Frame rate vs Resolution: a subjective evaluation of spatio-temporal perceived quality under varying computational budgets

K. Debattista[†1], K. Bugeja[2], S. Spina[2], T. Bashford-Rogers[3], V. Hulusic[4]

[1] WMG, University of Warwick, UK
[2] Dept. of Computer Science, University of Malta, Malta
[3] Computer Science and Creative Technologies, University of the West of England, UK
[4] LTCI, Télécom ParisTech , Université Paris-Saclay, France

Abstract

Maximising performance for rendered content requires making compromises on quality parameters depending on the computational resources available. Yet, it is currently unclear which parameters best maximise perceived quality. This work investigates perceived quality across computational budgets for the primary spatio-temporal parameters of resolution and frame rate. Three experiments are conducted. Experiment 1 (n = 26) shows that participants prefer fixed frame rates of 60 frames per second (fps) at lower resolutions over 30 fps at higher resolutions. Experiment 2 (n = 24) explores the relationship further with more budgets and quality settings and again finds 60 fps is generally preferred even when more resources are available. Experiment 3 (n = 25) permits the use of adaptive frame rates, and analyses the resource allocation across seven budgets. Results show that while participants allocate more resources to frame rate at lower budgets the situation reverses once higher budgets are available and a frame rate of around 40 fps is achieved. In the overall, the results demonstrate a complex relationship between frame rate and resolution's effects on perceived quality. This relationship can be harnessed, via the results and models presented, to obtain more cost-effective virtual experiences.

Categories and Subject Descriptors (according to ACM CCS):   I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation

1. Introduction

Real-time image-synthesis generates rendered content which can be visualised at interactive rates for a variety of applications including games, virtual reality and simulation. However, current levels of computational performance still require compromises to be made and this has an effect on the quality of the delivered experience. Quality in computer graphics may depend on many parameters, such as resolution, frame rate, geometric complexity, texture quality, shading etc. Computational resources have been increasing significantly, however, until both hardware and software in rendering systems can exceed thresholds of the human visual

system, decisions on what settings to apply to quality parameters are still required.

Rendering a scene at given quality settings requires a certain amount of computation, we refer to as computational cost. The resource delivering the content has computational performance limitations that shall be referred to as the computational budget. Quality parameters are likely to change when the same content is delivered on devices with different computational budgets, for example, the switch from a desktop to a laptop, or when upgrading a graphics card. When the computational cost provided by given settings is greater than the computational budget available these settings must be adjusted to meet the budget. However, it is unclear how and which quality parameters should be modified to obtain the best perceived quality. Would any given extra perfor-
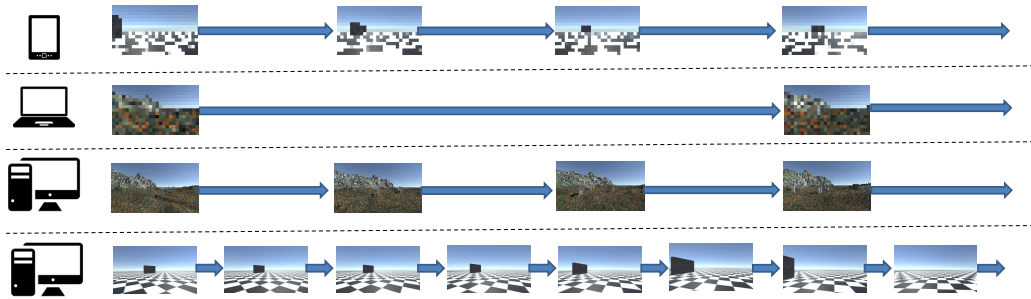
---

† K.Debattista@warwick.ac.uk

Figure 1: *Adjusting spatio-temporal parameters based on perceived quality and computational budget. Different frame rate and resolution settings are used by different resources to maximise perceived quality for different scenarios.*

mance be used more effectively to render more geometry, increment the frame rate, improve the physics simulations, raise the resolution or increase the quality of other parameters?

This work explores the effect of changing settings on perceived quality at different computational budgets. We hypothesise that the changes in quality settings vary depending on the computational budget available. In order to accomplish this, and since, to the best of our knowledge, this is the first study of its type, this work considers the primary spatio-temporal quality parameters of frame rate and resolution. Figure 1 demonstrates how different resources could render scenarios at different frame rates and resolutions. Typically, frame rate is kept fixed and resolution is adjusted to maintain this constant or vice-versa. This may not be the best use of the computation available, yet, it is unclear which combination of frame rate and resolution settings provide the best perceived quality for that resource.

The goal of this work is to identify the best settings to maximise perceived spatio-temporal quality. Via a novel methodology the spatio-temporal relationship is examined. The method permits participants to compare and directly allocate quality parameters over distinct computational budgets; it is demonstrated and validated across three experiments. The first two with locked frame rates and the third unlocked which permits the creation of a model that recommends the parameters required to achieve the best perceived quality given a computational budget. The main contributions of this work are:

- A methodology for analysing preferences in resource allocation for spatio-temporal quality settings that can be used to obtain the best frame rate and resolution combination for a given scenario.
- The finding that participants prefer 60 fps at the cost of lower resolutions than 30 fps at higher resolutions.
- The finding that at lower computational budgets more resources are allocated to frame rate than to resolution. Subsequently, when frame rates reach around 40 fps, alloca-

tion to resolution overtakes that to frame rate. Finally, allocation may begin to even out, see Figure 5.
- The presentation and validation of a model for predicting the spatio-temporal parameters that provide the best perceived quality given a computational budget.

2. Background and Related Work

Background and related work on frame rate and resolution and their interaction is initially presented and followed by an overview of constraint-based rendering.

2.1. Frame rates and resolutions

Frame rate and resolution constitute two main characteristics of modern displays and rendering systems. The spatial extent of the delivered content is based on the spatial resolution or number of pixels in a given frame and the temporal resolution on the frame rate, or number of frames computed per second. Resolution, particularly, lends itself to substantial increases and is advertised as one of the selling points of new consumer technologies. Current technology supports resolutions of partial HD ($1,280 \times 720$), (full) HD ($1,920 \times 1,080$ pixels), currently the de facto standard, while higher resolutions, such as WQHD ($2,560 \times 1,440$), and UHD ($3,840 \times 2,160$) are becoming popular. Ever since the introduction of sound tracks in 1930, the sound film was standardised at 24 fps [RM00]. Performance of different frame rates on tasks has been analysed in the past [CT07]; however, this is mostly for lower frame rates (≤ 30 fps). Higher frame rates can be considered beneficial for film [WAH*15], and interactive experiences, particularly virtual reality [ZRSK15] and video games [CC07]. McCarthy et al. [MSM04] compared frame rate and resolution for streamed video of sports events on small screen devices by reducing both and found reductions in resolution to be less desirable than reductions in frame rate. Cranley et al. [CPM06] analyse frame rate and resolution for video encoding via pairwise comparisons. Their study is similar in concept to ours but is limited to a maximum resolution of 176×144 and frame rate of 25 fps, a

fixed and relatively small choice of relationships and no interactivity.

In computer graphics and visualisation there has been some direct exploration of resolution and frame rate. Claypool and Claypool [CC07] investigated the effect of frame rate and resolution on first person shooters by presenting participants with a scenario in which frame rates and resolutions were set out of five frame rates and three possible resolutions. They found that higher frame rates improve performance for movement and shooting with frame rates up to 60 fps providing a significant improvement in performance. The tested three resolutions did not provide a significant difference in performance. Claypool and Claypool [CC09] extended the work to a different genre with other perspectives considered, and again found a general improvement in performance for frame rate and not resolution for the fixed scenarios tested. Didyk et al. [DER*10] ran a subjective evaluation in which participants rated perceived blur in animations rendered at 120 fps significantly higher when compared to 40 fps and also found participants' performance in a game improved significantly for 120 fps compared to 40 fps and 60 fps. Bracken and Skalski [BS09] found higher levels of presence in a video game at full HD resolutions of 1080i as opposed to 480p. Hänel et al. [HWHK16] compared frame rate and rendering quality in a volume renderer for visualisation and found an adaptive continuous adjustment method performed better than fixed quality or frame rate.

