times; for example, between 1/240 s and 1/120 s. Because the participant is experiencing (at most) a frame time equivalent to a 120 Hz frame rate, we included that constant frame rate in our experiment design.

The three VFT conditions sampled the same domain of frame times as the constant frame time condition but differed from one another in the amplitude (A) of the sine wave used to generate the sequence of frame times. The amplitude and the min and max frame duration from both the constant and variable frame time conditions are provided in Table II.

In total, there were 26 levels across the four conditions. Participants encountered each of the 26 levels 11 times totaling 286 trials. To account for order or carryover effects, we presented the 26 levels in random order within a single block (of 26 trials). Participants, therefore, encountered 11 blocks of 26 trials so that no block contained two trials of the same level. This experiment took approximately 1 hour to complete.

4) Participants: Participants were volunteers recruited from a single institution and have interest in gaming and graphics design. A total of 10 participants completed the experiment. However, one participant's data were excluded from analysis because their responses were overly noisy (i.e., the standard error of the mean for their response at each level was 2-3× as large as the other participants). All participants signed an informed consent form approved by the IRB.

D. Analysis and Results

When comparing the frame time sequences (levels) within the same condition, perceived smoothness monotonically decreased as frame duration increased. This relationship held across all four conditions tested in experiment 1 and is shown in Fig. 4 where perceived smoothness is plotted against time-weighted mean frame duration (the x-coordinate where vertical error bars project out in the right three panels). This shows that the findings from prior work [2] continue to hold even when frame time varies. The smoothness rating changed almost linearly as the time-weighted mean frame duration increased, with the slope being slightly less steep at the short frame times. Our reference stimuli, 15 and 360 Hz, effectively anchored perceived smoothness ratings with averages ranging from below 2 to above 6 in the 30 and 240 Hz CFT conditions.

Our primary analysis homes in on the descriptive statistic of the VFT sequences that best predict perceived smoothness across the three VFT conditions. Is it the time-weighted mean frame duration, the min or max frame duration, the frame durations corresponding to 95th or 99th percentiles, or some other percentile of the VFT sequence that predicts the participant's perceived smoothness? First, we fit a sigmoid function to the observed data in the CFT condition (leftmost panel in Fig. 4) using a least-square criterion,

$$f_{\alpha,x_0,\beta,C}(x) = \frac{\beta}{1 + e^{-\alpha(x-x_0)}} + C \tag{1}$$

yielding the following fit parameters: $\alpha = -116.983$, $\beta = 7.63$, $x_0 = 0.0306$, $C = -1.091$ with a mean squared error (MSE) of 0.0458.

We then overlaid the fitted sigmoid curve on the VFT sequences comprising the other three conditions, as depicted by the green curve superimposed over the observed data in each of the right three panels of Fig. 4. For each VFT sequence (horizontal lines in Fig. 4), one can visually search for the CFT yielding equivalent predicted smoothness by finding the point on the green curve with an equal smoothness rating[2]. In other words, we identify where the green curve intersects the horizontal (variably colored) line. An example of this procedure is visually depicted in 5, right. This equivalently smooth CFT tends to be closer to the longest frame duration

[2]The green curve interested all VFT sequences except for the (1/240-1/120s) level. We exclude this data point from analysis but revisit it in the discussion.

