

The Effect of Synthetic Shutter on Judder Perception—An HFR and HDR Data Set and a User Study

By Ianik Beitzel, Aaron Kuder, and Jan Fröhlich

Abstract

While high dynamic range (HDR) enhances the viewing experience by rendering a more life-like image, it amplifies motion artifacts due to the higher temporal contrast sensitivity of the human visual system at higher luminance levels and higher contrast ratios typical for HDR images. These motion artifacts are known as strobing or judder. They can be overcome by increasing temporal resolution, also known as higher frame rate (HFR). This article introduces an HFR and HDR data set for research on judder perception and presents a user study on motion portrayal. The data set is designed to visualize motion in HDR and includes versions of different frame rates. The user study compares the motion portrayal of four different synthetic shutters and two frame rates.

Keywords

High dynamic range (HDR), higher frame rate (HFR), motion judder

Motivation

Higher frame rate (HFR), along with high dynamic range (HDR), is intended to enhance the cinematic experience.¹ For example, Dolby Laboratories introduced Dolby Cinema, combining Dolby Atmos for sound and Dolby Vision for image. Dolby Vision raises the peak luminance in cinemas from 48 to 108 cd/m² and increases the dynamic range from 2000:1 to 1,000,000:1.² As of 2019, Dolby has opened more than 200 Dolby Cinemas worldwide.³ Samsung and Sony have introduced direct view LED screens that can present movies up to 500 cd/m² and also feature a dynamic range of 1,000,000:1.^{4,5} For television,

more than 50% of all ultrahigh-definition (UHD)-TVs sold in 2018 were HDR capable.⁶ While HDR enhances the overall viewing experience, greater luminance and contrast also amplify motion artifacts such as judder. These artifacts are most often perceived in shots where the camera, the scene content, or both are moving.

Motion artifacts like judder were already present in the earliest motion pictures. They actually contribute to the “cinematic look,” which is an integral part of the movie-going experience. However, with HDR, judder artifacts are often perceived as disturbing, as they are stressful to the human eye. This calls for methods to reduce said artifacts.

One approach is to increase the frame rate for the whole imaging pipeline from acquisition to presentation. In cinema, everything faster than the standard 24 frames per second (fps) is considered HFR. Peter Jackson’s *The Hobbit—An Unexpected Journey* was one of the first blockbusters acquired and presented in 48 fps. Shooting the film in HFR made the action and landscape shots look smoother, but also took away the “cinematic feel” for some viewers. The higher temporal resolution revealed the props as artificial objects. The attached rubber noses of the actors, for example, were more easily identified as rubber rather

than real human noses. This gave the film a videoesque touch instead of a “cinematic look.”⁷ Hence, for many viewers, doubling the frame rate in cinema to reduce motion artifacts was not ideal.

In response, this work evaluates whether motion artifacts, specifically judder, can be reduced without increasing the frame rate in the presentation. First, an HDR and HFR data set is introduced and the methods of the film acquisition and post-production of this

Motion artifacts like judder were present in the earliest motion pictures. They actually contribute to the “cinematic look,” which is an integral part of the movie-going experience. However, with HDR, judder artifacts are often perceived as disturbing, as they are stressful to the human eye. This calls for methods to reduce said artifacts.

data set are presented. Then, this data set is used to perform a study of whether judder motion artifacts can be reduced by longer synthetic shutters.

Methods

Acquisition

The test footage was recorded at Stuttgart Media University (HdM) in 2017 on an ARRI AMIRA camera using ARRI Ultra Prime lenses. The shots are designed to show off the differences between standard frame rate (SFR) and HFR at several luminance levels while taking into account the established picture language of film.

Different kinds of motion such as pans at various speeds as well as steady shots are captured. Also, the type of motion varies within the frame. Fast movements such as shaking a cocktail and the sparks of a campfire, and also slower movements like walking, are present. In some shots, multiple layers of action are captured. In addition to the protagonists, moving objects in the foreground and background are a challenge for temporal film presentation, especially sharp vertical elements with high contrast.

The focus of the test is on two scenes, a high key lit bar scene and a fireplace using available light. The bar is fully staged with actors and camera movement. The fireplace scene is designed to study motion portrayal of a small fast-moving object without a storyline.

The SFR footage is shot at 24 fps and a 180° shutter angle. The HFR footage is captured at 192 fps, and the 356° maximum shutter angle of the ARRI ARMIRA camera is used to capture as much light and motion as possible. The HFR shots build the foundation for the frame blending/synthetic shutter in post-production.

Both scenes are lit for HFR settings. The light setup is not changed when switching the camera to SFR, and neither the ISO nor the f-stop/aperture of the camera and lens is changed. Therefore, the shots in 24 fps are overexposed by two stops. Due to the dynamic range of the ARRI AMIRA, no important details are lost. All shots are recorded as QuickTime files in 2048 × 1152 RGB-pixels and coded in 12-bit LogC 3 ARRI Wide Gamut 3 using ProRes444 intraframe compression.