Rendering systems will frequently attempt to fix either resolution or frame rate while adapting the other. Wong and Wang [WW13] used a control engineering method to maintain fixed frame rates for real-time rendering. An alternative by Frey et al. [FSME14] made use of spatio-temporal control for a volume renderer based on spatial and temporal errors computed in real-time, however, this work is largely applicable to progressive rendering whereby the frame is refined over time unless a change in rendering context, such as user navigation, occurs.

2.2. Rendering with varying settings

Funkhouser and Séquin [FS93] described a system of rendering under timing constraints to maintain fixed frame rates for interactive rendering by attempting to maximise a benefit function while maintaining a cost function within a given constraint; this work focused primarily on level of detail methods, but the general framework can be applied to more generic aspects of rendering. Dumont et al. [DPF03] described rendering under constraints for global illumination using a hierarchical radiosity solution. Recently, Wang et al. [WYM*16] used power as the constraint in order to maximise visual quality for devices with limited power supplies. The above systems constrain over some resource but their formulation does not rely on quality that is based on data capture with real participants. Slater et al. [SSC10] permitted participants to modify the quality of scenarios directly

for a number of conditions such as direct lighting. However this was done in the context of identifying presence in an environment and was independent of constraints.

In contrast with the work outlined above, the work presented here gives participants the ability to choose the settings they prefer interactively, it compares the relationship of frame rate and resolution directly by varying both concurrently. Furthermore, it explores resolution and frame rate combinations based on computational costs which are then used to form a predictive model for parameter selection.

3. Motivation and Overview

Quality in rendering is a function of many parameters: resolution, frame rate, geometric complexity, texture quality, shading etc. Formally, the perceived quality Q of a delivered scenario is a function of modifiable parameters $p_1, \ldots, p_n$. The computational cost C of rendering a given scenario also depends on $p_1, \ldots, p_n$, and on the resource R as different hardware can be more/less adept at certain computations. In order to maximise the use of R, it would be ideal to obtain the highest $Q(p_1, \ldots, p_n)$ such that $C(R, p_1, \ldots, p_n)$ does not exceed the computational budget of the resource $B(R)$. This can be formulated as:

$$\max_{p_1, \ldots, p_n} Q(p_1, \ldots, p_n) \text{ s.t. } C(R, p_1, \ldots, p_n) \leq B(R). \quad (1)$$

$B(R)$ can be calculated by measuring the computational performance for that resource. $C(R, p_1, \ldots, p_n)$ can be computed statically before run time or interactively. Due to the complexity of Equation 1, this work explores the relationship between the two primary spatio-temporal parameters of frame rate and resolution.

3.1. Motivation for choice of frame rate and resolution as parameters

Perceived quality in terms of frame rate and resolution transcends graphics and permeates all of visual computing, for example in multimedia, video games, photography etc. Frame rate and resolution are, for the most part, directly related, doubling resolution or frame rate incurs the same general overall cost. Furthermore, most rendering algorithms such as Monte Carlo-based methods and rasterisation methods are a function of resolution. Many aspects of rendering such as shading, texturing, lighting are a function of resolution too. Admittedly, certain methods do not necessarily scale linearly over time and space (e.g. various caching methods [WRC88]) but certain aspects of these computations would still do. Exploring the perceived fidelity of a large set of quality settings in Equation 1 would require a combinatorial increase in the number of trials and attributes to measure. Furthermore, frame rate and resolution are linearly dependent making the cost C in Equation 1 relatively
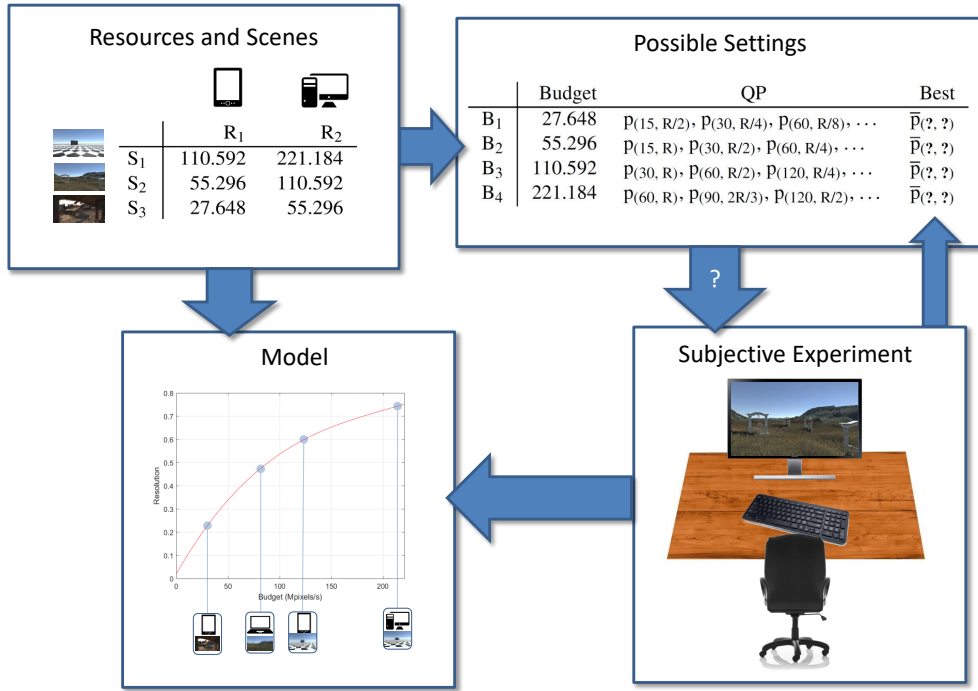
Figure 2: *Methodological Overview. Certain resources can render scenes at better quality than others, yet, it is unclear which settings result in the best perceived quality. A series of subjective experiments are carried out to identify the best perceived quality and present a model that identifies the best settings given the scene pixel complexity and the computational resource available.*

easy to compute as the number of pixels computable within a given time frame, assuming pixels in a given scenario on average have the same cost.

3.2. Frame rate vs resolution

When considering frame rate and resolution only, Equation 1 becomes:

$$\max_{p_{(fps,res)} \in QP} Q(p_{(fps,res)}) \text{ s.t. } C(R, p_{(fps,res)}) \leq B(R), \qquad (2)$$

where $p_{(fps, res)}$ is short for the parameters for frame rate and resolution (p_{fps}, p_{res}), and QP is the set of all $p_{(fps, res)}$ with the same cost $C(R, p_{(fps,res)})$. The main objective of this work is two-fold:

1. to identify the maximum perceived quality Q across varying computational budgets to be able to recommend a $\bar{p}_{(fps, res)}$ for that budget
2. to investigate whether the $\bar{p}_{(fps, res)}$ ratio changes across budgets.

For 2., our hypothesis ($\mathbf{h}_1$) is that the ratio does change across budgets. The experiments set out to identify this across a number of trials as:

$$\max_{p_{(fps,res)} \in QP_i} Q_i(p_{(fps,res)}) \text{ s.t. } C(p_{(fps,res)}) = B_i, \text{ for } i = 1, \ldots, m,$$

$$(3)$$

where participants are asked to identify the value of Q_i across m simulated budgets (henceforth budgets) $B_{1, \ldots, m}$ for each experiment. The $p_{(fps, res)} \in QP_i$ are chosen such that $\forall \, p_{(fps, res)} \in QP_i, C(p_{(fps, res)}) = B_i$ and the experiments identify the particular $\bar{p}_{(fps, res)}$ that produces the maximum Q_i for trial i. Furthermore, since simulated budgets are used R is removed from both sides of the equation.

