

Evaluation of the Impact of High Frame Rates on Legibility in S3D Film

Michael Marianovski
Dept. of Biology
York University
4700 Keele Street
Toronto ON M3J 1P3*

Laurie M. Wilcox
Dept. of Psychology
York University
4700 Keele Street
Toronto ON M3J 1P3†

Robert S. Allison
Dept. of Electrical Engineering and Computer Science
York University
4700 Keele Street
Toronto ON M3J 1P3‡

Abstract

There is growing interest in capturing and projecting movies at higher frame rates than the traditional 24 frames per second. Yet there has been little scientific assessment of the impact of higher frame rates (HFR) on the perceived quality of cinema content. Here we investigated the effect of frame rate, and associated variables (shutter angle and camera motion) on viewers' ability to discriminate letters in S3D movie clips captured by a professional film crew. The footage was filmed and projected at varying combinations of frame rate, camera speed and shutter angle. Our results showed that, overall, legibility improved with increased frame rate and reduced camera velocity. However, contrary to expectations, there was little effect of shutter angle on legibility. We also show that specific combinations of camera parameters can lead to dramatic reductions in legibility for localized regions in a scene.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation – Display Algorithms; I.3.3 [Computer Graphics]: Picture/Image Generation – Human Factors; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism – Human Factors.

Keywords: high frame rate, perception, acuity, blur, stereoscopic 3D

1 Introduction

In the 1920s cinema frame rate was standardized to facilitate stable audiovisual synchronization [Holman 2010]. The resulting 24 frames per second (fps) standard was a compromise that provided acceptable visual quality under technical and economic constraints; these trade-offs are no longer necessary in modern digital cinema. In fact it is now possible to capture and present cinema content at much higher frame rates (HFR). As outlined below, many believe that the enhanced fidelity provided by HFR capture/presentation has the potential to dramatically improve viewer experience. That is, by increasing the temporal sampling rate (for matched capture and display), content will appear 'crisper' or of higher resolution. Further, motion induced artefacts such as strobing and judder,

which are common in 24 fps content, will be reduced. With such potential gains, it seems obvious that content creators and distributors would adopt HFR protocols. This point of view is championed by long-time HFR advocates such as Doug Trumbull, creator of the 60fps, 70-mm Showscan™ format in the late 1970s. Trumbull has argued that HFR is essential to meeting the resolution levels that allow viewers to feel fully immersed in a film [Kaufman 2012]. It stands to reason that reduction of motion artefacts and blur will allow viewers to see more detail and make content more legible. This has implications for particular types of content such as advertising or sporting events. However, even though the technical impact of HFR on motion quality has been widely recognized, there have been few attempts to assess its impact empirically. The goal of this experiment was to evaluate the impact of frame rate on task performance using S3D movies. As outlined below, we consider variables that interact with frame rate to determine the degree of motion blur in film content: shutter angle and camera motion.

To create moving pictures, a dynamic scene is stroboscopically sampled to create a sequence of discrete still images. The fidelity of the motion sequence depends critically on the temporal sampling rate (frame rate). Sampling artefacts will be visible to the viewer if they fall within a range of spatio-temporal conditions identified as the 'window of visibility' [Watson 1986; Watson 2013]. Thus in the case of motion across a sensor, low frame rates are more likely to insufficiently sample the motion, resulting in jumps in apparent position, which break the sense of smooth motion (judder). Increasing the sampling rate pushes the potential artefacts outside the visible range, resulting in the percept of relatively smooth motion.

Motion blur is also affected by the exposure duration of each frame. While mechanical shutters are no longer in use, it is still conventional to refer to the exposure duration in terms of the equivalent shutter angle. The typical shutter angle used for 24 fps footage is 180°, which is equivalent to the shutter being 'open' for half of the frame. Under stroboscopic motion conditions, the degree of motion blur in an image is directly related to the length of time it is exposed, as the image moves across the camera sensor during this interval. The impact of exposure on motion blur is complicated by the fact that projection systems show each frame for a fixed flash duration, and these flashes are typically repeated two or three times per frame. This projector exposure may also impact perceived motion blur.

