# THE VISIBILITY OF MOTION ARTIFACTS AND THEIR EFFECT ON MOTION QUALITY

*Alex Mackin*⋆*, Katy C. Noland*†*and David R. Bull*⋆

⋆Department of Electrical and Electronic Engineering, University of Bristol, BS8 1UB, UK
{A.Mackin, Dave.Bull}@bristol.ac.uk
†BBC Research & Development, Centre House, 56 Wood Lane, London, W12 7SB, UK.
Katy.Noland@bbc.co.uk

## ABSTRACT

The visibility of motion artifacts in a video sequence e.g. motion blur and temporal aliasing, affects perceived motion quality. The frame rate required to render these motion artifacts imperceptible is far higher than is currently feasible or specified in current video formats. This paper investigates the perception of temporal aliasing and its associated artifacts below this frame rate, along with their influence on motion quality, with the aim of making suitable frame rate recommendations for future formats. Results show impairment in motion quality due to temporal aliasing can be tolerated to a degree, and that it may be acceptable to sample at frame rates 50% lower than those needed to eliminate perceptible temporal aliasing.

***Index Terms*—** High frame rates, temporal aliasing, motion artifacts, motion quality, motion blur.

## 1. INTRODUCTION

Video frame rates, higher than those conventionally used today, have been shown to lead to increases in perceived quality [1, 2, 3], due to a reduction in motion artifacts (motion blur and temporal aliasing) [4]. This has generated interest in the film, broadcast, streaming and virtual reality (VR) communities [5, 6]. The frame rates required to eliminate these motion artifacts [7, 8] can be far higher than those specified in the most recent video standard ITU-R BT.2020-2 (Rec. 2020) [9], which supports frame rates up to 120 Hz. To make more realistic and informed recommendations, it is necessary to understand how the visibility of blur and temporal aliasing vary with frame rate and, as the two types of motion artifacts are interdependent, it is essential to take both into account [7].

The visibility of motion blur has already been studied in some detail [4, 10, 11]. However it can be difficult to separate the effects of temporal filtering by the human visual system [12], and the blurring imposed by the display and camera. In this paper we study the perception of, and the impairment in motion quality due to temporal aliasing and its associated artifacts in isolation, achieved by using a stroboscopic display system, which simulates impulsive sampling through the use of a strobe light with a very small duty cycle (short on-time).

## 2. RELATED WORK

The region of perceptible frequencies, defined by the spatio-temporal contrast sensitivity function [13], is referred to by Watson et al. as the "Window of visibility" [14]. Perceptible temporal aliasing is eliminated when spectral replicas of a moving stimulus, due to sampling, fall outside this region [8]. The location at which a spectral replica enters the window of visibility determines whether temporal aliasing will be perceived as flicker, the appearance of perceptible periodic changes in luminance [15], or as strobing, which can manifest itself in various ways. Multiple imaging (banding) is where a stimulus appears at discrete spatially separated locations simultaneously [16], and is due to persistence of vision [17]). Non-smooth motion (judder) [18] depends on the characteristics of the stimulus [19], and the type of the display used [20]. Edge flicker occurs during smooth pursuit eye movements on high-persistence displays [21].

Previous research has shown that the visibility of, and perceived impairment in motion quality due to temporal aliasing artifacts is dependent on a number of factors [18]. These include: characteristics of the stimulus [4], luminance [8], shutter angle [11], eye movements [22] and speed of motion [23]. Previous research has however been limited by the capabilities of the display and the acquisition device used, restricting the range of frame rates that could be studied. We address these issues with the use of a stroboscopic display system with constant luminance, and a simple moving stimulus. We also ask subjects to fixate their gaze, as temporal aliasing artifacts are most visible in non-tracked motion [18].

## 3. EXPERIMENTAL SETUP AND METHODOLOGY

A white circular rigid card (27 cm diameter) with 1 cm long radial black lines was attached onto the tape reel of a Studer A307 tape recorder (see Fig. 1). The speed of the lines relative to the participant was adjusted by changing their radial position and/or the speed of the tape. The black lines had a Michelson contrast [24] of 0.87 when illuminated. A Monarch Instrument Nova-Strobe PBL strobe light was used to simulate temporal sampling at different frame (flash) rates.

**Fig. 1**: The white circular card with black lines (left), and an example of multiple imaging along with the vertical wire used for fixation (right).

