

A Luminance-aware Model of Judder Perception

ALEXANDRE CHAPIRO, ROBIN ATKINS, and SCOTT DALY, Dolby Laboratories, USA

The perceived discrepancy between continuous motion as seen in nature and frame-by-frame exhibition on a display, sometimes termed judder, is an integral part of video presentation. Over time, content creators have developed a set of rules and guidelines for maintaining a desirable cinematic look under the restrictions placed by display technology without incurring prohibitive judder. With the advent of novel displays capable of high brightness, contrast, and frame rates, these guidelines are no longer sufficient to present audiences with a uniform viewing experience. In this work, we analyze the main factors for perceptual motion artifacts in digital presentation and gather psychophysical data to generate a model of judder perception. Our model enables applications like matching perceived motion artifacts to a traditionally desirable level and maintain a cinematic motion look.

CCS Concepts: • Human-centered computing → Displays and imagers; • Computing methodologies → Graphics systems and interfaces; Perception;

Additional Key Words and Phrases: Frame rate, display technologies, motion perception, judder, high dynamic range, contrast, brightness, speed perception

ACM Reference format:

Alexandre Chapiro, Robin Atkins, and Scott Daly. 2019. A Luminance-aware Model of Judder Perception. *ACM Trans. Graph.* 38, 5, Article 142 (July 2019), 10 pages.

<https://doi.org/10.1145/3338696>

1 INTRODUCTION

In this work, we pursue a practical examination of motion artifacts in the modern display landscape. In vision science such artifacts can be classified in various ways based on distinct features of the perceivable error (Section 2.2)—we will simplify this discussion and refer to these effects collectively as *judder*. When excessive, it ruins the quality of motion in a video and makes the content difficult or unpleasant to watch. Details of moving objects may be lost completely, and certain scenes can become disturbing or even entirely un-watchable. This is especially true for camera pans, which are the focus of this work. Pans are characterized by a swivel motion, generally horizontal, such that the majority of the filmed scene appears to move in the direction opposite to the camera rotation.

Judder is widespread in modern display systems but not when looking at real-life objects, chiefly due to static presentation: That

Authors' addresses: A. Chapiro, R. Atkins, and S. Daly, Dolby Laboratories, 432 Lakeside Dr., Sunnyvale, CA 94085, USA; emails: alex@chapiro.net, ratki@dolby.com, sdaly@dolby.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2019 Association for Computing Machinery.

0730-0301/2019/07-ART142 \$15.00

<https://doi.org/10.1145/3338696>

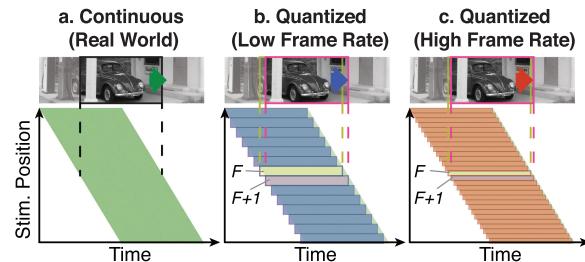


Fig. 1. This figure represents the horizontal movement of an image. Unlike the continuous case (left), the discrete static presentation (middle) of moving objects naturally leads to a mismatch within the human visual system. Instead of varying smoothly, object position jumps ahead or lags behind the continuous original, generating perceptual artifacts. Higher frame rates (right) reduce this artifact.

is, a display presents us with a quickly changing series of still frames. When images are updated at the traditionally used frequency (known as *frame rate*, historically set at 24 unique frames per second for cinema), there often remains a perceptible residual difference between a real-life object and its displayed counterpart (see Figure 1). When frame rates are increased, judder is reduced or even becomes imperceptible to viewers.

Although judder can be seen as undesirable if one aims to reproduce reality exactly, some audiences have become accustomed to a certain look when consuming cinematic content, including the 24 frame-per-second (FPS) presentation format. This appearance is strongly associated with artistic presentation, Hollywood or other higher-budget production, and professional quality. When the frame rate is increased, some viewers balk at the smooth motion, which is then perceived as an artifact itself (Giardina 2012). This is often referred to as the *soap opera effect*, associated with low-cost or home production. This reaction is generally not present when considering non-cinematic forms of entertainment (Wilcox et al. 2015). In sports, video games and low-budget productions, high frame rates have become standard and judder is expected to be minimized. Further, a dichotomy between users that prefer smooth motion and others who prefer a certain level of “cinematic” judder exists: When considering cinematic content, we see that judder is a unique artifact that we may want to preserve, within some threshold—not too high so as to not ruin the quality of the content but not too low or we risk making viewers unhappy with the content’s new look.

An additional degree of freedom in this problem is that motion artifacts are influenced by other factors associated with the displayed content and the screen being employed by the viewer. Several factors may affect judder (Daly et al. 2015; Larimer et al. 2001), such as content brightness, dynamic range, contrast, speed and direction of motion, shutter angle and shutter speed, eccentricity in the field of view of the observer, and, of course, frame rate. Interestingly, as display technologies evolve, mainly with the

introduction of high dynamic range televisions and projectors, the magnitudes of several of these factors (e.g., brightness, contrast, etc.) are changing, thus affecting their influence on judder. Feedback from artists and audiences shows that some movies that had desirable judder characteristics in a traditional cinema in the past may now appear to have too much judder when seen in high dynamic range (HDR) cinemas. This places new and restricting constraints on content creators in relation to what kind of scenes and motions can be present in their movies, effectively diminishing the HDR effect or enforcing only slow movement to be present. Our work aims to model judder perceptually, paving the way for tools that are able to quantify and control judder.

2 RELATED WORK

The differences between the natural continuous motion of an object and its discrete sampled counterpart result in perceived artifacts that we term judder. In this section, we point out that:

- Existing perceptual metrics are not sufficient to predict judder preference (Section 2.1, 2.2)
- Judder is likely to become a more prominent problem for novel display modes such as HDR (Section 2.3)
- Existing frame interpolation methods require additional information to produce perceptually pleasant results (Section 2.4)

2.1 Motion Perception

Motion perception has been widely studied (Nishida 2011), from the basics such as Fourier motion and the aperture problem to more advanced topics such as transparent motion (Mulligan 1992), motion aftereffects (Derrington and Badcock 1985), and biological motion (Johansson 1973). Other vision science work directed to distortions of practical interest include the motion sharpening effect, whereby blur levels of static frames are substantially less visible when the frames are displayed in motion (Bex et al. 1995; Takeuchi and De Valois 2005; Westerink and Teunissen 1995). The smoothness of sampled motion without specifically referring to the common term, judder, has been investigated including the effects of intermittent occlusion (Scherzer and Ekroll 2009), which can be applied to projector shutter designs.

Perceptual image difference metrics such as HDR-VDP (Mantiuk et al. 2011) compare two images by employing a pipeline modeling the human visual system. Video difference metrics are much less common. Work from Watson and Malo on the Standard Spatial Observer (2002) contains a video quality metric based on their existing model but does not provide implementation strategies. Aydin et al. (2010) extend their previous work on HDR image comparison to videos by employing a spatiotemporal contrast sensitivity function (STCSF) and a three-dimensional cortex transform. The *window of visibility* (Watson 2013; Watson et al. 1986) has been developed to model the interactions of the visual system with spatiotemporal aliasing distortions due to frame rate designs and more recently this work has been extended to a *pyramid of visibility* by showing that the STCSF is well approximated as a linearly separable function of its factors, including the log of adapting luminance (Watson and Ahumada 2016).