In the following section, the visual content and the major technical specifications of all scenes are summarized. More information about technical details such as T-stops and lenses can be found in the corresponding Bachelor Thesis⁸ and on the project website.⁹

Scene 1: Bar

The bar scene is set in the afternoon. The barkeeper walks through the entire room and takes drink orders from two female guests. Next, he prepares two cocktails and hands them over to the ladies. Then, a third guest approaches the bar and orders a drink.

As shown in **Fig. 1**, the scene shows a high contrast mixed light of daylight and tungsten. It starts off with the protagonist in front of a very dark background. As he walks through the room, the background changes

to a very bright and high key lit glass frontage. Such an extreme light setup was deliberately chosen to visualize the differences between SFR and HFR. For example, the judder artifacts increase as the mastering luminance rises.

Scene 2: Fireplace

The visual content of this scene is a fireplace burning at night. A piece of cardboard is used to fan the fire, and a piece of wood is thrown into the fire to create fast-moving sparks as shown in **Fig. 2**.

In contrast to the bar scene, only available light was used, which results in a very low key lit scene.

Content Post-production

Editing and Grading

The editing and grading of the two scenes are done for both SFR (24 fps) and HFR (192 fps). As the SFR footage is created by blending the HFR footage, the timing of the action and framing is identical in both frame rates. The edits happen at the exact same moment if the footage is played back side by side at the respective frame rates. The color grading is performed by a professional colorist on a Sony X300 monitor for the standard dynamic range (SDR) and HDR display scenarios listed in **Table 1**. All versions are adjusted to best convey the artistic intent but without taking the judder effect into account. If the 24 fps version of these scenes was graded for a commercial presentation, the contrasts would have been reduced in order to reduce the judder effect, compromising on the artistic intent with respect to the color grading. The original footage and the graded sequences are provided on the project website⁹ for further research.

Creation of Different Frame Rate Versions

In order to modify the motion portrayal, the 192 fps source frames are blended together to 24 or 48 fps by means of a virtual shutter function, as proposed by Tony Davis.¹⁰ We introduce a new frame blending function to control the smoothness of the synthetic shutter by means of three parameters. This shutter function is visualized in **Fig. 3**.

Three parameters control the shape of the shutter function:

- a = Attack
- r = Release
- e = Elongation

Attack describes the steepness of the shutter function's inbound slope. The release parameter describes the steepness of the shutter function's outbound slope. This can be compared to the attack and release phases in audio signal

TABLE 1. The different grading versions.

Title	Dynamic range	Peak luminance	Black level
SDR	2000:1	100 cd/m ²	0.05 cd/m ²
HDR	1.000.000:1	1000 cd/m ²	0.001 cd/m ²