To reduce the risk of photosensitive epilepsy, the lowest flash rate of the strobe was set to 60 Hz [25]. To ensure that impulsive sampling was simulated, the duty cycle of the strobe (akin to the shutter angle of a video camera) was set such that the distance traveled by the line within a flash was always smaller than the spatial acuity of the human eye, which we assumed to be 1 arc min as a worse case scenario [27]. This also ensures that the line doesn't appear to move during a flash. Any perceived blurring will be due to temporal filtering by the visual system [12], as no blurring is imposed by the strobe. Therefore, any perceived motion artifacts (compared to an alias-free reference) will be due to temporal aliasing.

Participants sat 60 cm away from the card, while the strobe was placed 20 cm away, which gave an illuminated region of approximately $10^\circ \times 8.5^\circ$. The strobe was the only source of light in the room. Participants fixated their gaze on a marker on a thin vertical wire in front of the card. Motion is assumed to be horizontal at this fixation point. Head movements were restricted by using a chin rest, while a viewing window ensured that only the illuminated region was visible.

Twenty-three participants from BBC Research & Development with an age range of 21–65 years took part in the experiment, and were screened for normal or corrected-to-normal vision. Prior to testing, each participant was given instructions related to the testing process, and took part in a brief training exercise, highlighting the temporal aliasing artifacts they were likely to perceive. A complete session lasted no longer than 30 minutes, and contained regular breaks to limit fatigue. Participants' scores were screened in accordance with the method outlined in ITU-R BT. 500-13 [26].

The target luminance of the strobe was 150 cd/m$^2$, however due to limitations of the strobe light, some variations in measured luminance levels was observed with respect to flash rate, as shown in Table 1. The stimulus, which is modelled as a square pulse, had a width of 20 arc min (0.35 cm). The stimulus speeds used in the experiment were: 10, 30, 50 and 70 °/s.

The reference flash rate used to represent alias-free motion was chosen to be 2000 Hz during informal experimentation.

**Table 1**: The measured luminance (cd/m$^2$) reflected from the card, averaged over 1 second, at the flash rates (Hz) used in the experiment.

| Flash Rate | 60 | 100 | 150 | 300 | 600 | 1000 | 2000 |
|---|---|---|---|---|---|---|---|
| Luminance | 158 | 154 | 151 | 149 | 151 | 145 | 168 |

The double-stimulus impairment scale (DSIS) [26] method was used to record subjective opinions, and involved participants being shown the reference flash rate for 5s before being shown a randomly chosen flash rate for a further 5s. Participants then rated their perceived impairment in motion quality compared to the reference. Participants also recorded whether they perceived blur, flicker, multiple imaging and/or non-smooth motion. Impairment was rated on a 5-point scale ranging from 'Imperceptible' (5) to 'Very annoying' (1) [26]. The reference flash rate was included as a test presentation.

## 4. RESULTS

Mean impairment scores (MIS) collected in the experiment are shown in Fig. 2. According to common practice [4, 10], we interpret a MIS of 4.5 as the critical frame rate, corresponding to an imperceptible difference compared with the reference, and a MIS of 3.5 as an acceptable frame rate. Both the critical and acceptable frame rates increase as the speed of the stimulus increases, as predicted by Watson [8]. Unacceptable impairment due to temporal aliasing artifacts was not observed at the lowest speed of of 10 °/s. All speeds had an imperceptible difference at 1000 and 2000 Hz compared to the reference, therefore these data points have been omitted from Fig. 2. We see the effect of diminishing returns with respect to impairment, except when the stimulus speed is 70 °/s.

Fig. 3 shows the percentage of participants who perceived individual motion artifacts at all tested frame rates. The point at which 50% of participants perceive an artifact is commonly used as a threshold for detection (indicated by the dotted line) [13, p. 8]. Flicker was only perceived at 60 Hz, due to it being the only tested frame rate below the critical flicker frequency of the human visual system [15]. Temporal aliasing artifacts (especially flicker) appear to conceal any blurring imposed by the visual system [12] at the lower frame rates tested, which was more likely to be visible at higher speeds. Non-smooth motion would be expected at the higher frame rates tested, by virtue of there being a perceptible displacement between samples. We postulate that the perception of smooth (apparent) motion at these frame rates is due to the appearance of multiple images [28], and/or amodal completion [29].