An overview of the methodology with an illustrative example that helps explain the motivation and process is shown in Figure 2. Without loss of generality, it is assumed that the cost and budget are measured in mega-pixels per second (Mpixels/s). The table labelled "Resources and Scenes" shows an example of the performance of two resources R_1 and R_2 across three scenes S_1, S_2, S_3. Scenes range in order of complexity from S_1 to S_3 and hence require further

computation to render and R_1 is half as powerful as R_2. For any given budget B_i and QP_i containing a series of settings $p_{(fps, res)}$ (see the table "Possible Settings" in Figure 2) that could be run at that budget, it is desirable that the winning $\bar{p}_{(fps, res)}$ is identified. The subjective evaluation provides this information. By fixing B_i to a fixed number of budgets the best $\bar{p}_{(fps, res)}$ is identified for each scenario and thus constitutes the best frame rate and resolution combination to be used under that budget. Furthermore, for unlocked frame rates (Experiment 3), the results are also used to build a model which can predict the right $\bar{p}_{(fps, res)}$ based on the input budget. While results from this work can be used directly, since the work presented is based on five scenarios, other scenarios which do not fit the mould presented can use the same proposed methodology to obtain a frame rate and resolution for that particular case.

3.3. Overview of experiments

The method is demonstrated with data captured and investigated over three experiments. Experiment 1 investigates the relation between frame rates and resolution for locked frame rates at 15, 30 and 60 fps across two simulated budgets ($m = 2$). Experiment 2 extends this to include higher frame rates (90 and 120 fps) and the inclusion of two more simulated budgets ($m = 4$). Experiment 3, further increases simulated budgets ($m = 7$), and also removes the fixed frame rate constraint permitting any choice of frame rate and resolution combination that amount to the same computational cost. The captured data is used to build a model when given a computational cost as outlined in Section 3.2 This work is limited to non-interactive environments as we wanted to control for screen latency, but, we suspect that interactivity will have an effect on results [CC09]. In the experimental description $p_{(x, yR)}$ denotes the parameters with x fps and yR resolution where $R = 2{,}560 \times 1{,}440$, $y \in (0, 1]$ constant, per trial. The 16:9 aspect ratio is maintained throughout all experiments.

4. Experiment 1: Exploring traditional settings

Experiment 1 was designed to investigate the difference in quality of distinct frame rate and resolution combinations with the same computational costs. For this experiment parameters consist of fixed frame rates up to 60 fps and a maximum resolution of $2{,}560 \times 1{,}440$. This serves to give an indication of which frame rate and resolution settings result in the best perceived quality and whether computational performance is best dedicated to increasing frame rate or resolution on current standard display devices.

4.1. Design

The design employed by this experiment was a forced choice evaluation. Two budgets, each one targeting a fixed computational cost where considered. The first, B_1, corresponds to the maximum resolution running at 15 fps or 55.296 Mpixels/s with three possible parameter options $p_{(15, R)}$, $p_{(30, R/2)}$ and $p_{(60, R/4)}$. As mentioned above, $R = 2{,}560 \times 1{,}440$, and therefore, R/2 corresponds to a resolution of $1{,}810 \times 1{,}018$ and R/4 to $1{,}280 \times 720$. A second budget B_2 consisted of 110.592 Mpixels/s with two possible parameters $p_{(30, R)}$ and $p_{(60, R/2)}$. Locked frame rates synchronised with the display were chosen to avoid out-of-sync duplicate frames that occur with v-sync and can cause judder, and/or screen tearing, an objectionable artifact, that would occur without v-sync if computation were not fast enough. In terms of Equation 3, $m = 2$, with $QP_1 = \{p_{(15,R)}, p_{(30,R/2)}, p_{(60,R/4)}\}$ and $QP_2 = \{p_{(30,R)}, p_{(60,R/2)}\}$.

The forced choice method was applied by presenting a looping animation to the participants and allowing them to switch among options in the experiment interactively, as many times as they wanted and take as long as they needed. B_2 was tested by allowing participants the choice of $p_{(30, R)}$ and $p_{(60, R/2)}$ via keyboard controls. B_1 was split into two computations. B_{1a} consisted of three trials of 2-way alternative forced choice pairwise comparisons between all the pairs $p_{(15, R)}$, $p_{(30, R/2)}$ and $p_{(60, R/4)}$ consisting in three trials per scene. B_{1b} was a three-way forced choice among $p_{(15, R)}$, $p_{(30, R/2)}$ and $p_{(60, R/4)}$ each one randomly assigned to keys 1 to 3 on the keyboard. B_{1a} and B_{1b} were used to check consistency of results in the methodological approach. In order to maintain control over settings and not increase the complexity of the method no anti-aliasing, both spatially and temporally, and/or blurring effects were applied. Adding these would require careful control over settings and would require a significant increase in test cases explored.

Budget was a within-participants independent variable. B_{1a}, B_{1b} and B_2 was shown to all participants for a total of five tests; three for B_{1a} and one each for B_{1b} and B_2. The order of these trials was randomised. Participants conducted these trials over five different scenes, which was also a within-participants variable. The order of scenes was also randomised as were key bindings to different parameter settings for each trial.

4.2. Materials

The five scenes chosen are shown in Figure 3. Names in brackets will be used to refer to the scene for the rest of the paper. They cover a number of typical scenarios, two outdoor scenes and two indoor scenes, and a wide variation in complexity from plain graphics in St to more complex scenarios like Corr, Rep and VV. Figure 4 shows box plots of the spatial and temporal perceptual information of the scenes calculated based on ITU P.910 recommendations [IL99]. There is a broad spatial variation and significant diversity in the temporal index, witnessed by the range and variance, even though means are fairly constant in the temporal domain.

Each scene was presented as a single infinitely looping

(a) Corridor (Cor)	(b) Pendula (Pen)	(c) Republique (Rep)	(d) Straightforward (St)	(e) Viking Village (VV)

Figure 3: *Scenes used in the experiments. Cor and VV, are courtesy of Unity technologies. Rep is courtesy of Camouflaj. All three are available from the Unity asset store. Pen and St were created in house.*

animation. The experiments were implemented in Unity and permitted a user to interactively change between budgets through use of the keyboard. Changes were instantaneous. Settings were allocated to '1' and '2' for the pairwise comparisons and '1', '2' and '3' for the 3-way comparisons. The animations were rendered directly at the chosen resolution of yR (R = 2,560 × 1,440, $y \in (0,1]$) and up-sampled via bilinear interpolation to R to be visualised as a full-screen render. Participants pressed 'enter' when they were satisfied with their choice. All other settings were kept constant for all rendered animations. Software was designed to randomly assign a parameter combination to a key, the order of the trials and the order of the presented scenes. While running the software displays the number of possible selections for each trial and the one being currently viewed in the top right corner.

Scenes were presented on a 27" Acer CB270HU display, running at 2,560 × 1,440 (maximum resolution of the display) at 60Hz with a maximum brightness of 350cd/m^2 and 6 ms response time. The machines used consisted of an Intel Core i7-5820K @ 3.30Ghz CPU and an Nvidia GTX 780, running under Windows 10.
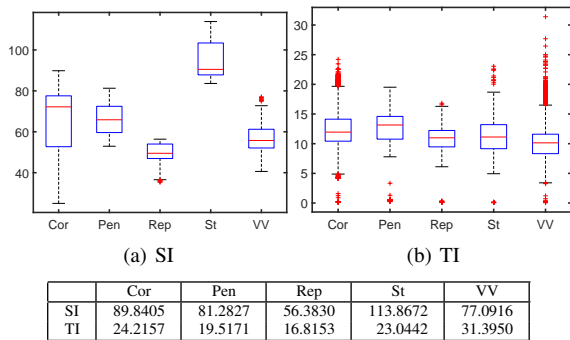
4.3. Participants and procedure

26 participants volunteered to take part in the experiment. Volunteers were composed of faculty, administrative staff and students from a university with an average age of 29.73. All exhibited normal or corrected-to-normal vision.