Most of these works have been geared toward threshold detection of temporal artifacts as modeled by the STCSF. By contrast, our work focuses on suprathreshold appearance and aims to provide a practical model that can be used to predict and compare distortions relevant to display applications.

2.2 Judder and Frame Rate

Subjective preference for higher frame rates was studied by Wilcox et al. (2015), who showed that in an experimental setting audiences preferred high frame rate clips to traditional 24Hz presentation. This work was later extended (Allison et al. 2016) to specifically target expert viewers with similar outcomes. A stronger preference for high frame rate presentation was found for action clips. Shutter angle was not found to play an effect on participants' choices. Although higher frame rates were preferred, the stimuli employed in the experiments were relatively simple short clips and may not fully evoke a cinematic sensation associated with high production value feature films, which may have reduced the propensity of the audience to be disturbed by the "soap opera effect."

Explorations specifically on the perception of judder include qualitative work on understanding what underlies the "film look" (Roberts 2002), including factors such as frame rate, object motion, and whole-frame motion (panning), as well as the causes within the eye as opposed to on-screen. The perception of double edges and their relation to source image sharpness is discussed.

Initial work on judder from a vision science perspective (Larimer et al. 2001) found that judder increases at both the high speeds (as expected), as well as the very low speeds. This result was confirmed in this work, as seen in Section 5 and 6. An initial investigation on fundamental judder perception (Daly et al. 2015) from a quantitative approach discussed four underlying components of judder: motion blur, non-smooth motion, multiple edges, and flickering. The relative weights of these distortions into an overall perception of judder are unknown, and individual variability is possible. The dominant effect was frame rate, followed by speed, but other factors had measurable effects. The study sampled few variations in a large set of perceptual factors to achieve a high level understanding of the space, and thus is not suitable as the basis of a visual model.

2.3 Display Technology

Although standard cinema specifications only require a luminance of 48cd/m^2 (DCI 2012), high dynamic range and ultra-bright displays are becoming ever more common and important as part of UHD television (Nilsson 2015). Major TV manufacturers including LG, Sony, Panasonic, Samsung, and TCL have released consumer-level HDR screens with peak luminance varying from 800 to 4000cd/m^2 , with yet brighter displays announced or demonstrated. With the introduction of OLED televisions into the market, ultra-high contrast display has also become ubiquitous.

Recent art has seen the effect of extended luminance and contrast explored for other perceptual factors such as color (Kim et al. 2009) and stereopsis (Didyk et al. 2012), and researchers have noted that judder is likely to become a more prominent problem for novel display modes such as HDR (McCarthy 2016; Noland 2016), confirmed by this work. Feedback from industry professionals shows

that a commonly used technique to combat judder introduced by increased brightness in novel displays consists of simply reducing content brightness during grading, which is undesirable.

In the context of limited transmission bandwidth, perceived quality tradeoffs between spatial and temporal resolutions have been shown to have a nonlinear relationship (Debattista et al. 2018).

2.4 Frame Interpolation

This work is done within the context of presentation on a modern display. Such displays are generally capable of showing images at a higher rate than the traditional 24Hz (e.g., 60, 120, or 144Hz). In this section, we discuss the problem of video frame rate re-sampling. These algorithms are occasionally concerned with the removal or preservation of judder, but lack the psychophysical data to make informed editing decisions.

There is a great variety of frame interpolation methods available. Traditional methods use motion flow vectors computed during video encoding (Zhai et al. 2005), derivatives of optical flow (Mahajan et al. 2009) or even phase-based motion estimation (Meyer et al. 2015). Because motion estimation often fails, especially when the movement is non-translational or when objects are not rigid, these methods may generate artifacts. Recent advances in machine learning have been applied to frame interpolation, both as a tool to compute better flow (Dosovitskiy et al. 2015), to generate entirely new frames at the cost of adding blur (Vondrick et al. 2016), or combine both techniques (Liu et al. 2017). While these methods can be used to obtain video at different frame rates, they do not solve the problem of finding a desirable judder pattern intrinsically.

Templin and colleagues developed a method to simulate arbitrary frame rates on displays with high refresh rates (2016). This method is meant for fine-tuned control of frame rate effects and was developed specifically to control judder but does not provide a definite answer for what the desirable characteristics of a video asset are in terms of judder or how to predict or match judder across different display technologies. This is precisely the goal of our work.

3 EXPERIMENTAL SETUP

To study judder through psychophysical trials, an experimental setup consisting of a display capable of reliably showing content at a large range of frame rates is necessary. To test the suitability of a display for our study, we measured its refresh rate and stability with a pair of Thorlabs PDA36A photo sensors that measure light in the visible wavelengths (350–780nm) at a frequency of up to 10MHz connected to an oscilloscope. A sample measurement can be seen in the supplementary material. To assess potential crosstalk between consecutive frames, the display was filmed with a high-speed camera while displaying temporal calibration targets.

After some experimentation, we found that our LCD displays were unable to show individual frames correctly at 120Hz due to slow rise times, resulting in incorrect frame brightness. Instead, our setup consists of an LG C7 2017 55" RGBW OLED TV. The display was calibrated to a standard CIE D65 ($x = 0.3127$, $y = 0.329$)

whitepoint using a Photo Research SpectraScan PR740 spectroradiometer perpendicular to the center point of the screen. The display operated at a resolution of 1920×1080.

We also found that typical video players are unable to reliably play at 120Hz without frame drops or repeats. To fix this we employed a custom OpenGL GPU implementation with multithreaded image loading from disk. To read uncompressed video frames, our application required a fast hard drive: For our resolution, 1920 * 1080 * 3 (RGB) * 1 byte (8-bit images) * 120Hz ≈ 750 MB/s, assuming a perfect system. Our server had a RAID-0 disk array providing >2GB/s read speeds and a Quadro M4000 GPU and was connected to the display via an HDMI 2.0 cable.

4 PSYCHOPHYSICAL EXPERIMENTS

Based on previous work discussed in Section 2.2 and feedback from industry content creators, we identified the main components of video that are believed to impact judder. Our goal is to examine these factors and generate a predictive model. We ran three psychophysical experiments measuring perceived judder.

4.1 Perceptual Factors Affecting Judder

The following are expected to be the strongest and most useful elements, present in all three studies:

Frame Rate. Expected to be the main factor affecting judder (Watson 2013), frame rate is in many ways the most malleable dimension for the manipulation of cinematic content as changes can be made without significantly altering the composition. Our experiments were performed on subsets of 24, 30, 60, and 120Hz. Frame rates under 24Hz are not widely employed in practice. In the practical cinematic scenarios targeted in this work (Fujine et al. 2007), rates over 120Hz are unlikely to exhibit judder, although motion artifacts could be visible at higher frame rates under different conditions, such as very high stimulus speed (Watson and Boff 1986).