During informal experimentation, and through conversations with participants, we noticed that multiple imaging manifests itself in a number of ways. When the spatial displacement between flashes (samples) is greater than the width of the line, as shown in Fig. 4 (a), distinct multiple images are

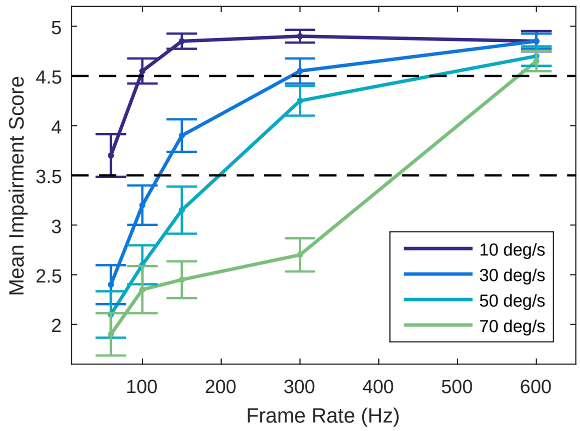



**Fig. 2**: The mean impairment scores for all the speeds and some of the frame rates used in the experiment. Error bars represent standard error of the mean.

perceived. If the displacement between the flashes is equal to the width of the line, the multiple images will interlock, and lead to the perception of a spatial smearing[1], as shown in Fig. 4 (b). This interlocking phenomenon also occurs when the spatial displacement between flashes is equal to $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$... the width of the line, which corresponds to frame rates of:

$$f_s = \left\{ \frac{kr}{W} : k \in \mathbb{Z}^+ \right\} \tag{1}$$

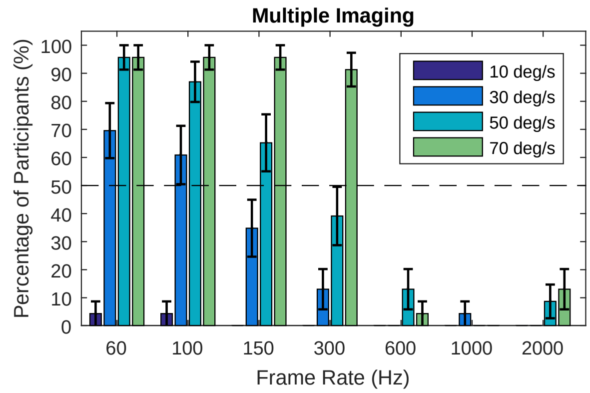
where $r$ is the speed of the line subtended on the retina (°/s) and $W$ is the subtended width of the line (°). This corresponding to flash rates of 60, 100, 150, 300 and 600 Hz at 10 °/s, and 150, 300 and 600 Hz at 50 °/s.

As the speed of the stimulus increases, the width of any blurring imposed by the visual system increases (for a stimulus with sharp edges) [30]. However, in Fig. 3 (b) there is an large increase in the perception of blur at 50 °/s compared to 70 °/s at 150 and 300 Hz. This can be explained by the spatial smearing that occurs at 50 °/s at these frame rates. 600 Hz appears to be above the critical frame rate for both 10 and 50 °/s, and therefore, either multiple imaging is no longer perceived, or the spatial smearing at this frame rate is indistinguishable from the blur imposed by the visual system.

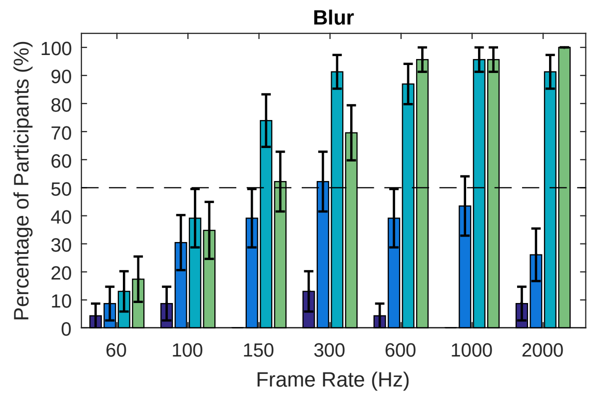
Participants informed us that temporal aliasing artifacts contributed more to the impairment in motion quality than the spatial smearing described above, especially when the reference also appears blurred at the higher speeds tested (Fig. 3 (b)), as they are perceptually similar. Therefore, the apparent outlier in Fig. 2 (300 Hz, 70 °/s), can be attributed to the fact that participants were more likely to perceive blur at 50 °/s (Fig. 3 (b)), which would inflate impairment scores at this point.