Participants were asked to sit at the allocated desk and gave their informed consent prior to commencement. They were shown a demonstration with two trials via the St scene and asked to "please select the setting you would choose to experience this virtual environment in". Following the demonstration they were given an opportunity to ask any questions, after which the experiment commenced and no interaction between the experimenter and the participant took place until debriefing once all trials were concluded.

4.4. Results

Descriptive statistics are presented in Table 1. Since frame rate and resolution are linearly related, only the frame rate results are presented. Mode is of particular importance here as it presents the most common selection and would act as the recommended setting in a model. These statistics demonstrate fairly consistent choices by participants and a clear trend to prefer 60 fps over the other choices in the overall in spite of the reduced resolution.

Test across scenes: Results are further analysed on choices made by participants based on the given choice per budget and per scene. For B_{1a} a table of aggregates of the trial choice from all three trials using pairwise comparisons was used to generate a table with winning frequencies for $p_{(15, R)}$, $p_{(30, R/2)}$ and $p_{(60, R/4)}$, with each entry $\in [0, 3]$. For B_{1b} a similar table was computed based on choices for $p_{(15, R)}$, $p_{(30, R/2)}$ and $p_{(60, R/4)}$. For B_2 the table was constructed on choices of $p_{(30, R)}$ and $p_{(60, R/2)}$.

Initially, for each of B_{1a}, B_{1b} and B_2 the choice of $p_{(fps, res)}$ was compared across scenes. For B_{1a} Friedman tests (non-parametric repeated measure ANOVA equivalents) were conducted for the main effect of scenes for the three possible choices ($p_{(15, R)}$, $p_{(30, R/2)}$ and $p_{(60, R/4)}$) and no significance ($p > 0.05$) was found for all three tests. For B_{1b} and B_2, due to participants only choosing one result per trial, the dichotomous nature of the results entailed the use of Cochran's Q test [Coc50] developed for such conditions.

	Cor	Pen	Rep	St	VV
SI	89.8405	81.2827	56.3830	113.8672	77.0916
TI	24.2157	19.5171	16.8153	23.0442	31.3950

Figure 4: *Spatial (SI) and temporal perceptual information (TI) computed based on ITU P.910 recommendations [IL99] overview. Table shows max SI ant TI as recommended in ITU P.910.*

For B_{1b} the main effect of scenes, no significance ($p > 0.05$) was found across all three possibilities. For B_2 Cochran's Q again reported no significance ($p > 0.05$) for both ($p_{(30, R)}$ and $p_{(60, R/2)}$).

Effect of parameters: The participants' choices were tested to see whether there was a significant difference in the choices made. This was done for each budget B_{1a}, B_{1b}, B_2. For these tests, due to the lack of significance in scene differences, results are presented by aggregate choices for each trial across all scenes. For B_{1a} and B_{1b} tests were initially conducted via the Friedman test and subsequently, when significance was identified, pairwise comparisons, with Bonferroni corrections, were conducted via Wilcoxon signed-rank tests (the non-parametric equivalent of the related samples t-test). B_2 used Wilcoxon directly as there were only two choices.

Table 2 summarises the results and comparisons. The main effect of parameters was reported to be significant, for B_{1a}, $\chi^2(2) = 46.32$, $p < 0.01$, for B_{1b}, $\chi^2(2) = 34.85$, $p < 0.01$. For both budgets B_{1a} and B_{1b} pairwise comparisons via Wilcoxon's test, and Bonferroni corrections were subsequently conducted to identify differences between all choices. Statistical differences between all three conditions for B_{1a} and B_{1b} were noted. For B_{1a} the ordering was $p_{(60, R/4)}$ (Mdn = 13.5), $p_{(30, R/2)}$ (Mdn = 6) and $p_{(15, R)}$ (Mdn = 0) and all changes were significant at $p < 0.01$. For B_{1b} the order was the same $p_{(60, R/4)}$ (Mdn = 4), $p_{(30, R/2)}$ (Mdn = 1) and $p_{(15, R)}$ (Mdn = 0), all significant at $p < 0.01$. For B_2 Wilcoxon's tests revealed significance between the two conditions for B_2, $p < 0.01$, $r = -0.59$, with $p_{(60, R/2)}$ (Mdn = 4.5) preferred to $p_{(30, R)}$ (Mdn = 0.5).

Tests across budgets: Finally, we were interested in whether the budget is affecting participants' decisions. In order to test for this, the choices users made were normalised by the allocated budget and compared across B_{1b} and B_2 with scores collapsed across scenes. B_{1a} was not used as its averages would be skewed because of the forced choice for $p_{(15, R)}$ vs. $p_{(30, R/2)}$ which would decrease the average. Wilcoxon's test was applied and a significant difference was found $p < 0.01$, $r = -0.62$. For B_{1b} more resources (Mdn = 0.098×10^{-5}) were allocated to frame rate than for B_2 (Mdn = 0.052×10^{-5}). This entails $\mathbf{h}_1$ holds in this case. However, it is difficult to draw significant conclusions from these two data points hence the following experiments explore more budget possibilities.

5. Experiment 2: Higher budgets and frame rates

Experiment 1 showed a preference for 60 fps across two possible budgets. However, it did not provide an indication of higher frame rates and computational budgets. Experiment 2 used a setup that permitted a frame rate of up to 120 fps which allowed for further budgets and potential parameters.

	Cor	Pen	Rep	St	VV	Overall
			Mean			
$B_{1a,1}$	30.00	28.27	30.00	28.85	29.42	29.31
$B_{1a,2}$	54.23	56.54	56.54	53.08	51.92	54.46
$B_{1a,3}$	58.27	56.54	58.27	58.27	60.00	58.27
B_{1b}	54.23	50.19	54.23	51.92	45.58	51.23
B_2	54.23	53.08	57.69	56.54	54.23	55.15
			Median			
$B_{1a,1}$	30.00	30.00	30.00	30.00	30.00	30.00
$B_{1a,2}$	60.00	60.00	60.00	60.00	60.00	60.00
$B_{1a,3}$	60.00	60.00	60.00	60.00	60.00	60.00
B_{1b}	60.00	60.00	60.00	60.00	60.00	60.00
B_2	60.00	60.00	60.00	60.00	60.00	60.00
			Mode			
$B_{1a,1}$	30.00	30.00	30.00	30.00	30.00	30.00
$B_{1a,2}$	60.00	60.00	60.00	60.00	60.00	60.00
$B_{1a,3}$	60.00	60.00	60.00	60.00	60.00	60.00
B_{1b}	60.00	60.00	60.00	60.00	60.00	60.00
B_2	60.00	60.00	60.00	60.00	60.00	60.00

Table 1: *Descriptive statistics for Experiment 1. $B_{1a,1}$, $B_{1a,2}$ and $B_{1a,3}$ represent $p_{(15, R)}$ vs $p_{(30, R/2)}$, $p_{(30, R/2)}$ vs $p_{(60, R/4)}$ and $p_{(15, R)}$ vs $p_{(60, R/4)}$ respectively.*

QP	B	Mpixels/s	Parameters		
QP_1	B_{1a}	55.296	$P_{(60, R/4)}$	$P_{(30, R/2)}$	$P_{(15, R)}$
QP_1	B_{1b}	55.296	$P_{(60, R/4)}$	$P_{(30, R/2)}$	$P_{(15, R)}$
QP_2	B_2	110.592	$P_{(60, R/2)}$	$P_{(30, R)}$	

Table 2: *Comparisons for Experiment 1. Order (left to right) denotes participant preference. All results are significantly different at $p < 0.05$. R = $2,560 \times 1,440$.*

5.1. Design

The overall design of Experiment 2 follows the design of Experiment 1. However, two new budgets were added to account for the higher frame rates. The new budgets correspond to combinations of the highest resolution with a frame rate of 15, 30, 60, 90 fps. The budgets and resulting trials are described in Table 4, here sorted by the final choices discussed below. In terms of Equation 3, $m = 4$ and $QP_{1...4}$ are shown in Table 4. For B_1 there are now five possible choices, four for B_2, three for B_3 and two for B_4. The design of this experiment follows an n-way forced choice for n $\in [2,5]$. No pairwise comparisons were used for this experiment due to the similarities in B_{1a} and B_{1b} in Experiment 1 and due to the large number of choices that would be required to compare all parameters. The choice of setting was modified via the keyboard and bindings were randomly assigned to settings. The same five scenes as in Experiment 1 were used, with four trials per scene. Both variables were randomly assigned.