Speed. As this work is mostly concerned with camera pans, we chose to parametrize this dimension based on panning speed, i.e., we assume the analyzed scene will have a unique dominant motion component. To select meaningful values, we turn to existing rules of thumb for pans in cinema (Burum 2007). In many cases, a 7s pan across the screen is considered the fast limit for a traditional shot to avoid motion artifacts. We chose to probe pans with speeds of 5, 7, 11, and 17s so we have information above, at, and below this limit. These speeds translate to approximately 2, 3, 4.7, and 6.6 deg/s based on the viewing distance of three picture heights, which corresponds to a horizontal 33° viewing angle for the whole screen. We expect higher speeds to translate into more perceived judder. For very slow speeds, the opposite may happen due to spatial position accuracy, especially at lower spatial resolutions (Larimer et al. 2001). These values are all within the region of smooth pursuit tracking (Li et al. 2010; Meyer et al. 1985). Smooth pursuit has a gain slip of about 0.9, and for wide field of view, catch-up saccades are used by the visual system to continue with smooth pursuit tracking. We expect that both smooth pursuit eye movements and occasional catch-up saccades may occur for some viewers, but that the rating assessment would be

accumulated during the smooth pursuit behavior during any given trial. In particular, any visibility during a saccade will be substantially suppressed.

Adapting Luminance. Known to influence spatiotemporal contrast sensitivity (de Lange Dzn 1958; Van Nes and Bouman 1967), adapting luminance has been modelled in image engineering (Barten 1989; Daly 1992). Following common practice in this field, we model adaptation luminance as the mean of the image's luminance in cd/m^2 . We expect sensitivity to rise as luminance increases, so it is reasonable to expect that judder may become more noticeable. We test mean luminance levels of 2.5, 10, and $40\text{cd}/\text{m}^2$. Note that image brightness varies based on scene content and the values presented here are taken as representative averages of mean luminance present in cinema and home theater for the first and second, while the final value is equally spaced but higher and represents a hypothetical high-luminance screen.

Additional factors are explored in some of our experiments:

Contrast. The STCSF operates on contrast, which is expected to be a factor. In this work, we used Michelson contrast, defined as $(I_{\max} - I_{\min})/(I_{\max} + I_{\min})$ to describe this factor, where I is the luminance image being used. Higher contrast is expected to lead to more perceived judder. We probed contrast values of 0.95, 0.75, and 0.5. Cinematic content traditionally strives to use as much of the available contrast as possible, and as such reducing contrast is usually not a feasible form of judder control.

Shutter Angle. More generally, motion blur reduces the retinal-image contrast of high spatial frequencies, potentially reducing judder. Shutter angles are traditionally described in degrees of arc with 360° representing a shutter that is open for the entire duration between frame captures. Smaller values represent fractional exposures: A 180° shutter means the sensor is exposed for half the available time. We explored shutter angle values of 360° , 180° , and 0° (the latter taken by convention when using an unaltered still image to simulate a camera pan). Blur is often used to control judder by content creators when no other options are available, but blurring strongly alters the scene leading to lost information.

Motion Direction. Previous research on judder (Daly et al. 2015) hypothesized that horizontal motions may produce more judder than vertical for some stimuli due to the better performance of horizontal smooth pursuit over vertical (Ke et al. 2013), but no data are provided to compare these results quantitatively.

To gauge whether directionality of motion is an integral part of judder perception, we ran a brief user study. The experimental procedure, stimulus, and analysis were identical to *Experiment 1* described in Section 4.2 and 5, but the only variables were *frame rate* at 24, 30, and 60 FPS and directionality, with the edge moving top-to-bottom (*vertical*) or left-to-right (*horizontal*). Thirteen users participated, with results shown in Figure 2. Frame rate was a significant factor ($F = 410.58, p \ll 0.01$). Direction was not found to be significant ($F = 2.24, p = 0.14$) with *horizontal* and *vertical* showing very similar behavior in terms of judder. In the rest of this work, we focus on horizontal movement, as this is the most common case in cinema and television.

4.2 Experimental Procedure

Participants were seated comfortably in a dark room. A chin rest was not employed as we felt that simulating natural viewing conditions would be more important than restricting head movements.

The experiment was performed with binocular viewing and natural pupil size. The display was placed at a standard three picture heights from the participants (Poynton 2012, p9).

As the goal of this work is to generate a transducer function from a high-dimensional space of parameters to judder, direct comparisons between conditions would have resulted in a prohibitively large number of trials. Instead, participants were instructed to rate stimuli in terms of judder on a scale of 0–9 (the former meaning no judder), with judder defined to participants as the absence of smooth motion. Stimuli were shown in three blocks of equal luminance (at mean luminance levels of 2.5, 10, or $40\text{cd}/\text{m}^2$) to allow participants to perform at a stable luminance adaptation state, with random block order and random presentation order of stimuli within each block for each participant. To allow the users to calibrate their responses to the dataset, a training session was conducted prior to each experiment that consisted of one of these three blocks chosen at random. Training results were excluded from further analysis. Results for each participant were normalized to use the entire 0–9 range to account for user variability.

Participants could observe each stimulus as many times as necessary to make their decision, after which responses were recorded using a standard keyboard. Each run consisted of 3 frame rates, 3 luminance levels, 4 speeds, and 1 additional parameter per experiment with 3 states, totaling 108 stimuli for the main experiment and an additional 36 for the training session. Participants took, on average, 30 minutes to complete the experiment, which was judged to be the longest period possible before fatigue.

4.3 Stimuli

As we expect higher spatio-temporal sensitivity to achromatic stimuli (Kelly 1983), images were presented in grayscale (see Figure 3).

Experiment 1. This experiment had the following variables:

- Frame rate = [30, 60, 120]Hz
- Mean luminance = [2.5, 10, 40]cd/m²
- Speed = [17, 11, 7, 5] second pan across the screen
- Michelson Contrast = [0.5, 0.75, 0.95]

We employed a traditional vision science stimulus termed “disembodied moving edge” (Savoy and Burns 1989) on a gray background. The edge moved across the screen left-to-right with

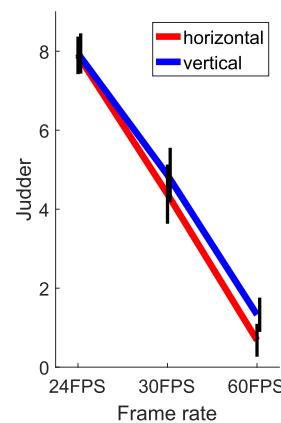


Fig. 2. Vertical bars depict 95% confidence intervals.

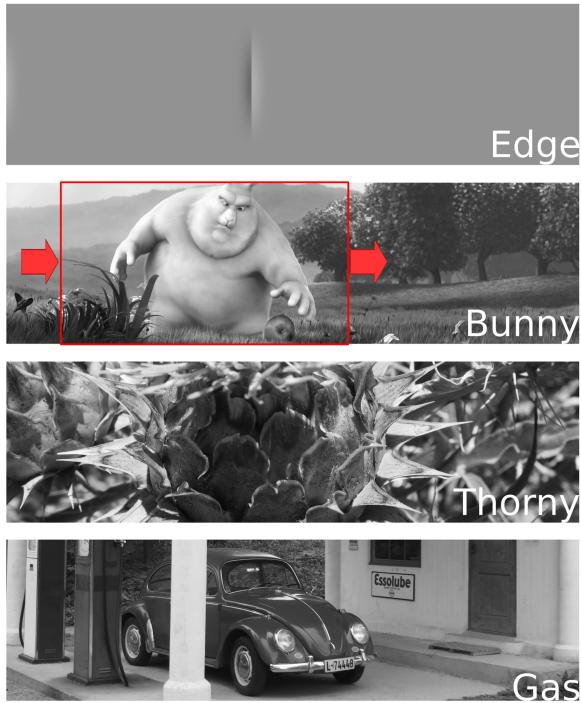


Fig. 3. Stimuli used in our experiments. In each trial, the displayed image represented by the red rectangle, pans across the scene horizontally at the desired speed.

the desired speed. The edge consisted of two vertical, adjacent columns of pixels with values M and m symmetric about the mean and calculated to match the desired contrast (see Appendix A), which then smoothly decayed toward the mean away from the edge.