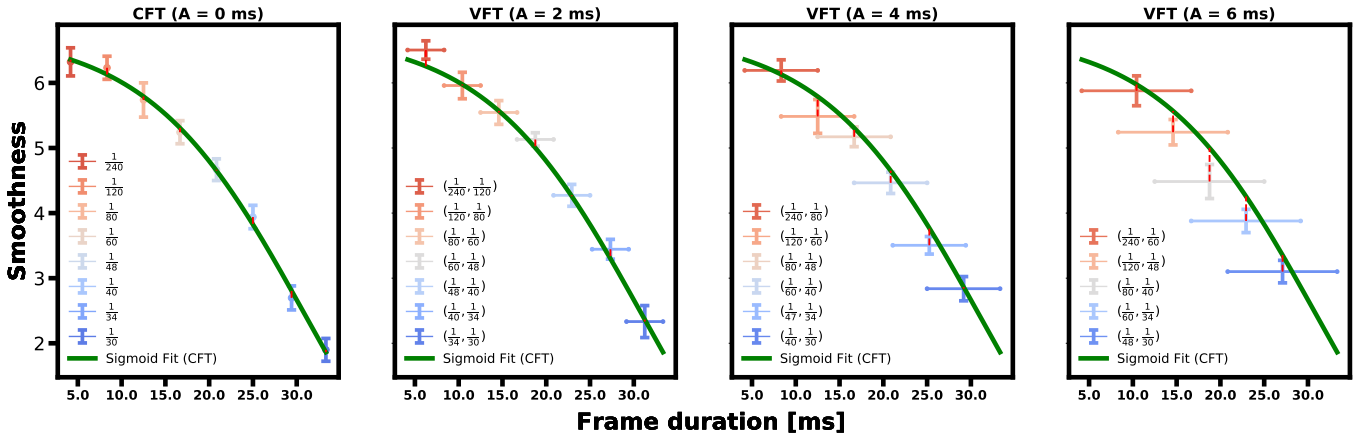


Fig. 4. Perception of smoothness plotted as a function of time-weighted mean frame duration for each of the four conditions tested in experiment 1. Participants utilized a rating scale from 1 (low) to 7 (high) to indicate their perception of smoothness during gameplay. (Left) Mean smoothness ratings across subjects for each of the eight frame duration levels tested in the constant frame time condition. Error bars represent the standard error of the mean. A sigmoid function (see Equation 1) is fit to the pooled participant smoothness scores for CFTs and is superimposed over all plots in the figure (green lines). We computed MSE as a measure of goodness of fit, and red dotted lines depict prediction errors of the fit with respect to observed data. (From middle left to right) 2/4/6 ms amplitude VFT condition: smoothness scores for each of the levels presented to participants are plotted against the time-weighted mean frame duration of each VFT sequence (i.e., the x-coordinate where vertical error bars intersect the horizontal lines). Legends include the range of frame times for each VFT sequence using a 1/x convention to help interpret results.

in the given VFT sequence (the right edge of the horizontal line) than the shortest frame duration in the sequence (the left edge of the same line).

In applying the logic from the above analysis to each of the three VFT conditions, we found that the amplitude of frame time variation had a systematic effect on the equivalently smooth CFT. When the amplitude of frame time variation was small (2 ms), the equivalently smooth CFT was not strongly biased toward the longer frame time in the VFT condition. In fact, it tended to cluster around the time-weighted mean frame duration of each VFT sequence (Fig. 4, middle left panel). When the amplitude of frame time variation was large (6 ms), the equivalently smooth CFT became strongly biased toward the longer frame time in the VFT sequence (Fig. 5 left). We can conceptualize this bias another way by considering the percentile rank of the equivalently smooth CFT in the distribution of frame times for each VFT sequence. The conversion from frame duration to percentile rank for the equivalently smooth CFT is exemplified in the inset plot of Fig. 5 for the sequence (1/48, 1/30).

To confirm this, we fit two sigmoids per condition. The first sigmoid was fit on the time-weighted mean frame duration of each VFT sequence (as was done in the CFT condition). For the second fit, we applied the following procedure separately to each of the 3 VFT conditions. We computed the percentile rank of the equivalently smooth CFT for each VFT sequence and took the mean percentile rank across the sequences. Next, we identified the frame duration that corresponded to this mean percentile rank in each VFT sequence. This set of frame duration times served as our regressor in the second sigmoid fit. Overall, we observe more accurate fits when using the mean percentile instead of the time-weighed mean of the frame time series (Table III). This indicates that, particularly for larger frame time variations, using a model based on a fixed (high) percentile of frame times better predicts perceived smoothness than time-weighted mean frame time.

IV. Discussion

A. Summary of Results

The results from our experiment demonstrate that VFT affects gamers' perception of smoothness similarly to CFT (Fig. 4). As the time-weighted mean frame time of the VFT sequence increases, the perceived smoothness of the overall experience of the targeting task decreases. However, as more variation is added to a constant frame time baseline (i.e., larger amplitude), our perception of smoothness becomes tethered to the frames presented on screen for longer periods of time (e.g., the sampled frames near the peak of the sine wave in Fig. 2). This is realized by our MSE analysis (table III) of the VFT conditions. We found that higher percentiles of the frequency distributions for each VFT sequence served as better predictors of smoothness than the time-weighted mean frame duration (A=2ms and A=6ms conditions). For example, we saw a 20% reduction in MSE when we regressed the smoothness scores of participants in the A = 6 ms condition on the frame duration corresponding to the 70th percentile of each of the five VFT sequences as opposed to the time-weighted mean frame duration. It should be noted that we fit four free parameters of the sigmoid function to (essentially) five data points (the means smoothness scores across all participants for the VFT sequences) so some overfitting is to be expected. Other models (e.g., simple linear regression) may better predict the smoothness in the VFT sequences with small variation in frame timing, especially when the VFT sequence oscillates

