https://doi.org/10.2352/ISSN.2470-1173.2020.11.HVEI-234

Psychophysics Study on LED Flicker Artefacts for Automotive Digital Mirror Replacement Systems

Nicolai Behmann, Holger Blume; Institute of Microelectronic Systems, Leibniz University Hannover; Hannover, Germany

Abstract

LED flicker artefacts, caused by unsynchronized irradiation from a pulse-width modulated LED light source captured by a digital camera sensor with discrete exposure times, place new requirements for both visual and machine vision systems. While latter need to capture relevant information from the light source only in a limited number of frames (e.g. a flickering traffic light), human vision is sensitive to illumination modulation in viewing applications, e.g. digital mirror replacement systems.

In order to quantify flicker in viewing applications with KPIs related to human vision, we present a novel approach and results of a psychophysics study on the effect of LED flicker artefacts. Diverse real-world driving sequences have been captured with both mirror replacement cameras and a front viewing camera and potential flicker light sources have been masked manually. Synthetic flicker with adjustable parameters is then overlaid on these areas and the flickering sequences are presented to test persons in a driving environment. Feedback from the testers on flicker perception in different viewing areas, sizes and frequencies are collected and evaluated.

Introduction

Conventional side and rearview mirrors in vehicles will in future be replaced by smart camera-display combinations. Benefits include an increased viewing experience through object detection overlays and the same view from all viewing angles, as well as a reduced wind resistance. However, illumination modulated light sources cause disturbing artefacts in the reproduction of the scene. Machine vision algorithms are effected in terms of non-captured traffic lights or variable speed limit signs. However, while flicker mitigation systems in this case need to ensure the perception of at least one frame with the information encoded in the light, human vision is sensitive to illumination modulation.

In order to quantify the effect of flicker on visual perception from humans, KPIs need to be standardized and correlated to psychophysics studies. Those are necessary for the government to set margins for legal admission of mirror replacement systems and for OEMs to compare different flicker mitigation systems. Moreover, these KPIs are needed to evaluate the effect of flicker mitigation algorithms, e.g. presented by the authors in [1].

In this paper, we present a novel psychophysics study on the human visual perception of flicker in front- and side mirror applications. Therefore, real-world driving sequences have been captured in diverse lighting and environmental conditions. Manually, these sequences have been segmented in potentially flickering pixels on different light sources (e.g. traffic light, advertisement lights and daytime headlights). In a static in-cabin car demonstrator, test persons are presented to those real-world sequences, overlaid with synthetic flicker artefacts (varying flicker frequency, modulation and size). The perception of flicker and subjective impression on the disturbance of different flicker settings on different screens is recorded.

The remainder of the paper is structured as follows: First, previous work from literature related to psychophysics studies on flicker are presented. Subsequently, our capture and testing setup is described in details. The results from the studies are presented and discussed in the following chapter, before concluding this paper.

Related Work

The IEEE P2020 Working Group on Automotive Image Quality [2] implements a subgroup working on KPIs for LED flicker artefacts in both visual and machine vision use cases. First KPIs for flicker detection and modulation amplitude have been drafted. Additionally, a first psychophysics study on area flicker is presented and at the moment of writing in progress. In their study, testers are presented three different flickering videos with varying flicker edge sharpness, frequency and contrast. Testers are then requested to rate the video-to-be-assessed in regards to the two other fixed reference flickering videos. However, this study lacks practical relevance for most flickering light sources in real-world driving situations. The investigated area flickering mostly originates from modulated street lamps illuminating the scene. But today, the majority of flicker artefacts arise out of local, direct pointing LED lights (daytime head- and taillights, fuel station price tags, marketing banners).

First studies on flicker have been performed by Brown [3] and Kelly [4] 50 years ago regarding flicker in cinematic applications. Using a sinusoid light source, the critical flicker perception frequency is evaluated for large area illumination. Later, these studies have been extended from the temporal-frequency domain to the spatial-frequency domain. While these studies precisely examine the flicker sensitivity of the human visual cortex for different temporal and spatial frequencies and intensities, as well as flicker locations in the fovea area, their studies exclude local point light sources that dominate the LED flicker light sources nowadays.

The methods for the first assessment of our psychophysics study base on methods proposed by Ives [5] and Kelly [6]. Starting with a high frequency flicker (above 70 Hz of the 1. order wave) [7], which is not perceived by the human as flicker, the critical flicker detection frequency is found by lowering the frequency on different illumination levels [8]. The following evaluation of real world driving sequences with overlaid artificial flicker base on recommendation of the ITU [9] for procedures and environmental conditions for the subjective assessment of the quality of television images. In this case, the ability of the human visual perception to retain quality under non-optimum conditions that

Figure 1: Side mirror camera setup.

relate to transmission (Impairment Assessments) are conducted.

LED Flicker Dataset Acquisition

The psychophysics study on human visual perception of flicker artefacts for automotive viewing applications requires the acquisition of a large scale dataset of real-world driving sequences. Additionally, the evaluation has to be reproducible in the lab for all test persons. Therefore, no real driving is conducted, but sequences are captured and prepared before presenting to the testers audience.