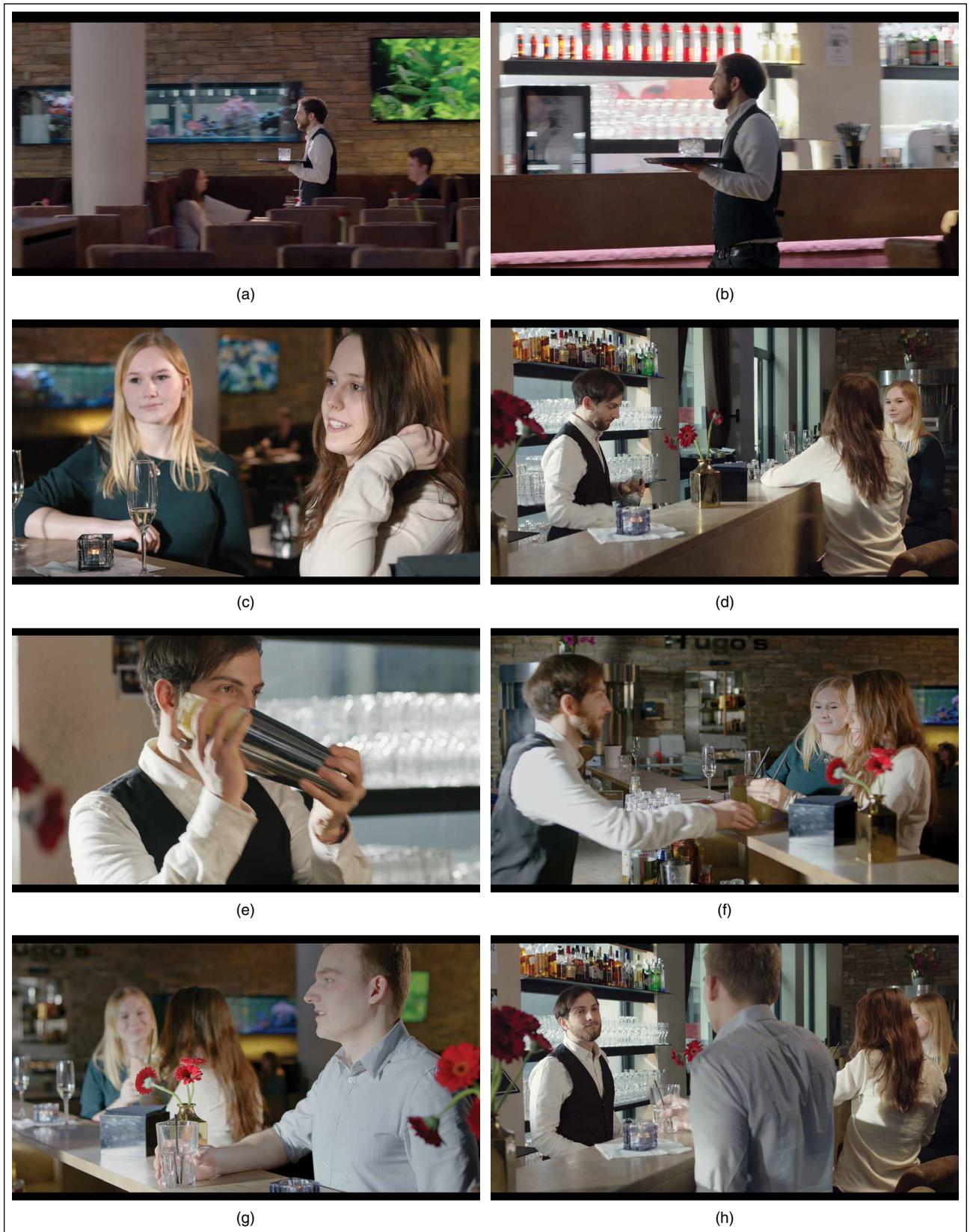


FIGURE 1. Stills of the bar scenes. (a) Bar 1. (b) Bar 2. (c) Bar 3. (d) Bar 4. (e) Bar 5. (f) Bar 6. (g) Bar 7. (h) Bar 8.

processing, where the attack phase alters the input level and the release phase alters the output level of the signal. The elongation controls the length of the shutter function.

The final shutter function is implemented by means of two hyperbolic tangents, one in forward orientation and the other reflected over the y -axis.



FIGURE 2. Stills of the fireplace scenes. (a) Fire medium shot. (b) Fire close up.

The steepness of the slope translates into attack and release, with the attack being the slope of the first tanh and release that of the second tanh. The elongation is the offset of the function. The bigger the offset value, the further apart the hyperbolic tangents slide and the longer the elongation of the shutter becomes. The offset is used in both tangents, assuring an equal shift of the function. In equation (1), attack, release and elongation are abbreviated as “a,” “r,” and “e,” respectively

$$f(x) = \tanh\left(\frac{x+e}{a}\right) + \tanh\left(\frac{-x+e}{r}\right) \quad (1)$$

Figure 3 shows an example synthetic shutter for creating a 24 fps version from the 192 fps master. The x-axis is divided into eight sections between each major tick. One major tick represents one 1/24 frame. Eight 1/192 frames yield one 1/24 frame. By dividing the major ticks into 1/8th, one 1/24 frame is represented by the space in between the major ticks (e.g., from -1 to 0). 1/192th frames, 1/24th frames, and generally the time are represented by the x-axis.

The y-axis indicates the weight of each frame. Regardless of how many frames are blended, the sum of weights is normalized to one to keep the luminance of the resulting version the same as the original.

The red box in Fig. 3 represents a 180° box shutter. To create a 180° box shutter, the frame has to be exposed half of the time. In this case, the output signal is 24 fps, occurring from 192 fps material. As stated before, one 1/24th frame lies between two major ticks, or eight 1/192th frames. Exposing the frame half of the time means combining half of the 1/192th frames to one 1/24th frame. Therefore, four 1/192th frames are combined, and the other four are discarded. The red box indicates that the shutter is open (i.e., exposure), and the white gap in between indicates a closed shutter.

Creating a synthetic 180° shutter depends on the frame rate of the source material and, subsequently, on the target frame rate of the output. This 180° shutter is also called a box shutter and has been the shutter of choice for most cinematic productions, being one of the key factors of the “film look.”

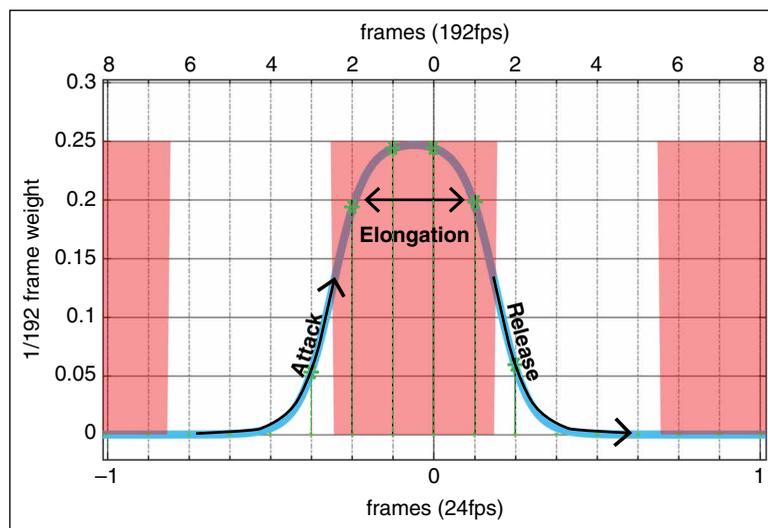


FIGURE 3. Attack, release, and elongation parameters visualized. The red box visualizes four classic 24 fps 180° shutter. The blue line shows our shutter function, sampled at the green dots for rendering 24 fps from the original 192 fps data.