Experiment 2. This experiment had the following variables:

- Frame rate = [30, 60, 120]Hz
- Mean luminance = [2.5, 10, 40]cd/m²
- Speed = [17, 11, 7, 5] second pan across the screen
- Images = [bunny, thorny, gas]

In this experiment, we chose to employ still image pans to provide participants with more realistic stimuli. A 1920×1080 window of a larger image is displayed and pans across the screen at the desired speed. Three images were selected, depicting a rendered scene from Big Buck Bunny¹ and two natural scenes showing a thorny cactus flower with strong diagonal features and a gas station with strong vertical lines.

Experiment 3. This experiment had the following variables:

- Frame rate = [24, 30, 60]Hz
- Mean luminance = [2.5, 10, 40]cd/m²
- Speed = [17, 11, 7, 5] second pan across the screen
- Shutter Angle = [0°, 180°, 360°]

¹Big Buck Bunny, copyright Blender Foundation <http://www.bigbuckbunny.org>.

In this experiment, we used only the *thorny* image from Experiment 2 but were interested in exploring the effect of shutter angle on judder. As the original content is a still image, we consider its shutter angle to be 0° by convention. Knowing the speed and frame duration for each stimulus, we computationally simulated shutter angles of 180° and 360° to see if photorealistic motion blur values (i.e., as would have been shot with a real-life camera) for image pans could be used as a viable strategy to mitigate judder. Although stronger blurring is possible and often employed as a tool in production, we chose not to include it in this test as it would significantly alter the scene. In addition, we explored lower frame rates introducing 24Hz and removing 120Hz from the set.

5 EXPERIMENT RESULTS

The results of experiments 1 to 3 were analyzed using linear mixed regression, treating participants as random effects to capture differences in their baseline rates of judder detection. The resulting p-values are reported below. Effect sizes are discussed in Section A.2.

Experiment 1. Twelve participants took part. As expected from Section 4.1, significant factors included *mean luminance* ($F = 53.68, p \ll 0.01$), *frame rate* ($F = 633.07, p \ll 0.01$) and *speed* ($F = 36.34, p \ll 0.01$). *Contrast* was not found to be a significant factor ($F = 0.54, p = 0.46$), possibly due to all stimuli being of suprathreshold contrast and contrast constancy effects (Georges and Sullivan 1975).

The results of this experiment can be seen in Figure 4; for simplicity, the plotted data have been averaged over the contrast dimension and participants. By comparing the three plots, we note that frame rate has a powerful effect on mitigating judder, with results at 120 and 60Hz showing little perceived judder, while 30Hz stimuli were all perceived with high levels of judder. A clear trend from the 30Hz plot is that, at this frame rate, judder increases uniformly with luminance. In addition, speed has a nearly linear effect on perceived judder.

Experiment 2. Eleven participants took part. Once again, *mean luminance* ($F = 52.09, p \ll 0.01$), *frame rate* ($F = 849.78, p \ll 0.01$), and *speed* ($F = 40.26, p \ll 0.01$) were found to be significant factors. Type of *image* shown ($F = 1.62, p = 0.20$) was not significant.

The results can be seen in Figure 5, averaged out over the three different image types and participants for simplicity. The product of this experiment is similar to *experiment 1*, which validates our expectation that judder perception for more realistic stimuli still follows the expected trends.

Experiment 3. Ten participants took part. Significant factors included *mean luminance* ($F = 249.77, p \ll 0.01$), *frame rate* ($F = 1017.50, p \ll 0.01$), and *speed* ($F = 154.92, p \ll 0.01$). *Shutter angle* was not found to be a significant factor ($F = 0.23, p = 0.63$).

The results can be seen in Figure 6, averaged out over *shutter angle* and users, for simplicity. In this experiment, we introduced a lower-frame-rate condition, 24Hz, and removed the highest frame rate from previous cases, as its results contained almost no information (120Hz). As this results in a different scaling of the responses, data were normalized such that the blocks overlapping with experiments 1 and 2 have matching means before generating the model described in Section 6. Note that, as previously, little judder is experienced at 60Hz, while results for 30Hz look similar to

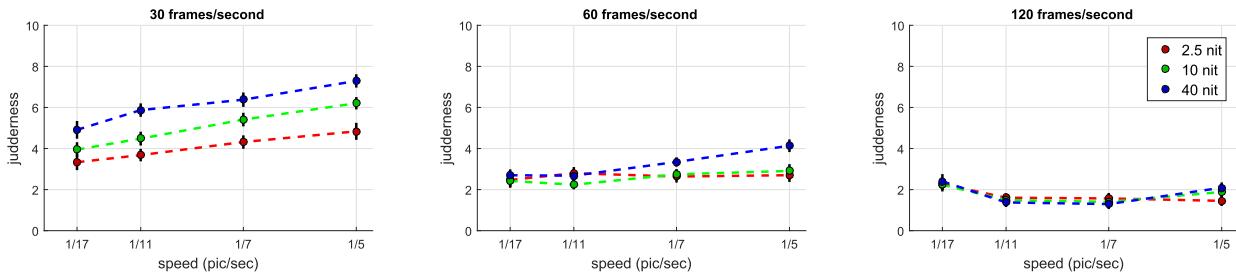


Fig. 4. Results for experiment 1 (moving edge), averaged over participants and contrasts. Vertical lines depict standard error over all samples. Results for 120 (right) and 60 FPS (mid) show little judder. Thirty FPS (left) appeared considerably distorted—judder increases almost linearly with speed, and there is a neat separation between luminance levels (plotted in red, green, and blue), with higher luminances considered to have more judder.

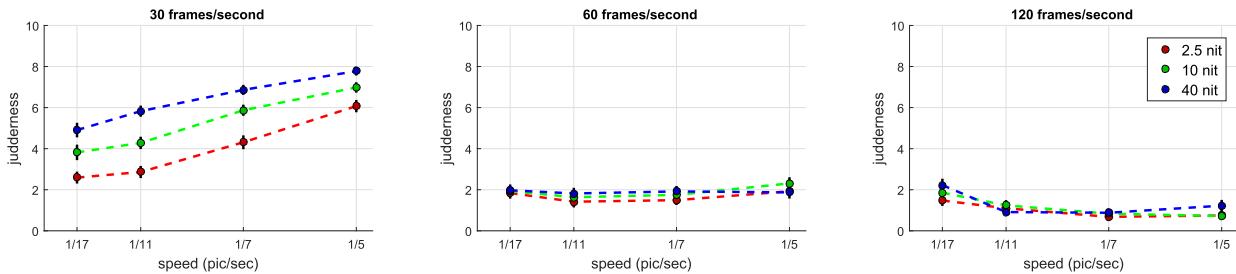


Fig. 5. Results for experiment 2 (panning complex images), averaged over participants and images. Vertical lines depict standard error over all samples. Results are similar to experiment 1, with 120 (right) and 60 FPS (mid) not showing much judder. Thirty FPS (left) continues to present a positive and clearly separable correlation of judder with speed and luminance.