If $k$ in Eq. 1 is greater than 1 and non-integer, there is the perception of smearing with distinctive spokes (Fig. 4 (c)).

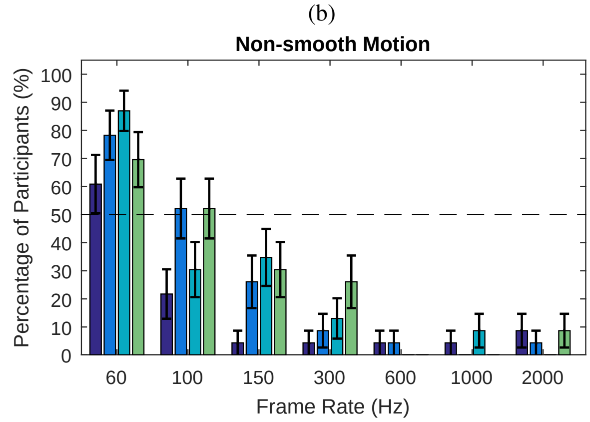
[1]This spatial smearing is perceptually similar to motion blur.




(a)




(b)




(c)

**Fig. 3**: Percentage of participants who perceived individual motion artifacts for all test conditions. Error bars represent standard error of the mean.

Fig. 5 shows a comparison between: the critical and acceptable frame rates predicted from our experiments (interpolating Fig. 2 to find where MIS equals 4.5 and 3.5 respectively), and previously reported results from Watson et al. [14] (participants ABW and JEF) and Bex et al. [16]. A linear relationship exists between stimulus speed and the critical frame rate. Each of the experiments used different stimulus widths and luminance levels, and may explain some of the disparity between results.

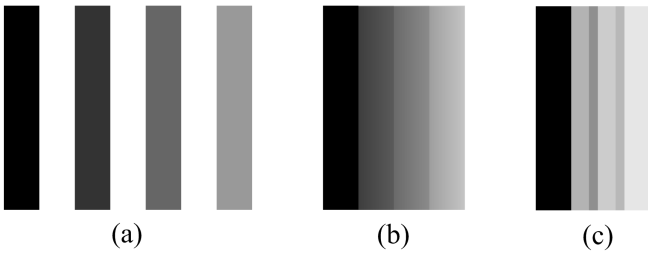



**Fig. 4**: The manifestations of multiple imaging, where the spatial displacement between samples is equal to: (a) 2, (b) 1 and (c) $\frac{3}{4}$ the line width.

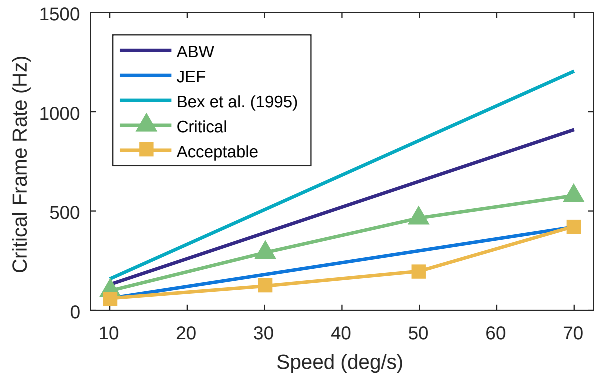



**Fig. 5**: A comparison between previously reported reported data, and our experimental results.

## 5. DISCUSSION

Our results show that it is possible to temporally sample at frame rates lower than the critical frame rate, while maintaining an acceptable level of motion impairment. Using our results we can estimate acceptable frame rates for sample video content. Pan speeds of up to 1.3 screen widths per second have been found in UHD-1 footage of athletics shot during the Commonwealth Games 2014 [31] (motion speeds will vary for other types of content). This corresponds to a speed of 30 °/s for the median viewing distance (2.63 m) and ideal screen size (48") of a sample of UK residents[32]. Critical frame rates at this speed would be around 280 Hz, with an acceptable frame rate of about 140 Hz, a 50% reduction. This conjecture discounts any influence that the characteristics of video content may have on the perceived impairment of motion quality. These calculated frame rates are dependent on motion speed, and highlights how the need for higher frame rates will be both content-dependent, and dependent on the display e.g. the speed subtended on the retina will increase with larger screens, and therefore higher frame rates will be required.