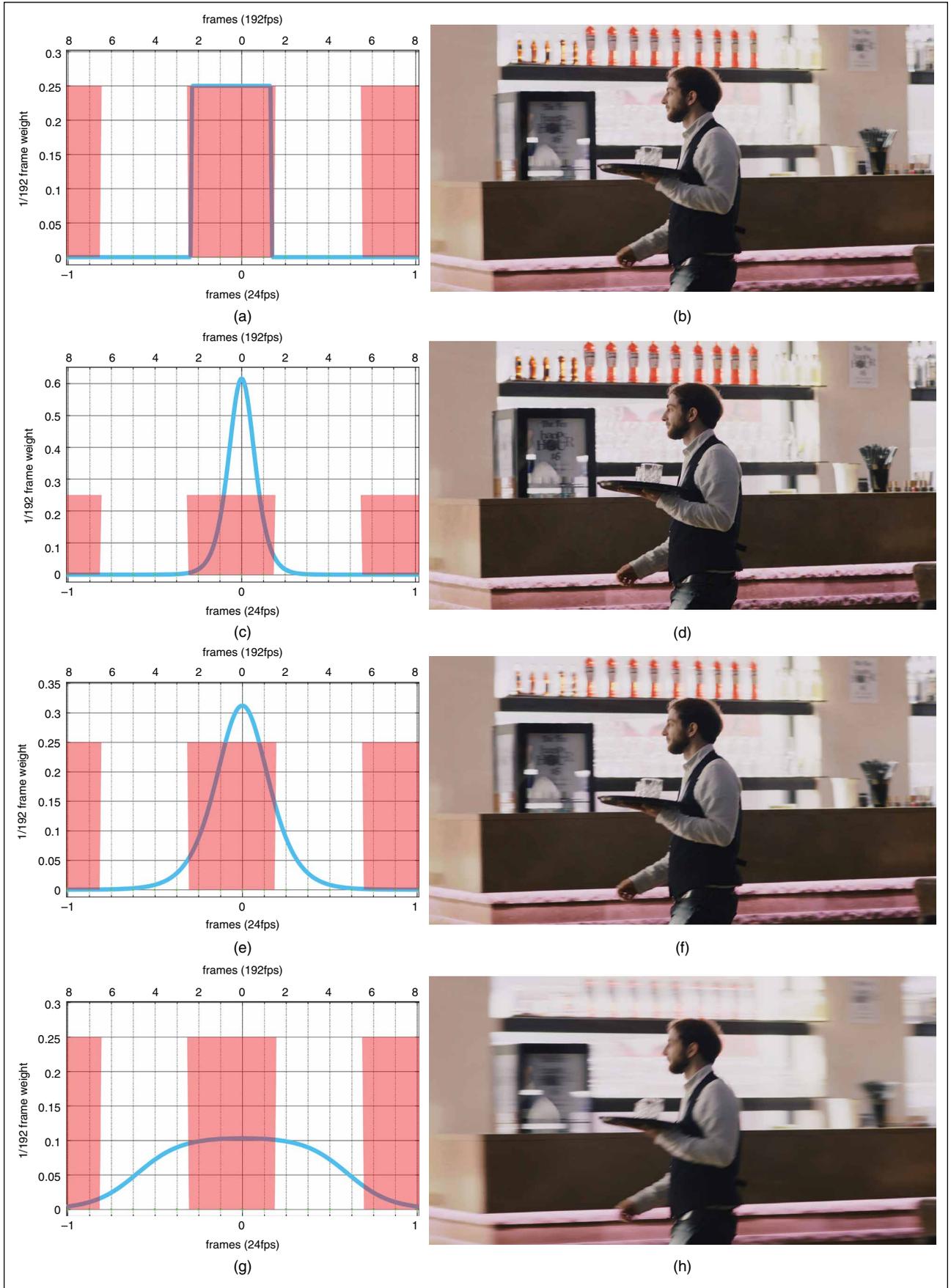


FIGURE 4. Visualization of the shutter length from the user study. (a) and (b) 180° shutter. (c) and (d) Short length synthetic shutter. (e) and (f) Medium length synthetic shutter. (g) and (h) Long length synthetic shutter.