The impact of both frame rate and shutter angle are modulated by image velocity. Motion artefacts caused by rapidly moving cameras or objects can be reduced by employing higher frame rates and smaller shutter angles. However, although the physical constraints of stroboscopic image sampling can be modeled, and blur can be quantified, there is evidence that these calculations do not directly predict *perceived* blur. For instance, Burr and Morgan (1997) showed that moving objects are perceived as less blurry than would be expected from predicted motion blur. Other researchers have

*e-mail: marimich@yorku.ca

†e-mail: lwilcox@yorku.ca

‡e-mail: allison@cse.yorku.ca

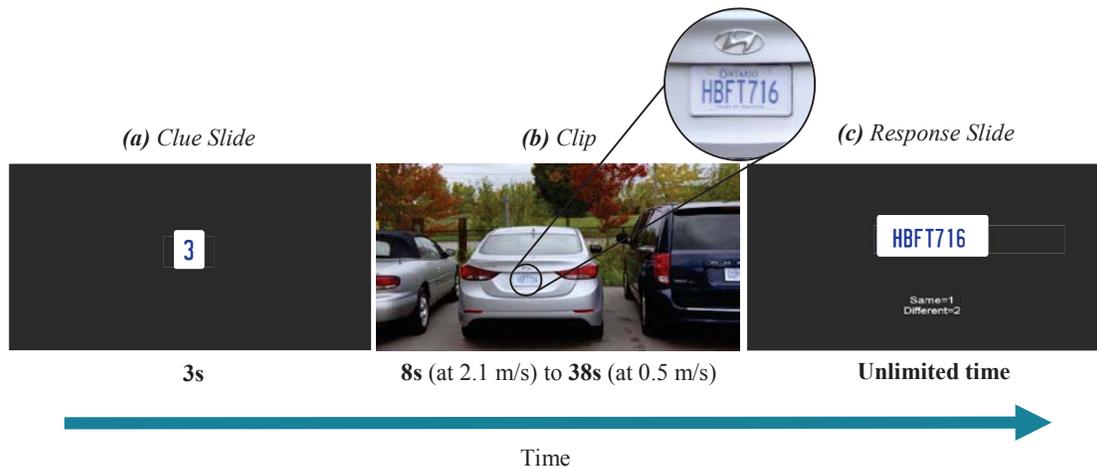


Figure 1: A timeline of the experimental procedure. The cue slide was displayed for 3 s followed by presentation of the clip which lasted 8 to 38 s depending on the camera speed. Lastly, the response slide was displayed until all participants responded.

contrasted perceived blur in static and moving images that contain equivalent blur to show that moving objects appear sharper than equivalently blurred static objects [Bex et al. 1995]. These studies create a strong foundation for understanding the effect of de-blurring on visual acuity. However, in all cases simple stimuli such as moving bars or sinusoidal gradients were presented in 2D. It is not clear if the observed results apply to more realistic situations, for instance with natural scenes presented in S3D. Here we use S3D footage filmed under conditions of varying motion and blur created by manipulating frame rate, shutter angle and camera speed to quantitatively assess their impact on a letter legibility task.

2 Methods

2.1 Participants

Participants were recruited from the Department of Psychology Undergraduate Research Participant Pool at York University and received course credit as compensation. All participants provided informed consent prior to the experiment and wore their prescribed corrective glasses/lenses. Prior to the experiment all participants completed an acuity test (a modified tumbling E chart) and a test of stereopsis in which they had to identify letters presented in a random dot stereogram. A total of 73 observers participated in the study; five were excluded due to either poor performance on the acuity (minimal angular resolution greater than 1.0 minute of arc) or stereoacuity task (disparity threshold greater than 1.5 minutes of arc) or for failing to respond to all the trials. Thus, the results of 68 viewers were included (36 in the first session, 32 in the second).

2.2 Apparatus

Capture: The stimuli were filmed by a professional film crew at Sheridan College’s Screen Industries Research and Training (SIRT) Centre. The left and right cameras used for filming the stereoscopic 3D footage were Sony® F65 cameras with 19-90 mm T2.6 Fujinon® Cine Zoom Lenses. The cameras were mounted on a third generation Stereo 3D Tango beam splitter rig. The footage was color corrected on the Quantel Pablo Rio.