Measured critical and acceptable frame rates may be lower when an actual camera shutter and display are employed, because temporal integration in the camera and display attenuates higher spatio-temporal frequencies within the signal. This may mask some temporal aliasing artifacts by introducing blur, highlighting why the perception of blur and temporal aliasing must be considered together when choosing suitable frame rates, as they are interdependent [7]. The type of display may also play a role, as the perception of temporal aliasing artifacts, specifically non-smooth motion and edge flicker may be exacerbated in hold-type displays, in part due to retinal slip [20]. For real video content, we must also consider the effects of spatio-temporal masking (e.g. for static and dynamic textures) and complex stimuli featuring semantic information, as these may affect the visibility of motion artifacts. A study using video content, suggested a frame rate of 250 Hz [4] to eliminate motion artifacts (although the range of frame rates studied was limited by the display), which is far lower than some of the frame rates in our experiment, highlighting the effects mentioned in this paragraph.

For the speeds tested, the critical and acceptable frame rates for temporal aliasing artifacts were below 700 Hz, the proposed frame rate to remove blur imposed by the camera shutter with Nyquist sampling in [7]. This means that, with realistic sampling using a camera shutter, the frame rate can be reduced to the acceptable frame rate for temporal aliasing, and sharp images maintained with a short shutter duration.

## 6. CONCLUSIONS

We have measured the perceived impairment in motion quality due to temporal aliasing artifacts in isolation, over a range of frame rates and stimulus speeds. Results show that we can tolerate some degree of impairment in motion quality due to these artifacts, and that we may be able to sample at frame rates 50% lower than those needed to eliminate temporal aliasing. We have also shown how the visibility of individual temporal aliasing artifacts varies with frame rates, and that temporal aliasing (e.g. multiple imaging) can be perceived in various ways. Predicted critical frame rates from our experimental results are slightly different than previously reported results, although a finer granularity of frame rates and investigation into the effects of stimulus size would be needed before a proper comparison could be made. Further investigation is also required into the perception of, and the effect on motion quality due to the interlocking phenomena outlined in this paper.

## 7. ACKNOWLEDGEMENTS

The authors acknowledge BBC Research & Development and EPSRC for funding through an iCASE award. We would like to thank Tony Richardson, Andrew Mason, Paul Cooper and Andrew McParland for helping us design and set-up up the experiments, as well as Graham Thomas and Andrew Cotton for their invaluable advice, comments and feedback.

All the experiments were performed at BBC Research & Development in London, while further analysis of the results was completed at the University of Bristol.

## 8. REFERENCES


[1] Y. Ou, Z. Ma, T. Liu, and Y. Wang, "Perceptual quality assessment of video considering both frame rate and quantization artifacts," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 21, no. 3, pp. 286–298, 2011.

[2] M. Emoto, Y. Kusakabe, and M. Sugawara, "High-frame-rate motion picture quality and its independence of viewing distance," *Journal of Display Technology*, vol. 10, no. 8, pp. 635–641, 2014.

[3] A. Mackin, F. Zhang, and D. Bull, "A study of subjective video quality at various frame rates," in *Image Processing (ICIP), 2015 22nd IEEE International Conference on*, 2015.

[4] Y. Kuroki, T. Nishi, S. Kobayashi, H. Oyaizu, and S. Yoshimura, "A psychophysical study of improvements in motion-image quality by using high frame rates," *Journal of the Society for Information Display*, vol. 15, no. 1, pp. 61–68, 2007.

[5] M. Armstrong, D. Flynn, M. Hammond, S. Jolly, and R. Salmon, "High frame-rate television," *BBC Research & Development White Paper 169*, 2008.

[6] D. Lanman, H. Fuchs, M. Mine, I. McDowall, and M. Abrash, "Put on Your 3D Glasses Now: The Past, Present, and Future of Virtual and Augmented Reality," in *ACM SIGGRAPH 2014 Courses*, 2014, pp. 12:1–12:173.

[7] K. Noland, "The application of sampling theory to television frame rate requirements," *BBC Research & Development White Paper 282*, 2014.

[8] A. Watson, "High frame rates and human vision: A view through the window of visibility," *SMPTE Motion Imaging Journal*, vol. 122, no. 2, pp. 18–32, 2013.