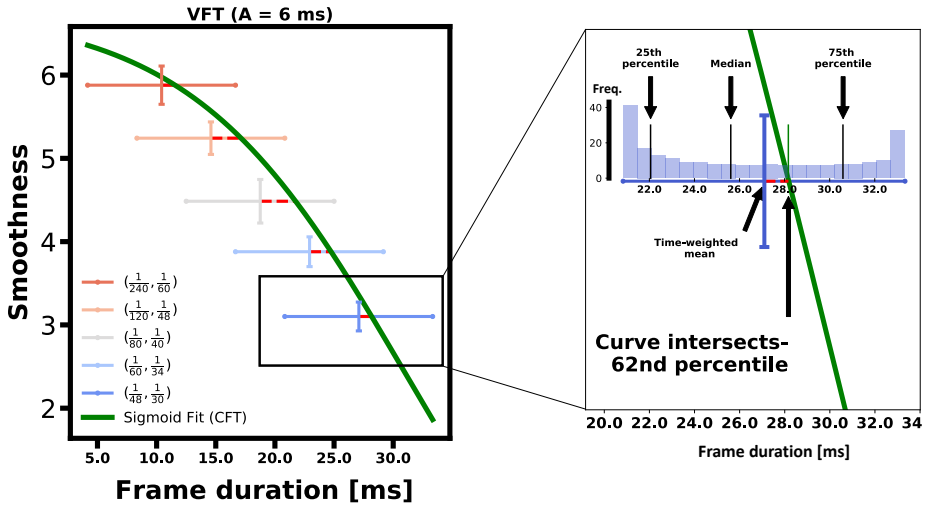


Fig. 5. Perception of smoothness in the 6 ms VFT condition (A = 6 ms). (Left) The sigmoid fit to the constant frame duration data (green curve) is overlaid on top of horizontal bars representing the range of frame times for each of the five levels in the A = 6 ms VFT condition. Horizontal dashed-red lines extend from the time-weighted mean frame duration of each VFT sequence to the intersection point of the green curve and the horizontal lines. We consider these intersection points as equivalently smooth CFTs for corresponding VFT sequences. The inset figure to the right depicts the frequency distribution of frame times in the VFT sequence oscillating between (1/48, 1/30) frame times. Vertical bars overlaid on the distribution highlight the 25th, median (50th), and 75th percentiles and the time-weighted mean frame duration. Additionally, the last vertical bar highlights the percentile rank of the frequency distribution corresponding to the equivalently smooth CFT.

TABLE III
COMPARISON OF SIGMOID MSE FOR TIME-WEIGHTED MEAN AND PERCENTILE-BASED FITS. NOTE THAT COMPARISONS SHOULD OCCUR WITHIN CONDITIONS, NOT ACROSS THEM, AS THE NUMBER OF POINTS FIT BY THE 4-PARAMETER SIGMOID (SEE EQ. 1) WAS DIFFERENT BETWEEN CONDITIONS.

Condition	MSE Weighted Mean	MSE Percentile	Reduction in MSE
2 ms	0.0351	0.0334	5%
4 ms	0.0532	0.0537	-0.95%
6 ms	0.0053	0.0043	19.7%

between 1/240 and 1/120 s because it did not intersect the green curve in Fig. 4. Nonetheless, this does not preclude our conclusion that frames persisting on screen for the longest period of time in the VFT sequences have a more pronounced influence on the perception of smoothness in gameplay than the frames near the time-weighted mean frame duration of the VFT sequence as our results are in line with [3].

One important takeaway from this study is that a gamer's perception of smoothness is mediated by frame duration but second order statistics (i.e., the variability in frame time) potential act as a third variable moderating the relationship between the perception of smoothness during gameplay and the time-weighted mean frame duration of a VFT sequence. As shown in Fig. 4, moving from the leftmost panel (A=0) to the rightmost panel (A=6ms), we note that the variance in the perceived smoothness scores decreases across the different VFT levels within a single condition.