Testing: The clips were arranged in a playlist in a pre-specified order and projected using a Christie® Solaria 4220 Digital Cinema Projector with a resolution of 4096 x 2160 pixels and a throw

distance of 24.6 m. The projector flash rate protocol was fixed at a triple flash for the 24 and 48 fps clips (on each frame six images were flashed, 3 repetitions of the left image and 3 repetitions of the right image), and double flash for each eye for the 60 fps clips. The stimuli were projected onto a Da-lite White Screen with a width of 6.7 m and a height of 3.8 m.

Participants sat in front of the screen in five rows. The first row was positioned 7.3 m from the screen and each subsequent row was offset by 1 m so that the back row was 11.3 m from the screen. Each row had seven or eight chairs. Viewers wore Xpand 3D® Cinema shutter glasses.

Participants recorded their responses using ResponseCard LCD electronic clickers made by Turning Technologies, LLC. A corresponding receiver and polling software (TurningPoint® 5) was used to record responses.

2.3 Stimuli

Clip: The test stimuli consisted of a series of film clips, organized in a playlist as described below. Each test shot consisted of the same scene: a row of six cars with their rear license plates aligned and facing the camera (Figure 1b). The camera moved laterally along a track positioned behind the cars, starting from the left side mirror of the leftmost car and ending with the right side mirror of the rightmost car. All factorial combination of the frame rate (24, 48 and 60 fps), shutter angle (90, 180 and 358°) and camera speed (0.5, 0.9, 1.5 and 2.1 m/s) were filmed, resulting in 36 conditions. Each condition was filmed twice for a total of 72 shots. In each shot a unique set of license plates, prepared and ordered in advance, were mounted to the cars using hidden magnets (see Figure 2 to compare sample frames from clips depicting the effects of speed, frame rate and shutter angle).

Due to a technical mishap during filming all eight shots at 60 fps and 90° shutter angle had to be discarded (4 speeds * 2 takes). Thus a total of 32 testable conditions were available for experimentation.

License plates: The license plates used in this experiment were specially designed and produced by a prop company. They were modified replicas of Ontario, Canada official license plates. Each

plate comprised a unique combination of four letters followed by three numbers with constant character spacing (Figure 1b).

Cue slide: A mid-grey ‘cue’ slide was created for each trial and was displayed prior to the test clip. This slide identified the target car and license plate for the participant. The cue consisted of a number between 1 and 6 (corresponding to the target car) in a white rectangle at the center of the screen (Figure 1a).

Response slide: The response slide was similar to the cue slide, but contained a series of four letters and three numbers with a font similar to that used in the license plates (Figure 1c). The characters on the response slide were the same as those in the cued license plate on half the trials (‘same’ trial), on the remaining trials a single letter was substituted (‘different’ trial). To make the task manageable, the character substituted was always one of the initial four letters and never the numbers; the participants were made aware of this rule. Both the cue and response slide content was displayed at the screen plane (zero binocular parallax).

2.3.1 Procedure

Participants were transported to the SIRT studio by bus. Two groups of participants (39 and 34 participants) were tested on separate days. Prior to testing, participants were given a brief introduction to the experiment and were informed of the general objective. Participants then gave their written informed consent and received response sheets (for demographic information as well as the acuity and stereoacuity test).

At the studio, viewers were seated and their acuity and stereoacuity were assessed. At the outset, the participants completed three practice trials to familiarize themselves with the task. Each experimental trial was preceded by a cue slide which contained a number between 1 and 6 that corresponded to the order in which the cars entered the clip. The experimenter stated this number verbally to help avoid confusion (since the cars were not explicitly numbered in the clips). Viewers were asked to attend to this plate while watching the clip so they could indicate whether the letters on the subsequent response slide were the same or different. Participants were prompted to respond using the electronic clicker by clicking “1” for same and “2” for different. Trial duration