5.2. Materials

The same scenes and animations as in Experiment 1 are used. The software also accounted for further key-bindings, up to five, depending on the trial. Hardware was changed to provide support for higher frame rates. A 27" Acer AB271HU display running at $2,560 \times 1,440$, with a brightness of 350cd/m^2 and a latency of 4 ms was used. The monitor used was a G-sync monitor and G-sync was enabled

allowing for unrestricted choice of frame rate without duplicate pixels or screen tearing. The PC consists of an Intel Core i7-6700K @ 4.0Ghz and an Nvidia GTX 1080 GPU running under Windows 10.

5.3. Participants and procedure

24 participants volunteered for this experiment. They exhibited normal or corrected-to-normal vision. Participants for Experiment 2 were mostly drawn from faculty, administrative staff and students of a university; this university is located in a different country from the one in Experiment 1. Average participant age was 32.5. The procedure followed Experiment 1 except for the choice of further options in the experiment.

5.4. Results

Descriptive statistics are shown in Table 3. As with Experiment 1 all results showed preference for higher frame rates, and mode and median results present a strong preference for 60 fps choices for all budgets that included that option. B_1 and B_2 re-enforce findings from Experiment 1. The Rep scene shows an increased choice in frame rate at higher budgets, not exhibited by the other scenes.

Test across scenes: Results for the main effect of scenes were conducted similarly to Experiment 1. The number of tests conducted depended on the budget. Since all budgets use n-way forced choice the results per scene are dichotomous, hence Cochran's Q-test was used for all of them. For B_1, B_2 and B_3 no significant difference was observed across scenes, $p > 0.05$. For B_4 a significant difference was observed, $p < 0.05$. As expected from the descriptive statistics it is Rep that was different at the higher budget, pairwise comparisons show a difference between Rep and VV and Rep and Cor. No other differences are noted.

Effect of parameters: As with Experiment 1, results for chosen parameters were computed via aggregating parameter choices across the scenes. Results for each budget (with more than two parameters compared, all except B_4) were first analysed via a Friedman test and subsequently pairwise comparisons applied. Table 4 shows ordering for all parameters tested for each budget and shows groupings for those parameters that did not provide significant differences.

B_1 produced a significant effect $\chi^2(4) = 44.31$, $p < 0.01$. Results for pairwise comparisons using Wilcoxon signed-rank test are shown in Table 4 showing preference for $p_{(60, R/4)}$ (Mdn = 3), significantly different from other conditions tested. The other parameters are grouped into two overlapping groups as shown via coloured groups (entailing no significant difference) in Table 4. Medians were $p_{(30, R/2)}$ (Mdn = 1), $p_{(90, R/6)}$ (Mdn = 0.5) and $p_{(15, R)}$ (Mdn = 0).

B_2 also produced a significant effect $\chi^2(3) = 13.00$, $p < 0.01$. Colour groupings are shown in Table 4. Medians were

$p_{(60, R/2)}$ (Mdn = 2), $p_{(90, R/3)}$ (Mdn = 1), $p_{(30, R)}$ (Mdn = 1) and $p_{(120, R/4)}$ (Mdn = 1). B_3 also resulted in a significant effect $\chi^2(2) = 28.17$, $p < 0.01$. Two groups are identified $p_{(60, R)}$ (Mdn = 3.5), $p_{(120, R/2)}$ (Mdn = 1), $p_{(90, 2R/3)}$ (Mdn = 0). B_4 was tested with Wilcoxon signed test as only two conditions were present. It was found to be significant at $p < 0.05$, $r = 0.33$, $p_{(90, R)}$ (Mdn = 3) was preferred to $p_{(120, 3R/4)}$ (Mdn = 2).

Results confirm conclusions from Experiment 1, 60 fps is the preferred frame rate again, even for higher computational budgets. In Experiment 1, 60 fps was the highest frame rate and it was always chosen, here it is no longer the highest and it is still chosen, clarifying that perceived quality is not just a function of frame rate. At higher computational budgets while there is a preference for 90 fps the effect is not that pronounced as shown by the results in B_3, no significant difference, and B_4 where there is a difference but not as distinct as with other frame rates and budgets.

Tests across budgets: Finally, as with Experiment 1, the main effect of budget was tested. The data was normalised by budget across all budgets. A Friedman test gave significance for the effect of budget $\chi^2(3) = 63.79$, $p < 0.01$. As with Experiment 1, B_1 allocates more resources (Mdn = 0.11 $\times 10^{-5}$) to frame rate than for B_2 (Mdn = 0.065 $\times 10^{-5}$). B_3 (Mdn = 0.033 $\times 10^{-5}$) and B_4 (Mdn = 0.030 $\times 10^{-5}$) follow in that order. Pairwise comparisons with Bonferroni corrections showed significant differences between B_1 and the rest, B_2 and the rest but no significant difference between B_3 and B_4. Again, $\mathbf{h}_1$ holds. These results show a trend in decrease of frame rate allocation in budgets; lower budgets will be investigated in the following experiment.

	Cor	Pen	Rep	St	VV	Overall
			Mean			
B_1	56.25	56.25	66.88	51.88	49.38	56.13
B_2	68.75	83.75	82.50	66.25	60.00	72.25
B_3	70.00	78.75	83.75	71.25	70.00	74.75
B_4	101.25	97.50	108.75	100.00	96.25	100.75
			Median			
B_1	60.00	60.00	60.00	60.00	60.00	60.00
B_2	60.00	90.00	75.00	60.00	60.00	60.00
B_3	60.00	60.00	75.00	60.00	60.00	60.00
B_4	90.00	90.00	120.00	90.00	90.00	90.00
			Mode			
B_1	60.00	60.00	60.00	60.00	60.00	60.00
B_2	60.00	60.00	60.00	60.00	60.00	60.00
B_3	60.00	60.00	60.00	60.00	60.00	60.00
B_4	90.00	90.00	120.00	90.00	90.00	90.00

Table 3: *Descriptive statistics for Experiment 2.*

6. Experiment 3: Unlocking frame rates

This experiment frees the constraint of having fixed frame rates. While not practical for the majority of displays, the current support of adaptive sync/G-sync/FreeSync in modern displays enables the study of unlocked frame rates without the constraints of fixed frame rates that occur with v-sync or the screen tearing that occurs without. Furthermore, lower computational budgets were explored and the data was used

QP	B	Mpixels/s	Parameters				
QP_1	B_1	55.296	P(60, R/4)	P(30, R/2)	P(90, R/6)	P(15, R)	P(120, R/8)
QP_2	B_2	110.592	P(60, R/2)	P(90, R/3)	P(30, R)	P(120, R/4)	
QP_3	B_3	221.184	P(60, R)	P(120, R/2)	P(90, 2R/3)		
QP_4	B_4	331.776	P(90, R)	P(120, 3R/4)			

Table 4: *Comparisons for Experiment 2. Order (left to right) denotes participant preference. Circles show results that are significantly different at p < 0.05. R = 2,560 × 1,440.*

to build a model that predicts the best quality parameters given a computational budget.