[9] ITU-R Rec, "BT. 2020-2, Parameter values for UHDTV systems for production and international programme exchange," *International Telecommunications Union. Geneva*, 2015.

[10] M. Sugawara, K. Omura, M. Emoto, and Y. Nojiri, "Temporal sampling parameters and motion portrayal of television," *SID Symposium Digest of Technical Papers*, vol. 40, no. 1, pp. 1200–1203, 2009.

[11] L. Wilcox, R. Allison, J. Helliker, B. Dunk, and R. Anthony, "Evidence that viewers prefer higher frame-rate film," *ACM Transactions on Applied Perception*, vol. 12, no. 4, pp. 15, 2015.

[12] S. Hammett, M. Georgeson, S. Bedingham, and G. Barbieri-Hesse, "Motion sharpening and contrast: Gain control precedes compressive non-linearity?," *Vision Research*, vol. 43, no. 10, pp. 1187 – 1199, 2003.

[13] P. Barten, *Contrast sensitivity of the human eye and its effects on image quality*, SPIE press, 1999.

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

[15] M. Emoto and M. Sugawara, "Critical fusion frequency for bright and wide field-of-view image display," *Journal of Display Technology*, vol. 8, no. 7, pp. 424–429, 2012.

[16] P. Bex, G. Edgar, and A. Smith, "Multiple images appear when motion energy detection fails," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 21, pp. 231–238, 1995.

[17] J. Farrell, M. Pavel, and G. Sperling, "The visible persistence of stimuli in stroboscopic motion," *Vision Research*, vol. 30, no. 6, pp. 921–936, 1990.

[18] S. Daly, N. Xu, J. Crenshaw, and V. Zunjarrao, "A psychophysical study exploring judder using fundamental signals and complex imagery," in *SMPTE Conferences*. Society of Motion Picture and Television Engineers, 2014, vol. 2014, pp. 1–14.

[19] D. Burr, J. Ross, and M. Morrone, "Smooth and sampled motion," *Vision Research*, vol. 26, no. 4, pp. 643–652, 1986.

[20] C. Poynton, "Motion portrayal, eye tracking, and emerging display technology," in *SMPTE Conferences*. Society of Motion Picture and Television Engineers, 1996, vol. 1996, pp. 192–202.

[21] J. Larimer, J. Gille, and J. Wong, "Judder-induced edge flicker in moving objects," in *SID Symposium Digest of Technical Papers*. Wiley Online Library, 2001, vol. 32, pp. 1094–1097.

[22] S. Daly, "Engineering observations from spatiovelocity and spatiotemporal visual models," in *Photonics West '98 Electronic Imaging*. International Society for Optics and Photonics, 1998, pp. 180–191.

[23] D. Hoffman, V. Karasev, and M. Banks, "Temporal presentation protocols in stereoscopic displays: Flicker visibility, perceived motion, and perceived depth," *Journal of the Society for Information Display*, vol. 19, no. 3, pp. 271, 2011.

[24] A. Michelson, *Studies in optics*, Courier Corporation, 1995.

[25] C. Binnie, J. Emmett, P. Gardiner, A. Wilkins, G. Harding, and D. Harrison, "Characterizing the patterned images that precipitate seizures," *SMPTE*, vol. 111, pp. 32–39, 2002.

[26] ITU-R Rec, "BT. 500-13, Methodology for the subjective assessment of the quality of television pictures," *International Telecommunications Union. Geneva*, 2012.

[27] J. Drewery and R. Salmon, "Tests of visual acuity to determine the resolution required of a television transmission system," *BBC Research & Development White Paper 92*, 2004.

[28] T. Scherzer and V. Ekroll, "Occlusion improves the interpolation of sampled motion," *Vision research*, vol. 62, pp. 17–25, 2012.

[29] A. Michotte, G. Thines, and G. Crabbé, "Amodal completion of perceptual structures," *Michotte's experimental phenomenology of perception*, pp. 140–167, 1991.

[30] S. Hammett, M. Georgeson, and A. Gorea, "Motion blur and motion sharpening: temporal smear and local contrast nonlinearity," *Vision research*, vol. 38, no. 14, pp. 2099–2108, 1998.

[31] M. Pindoria, "Higher frame rate television for future broadcasts at the Commonwealth Games," BBC Research & Development Blog, August 2015.

[32] K. Noland and L. Truong, "A survey of UK television viewing conditions," *BBC Research & Development White Paper 287*, 2015.