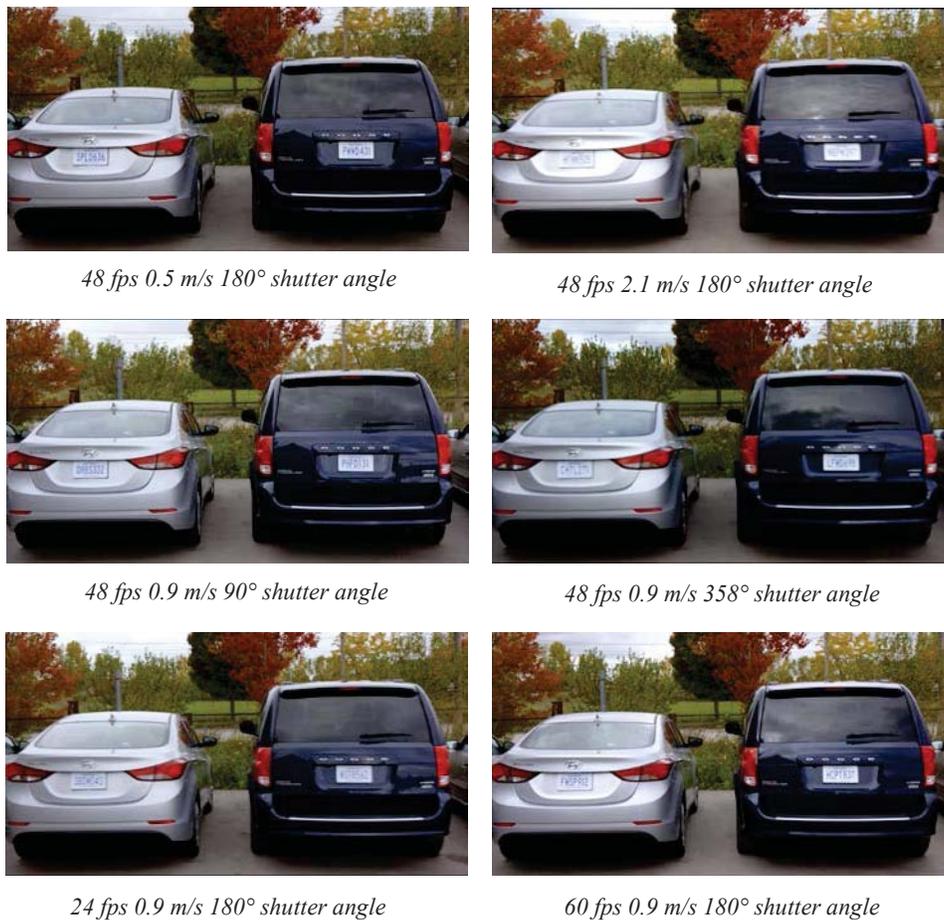


Figure 2: Sample frames taken from stimuli used in the experiment. Each panel shows a single frame from a given clip with a unique set of camera parameters (frame rate, camera speed and shutter angle) producing varying levels of blur. The upper row illustrates the effect of camera speed at a fixed frame rate (48fps) and shutter angle (180°). The middle row shows the effect of shutter angle with a fixed frame rate (48fps) and speed (0.9 m/s). The lower row illustrates the effect of frame rate with a fixed speed (0.9 m/s) and shutter angle (180°).