6.1. Design

In this study participants were allowed to allocate any frame rate resolution combination to fit within their given budget. Unlike the QP contents from previous experiments, $QP_{1...7}$ do not consist of a small discrete number of $p_{(fps, res)}$ but a continuous set of options. This was achieved via a slider at the bottom of the screen that participants could freely manipulate across a single dimension that modifies both parameters concurrently. The direction was set randomly for each trial. The slider permitted direct change of the frame rate or the resolution in a perceptual uniform space. It could be set in a perceptually uniform mode for frame rate or resolution. When the slider is set to perceptually uniform frame rate, a change on the slider corresponds to a change in the perceptually uniform space of frame rate and the resolution is changed accordingly. Similarly, when the slider is set to a perceptually uniform resolution mode slider changes correspond to perceptually uniform changes for resolution. The perceptually uniform spaces were constructed via a psychophysics experiment discussed in Section 6.2. Each budget for each scene consisted of two runs, one for each slider.

Seven budgets were used for this task and can be seen in Table 5. These included three smaller budgets not used in the previous two experiments, and an added intermediate budget B_5 at 82.944 Mpixels/s not explored earlier. In total each participant ran 5 (scenes) × 7 (budgets) × 2 (sliders) = 70 conditions.

B_1	B_2	B_3	B_4	B_5	B_6	B_7
6.912	13.824	27.648	55.296	82.944	110.592	221.184

Table 5: *Budgets in Mpixels/s for Experiment 3.*

6.2. Defining a perceptually uniform space for the sliders

Since a linear slider change would entail very quick changes for lower frame rates/resolutions, for example, for possible frame rates of 1 to 120 the changes across half the slider (60-120) may appear largely imperceptible to the participants while big changes may occur for the initial half. Ideally, the sliders would provide perceptually uniform changes for

the participants. In order to achieve this a psychophysics experiment was run to construct a perceptually uniform space across both frame rate and resolution.

Design: Two experiments were run, one for resolution and one for frame rate. All aspects of the design were the same except the change in stimuli. A one-way forced choice psychophysics design was employed. Stimuli were presented as a series of seven second animations and at some random point, between the 2 to 4.5 second bracket, the resolution/frame rate could change (increase) and if the participant detected the change they recorded the event by pressing the space bar. The goal was to complete a fully uniform space, hence six points were chosen at which a psychometric curve would be generated for both frame rate and resolution; these points are termed anchors. Six anchors were chosen at 12, 15, 30, 45, 60 and 90 fps for frame rate and at 0.1R, 0.125R, 0.25R, 0.375R, 0.5R and 0.75R for resolution. Changes occurred at varying degrees of the stimulus δ frame rate/resolution. Initially changes were relatively large but they were decreased based on a uniform distribution in $\log_{10}$ from up to double the anchor to the actual stimulus value (e.g. for the anchor at 30 the trials would commence by increasing up to 60 and the δ increase would decrease in exponential steps getting the total frame rate to 30 if successful attempts are recorded). For simplicity, all changes were conducted from the anchor to a higher frame rate/resolution. A 3-up 1-down approach was employed; successful trials decreased δ by one step, an unsuccessful one sent it up by three steps. Each participant performed 20 trials for each of the chosen anchors per scene. The test was conducted via animations to recreate the conditions found in Experiment 3. For the temporal tests the resolution was fixed at 2,560 × 1,440, for spatial tests the frame rate was fixed at 60 fps.

Materials: Two scenes, Pen and St, were used. Two random animations were used for each scene. The experimental software was implemented in Unity. The application produced the next trial according to the previous result.

Participants: Eight participants conducted the experiment. They exhibited normal or corrected-to-normal vision.

Results Data at each anchor point was aggregated across both scenes and all participants, and used to construct a psychometric curve via a success rate vs δ resolution/frame rate plots. The data was used to fit a psychometric curve (via $\frac{1}{1+e^{-a(x-b)}}$) at each of the six anchors and the threshold at 75% success rate was calculated as the threshold for each an-

	Cor	Pen	Rep	St	VV	Overall
Mean (fps)						
B_1	22.82	19.14	28.76	25.28	22.54	23.71
B_2	27.80	24.78	31.92	36.02	27.00	29.50
B_3	34.56	34.78	42.10	38.50	34.94	36.98
B_4	49.58	43.46	50.36	45.58	42.36	46.27
B_5	48.54	56.34	62.26	52.94	50.54	54.12
B_6	62.06	62.02	74.78	52.58	55.66	61.42
B_7	87.90	91.10	88.08	83.22	84.84	87.03
Median (fps)						
B_1	22.50	20.50	28.50	24.00	23.00	23.00
B_2	26.00	25.00	30.50	34.00	28.50	28.50
B_3	36.00	33.00	41.50	36.50	33.00	36.00
B_4	54.00	47.00	47.00	41.50	38.00	47.00
B_5	51.00	54.50	59.50	46.50	50.00	51.00
B_6	62.00	55.00	75.00	46.50	52.00	55.00
B_7	90.00	87.50	90.00	85.50	90.00	90.00

	Cor	Pen	Rep	St	VV	Overall
Mean (resolution as a factor of R = 2,560 × 1,440)						
B_1	0.17	0.16	0.13	0.12	0.15	0.15
B_2	0.22	0.21	0.19	0.16	0.20	0.20
B_3	0.30	0.28	0.24	0.26	0.31	0.28
B_4	0.40	0.43	0.37	0.40	0.43	0.40
B_5	0.58	0.48	0.47	0.51	0.55	0.52
B_6	0.60	0.62	0.51	0.68	0.65	0.61
B_7	0.74	0.70	0.74	0.78	0.77	0.74
Median (resolution as a factor of R = 2,560 × 1,440)						
B_1	0.09	0.09	0.07	0.09	0.08	0.09
B_2	0.15	0.16	0.13	0.11	0.14	0.14
B_3	0.21	0.24	0.19	0.22	0.23	0.22
B_4	0.33	0.42	0.35	0.38	0.40	0.38
B_5	0.52	0.45	0.44	0.49	0.47	0.47
B_6	0.58	0.63	0.46	0.71	0.63	0.63
B_7	0.74	0.74	0.75	0.76	0.75	0.75

Table 6: *Descriptive statistics for Experiment 3.*

chor. The respective spatial and temporal anchors and their threshold were then used to fit the spatial vs threshold curve and temporal vs threshold curve. Two points at 0 and at R for spatial and 120 fps for temporal were added to enable use with the slider up to a maximum value and fit into a uniform space. The spatial curve was fit to $0.04566e^{3.1x}$ with an $r^2_{adj} = 0.99$ (r^2_{adj} indicates how predictive the model is in terms of the population it was sampled from, typically as $r^2_{adj} = 1 - (1 - r^2)\frac{n-1}{n-p-1}$, where n is number of samples and p number of predictor variables) and the temporal curve to $0.9755x^2 + 0.09821x + 0.009394$ with an $r^2_{adj} = 0.93$. These fittings were chosen as these functions provided the best fit and were easily invertible.

This method could be used for making changes to the settings without users perceiving differences, but we hesitate to use it in this fashion for now, since we do not have large numbers of trials to use it in confidence. However, for the use in sliders, where detection of the perceptual change is used primarily for guidance and is not crucial, it should be sufficient.

6.3. Materials

The same five scenes as in the first two experiments were used for this experiment. The only difference was in the use of the sliders. As mentioned above, the sliders were designed to produce perceptually uniform changes to both resolution and frame rate. Sliders were controlled by either mouse, or the left and right arrow keys for fine movement and up and down for coarser movement. The initial position was determined individually for each trial by a randomly weighted Gaussian centred around the middle of the slider. The same display and hardware was used as in Experiment 2 and G-sync was again enabled.

6.4. Participants and procedure

25 participants volunteered for this experiment. The average age was 31.72. They were comprised from members of staff and students from the same university as in Experiment 2.