B. Limitations

Our work, utilizing a simple sine wave sampling scheme, is a best-effort attempt to capture the effects of always changing VFT without substantial investment in profiling existing applications and their frame time (interactions). Developing and profiling more ecologically valid frame time variation approaches an invaluable contribution left to future work. Furthermore, we had a relatively small sample size. Future work can include a larger and more diverse sample of participants and use more objective/rigorous measurement techniques (e.g., psychophysics) to replicate and extend our findings.

1) Gaming beyond the virtual sandbox: The goal of our study was to isolate the effects of VFT on perception of smoothness As a result, we selected controlled tasks and procedures in this experiment that maintained high internal validity while sacrificing external validity. Real-world gaming scenarios vary considerably with respect to our study. For example, the graphics are much richer and gamers engage in tasks for longer periods of time. Thus, our results can inform future work about the moment-to-moment effects of VFT on perceptual smoothness.

However, our results are limited in predicting the gamer experience in events that span longer periods of time such as traveling from a respawn location to the central battlefield where all the action occurs in a game series like *Call of Duty*. Moreover, it is unclear how VFT impacts perceptual experience over longer periods of gameplay. It could be the case that people adapt to the perceptual artifacts caused by VFT and do not experience it after 30 minutes of gaming.

C. Suggestions for Researchers

Researchers studying frame rate or time as an independent variable in perceptual contexts should carefully consider the effects of variation over time. Systems used to conduct experiments should have appropriate hardware specifications to guarantee they meet or exceed the desired frame rate and potentially inject additional delay to regularize frame times to as close to constant as possible. When reporting frame times as a single period, or frame rate, investigators should report variational statistics demonstrating how well controlled their constant frame rate/time was. When frame rates or times are not sufficiently well controlled, we have demonstrated that they significantly affect the perception of smoothness, with the mean frame time mispredicting the equivalent perception of a constant frame time condition.

We note that non-VRR systems with VSync disabled provide an experience similar to VRR, but with spatial artifacts (*tears*) present in each frame. These tears come from a newly rendered frame buffer being swapped in during scan out of the previous frame buffer, in some cases resulting in an screen space discontinuity. The impact of these tears comes in part from where they are located on screen and also the relationship of frame time and display update period. Less impacts of tearing we suspect that a study of VFT with VSync off would yield similar results to ours.

D. Suggestions for Gamers

Gamers should be mindful of additional dynamic workloads and other sources of VFT in their systems. For this reason, they are uniquely well suited to evaluate the influence of variable frame timing on their gameplay as the effects seem primarily perceptual. When VFT becomes bothersome to gamers, they have several options for improving the regularity of frame times. Traditional techniques like buffering larger numbers of frames and enabling VSync can help in some cases, but at the cost of additional latency and increased VFT in certain non-performant cases. Alternatively, gamers can use VRR or approaches like in-game frame rate limiters and adjust settings to reduce additional compute workloads. Here the primary benefit of VRR is that, unlike more traditional display techniques, it can target constant frame rates other than traditional display refresh rates without exotic display configuration (i.e., 60, 120, 144, 240 Hz). This allows the gamer to better tailor their ideal, consistent refresh rate to the system capabilities by running at 100 Hz instead of alternating between 1/60 s and 1/120 s frames.

V. Conclusions

This work focused on the influence of variable frame timing on human perception of smoothness in FPS-style gameplay. It considered a wide range of constant frame time sequences from 1/30 to 1/240 s and variable frame time sequences that oscillated between those typical frame times. Across the frame times tested, this user study demonstrates that large deviations from the average frame time of the VFT sequence negatively impact the perception of smoothness. Specifically, as the variability in frame time within a sequence increased (i.e., larger amplitude), the perception of smoothness decreased relative to a CFT with a similar mean frame time.

References

[1] V. S. Ramachandran and S. M. Anstis, "The perception of apparent motion," *Scientific American*, vol. 254, no. 6, pp. 102–109, 1986.

[2] G. Denes, A. Jindal, A. Mikhailiuk, and R. K. Mantiuk, "A perceptual model of motion quality for rendering with adaptive refresh-rate and resolution," *ACM Trans. Graph.*, vol. 39, no. 4, jul 2020. [Online]. Available: https://doi.org/10.1145/3386569.3392411

[3] S. Liu, A. Kuwahara, J. J. Scovell, and M. Claypool, "The effects of frame rate variation on game player quality of experience," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2023, pp. 1–10.