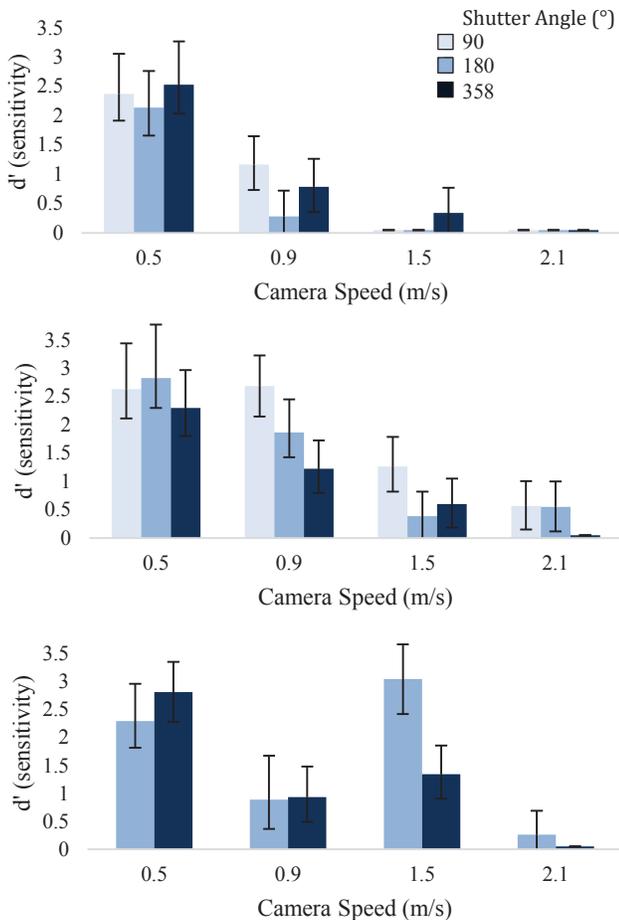


Figure 3: Here d' values are plotted for each camera speed tested in the experiment. Shutter angles 90°, 180° and 358° are represented by colours light, medium and dark respectively. Each panel represents a different frame rate ((a) 24 fps, (b) 48 fps, and (c) 60 fps). Error bars represent 95% confidence intervals.

necessarily varied with camera speed with viewing times of 8 and 38 seconds at 2.1 m/s and 0.5 m/s respectively. A total of 64 trials were tested in pseudorandom order in each 60-minute session. At the end of each session, response sheets were collected and participants were debriefed and returned by bus back to York University.

2.3.2 Randomization

Strict conditions were followed to create the ordered playlist. As we had two clips (takes) for each set of unique parameters, one take was presented as a *same* trial and one as a *different* trial to create an equal number of each. On *different* trials each letter position (1 to 4) was changed an equal number of times. Since the original license plates had four unique letters (no repeats), the changed letter was selected so that it would not be a duplicate of the three unchanged letters. This randomization procedure was repeated for both sessions.

3 Results

Results from the two sessions were compiled and d' was computed; this measure of sensitivity is relatively free from the biasing effects of criteria [Swets 1964]. By treating the participants as a sample from a normal population it was possible to sum correct and incorrect responses across participants and across sessions. Using these sums, hit rate (proportion of ‘same’ responses on *same* trials) and false alarm rate (proportion of ‘same’ responses on *different* trials) were calculated for each condition. In this manner d' was computed for each of the 32 conditions (Figures 3a, b and c). For each condition a 95% confidence interval was calculated by sampling the hit rate and false alarm rate from a binomial distribution 10,000 times. The 0.025 and 0.975 quantiles of sampled d' values were used as the limits of the confidence intervals. To evaluate the differences between conditions we computed the differences between d' values for pooled data and computed 95% confidence intervals (as described above). A Holm-Bonferroni correction was applied to control for the type 1 error rate resulting in a critical p value of 0.0167. The analysis confirmed the patterns described below, with a significant effect of speed ($p < 0.001$ for all comparisons) and a small but significant effect of shutter angle ($p < 0.0167$) between 90° and the remaining shutter angles, but no difference between 180° and 358°.

3.1 Frame Rate and Camera Speed

Figure 3a, b and c demonstrates the effect of camera speed with performance in general declining with increasing speed. The effect of camera speed is modulated by frame rate, with overall poorer performance at 24 fps relative to 48 and 60 fps. At all frame rates, performance was at chance ($d' = 0$) at the highest speed (2.1 m/s). At the lowest speed tested here (0.5 m/s), performance was essentially maximized (d' near 2.5) for all frame rates. The positive impact of frame rate on legibility was most evident at a camera speed of 1.5 m/s where at 24 fps observers were at chance, and improved with increasing frame rate up to 60 fps. Observers also improved between 24 fps and 48 fps in the 0.9 m/s condition. After the first test session we noted that performance in the 0.9 m/s condition at 60 fps was surprisingly poor. Our response and evaluation of this result is outlined in detail in section 3.3 below.

3.2 Shutter angle

Most of the frame rate and speed combinations tested here showed no effect of shutter angle on performance. Of 32 conditions only three showed a moderate impact of shutter angle (though in the predicted direction). That is, at 48 fps – 0.9 m/s and 1.5 m/s and 60 fps - 1.5 m/s discrimination performance was better when the shutter angle was smaller (shorter exposure). The relatively weak effect of shutter angle is discussed in more detail in the Discussion section.