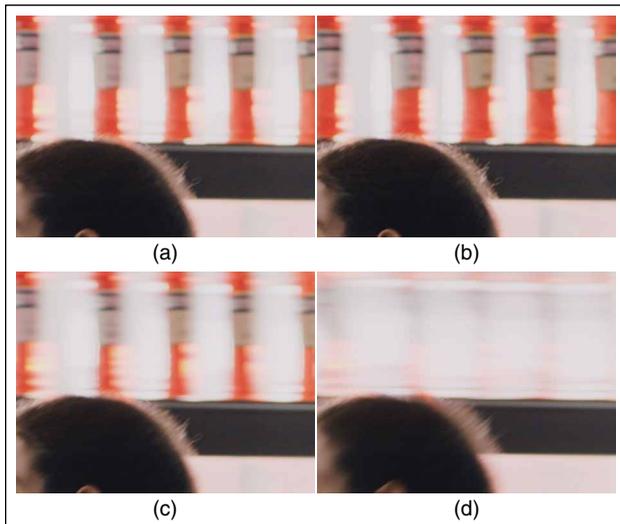


FIGURE 5. Zoomed-in view of one 24 fps frame rendered by the four different shutters. (a) 180° shutter. (b) Short length synthetic shutter. (c) Medium length synthetic shutter. (d) Long length synthetic shutter.

Synthetic Shutter Shapes Examples

After preselecting the different shutter shapes, three different shutters with shapes that resemble a Gaussian distribution are chosen for the visual survey. The “*a*” and “*r*” values are set to the same value to achieve a symmetric shape. An advantage of this shutter shape compared to a normal Gaussian distribution is that the weight is distributed more evenly between the mid-frame and its abutting frames. Blending by a Gaussian distribution function would emphasize the mid-frame, resulting in a higher noise level in the final frame blend.

The still frames that result from the three parameter sets are shown in **Figs. 4** and **5**. First, a 180° box shutter is given as a reference [(**Figs. 4(a)** and **5(a)**); the other three images show synthetic shutters according to *a*, *r*, and *e* parameters from **Table 2**. The box shutter is closest to the medium-soft synthetic shutter illustrated in **Figs. 4(e)**, **4(f)**, and **5(c)**, since the medium-soft synthetic shutter lays about 50% emphasis on the four frames that would have been blended by the box shutter and adds three more frames with less weight.

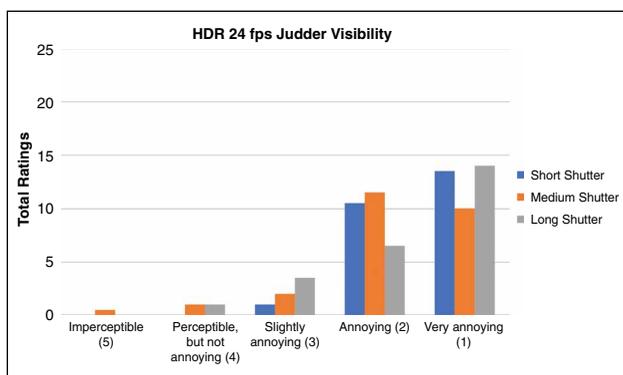


FIGURE 7. HDR, 24 fps sequences with the three synthetic shutters applied.

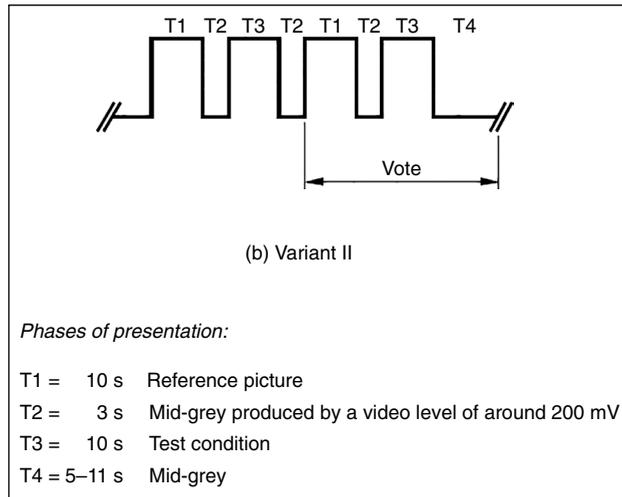


FIGURE 6. Presentation structure of test material. Illustration from “ITU-R BT.500-13.”¹¹

The three synthetic shutters differ in the number of frames being blended. Shutter 4c blends three frames, shutter 4e blends seven frames, and shutter 4g blends 17 frames. This was achieved by adjusting *attack* and *release* as well as the *elongation*. The more the frames are blended, the blurrier the still becomes, which can be seen in **Fig. 4**. Blending 17 frames yields a lot of motion blur, rendering the sequence very soft.

Visual Survey

In order to evaluate if judder artifacts can be reduced by frame blending, a visual survey is set up. The study is based on the International Telecommunication Union - Radiocommunication (ITU-R) Recommendation BT.500-13¹¹ and the International Telecommunication Union-Telecommunication (ITU-T) Recommendation P.800.1.¹² The goal of this survey is to determine how different shutter shapes affect the viewing experience related to judder artifacts. In this survey, the sequences, graded for 100 cd/m² and 1000 cd/m² peak luminance, are shown rendered with the four different shutter shapes introduced earlier. Participants are seated at a 1.5 screen height viewing distance in front of the screen and are asked to maintain this distance.