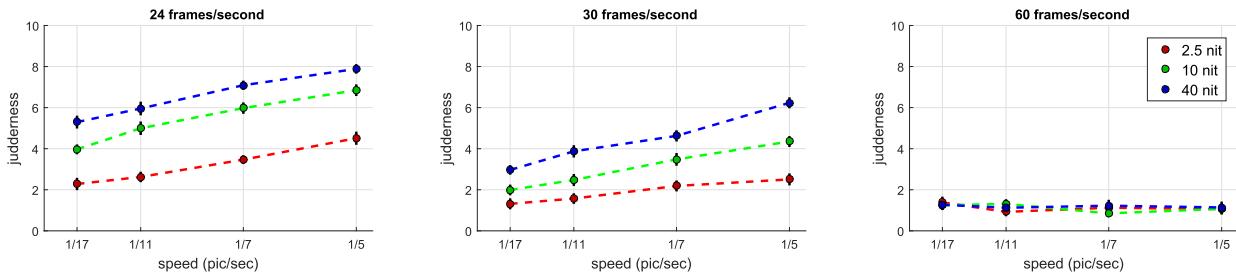


Fig. 6. Results for experiment 3 (panning thorny cactus image, with shutter angle variations), averaged over participants and shutter angles. Vertical lines depict standard error over all samples. Note that here we probe different frame rates. Sixty FPS (right) shows little judder, while 30 FPS (mid) exhibits similar behavior to experiments 1 and 2. Twenty-four FPS (left) shows even stronger judder responses, with a larger separation between the first luminance level of $2.5\text{cd}/\text{m}^2$ and the remaining levels, indicating that an increase in brightness has a strong impact on perceived judder.

earlier studies. At 24Hz, a stronger difference between the lowest luminance level (targeting standard dynamic range cinema) and the higher luminances is seen.

Variables that were not found to be significant were not included in our perceptual model of judder, although this result does not necessarily indicate that contrast, shutter angle, or type of image will never affect judder. It is possible, for example, that stronger distortions such as non-photorealistic blurring above 360° could significantly affect judder perception, but such values were not tested in this work, as they would strongly alter the look of the scene and thus do not lead to a satisfactory solution to mitigate judder.

6 JUDDER MODEL

To obtain an easily expressible model of judder J based on the most important factors, mean luminance L , frame rate F , and speed S , we combine the results of experiments 1–3 to design a single empirical model. We fit the perceptual data provided by user study participants into a second degree polynomial model:

$$J = P(\alpha(F), \beta(L), S), \quad (1)$$

where α and β are nonlinearities employed in perceptual modeling; specifically, for luminance we employ α the logarithm function and for frame rate β is the multiplicative inverse, i.e., we model

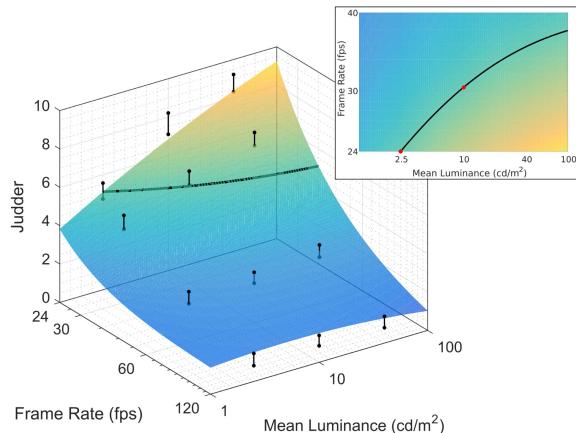


Fig. 7. A rendering of our judder model for a speed of 1/7 pictures per second. Vertical bars depict 95% confidence intervals. The horizontal line is an isoline with constant predicted judder, plotted separately in the inset. Note that a cinema reference of $2.5\text{cd}/\text{m}^2$ and 24 FPS, has an equivalent predicted judder for $10\text{cd}/\text{m}^2$ of roughly 31 FPS.

on frame duration. For details on the resulting function, please see Section A.3.

This is an excellent fit to the psychophysical data, with a mean absolute error of 0.24 (equivalent to 9.4%) between measured and predicted judder at the probed points. To present the reader with an error metric that relates to physical quantities, we also computed the mean error in the log-luminance domain (to avoid under representing errors in low-luminance conditions). Given N as the number of measured conditions, $O(i)$ being the observed means for each condition and $M(i)$ values predicted by our model, we calculate the error E as

$$E = \sum_{i=1}^N \frac{|\log(O(i)) - \log(M(i))|}{\log(O(i))} / N, \quad (2)$$

resulting in E at approximately 1.37%. A visualization of this model for a speed of 1/7 pictures per second can be seen in Figure 7 (left). A rendering of the overlaid surfaces for various speeds and a cross section of the resulting shapes for a given frame rate can be seen in Figure 8.

Finally, our model allows us to compute iso-lines over which perceived judder is expected to remain constant—this is a good opportunity to explore the hitherto unknown perceptual consequences of changes in judder due to different display technologies. In Figure 7 (right), we see that the predicted frame rate necessary to match the judder of a 24Hz, $2.5\text{cd}/\text{m}^2$ stimulus at 1/7 pic/s (a typical cinema frame rate and brightness at traditionally peak acceptable speed) displayed at $10\text{cd}/\text{m}^2$ is approximately 31Hz. This is significantly lower than the frame rates available in modern home theater displays. In addition, we calculate that if operating on speed, rather than frame rate, the judder of a traditionally fast 5s pan at $2.5\text{cd}/\text{m}^2$ is equivalent to that of an extremely slow 17s pan at a higher luminance of $15\text{cd}/\text{m}^2$. This further demonstrates that novel display technologies will require content creators to adapt their processes to deal with the newly introduced increase in perceptual

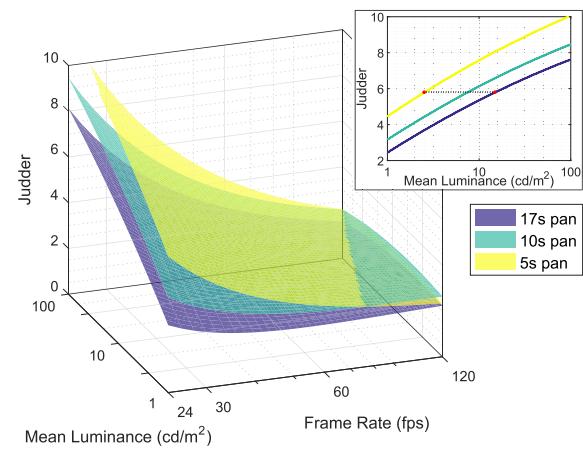


Fig. 8. On the left, we show our judder model plotted for a variety of speeds: 5, 10, and 17s pans. The plot is rotated for clarity. On the right, we see a cross section of various speeds projected on the luminance plane.

sensitivity, as the increase in brightness will incur prohibitive judder even for slow moving objects unless frame rates are increased.

7 VALIDATION

To test the validity of our model with more complex stimuli, we ran a Validation experiment. Using the same setup as the experiments described in Section 3, 15 participants were shown short reference videos at a lower frame rate and brightness. They were then presented with another version of the same video, but at a higher frame rate. The experimental task was to find the best match in terms of judder by controlling the mean luminance of the test stimulus through logmean offsets. Five stimuli were employed, two shots from *Nocturne*,² dubbed *Ballet* and *Kids*, and one shot from *Big Buck Bunny*. All stimuli contained camera pans and were likely to generate some perceivable judder and can be seen in the accompanying video. The details for each of the five tested cases can be seen in Table 1, with additional details presented in Appendix A.4.