3.3 60 fps-0.9 m/s

As mentioned above, the results of our first test session showed that performance in the 60 fps condition at 0.9 m/s was much worse than anticipated (for both shutter angles). Upon closer examination we found that the first car was coincidentally selected as the target car for both those conditions. We hypothesized that an attention lapse at the beginning of the clip might have reduced sensitivity. We ensured that in the second session the first car was not the target for the combination of 60 fps and 0.9 m/s. Nevertheless, performance was also poor for this condition in the second session.

We then meticulously reviewed the stimuli and verified that the playlist and clips were correct. During this process we noted that the target plates in this condition contained the letters M and W.

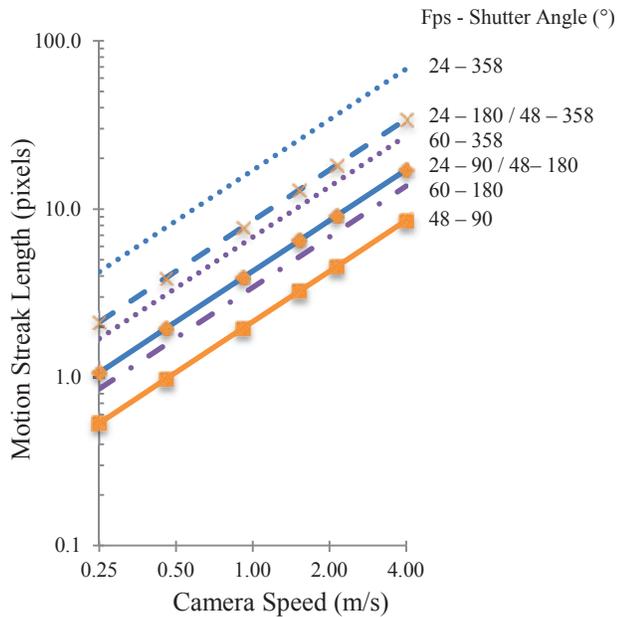


Figure 4: Theoretical length of motion streaks (amount of blur) for each of the frame rate, camera speed and shutter angle conditions tested. Note that the predicted blur is essentially the same for i) 24 fps – 180° SA and 48 fps – 358° SA and ii) 24 fps – 90° SA and 48 fps – 180° SA.

We hypothesized that the unexpectedly low performance was due to spatio-temporal aliasing created by the unique combination of the camera parameters (frame rate, shutter angle and camera speed). This aliasing would make it particularly difficult to discriminate the letters 'M' and 'W' during horizontal motion due to their repeating vertical structure. An analysis of individual responses across the entire dataset showed that participants tended to respond 'different' on same trials that contained an M or a W. A follow up experiment was conducted in which 8 new participants were shown the same set of clips, but were asked to report a single letter on a given license plate (according to a pre-cue). The data from this follow-up study confirmed that, on average, participants were worse at discriminating Ms and Ws than other randomly chosen letters. Further this effect was particularly pronounced for camera speeds of 0.9 m/s at 60 fps with a shutter angle of 358°.

4 Discussion

We used a letter legibility task to objectively assess the effects of frame rate, shutter angle and camera speed. We expect that this task will generalize well to perception of pattern detail and texture as well as text in a cinema context. In general, the influence of these variables on our task is predictable from their theoretical impact on motion blur. For instance, it was expected that performance would improve with increasing frame rate (increasing temporal resolution) and worsen with increasing shutter angle (increasing blur) and increasing camera speed. Further, we expected that these variables would interact, such that the poorest performance would be found at the lowest frame rate, largest shutter angle and highest velocity (see Figure 4 for theoretical blur predictions).

On the whole, the frame rate and motion speed predictions were upheld by our results. That is, for a given speed, performance was better at higher frame rates. Also performance was best at lower

speeds for all frame rates (the only exception was the 60 fps-0.9 m/s condition which is discussed below).

We anticipated that exposure duration, or shutter angle, would have a strong impact on performance. However we found only a weak effect of this variable, and we suspect that this reflects several factors. First, at the slowest camera speed in all three frame rate conditions, performance is at a maximum with d' at or near 2.5; it is likely that at this speed the task is trivially easy so none of the variables impacts legibility. It is clear that the task quickly becomes challenging; for all remaining speeds at 24 fps produce a floor effect, which again obscures any impact of shutter angle. Chance levels of performance are also seen at the highest speed at 48 and 60 fps making it difficult to see any impact of shutter angle. In the remaining conditions, we find that increasing shutter angle degrades performance as expected given the associated increase in motion blur (48fps at 0.9 m/s and 1.5 m/s; 60 fps at 1.5 m/s).