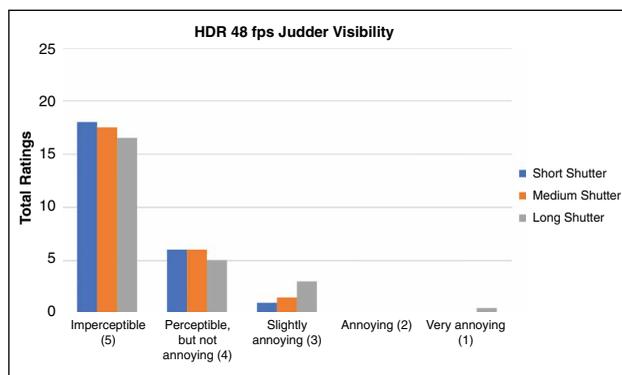


FIGURE 8. HDR, 48 fps sequences with the three synthetic shutters applied.

TABLE 2. The different impaired sequences.

Frame rate in fps	Shutter shape name	Parameters			Luminance in cd/m ²
		a	r	e	
24	Short	0.1	0.1	0.01	100
24	Medium	0.2	0.2	0.01	100
24	Long	0.25	0.25	0.6	100
24	Short	0.1	0.1	0.01	1000
24	Medium	0.2	0.2	0.01	1000
24	Long	0.25	0.25	0.6	1000
48	Short	0.1	0.1	0.01	100
48	Medium	0.2	0.2	0.01	100
48	Long	0.25	0.25	0.6	100
48	Short	0.1	0.1	0.01	1000
48	Medium	0.2	0.2	0.01	1000
48	Long	0.25	0.25	0.6	1000

Blackmagic Design Resolve is used as the playback system, outputting a 10 bit 2048 × 1080 pixel RGB 444 signal using the SMPTE ST 2084 nonlinearity and ITU BT.2100 primaries. The images are displayed on an EIZO ColorEdge PROMINENCE CG3145 monitor. This monitor uses a dual LCD-panel with an RGB LED backlight, reaching peak luminance of 1000 cd/m². It has a contrast ratio of 1,000,000:1. The screen diagonal is 31.1 in.¹³

The participants of this study included both experienced and novice viewers. Experienced viewers are professionals encountering motion artifacts on a daily basis. In total, 25 persons were tested, ranging from 23 to 55 years of age. All participants have normal vision or corrected normal vision.

The test method used is based on the ITU BT.500 double-stimulus impairment scale method. In this method, the participant is presented with an impaired sequence shortly after an unimpaired sequence as illustrated in Fig. 6. This method was adopted since

the reference sequence (the supposedly unimpaired one) already contains judder artifacts. Subsequently, the participant sees a sequence with a modified shutter, shortly after the reference sequence. This process is cycling, meaning that each time the participant will watch the reference sequence and then the impaired sequence. This resets the viewer’s visual habituation and makes it easier to rate the impaired sequence based on the reference.

The basic concept of judder was explained before the subject would see the training/reference sequence for the first time. The judder was pointed out while watching the training/reference sequence. Before watching the reference, the basic chronology of the timeline was described, preventing confusion during the actual test. Furthermore, the viewers were asked to focus on the whole scene, not just the actors. Between presentations of the clips, the users were given a questionnaire in which they rated the relevant sequences.

The reference is computed with the 180° box shutter since this is the cinematic standard shutter used most commonly in modern-day cinematography to create the “film look.”¹⁴ As can be seen in Table 2, the sequences tested in the study contain all possible combinations of two luminance levels, two different frame rates, and three different shutters, totaling 12 test sequences.

First, the reference sequence is shown to the participants, followed by the impaired sequence with a three second gray screen in between. After the impaired sequence, a seven second gray screen is shown, which allows time for the participant to rate the impaired sequence based on the reference sequence. To prevent systematic hysteresis errors in the evaluation, the order of the clips is randomized per participant.

Results

Figures 7 and 8 show the distribution of judder ratings from imperceptible to very annoying for the HDR sequences. It can be observed that increasing the frame rate has the strongest impact on the judder artifact. Horizontally, the five rating categories are displayed, vertically, with the number of votes per category.

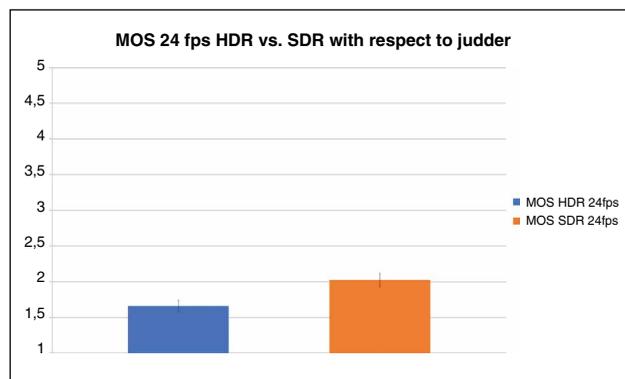


FIGURE 9. Mean opinion score comparing SDR and HDR at 24 fps.