We computed luminance predictions to match judder between test and reference using our metric as described in Equation (1). Although no other judder metric exists, we adapted the well-known Ferry-Porter law for flicker fusion thresholds (Tyler and Hamer 1990) to predict temporal artifact sensitivity. It is normally expressed as

$$CFF(L) = a * \log(L) + b, \quad (3)$$

where a and b are known constants and L is the mean luminance. If we introduce the simplifying assumption that the critical flicker fusion rate (CFF) is linearly correlated through a factor M with judder-sensitivity, then we can obtain a log-luminance equivalence like the one queried in this experiment. Denoting F_a and F_b as the two frame rates and L_a , L_b as the luminances:

$$F_a = M * CFF(L_a) = M(a * \log(L_a) + b), \quad (4)$$

²Nocturne, copyright Netflix, is a production shot entirely in high frame rate (<https://www.netflix.com/title/80233282>).

Table 1. Experimental Conditions for Validation Study

Validation Experiment	A	B	C	D	E
Stimulus	bunny	ballet	ballet	kids	kids
Ref. Lum. (cd/m^2)	2.5	2.5	2.5	2.5	2.5
Ref. frame rate	24	24	30	24	30
Test frame rate	30	30	40	30	40
Speed (second pan)	8s	4.4s	4.4s	6.4s	6.4s

$$F_b = M * \text{CFF}(L_b) = M(a * \log(L_b) + b). \quad (5)$$

Solving for L_b , we obtain the matching luminance prediction:

$$L_b = 10^{(aF_b \log(L_a) + b(F_b - F_a)) / (aF_a)}. \quad (6)$$

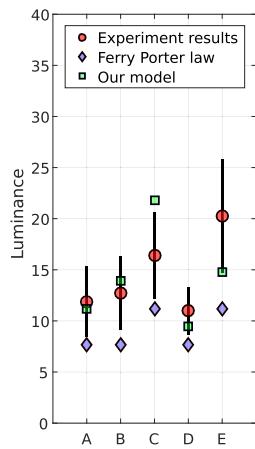


Fig. 9. The horizontal axis shows the experimental conditions as seen in Table 1. The vertical axis shows the luminance selected by users to match judder compared to model predictions. Vertical bars show 95% confidence intervals.

The results of this experiment can be seen in Figure 9, along with the predicted luminance match given by our model with the appropriate pan speed. Participants found the assignment to be challenging, as it included a cross-dimensional matching task across different image brightnesses. The prediction given by the static Ferry-Porter law only takes mean luminance into account and, even with the free parameter M , underestimates the luminance required to match our stimuli in every case: Using Equation (2) we obtain a mean luminance error of 17.1%. Our model achieves a mean error of 6.51%, improving the prediction in every case and falling within the 95% confidence interval in all but one case.

7.1 Applications to Content Creation

An application of our model is flagging content that is likely to produce high levels of judder in post production. Given the content and a target display, a shot boundary detector is applied to segment the video into shots. We recover the frame rate of the video and, for each shot, we calculate the mean luminance and translation speed. We proceed to apply our model as presented in Section 6. We demonstrate this technique by applying our model to the entire Big Buck Bunny short, which is composed of 128 shots. Shots that contain speed values less than the minimum explored in our user studies (i.e., a 17s pan or 2 deg/s average) were excluded from the computation, and judder prediction was set to zero. In this case, we simulate a cinema projector with a maximum and minimum brightness of 50 and $0.05cd/m^2$, respectively, and a gamma of 2.4 with a user sitting at three picture heights from the screen. Figure 10 shows the resulting judder prediction.

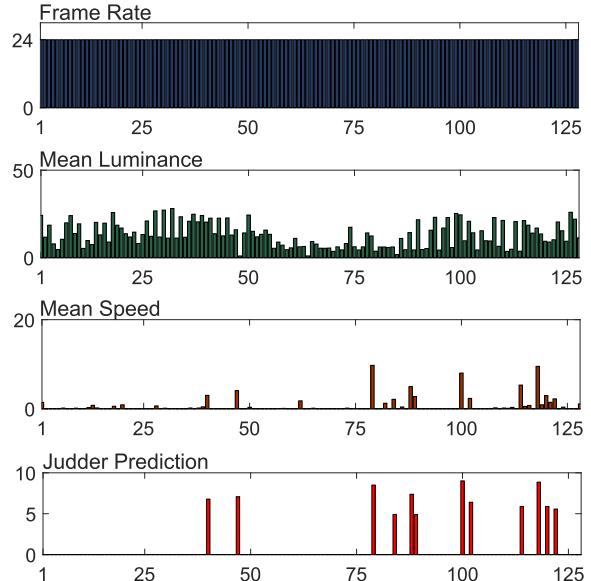


Fig. 10. We estimated perceived judder for each shot in Big Buck Bunny. From top to bottom, we show the parameters used in our model—frame rate, mean luminance, and mean speed, and, finally, our model’s prediction. This system can help flag problematic shots within a feature film at an early stage during post processing.

Similarly to the task of users in the validation experiment described in Section 7, content flagged as having excessive judder can be tone-mapped to a lower mean luminance range automatically. As an example, we selected shot 47 of Big Buck Bunny that showed a judder value of 7.1 in Figure 10. We can adapt the 7s rule as referenced in Section 4.1 to produce an estimate of a high boundary for acceptable judder: We set speed to the eponymous 7s pan, frame rate to a standard 24Hz, and mean luminance to $2.5cd/m^2$. The resulting value is 5.1 on our judder scale. We proceed to fix the speed and frame rate of the selected shot, and model judder for various mean luminance levels. We find that in order for this shot to have a predicted judder of 5.1, its mean luminance must be reduced from 16 to $3.1cd/m^2$ or approximately 0.72 log units. The judder scale for this shot, as well as a sample image tone mapped using a simple log-luminance offset, are shown in Figure 11. A comprehensive review of modern tone mapping methods was presented by Eilertsen et al. (2017).

8 CONCLUSIONS, LIMITATIONS, AND FUTURE WORK

We proposed the first model that predicts the magnitude of perceived judder for a given video. We gathered psychophysical data on a number of relevant factors using simple stimuli and later demonstrated that these measurements are also valid for complex scenes. Our experiments generated valuable data that can provide guidance to content creators considering emerging display technologies, such as necessary updates to existing rules of thumb in cinematography. This information can also be useful to display manufacturers and content providers when considering technical requirements for consumer technologies. Finally,



Fig. 11. On the left, judder predictions according to our model for this shot for different mean luminances. On the right, the input content with a mean luminance of 7.1 is tone mapped to $5.1\text{cd}/\text{m}^2$, reducing predicted judder.

perceptual judder modeling is crucial for frame rate re-sampling applications to avoid judder or other undesirable artifacts such as the “soap opera effect.”

Our work has certain limitations: The first is that our model is only valid for clips that contain relatively stable motion with a single dominant component, such as a camera pan. The “ballet” clip used in Section 7, for example, contains a speed ramp-up that may have caused our results to deviate from the data gathered from users as they may have graded on the slower or faster portion of the clip. Clips where little to no motion is detected are not expected to contain visible judder and are not supported by our model.

Extending Fourier-based spatio-temporal video difference prediction approaches (Section 2.1) to supra-threshold differences for judder would be an interesting task. Such algorithms require lengthy computation times per video, however, making them impractical for common real-time applications.