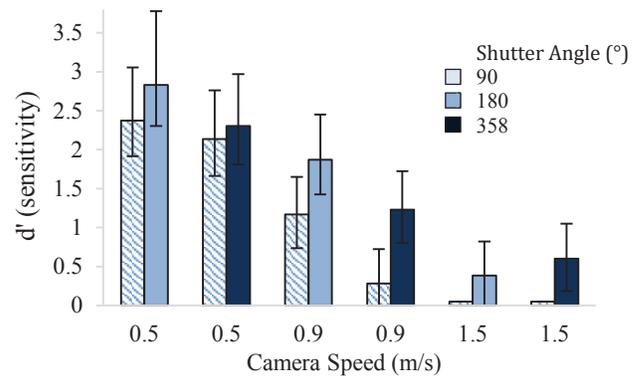


Figure 5: Results for equivalent exposure conditions in 24 fps (striped fill) and 48 fps (solid fill) clips are re-plotted from Figure 3. At each speed, conditions with theoretically equivalent exposure are plotted in pairs with shutter angles 90°, 180° and 360° indicated with light, medium and dark fill respectively. Error bars represent 95% confidence intervals.

Another way to evaluate the effect of shutter angle in our study is to compare conditions with theoretically equivalent exposure. For example, content shot at 24 fps with a 90° shutter angle should theoretically be equivalent to 48 fps at 180° because the duration per frame is halved but the shutter angle is doubled. We have plotted the series of equivalent exposure conditions for comparison in Figure 5, excluding the 2.1 m/s condition where the task is too difficult. It is clear from this figure that there is a difference in performance, and, in spite of the fact that the exposure is the same, performance is consistently superior in the 48 fps conditions.

The limited impact of shutter angle on legibility in our experiment remains somewhat puzzling, particularly given that we employed state-of-the-art capture and projection technologies. One factor that may play a role is the flash protocol used during presentation of the clips. As noted previously, S3D content is typically flashed multiple times per eye in alternation to prevent flicker. In our study we used a triple-flash protocol for both the 24 and 48 fps conditions, and a double flash protocol for the 60 fps presentation. Because we did not vary the inter-flash duration, the length of each flash was longer for the 60 fps compared to the 48 fps clips, although the overall projector exposure per frame was shorter in the 60 fps case. The effects of projector exposure should depend on eye movements [Banks et al. 2012]. However, even in the worst case, given that the flash durations are much shorter than the camera

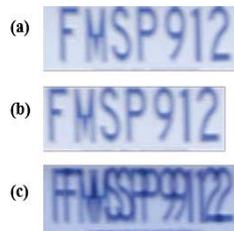


Figure 6. Simulation of the impact of the 0.9 m/s velocity and 60 fps frame rate on legibility of the character 'M' using images of plates from the footage used here. Panel a and b depict the plate in consecutive frames, and c shows them overlaid as they would appear to the viewer. Note that the M becomes ambiguous, and is easily mistaken for a W.

shutter duration we expect that they would have a weaker impact on motion blur.

4.1 60 fps-0.9 m/s

As noted above we found surprisingly poor performance on our legibility task in the 60 fps-0.9 m/s condition. In this case, letters with a vertically oriented periodicity are particularly susceptible to aliasing. Let us approximate the left and right sides of the letter, along with the middle v-shaped part (these are near vertical for M and W in this font), as repeated vertical bars with even spacing. We find that at 0.9 m/s and 60 fps, the vertical elements on a given frame move an amount almost exactly equal to the spacing of the vertical elements. The aliasing is easiest to appreciate in the frequency domain (for simplicity consider a repeating pattern of such elements or a vertically oriented grating). On successive frames the moving bars of the grating would always land on positions previously occupied by bars and the image of the bars would not change. The temporal frequency of such a change is near the frame rate and thus is double the Nyquist rate and strongly aliased (Figure 6).