[4] G. A. Slavenburg, M. Janssens, L. Lucas, R. J. Schutten, and T. Verbeure, "Variable refresh rate displays," *SID Symposium Digest of Technical Papers*, vol. 51, no. 1, pp. 669–672, 2020.

[5] E. Callway, "Variable frame rate technology — change is good!" in *SMPTE 2018*. New York, NY, USA: IEEE, 2018, pp. 1–11.

[6] C. H. Poth, R. M. Foerster, C. Behler, U. Schwanecke, W. X. Schneider, and M. Botsch, "Ultrahigh temporal resolution of visual presentation using gaming monitors and g-sync," *Behavior research methods*, vol. 50, no. 1, pp. 26–38, 2018.

[7] S. H. Ko, "A-24 db in-band noise-immunity mutual capacitance readout system for variable refresh rate of active-matrix oled display," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 4, pp. 1126–1135, 2019.

[8] B. You, H. Nam, and H. Lee, "Image adaptive refresh rate technology for ultra low power consumption," *SID Symposium Digest of Technical Papers*, vol. 51, no. 1, pp. 676–679, 2020. [Online]. Available: https://sid.onlinelibrary.wiley.com/doi/abs/10.1002/sdtp.13958

[9] R. K. Mantiuk and G. Denes, "Adaptive refresh rate and resolution: Exploiting the limitations of spatio-temporal vision," *SID Symposium Digest of Technical Papers*, vol. 52, no. 1, pp. 465–468, 2021. [Online]. Available: https://sid.onlinelibrary.wiley.com/doi/abs/10.1002/sdtp.14718

[10] A. Jindal, K. Wolski, K. Myszkowski, and R. K. Mantiuk, "Perceptual model for adaptive local shading and refresh rate," *ACM Transactions on Graphics (TOG)*, vol. 40, no. 6, pp. 1–18, 2021.

[11] M. Riahi and B. A. Watson, "Am i playing better now? the effects of g-sync in 60hz gameplay," *Proc. ACM Comput. Graph. Interact. Tech.*, vol. 4, no. 1, apr 2021. [Online]. Available: https://doi.org/10.1145/3451269

[12] B. Watson, R. Shrivastava, and A. Gavane, "The effects of adaptive synchronization on performance and experience in gameplay," *Proc. ACM Comput. Graph. Interact. Tech.*, vol. 2, no. 1, jun 2019. [Online]. Available: https://doi.org/10.1145/3320286

[13] A. Watson and A. Ahumada, "The pyramid of visibilty," *Journal of Vision*, vol. 16, no. 12, pp. 567–567, 2016.

[14] S. Daly, N. Xu, J. Crenshaw, and V. J. Zunjarrao, "A psychophysical study exploring judder using fundamental signals and complex imagery," *SMPTE Motion Imaging Journal*, vol. 124, no. 7, pp. 62–70, 2015.

[15] P. V. Johnson, J. Kim, D. M. Hoffman, A. D. Vargas, and M. S. Banks, "Motion artifacts on 240-hz oled stereoscopic 3d displays," *Journal of the Society for Information Display*, vol. 22, no. 8, pp. 393–403, 2014.

[16] A. Chapiro, R. Atkins, and S. Daly, "A luminance-aware model of judder perception," *ACM Trans. Graph.*, vol. 38, no. 5, jul 2019. [Online]. Available: https://doi.org/10.1145/3338696

[17] A. B. Watson, A. J. Ahumada, and J. E. Farrell, "Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays," *JOSA A*, vol. 3, no. 3, pp. 300–307, 1986.

[18] J. Spjut, B. Boudaoud, K. Binaee, A. Majercik, M. McGuire, and J. Kim, "Firstpersonscience: Quantifying psychophysics for first person shooter tasks," in *UCI Esports Conference*. Irvine: UCI, 2019.

[19] B. Boudaoud, J. Spjut, and J. Kim, "Firstpersonscience: An open source tool for studying fps esports aiming," in *Special Interest Group on Computer Graphics and Interactive Techniques Conference Talks*. New York, NY, USA: Association for Computing Machinery, 2022, pp. 1–2.

[20] ap, "Visual sorbet," Dec 2015. [Online]. Available: https://www.shadertoy.com/view/Xdd3D7