They exhibited normal or corrected-to-normal vision. The procedure followed the first two experiments except the introductory demonstration showed participants how to use the sliders.
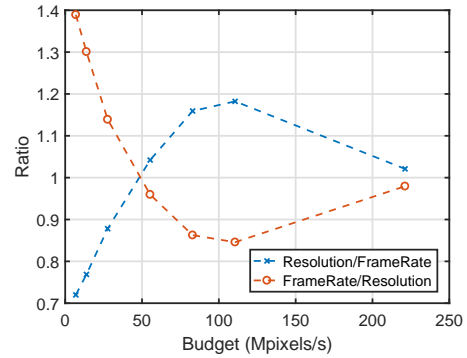

Figure 5: *Ratio of allocated resources. The ratio is the average fps (normalised by 120) to resolution (normalised by R) in red and resolution over fps in blue. A value of 1 would indicate an equal allocation to fps and resolution.*

6.5. Results

Table 6 shows descriptive statistics for both the median and mean of the overall choices across scenes for all the budgets. Notice how the overall budget for fps for B_1 is very close to the generally accepted standard of 24fps. At B_1 this comes at the cost of a significantly reduced resolution of 0.15R.

Tests across scenes: Data was collapsed across the two slider trials to have one result per participant per scene per trial. Friedman tests were conducted per budget. These resulted in a significant effect $p < 0.05$ for B_1, B_2, B_5 and B_6. Pairwise comparisons with Bonferroni corrections were conducted for these cases. For B_1, Pen was significantly different from Rep and St. For B_2 Rep was significantly different from Cor. For B_5 no differences were found. For B_6 Rep was different from St and VV. For all cases but one, the difference involved Rep as found in Experiment 2.

Tests across budgets: The higher number of budgets across a wider range gives a clearer indication of how perceived quality could be changing with budget allocation as Figure 5 illustrates. Here, the x-axis shows the budget and the y-axis the ratio of the average fps (normalised by 120) to resolution (normalised by R) in red and resolution over fps in blue. A value of 1 would indicate an equal allocation to fps and resolution. This suggests that initially an allocation towards higher frame rates was preferred, subsequently, allocation to resolutions were preferred and following this a reduction to a possibly even allocation is seen. Friedman's test found a significant effect of normalised budgets for data collapsed across scenes and slider, $\chi^2(6) = 117.05$, $p < 0.01$. Pairwise comparisons via Wilcoxon tests found significant differences across all budgets. Once again, there was a difference between the resource allocation across all budgets as originally hypothesised. The order and medians of the budgets were: $B_1 = 0.3660 \times 10^{-5}$, $B_2 = 0.2148 \times 10^{-5}$, $B_3 = 0.1331 \times 10^{-5}$, $B_4 = 0.0883 \times 10^{-5}$, $B_5 = 0.0663 \times 10^{-5}$, $B_6 = 0.0550 \times 10^{-5}$ and $B_7 = 0.0387 \times 10^{-5}$.

6.5.1. Model

The data lends itself to creating a model for predicting a given $p_{(fps, res)}$ based on a given computational budget. The computational budget can be calculated statically based on average pixel cost for a scenario on a given platform or can be assessed dynamically, for example based on previous frames. To calculate this curve, data was averaged per trial for each of the participants for fitting the model. A cubic polynomial fit was found to be most appropriate. The model, for budget x as input in Mpixels/s, was fit to the data as follows:

$$0.017x^3 - 0.105x^2 + 0.303x + 0.407, \qquad (4)$$

with $r_{\text{adj.}}^2 = 0.93$. The model predicts resolution as a multiple of R and frame rate is set accordingly. Figure 6(a) shows the participant choices and the fitting. Note some data which may be considered an outlier with high resolution values (belonging to the same participant) was left in for completeness.

6.5.2. Model Validation

Equation 4 could be computed for a limited number of budgets as done above. To test whether the fitting is robust, a k-fold validation was run. k-fold validation is an efficient method that permits the evaluation of a model by removing some samples from the model computation and then testing this newly calculated model with the removed samples. This makes it is possible to evaluate the effectiveness of the model without requiring further expensive data collection. In the approach proposed here, 1-fold or leave-one-out, cross validation is applied. In particular, one of the budget points is left out and the difference between that budget and the prediction of that same budget based on a model computed from the other budgets is calculated. This is conducted across all budgets and the error is averaged. Effectively, for all budgets

except the last, $\hat{B}_i$ is predicted for a model computed with all budgets except B_i and the absolute difference is calculated. For example, $\hat{B}_1$ is predicted from a model computed from B_2 to B_7 and compared with the original B_1. In the overall, this difference was 0.012 in terms of resolution, indicating that the model predicts different budgets well. The difference at B_6 was largest at 0.054 as there are fewer points in that area. Finer grain computations would probably help generate a model with better prediction accuracy. Figure 6(b) shows a superimposed graph for all curves computed at different budgets.

Similarly, how well scenes can predict others was explored. Again a model is built, fitting a cubic polynomial as above, without information of one of the scenes and the resulting model was compared with that of the excluded scene and the absolute difference calculated. The average difference was found to be 0.030. Unsurprisingly, it is Rep which provides the greatest error. It deviated from the model predicted by the four other scenes with an error of 0.049. Figure 6(c) shows the predicted models for each group of scenes excluding the one named in the legend.

7. Discussion

Our hypothesis, h_1, that resource allocation varies with the amount of budget available clearly holds true for the three experiments tested. Significant differences in budget allocation were found for all three experiments. Figure 5 illustrates this complex relationship. This indicates the need for careful allocation of resources is required for interactive rendering. Harnessing this relationship, for example with the model proposed in Equation 4, produces a better use of resources, potentially delivering higher quality virtual experiences. The model itself predicts the frame rate/resolution parameter based on the amount of resources present, determined via the amount of Mpixels/s a resource can compute, although this could be changed to other metrics such as pure instruction counts if necessary. The model is fairly robust and can be used across scenes and budgets as witnessed, for the tested scenes and budgets, in the cross-validation results.

Results across experiments are reasonably consistent. Settings and budgets across experiments will be prefixed with the experiment number for clarity. Experiment 1 showed the preference for 60 fps over 30 fps even at resolutions reduced to R/4 compared to R. This result was supported by Experiment 2 whereby the results for $E1B_1$ and $E1B_2$ were replicated by $E2B_1$ and $E2B_2$ respectively; these were computed by completely different participants at different universities. Furthermore, the choices in $E2QP_1$ and $E2QP_2$ were larger with 5 and 4 potential choices respectively and the setting at 60 fps was still considered preferable. $E2B_3$ also shows a preference for 60 fps but in this case 60 fps also corresponds to the highest resolution for that budget. Experiment 3 echoes the results of Experiment 2. This can be seen in $E2B_1$ vs $E3B_4$ and $E2B_2$ vs $E3B_6$. Interestingly, $E3B_7$ rises

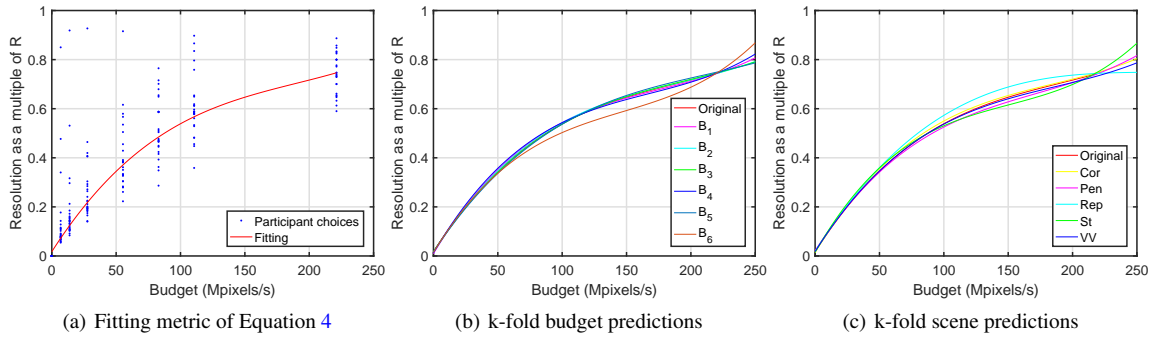(a) Fitting metric of Equation 4   (b) k-fold budget predictions   (c) k-fold scene predictions

Figure 6: *Fitting and k-fold cross validation for budget and scenes. For (b) and (c) legend denotes model built without the labeled budget/scene.*

higher than $E2B_3$. In Experiment 2 the results may be biased by the fixed choice of 60 fps while Experiment 3 may be providing a truer expected value for those budgets.