Thus, we have attributed this result to spatio-temporal aliasing caused by the combination of speed, frame rate and typeface characteristics. The fact that these variables interacted to degrade legibility so dramatically provide a strong caution to content creators. This is particularly true for content that depicts critical detail or text at a particular scale in a scene. Combinations of camera/object speed, shutter angle and the scale of the detail can result in substantially degraded imagery. Importantly, this loss of quality will be region-specific, some areas in a scene may not be affected while other regions are severely distorted.

4.2 De-blurring

Although this experiment was not designed to investigate the effect of perceptual de-blurring on legibility, our results do speak to that literature [Bex, Edgar and Smith 1995; Burr and Morgan 1997]. These studies have shown that, for simple targets, blurred stimuli in motion appear sharper than predicted by calculations of physical blur. However, for these stimuli this is a purely perceptual phenomenon and does not result in improved spatial resolution. Is the same true for naturalistic images presented in S3D? One way to evaluate this is to compare d' for clips that theoretically should have equivalent blur, but different rates of camera motion. More specifically, for frame rates of 24 and 48 fps we compare 0.5 m/s-180° shutter angle with 0.9 m/s-90° shutter angle. If de-blurring operations for naturalistic images improve performance, then we should find an effect of speed, where d' is higher for clips shot with faster camera motion. The rationale for this is that the visual system

primarily accommodates for motion based blur [Barlow 1958]. Instead, consistent with Bex et al's [1995] study, our results show that the task is more difficult at higher speeds and that motion degrades performance even though the blur should be equivalent.

4.3 4K Resolution

The content used here was captured and presented at 4K resolution, for all test conditions. It is notable that for the conventional cinema frame rate (24 fps) performance dropped steeply at speeds as low as 0.9 m/s (walking speed). Only at the slowest speed was legibility at acceptable levels at this frame rate. This illustrates that the benefits afforded by using higher resolution display systems are lost when the content is captured and displayed at 24 fps. Clearly decisions regarding increased display resolution need to be made in conjunction with the availability of adequate frame rates.

5 Conclusion

Performance in the experiments presented here was largely dependent on the amount of physical blur as predicted by manipulation of camera parameters. The exception to this occurred when specific conditions led to aliasing and other motion artifacts which degraded performance. From this it can be concluded that higher frame rates are not always preferable. Instead, filmmakers must consider several variables (content, frame rate, shutter angle, camera or subject motion, and presentation protocol) in conjunction when choosing filming conditions. In addition, we have shown that, in general, higher frame rates and smaller shutter angles should be used when the goal is to emphasize details in moving images.

Acknowledgements

We would like to acknowledge Roy Anthony for his valuable input to the experimental design, and our collaborators at SIRT, including Jim Hagarty, Bert Dunk and John Helliker, for support in capturing, and displaying experimental content. Also we thank Robin Archer for his technical support, and Aishwarya Sudhama, Alexandra Seay, Anna Kiseleva and Megan Goel for their assistance with data collection and organization.

References

- BANKS, M. S., READ, J. C. A., ALLISON, R. S., & WATT, S. J. 2012. Stereoscopy and the Human Visual System. *SMPTE Motion Imaging Journal*, 121, 24–43.
- BARLOW, H.B. 1958. Temporal and spatial summation in human vision at different background intensities. *The Journal of Physiology* 141, 337-350.
- BEX, P.J., EDGAR, G.K. AND SMITH, A.T. 1995. Sharpening of drifting, blurred images. *Vision Research* 35, 2539-2546.
- BURR, D.C. AND MORGAN, M.J. 1997. Motion deblurring in human vision. *Proceedings of the Royal Society B: Biological Sciences* 264, 431-436.
- HOLMAN, T. 2010. *Sound for film and television*. Taylor & Francis.
- KAUFMAN, D. 2012. High Frame Rate Cinema. *Creative Communities of the World*. https://library.creativecow.net/kaufman_debra/High-Frame-Rate-Cinema/1
- SWETS, J.A. 1964. *Signal detection and recognition by human observers: contemporary readings*. Wiley.
- WATSON, A.B. 1986. Temporal sensitivity. *Handbook of perception and human performance* 1, 6-1.

WATSON, A.B. 2013. High frame rates and human vision: A view through the window of visibility. *SMPTE Motion Imaging Journal* 122, 18-32.