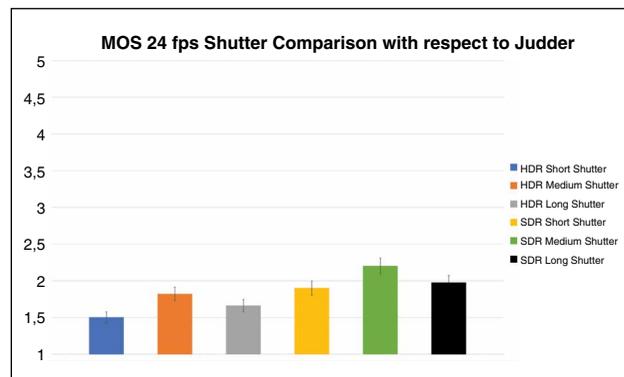


FIGURE 10. Mean opinion score for the three synthetic shutters in 24 fps.

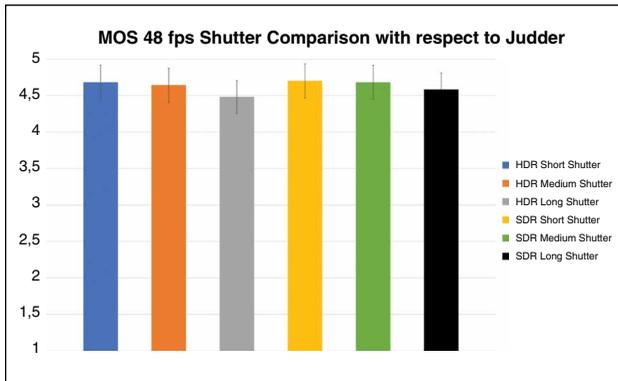


FIGURE 11. Mean opinion score for the three synthetic shutters in 48 fps.

The ratings are extremely one-sided, with most of the votes being “imperceptible” or “perceptible, but not annoying.” **Figures 9–11** show the mean opinion score (MOS) of all 24 fps sequences in HDR and SDR. The results are distributed between “very annoying” and “annoying.” In terms of motion judder, SDR is preferred by the participants. The MOS between SDR and HDR differs about 0.5 points.

Furthermore, **Figs. 9** and **10** support the fact that the spatiotemporal contrast sensitivity of the human visual system is reduced by lower luminance levels. Reducing the spatiotemporal contrast sensitivity subsequently reduces the perception of judder.

Figure 11 shows the comparison of the MOS of 24 fps shutters and the 48 fps shutter. With 48 fps, the participants perceive almost no judder artifacts regardless of the shutter type, while the judder artifacts are perceived regardless of the shutter type in 24 fps.

Hence, doubling the frame rate reduces judder artifacts significantly, and in most cases, judder artifacts disappear completely.

Conclusion

It is shown that moving from SDR to HDR increases the judder in the 24 fps presentation. The evaluation of the visual survey shows that judder cannot be significantly reduced by rendering softer synthetic shutter shapes at a traditional 24 fps frame rate. Even using a very long shutter still only results in a very slight reduction of judder but renders an extremely soft and, therefore, a non-cinematic motion feel. The computational effort and the increase of data rates to implement such a shutter do not justify the marginal benefit. Increasing the frame rate to 48 fps eliminates judder artifacts for almost all participants. Hence, to effectively reduce judder artifacts, a frame rate increase cannot be avoided. As a result, future research should look into the question of how to maintain the cinematic look at higher frame rates. The data set presented in this article is intended to help further investigation of this question. It can be downloaded from the project website.⁹

Acknowledgment

The Stuttgart Media University (HdM) and Arnold and Richter Cinetechnik (ARRI), Munich, Germany provided technical equipment for the shoot, postproduction, and user study. Stefan Grandinetti supervised content creation. We are thankful to all fellow students and friends who supported creating the test footage by helping in the shoot and acting. We would also like to thank the participants of the visual study, Florian “Utsi” Martin for grading the test footage, and the reviewers and editor for their helpful comments regarding this article.