Overall results demonstrate a clear preference for frame rates in excess of 30 fps. Over all budgets, tested across all experiments, only $E3B_1$ and $E3B_2$ had lower averages than 30 fps and $E3B_2$ was at 29.5 fps, so relatively close. Participant behavior over 30 fps is interesting. Experiment 3's ability to allocate frame rates freely demonstrated that balancing of resources towards higher resolutions occurs around 40 fps. With traditional displays, due to the issues of either using v-sync, which would result in unsynchronised frames, or without v-sync resulting in screen tearing, 60 fps is recommended unless resolutions will suffer significantly. However, even at resolutions of a quarter, preference has been shown for 60 fps. Further work will be required to measure the impact of rendering at non-synchronised frame rates on traditional displays and the effects of screen tearing on perceived fidelity. The ability of the adaptive sync/G-Sync/FreeSync's capabilities to unlock frame rates may make them very suitable for more balanced frame rate and resolution settings as shown by the results in Experiment 3. Furthermore, they can make the best use of the proposed framework via the model in Equation 4 to provide a higher perceived quality of experience.

7.1. Limitations and Future Work

This work investigated how to achieve the best relationship between different quality settings focusing on frame rate and resolution. Clearly these are not the only parameters that define quality, yet, they are probably the primary two, and exploring their relationship requires controlling for the other parameters. Furthermore, they make the solving of Equation 1 possible without significantly increasing complexity. A linear binding between these two parameters has also been assumed, and while this is relatively common it is not the case for all algorithms. The proposed methodology would require

a re-work to generalise over uneven spatial to temporal costs, but we believe their relationship could be studied and modeled by extending the psychophysics experiments from Section 6.2 into two dimensions thus constructing a psychometric surface which is independent of computational budget; this can then be used to modify their relationship via scaling either setting when required. However, such a study would require a large number of participant trials to be conducted. As discussed earlier, quality is a function of many parameters, and future work will also look into expanding the number of parameters, to eventually provide a more general solution to Equation 1.

The ability to generalise over vastly diverging scenarios may be an issue when trying to provide models of prediction; some scenarios, for example highly detailed scenes, may show a stronger preference for higher resolution than reported here. Further, significant data capture would be required to provide models that also depend on, possibly, complex scenario characteristics. This work has demonstrated an amount of robustness with the model being able to predict the right settings for budgets and scenes that it was not developed for, however, we do not claim it works for all scenarios, and conclusions drawn from this work are dependent on the stimuli used. Nonetheless, the methodology presented can be adopted and applied to specific scenarios. For example, creators of a specific game level or virtual reality application can identify the best perceived parameters by running participants on that scenario. Further data about the scenarios can potentially be used for more complex multivariate regression models; although the amount of data capture required may be significant. Importantly, the difference in resource allocation across budgets observed in these five tests demonstrates that careful allocation of resources should be considered for any scenario.

The presented work was conducted with passive viewing only, as we wanted to control for this factor. Future work will look into the effect of interaction and task performance.

Based on previous work [CC07], we suspect an increased preference for higher frame rates would be observed.

Acknowledgements

References

[BS09] BRACKEN C. C., SKALSKI P.: Telepresence and video games: The impact of image quality. *PsychNology Journal 7*, 1 (2009), 101–112. 3

[CC07] CLAYPOOL K. T., CLAYPOOL M.: On frame rate and player performance in first person shooter games. *Multimedia systems 13*, 1 (2007), 3–17. 2, 3, 13

[CC09] CLAYPOOL M., CLAYPOOL K.: Perspectives, frame rates and resolutions: it's all in the game. In *Proceedings of the 4th International Conference on Foundations of Digital Games* (2009), ACM, pp. 42–49. 3, 5

[Coc50] COCHRAN W. G.: The comparison of percentages in matched samples. *Biometrika 37*, 3/4 (1950), 256–266. 6

[CPM06] CRANLEY N., PERRY P., MURPHY L.: User perception of adapting video quality. *International Journal of Human-Computer Studies 64*, 8 (2006), 637–647. 2

[CT07] CHEN J. Y., THROPP J. E.: Review of low frame rate effects on human performance. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 37*, 6 (2007), 1063–1076. 2

[DER*10] DIDYK P., EISEMANN E., RITSCHEL T., MYSZKOWSKI K., SEIDEL H.-P.: Perceptually-motivated real-time temporal upsampling of 3d content for high-refresh-rate displays. In *Computer Graphics Forum* (2010), vol. 29, Wiley Online Library, pp. 713–722. 3

[DPF03] DUMONT R., PELLACINI F., FERWERDA J. A.: Perceptually-driven decision theory for interactive realistic rendering. *ACM Transactions on Graphics (TOG) 22*, 2 (2003), 152–181. 3

[FS93] FUNKHOUSER T. A., SÉQUIN C. H.: Adaptive display algorithm for interactive frame rates during visualization of complex virtual environments. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques* (1993), ACM, pp. 247–254. 3

[FSME14] FREY S., SADLO F., MA K.-L., ERTL T.: Interactive progressive visualization with space-time error control. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 2397–2406. 3

[HWHK16] HÄNEL C., WEYERS B., HENTSCHEL B., KUHLEN T. W.: Visual quality adjustment for volume rendering in a head-tracked virtual environment. *IEEE Transactions on Visualization and Computer Graphics 22*, 4 (2016), 1472–1481. 3

[IL99] INSTALLATIONS T., LINE L.: Subjective video quality assessment methods for multimedia applications. *Networks 910* (1999), 37. 5, 6

[MSM04] MCCARTHY J. D., SASSE M. A., MIRAS D.: Sharp or smooth?: comparing the effects of quantization vs. frame rate for streamed video. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2004), ACM, pp. 535–542. 2

[RM00] READ P., MEYER M.-P.: *Restoration of motion picture film*. Butterworth-Heinemann, 2000. 2

[SSC10] SLATER M., SPANLANG B., COROMINAS D.: Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. *ACM Transactions on Graphics (TOG) 29*, 4 (2010), 92. 3

[WAH*15] WILCOX L. M., ALLISON R. S., HELLIKER J., DUNK B., ANTHONY R. C.: Evidence that viewers prefer higher frame-rate film. *ACM Transactions on Applied Perception (TAP) 12*, 4 (2015), 15. 2

[WRC88] WARD G. J., RUBINSTEIN F. M., CLEAR R. D.: A ray tracing solution for diffuse interreflection. *ACM SIGGRAPH Computer Graphics 22*, 4 (1988), 85–92. 3

[WW13] WONG G., WANG J.: *Real-time rendering: Computer graphics with control engineering*. CRC Press, 2013. 3

[WYM*16] WANG R., YU B., MARCO J., HU T., GUTIERREZ D., BAO H.: Real-time rendering on a power budget. *ACM Transactions on Graphics (TOG) 35*, 4 (2016), 111. 3

[ZRSK15] ZIELINSKI D. J., RAO H. M., SOMMER M. A., KOPPER R.: Exploring the effects of image persistence in low frame rate virtual environments. In *2015 IEEE Virtual Reality (VR)* (2015), IEEE, pp. 19–26. 2