Although this work focused chiefly on judder arising from camera pans, as this is the most common scenario, an interesting avenue for future work would be to generalize the model to separate objects on the scene. User attention can be predicted using a visual saliency model and object speed could be estimated through optical flow techniques.

Our studies explored an overall sense of judder, but there are certainly different perceptual components to this artifact that contribute to the whole as subcomponents. At this stage of inquiry, we felt the strongest need would be to address the overall perception of judder. A logical next step would be to study the perceptual components of judder as their understanding may be useful to algorithm design. We hope our work encourages the community to further explore spatio-temporal artifacts and make all our experimental data fully available for analysis.

A APPENDIX

A.1 Disembodied Edge Calculation

Given a desired Michelson contrast c and mean L_a , we calculate the maximum and minimum values M and m as $M = L_a(1 + c)$, $m = L_a(1 - c)$. As a consequence of this, the resulting Michelson contrast is exactly $(L_a + L_a * c - L_a + L_a * c)/(L_a + L_a * c + L_a - L_a * c) = 2L_a * c/(2L_a) = c$. If one of these values does not fit within the dynamic range of the display being used, then both values can be shifted using a multiplier k so that $M = M * k$, $m = m * k$ to accommodate these practical constraints.

The maximum and minimum values computed in the previous step are set to fall off to the background level L_a in a smooth fashion. In our application, we employed a screen with horizontal resolution of 1,920 pixels and found a visibly acceptable smoothness to be achieved using a Gaussian falloff with standard deviation of 45 pixels.

A.2 Effect Size

Effect size can be calculated from an arbitrary scale such as the judder scale introduced by our article. We include the calculation of a Pearson’s correlation coefficient in the supplementary material containing the raw data for experiments 1–3. Following the guidelines set by “A Power Primer” (1992), we find that in all three experiments the independent variables *speed* and *luminance* had a small to medium strength effect on judder, *frame rate* had a strong effect, and the fourth variable (contrast, image type, and shutter angle, respectively) had an effect less than small.

A.3 Judder Model

Below, on the left, are the coefficients for our judder model as described in Section 6. On the right, the power of the appropriate term. This information is also available in the supplementary material. Note that our model takes as input speed in pixels per second on a screen with a horizontal resolution of 1,920×1,080 at three picture heights’ distance, subtending a 33° field of view. If a speed S is given in degrees per second, then a conversion function for speed should be used as follows: $\gamma(S) = \frac{S * 1920}{33}$. In addition:

$$\alpha(F) = 1/F, \quad (7)$$

$$\beta(L) = \log_{10}(L), \quad (8)$$

Coef.	$\alpha(F)$	$\beta(L)$	S
1620.37	2	0	0
86.81	1	1	0
0.32	1	0	1
-95.55	1	0	0
0.32	0	2	0
0.00	0	1	1
0.48	0	1	0
1.01	0	0	2
0.00	0	0	1
2.35	0	0	0

A.4 Validation Data

To compute a judder prediction using our model for an arbitrary scene, a speed measurement is necessary. As we are mostly concerned with panning scenes we expect the clip to have a strong main motion component due to the camera motion. We computed a frame-to-frame best-fit translation, which is pooled over all frames, filtered (speeds below 1 deg/s fall short of the smooth pursuit range (Meyer et al. 1985) and are filtered out), and averaged to compute the speed. Note that this procedure is only valid for clips that contain relatively stable motion. We obtained a mean panning speed of approximately 8s for *Bunny*, a fast 4.4s for *Ballet*, and 6.4s for *Kids*.

ACKNOWLEDGMENTS

The authors thank Shane Ruggieri for helping prepare the video for the article and Sema Berkten for final edits. We thank Shane, Thad Beier, and Jim Crenshaw for the insightful discussions on industry practices relating to judder management and motion in general. We thank Thomas Wan and the immersive experience team for help with the experimental setup and hardware. We thank the user study participants for their efforts. We thank Timo Kunkel and Sema for help with style and Seth Frey for help with the statistical analysis.