References

1. J. Froehlich et al., “Creating Cinematic Wide Gamut HDR-Video for the Evaluation of Tone Mapping Operators and HDR-Displays,” *Proc. SPIE*, 9023:9023 – 9023, 2014.
2. Dolby Laboratories, “Dolby Vision White Paper,” 2016. Accessed October 10, 2019. [Online]. Available: <https://www.dolby.com/us/en/technologies/dolby-vision/dolby-vision-whitepaper.pdf>
3. Dolby Laboratories, “Dolby Drives Global Momentum at CinemaCon 2019,” Mar. 2019. Accessed October 10, 2019. [Online]. Available: <https://news.dolby.com/en-WW/172164-dolby-drives-global-momentum-at-cinemacon-2019>
4. C. Chinnock, “Samsung Launches First U.S. LED Cinema Screen along with Post Production Facility,” 2018. Accessed October 10, 2019. [Online]. Available: <https://www.displaydaily.com/article/display-daily/samsung-launches-first-u-s-led-cinema-screen-along-with-post-production-facility>
5. Sony Corporation, “Sony at CinemaCon 2018: Re-inventing the Movie-Going Experience,” 2018. [Online]. Available: https://pro.sony/en_AL/press/cinemacon-2018-re-inventing-movie-going-experience
6. R. Larsen, “Annual 4k TV Sales to Surpass 100 Million This Year,” Dec. 2018. Accessed October 10, 2019. [Online]. Available: <https://www.flatpanelshd.com/news.php?subaction=showfull&id=1544074043>
7. R. Collin, “The Hobbit—An Unexpected Journey, Film Review,” Sept. 2012. Accessed October 10, 2019. [Online]. Available: <https://www.telegraph.co.uk/culture/film/filmreviews/9730525/The-Hobbit-An-Unexpected-Journey-movie-review.html>
8. A. Kuder, “High Frame Rate (HFR)—Datensatz unter dem Aspekt der Bewegungswahrnehmung,” Bachelor thesis, Stuttgart Media University (HdM), Apr. 2017. [Online]. Available: <https://hdms.bsz-bw.de/frontdoor/index/index/docId/6537>
9. I. Beitzel et al., “The HdM-HFR-2017 Data Set,” Visual Media Lab, Stuttgart Media University (HdM), 2019. Accessed October 27, 2019. [Online]. Available: <https://www.hdm-stuttgart.de/vmlab/hdm-hfr-2017/>
10. T. Davis, “Rethinking Frame Rate and Temporal Fidelity in a Cinema Workflow,” *SMPTE Mot. Imag. J.*, 126(8), Oct. 2017. Available: <https://doi.org/10.5594/JMI.2017.2733098>
11. International Telecommunications Union - Radiocommunication (ITU-R), Recommendation BT.500-13, “Methodology for the Subjective Assessment of the Quality of Television Pictures,” 2012. [Online]. Available: <https://www.itu.int/rec/R-REC-BT.500>
12. International Telecommunications Union - Radiocommunication (ITU-R), Recommendation ITU-T P.800.1, “Mean Opinion Score (MOS) Terminology,” Jul. 2016. [Online]. Available: <https://www.itu.int/rec/T-REC-P.800/>
13. EIZO Europe GmbH, “EIZO ColorEdge PROMINENCE CG3145 Datasheet,” Jul. 2018. Accessed October 10, 2019. [Online]. Available: https://www.eizoglobal.com/support/db/files/catalogs/ce/ColorEdge_PROMINENCE.CG3145.pdf
14. S. Daly et al., “A Psychophysical Study Exploring Judder Using Fundamental Signals and Complex Imagery,” 2014. [Online]. Available: <https://doi.org/10.5594/j18616>

About the Authors



Ianik Beitzel is a colorist and image technician at Studio Hamburg Postproduction GmbH, Hamburg, Germany. Prior to working at Studio Hamburg, Beitzel studied audiovisual media at Stuttgart Media University (HdM), Stuttgart, Germany where he graduated with a master's degree. He wrote his thesis in collaboration with Arnold & Richter Cinetechnik GmbH & Co. Betriebs KG (ARRI), Munich, Germany, where he did his research on synthetic shutter and its effect on judder reduction in motion pictures.



Aaron Kuder is a conforming engineer at ARRI Media, Munich, Germany. After graduating from Stuttgart Media University (HdM) with a bachelor thesis on creating and comparing cinematic test footage in standard frame rate and high frame rate, he is actively participating in research on higher frame rate cinema. Most recently, he presented his work at the HPA Tech Retreat, Palm Desert, CA, and the European

Federation of Cinematographers (IMAGO) Conference, Brussels, Belgium.



Jan Fröhlich is a professor of motion picture engineering at Stuttgart Media University (HdM). His research interests include high dynamic range (HDR) and wide color gamut (WCG) image encoding, color rendering, and camera metrology. Before becoming a professor, he worked as senior image scientist at Arnold & Richter Cinetechnik GmbH & Co. Betriebs KG (ARRI), Munich, Germany, and for Dolby Laboratories, Sunnyvale, CA, contributing to the development of the ITU Rec. 2100 ICtCp color space and the Dolby Vision HDR and WCG content distribution platform. Before starting his PhD, he was a technical director at CinePostproduction GmbH, Berlin, Cologne, Hamburg, and Munich, Germany. Fröhlich is a member of SMPTE, the Society for Imaging Science and Technology (IS&T), Fernseh-und Kinotechnische Gesellschaft (FKTG), and the German Society of Cinematographers (BVK).

A contribution received from the authors July 2019. This paper was presented at the Hollywood Professional Association Tech Retreat 2019, Palm Desert, CA, 11–15 February 2019. Copyright © 2020 by SMPTE.





subscribe today:
soundcloud.com/smpteconnect