REFERENCES

- Robert S. Allison, Laurie M. Wilcox, Roy C. Anthony, John Helliker, and Bert Dunk. 2016. Expert viewers' preferences for higher frame rate 3D film. *J. Imag. Sci. Technol.* 60, 6 (2016), 60402–1.
- Tunc Ozan Aydin, Martin Čadík, Karol Myszkowski, and Hans-Peter Seidel. 2010. Video quality assessment for computer graphics applications. *ACM Trans. Graph.* 29, 6 (2010), 161.
- Peter G. J. Bartlen. 1989. The square root integral (SQR): A new metric to describe the effect of various display parameters on perceived image quality. In *Human Vision, Visual Processing, and Digital Display*, Vol. 1077. International Society for Optics and Photonics, 73–83.
- P. J. Bex, G. K. Edgar, and A. T. Smith. 1995. Sharpening of drifting, blurred images. *Vis. Res.* 35, 18 (1995), 2539–2546.
- Stephen H. Burum. 2007. *American Cinematographer Manual*. Vol. 1. American Cinematographer.
- Jacob Cohen. 1992. A power primer. *Psychol. Bull.* 112, 1 (1992), 155.
- Scott Daly, Ning Xu, James Crenshaw, and Vikrant J Zunjarrao. 2015. A psychophysical study exploring judder using fundamental signals and complex imagery. *SMPTE Motion Imag.* 124, 7 (2015), 62–70.
- Scott J. Daly. 1992. Visible differences predictor: An algorithm for the assessment of image fidelity. In *Human Vision, Visual Processing, and Digital Display III*, Vol. 1666. International Society for Optics and Photonics, 2–16.
- DCI. 2012. Digital Cinema Initiatives, LLC. High Frame Rates Digital Cinema Recommended Practice. Retrieved from http://www.dcmovies.com/Recommended_Practice/.
- H. de Lange Dzn. 1958. Research into the dynamic nature of the human Fovea-cortex systems with intermittent and modulated light. II. Phase shift in brightness and delay in color perception. *J. Opt. Soc. Am.* 48, 11 (1958), 784–789.
- Kurt Debattista, Keith Bugeja, Sandro Spina, Thomas Bashford-Rogers, and Vedad Hulusic. 2018. Frame rate vs resolution: A subjective evaluation of spatiotemporal perceived quality under varying computational budgets. In *Computer Graphics Forum*, Vol. 37. Wiley Online Library, 363–374.
- Andrew M. Derrington and David R. Badcock. 1985. Separate detectors for simple and complex grating patterns? *Vis. Res.* 25, 12 (1985), 1869–1878.
- Piotr Didyk, Tobias Ritschel, Elmar Eisemann, Karol Myszkowski, Hans-Peter Seidel, and Wojciech Matusik. 2012. A luminance-contrast-aware disparity model and applications. *ACM Trans. Graph.* 31, 6 (2012).
- Alexey Dosovitskiy, Philipp Fischer, Eddy Ilg, Philip Hausser, Caner Hazirbas, Vladimir Golkov, Patrick van der Smagt, Daniel Cremers, and Thomas Brox. 2015. Flownet: Learning optical flow with convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision*, 2758–2766.
- Gabriel Eilertsen, Rafal K. Mantiuk, and Jonas Unger. 2017. A comparative review of tone-mapping algorithms for high dynamic range video. *Computer Graphics Forum* (2017). DOI: <https://doi.org/10.1111/cgf.13148>
- Toshiyuki Fujine, Yuhji Kikuchi, Michiyuki Sugino, and Yasuhiro Yoshida. 2007. Real-life in-home viewing conditions for flat panel displays and statistical characteristics of broadcast video signal. *Jpn. J. Appl. Phys.* 46, 3S (2007), 1358.
- M. A. Georgeson and G. D. Sullivan. 1975. Contrast constancy: Deblurring in human vision by spatial frequency channels. *J. Physiol.* 252, 3 (1975), 627–656.
- Carolyn Giardina. 2012. Peter Jackson responds to "Hobbit" footage critics, explains 48-frames strategy. *The Hollywood Reporter* (2012).
- Gunnar Johansson. 1973. Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* 14, 2 (1973), 201–211.
- Sally R. Ke, Jessica Lam, Dinesh K. Pai, and Miriam Sperling. 2013. Directional asymmetries in human smooth pursuit eye movements. *Invest. Ophthalmol. Vis. Sci.* 54, 6 (2013), 4409–4421.
- D. H. Kelly. 1983. Spatiotemporal variation of chromatic and achromatic contrast thresholds. *J. Opt. Soc. Am.* 73, 6 (1983), 742–750.
- Min H. Kim, Tim Weyrich, and Jan Kautz. 2009. Modeling human color perception under extended luminance levels. In *ACM Transactions on Graphics (TOG)*, Vol. 28. ACM, 27.
- James Larimer, Jennifer Gille, and James Wong. 2001. 41.2: Judder-induced edge flicker in moving objects. In *SID Symposium Digest of Technical Papers*, Vol. 32. Wiley Online Library, 1094–1097.
- Feng Li, Jeff B. Pelz, and Scott J. Daly. 2010. Effects of stimulus size and velocity on steady-state smooth pursuit induced by realistic images. In *Human Vision and Electronic Imaging XV*, Vol. 7527. International Society for Optics and Photonics, 752717.
- Ziwei Liu, Raymond Yeh, Xiaou Tang, Yiming Liu, and Aseem Agarwala. 2017. Video frame synthesis using deep voxel flow. In *Proceedings of the International Conference on Computer Vision (ICCV'17)*, Vol. 2.
- Dhruv Mahajan, Fu-Chung Huang, Wojciech Matusik, Ravi Ramamoorthi, and Peter Belhumeur. 2009. Moving gradients: A path-based method for plausible image interpolation. In *ACM Transactions on Graphics (TOG)*, Vol. 28. ACM, 42.
- Rafal Mantiuk, Kil Joong Kim, Allan G. Rempel, and Wolfgang Heidrich. 2011. HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. In *ACM Transactions on Graphics (TOG)*, Vol. 30. ACM, 40.
- Sean T. McCarthy. 2016. How independent are HDR, WCG, and HFR in human visual perception and the creative process. *SMPTE Motion Imag.* 125, 4 (2016), 24–33.
- Craig H. Meyer, Adrian G. Lasker, and David A. Robinson. 1985. The upper limit of human smooth pursuit velocity. *Vis. Res.* 25, 4 (1985), 561–563.
- Simone Meyer, Oliver Wang, Henning Zimmer, Max Grosse, and Alexander Sorkine-Hornung. 2015. Phase-based frame interpolation for video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'15)*. IEEE, 1410–1418.
- Jeffrey B. Mulligan. 1992. Motion transparency is restricted to two planes. *Invest. Ophthalmol. Vis. Sci.* 33, 4 (1992), 1049.
- Mike Nilsson. 2015. Ultra high definition video formats and standardisation. *BT Media and Broadcast Research Paper* (2015). Technical Report.
- Shin'ya Nishida. 2011. Advancement of motion psychophysics: Review 2001–2010. *J. Vis.* 11, 5 (2011), 11–11.
- K. C. Noland. 2016. High frame rate television: Sampling theory, the human visual system, and why the Nyquist-Shannon theorem does not apply. *SMPTE Motion Imag.* 125, 3 (Apr. 2016), 46–52. DOI: <https://doi.org/10.5594/JMI.2016.2531998>
- Charles Poynton. 2012. *Digital Video and HD: Algorithms and Interfaces*. Elsevier.
- Alan Roberts. 2002. The film look: It's not just jerky motion. *BBC R&D White Paper WHP 53* (2002).
- R. Savoy and M. Burns. 1989. Isolated cone classes and the disembodied edge: New stimuli for psychophysics and neurophysiology. *Invest. Ophthalmol. Vis. Sci.* 30 (1989), 220.
- Tom R. Scherzer and Vebjørn Ekroll. 2009. Intermittent occlusion enhances the smoothness of sampled motion. *J. Vis.* 9, 10 (2009), 16–16.
- Tatsuto Takeuchi and Karen K. De Valois. 2005. Sharpening image motion based on the spatio-temporal characteristics of human vision. In *Human Vision and Electronic Imaging X*, Vol. 5666. International Society for Optics and Photonics, 83–95.
- Krzysztof Templin, Piotr Didyk, Karol Myszkowski, and Hans-Peter Seidel. 2016. Emulating displays with continuously varying frame rates. *ACM Trans. Graph.* 35, 4 (2016), 67.
- Christopher W. Tyler and Russell D. Hamer. 1990. Analysis of visual modulation sensitivity. IV. Validity of the Ferry–Porter law. *J. Opt. Soc. Am. A* 7, 4 (1990), 743–758.
- Floris L. Van Nes and Maarten A. Bouman. 1967. Spatial modulation transfer in the human eye. *J. Opt. Soc. Am.* 57, 3 (1967), 401–406.
- Carl Vondrick, Hamed Pirsiavash, and Antonio Torralba. 2016. Generating videos with scene dynamics. In *Advances in Neural Information Processing Systems*, 613–621.
- Andrew Watson and Albert Ahumada. 2016. The pyramid of visibility. *J. Vis.* 16, 12 (2016), 567–567.
- Andrew B. Watson. 2013. High frame rates and human vision: A view through the window of visibility. *SMPTE Motion Imag.* 122, 2 (2013), 18–32.
- Andrew B. Watson, Albert J. Ahumada, and Joyce E. Farrell. 1986. Window of visibility: A psychophysical theory of fidelity in time-sampled visual motion displays. *J. Opt. Soc. Am. A* 3, 3 (1986), 300–307.
- Andrew B. Watson and K. Boff. 1986. Temporal sensitivity. *Handbook of Perception and Human Performance* 1 (1986), 6–1.
- Andrew B. Watson and Jesus Malo. 2002. Video quality measures based on the standard spatial observer. In *Proceedings of the 2002 International Conference on Image Processing*, Vol. 3. IEEE, III–III.
- Joyce H. D. M. Westerink and Kees Teunissen. 1995. Perceived sharpness in complex moving images. *Displays* 16, 2 (1995), 89–97.
- Laurie M. Wilcox, Robert S. Allison, John Helliker, Bert Dunk, and Roy C. Anthony. 2015. Evidence that viewers prefer higher frame-rate film. *ACM Trans. Appl. Percept.* 12, 4 (2015), 15.
- Jiefu Zhai, Keman Yu, Jiang Li, and Shipeng Li. 2005. A low complexity motion compensated frame interpolation method. In *Proceedings of the IEEE International Symposium on Circuits and Systems 2005 (ISCAS'05)*. IEEE, 4927–4930.

Received September 2018; revised May 2019; accepted May 2019