Camera Setup

In order to properly simulate the driving of a car, one front and two side mirror cameras are used for the acquisition and will be reproduced on a separate display during the study. The cameras use a 2 MP $[1920 \times 1080]$ global shutter CMOS sensor (1 / 1,23) with HDR (72 dB) from cinematic cameras. Pixel size is 5,86 m. Exposure time is maximized by the auto exposure algorithm, in order to mitigate flicker artefacts while capturing. However, exposure time is manually clipped for high speed scenes to minimize motion blur artefacts. The lens has a horizontal field of view (FOV) of 89. The front camera behind the windshield captures the scene in front of the car. All cameras are synchronized and capture at a fixed framerate of 25 Hz. The camera and recording setup is presented in Figure 1.

Scenes and Environments

The dataset is requested to represent a majority stake of common situations. The captured driving maneuvers are listed in the Table 1. While there is no natural flickering in the windshield, flickering is also overlaid in the central front screen. In this way the human visual perception of flickering in this area, e.g. for future back mirror replacement systems, is assessed. Driving maneuvers in the psychophysics flicker study dataset. Flickering objects are separated in front screen (F) and left (L) and right (R) digital side mirror.

Additional attention was paid to other road users with potentially flickering lights. This includes following cars with flickering light in the same lane, overtaking cars and cars reappearing in the mirror after a sharp turn. LED-based head- and taillights were captured with both the front and side-mirror cameras. Beside vehicle mounted light sources, static LED lights from two different fuel station price tags (white and red light), as well as different advertisement banners were recorded.

Table 1: Sequences used for the study. Size of the flickering area is indicated with small (S), medium (M), large (L).

ID	Flickering Objects	Environm.
A	F: variable speed limit sign (S-M) R: - L: headlights, overtaking car (M-L)	Highway Night Rain
C	F: 2x digital advertisement (L) R: digital advertisement (L-M) L: -	Urban Night Rain
D	F: - R: headlights of following car (S) L: headlights of following car (S)	Urban Daytime Sunny
E	F: - R: following truck, lane change (S) L: following truck, lane change (S)	Highway Daytime Sunny
F	F: traffic lights (M-S) R: - L: -	Urban Daytime Sunny
X	F: - R: - L: fuel station price tag (L)	Urban Daytime Sunny
Y	F: digital advertisement (M) R: 3x digital advertisement (M-L) L: -	Highway Night Rain

(a) Crop from recorded sequence. (b) Generated Mask (green).

Figure 2: Source image and light source mask for urban single lane sequence.

Light Source Masking

In order to generate flicker artefacts caused by controllable parameters, potentially flickering lights are masked in the recorded scenes on pixel level and can later be overlaid with synthetic flicker. Therefore, an image-processing based algorithm for assisted labeling has been developed. First, a human annotator identifies potentially flickering light sources in the first frame of the sequence and uses as watershed transformation based selection tool to mask the light source itself and reflections, if available. Each light source is assigned to a unique id, which allows the simulation of different flicker patterns on each light source. Subsequently, a dense optical flow in combination with a watershed transformation is used to warp the mask to subsequent frames. The annotator verifies the flicker mask for each frame.

Psychophysics Study

To the best knowledge of the authors, there has been no study on human visual perception of digital mirrors yet. However, in the ITU-R BT.500-14 recommendations [9], procedures and envi-

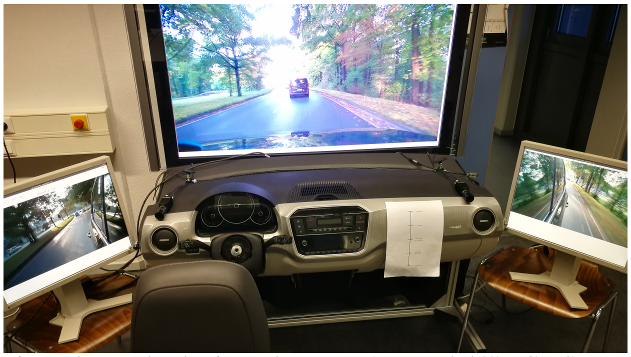


Figure 3: Psychophysic study setup. A central 55" LCD Screen replays the central camera, while the left and right 21" monitor act as the digital side mirror replacement system in the open car environment. The evaluation scale is printed on a paper.

ronmental conditions for the subjective assessment of the quality of television images are defined. Therefore, this psychophysics study is closely conducted according to these recommendations. A test group of 20 persons, of which 15 are male and 5 are female, have performed the study. All test persons have a driving license, five of them wear glasses.

Study Setup

An in-car atmosphere is generated with a real dashboard, on which three monitors are mounted.

A 55" central monitor ($[1920 \times 1080]$, $500 cd/m^2$, $178°$ FOV (H,V)) replays the video, acquired with the front view camera. Two smaller 21" monitors ($[1600 \times 1200]@60Hz$, $300 cd/m^2$, $178°$ FOV (H,V)) on the left and right of the test person are used to simulate the digital side mirror replacement system. These monitors are mounted on the height of the dashboard. The environmental light in the room is controlled and fixed for all studies.

Test Person Task

At the beginning of the psychophysics study, the test person is introduced to the testing environment, study purpose and duration. Subsequently, four different sequences of flickering are presented to the test person (acquired with different rolling shutter cameras in different scenarios). Additionally, one video with synthetic lights are used to highlight the difference between flicker and blinking. The testing part of the psychophysics study is then subdivided in three tasks, which are described in detail below. The latter two tests use the same rating scale for subjective perception:

5 imperceptible
4 perceptible, but not annoying
3 slightly annoying
2 annoying
1 very annoying

1. Critical flicker frequency assessment

First, the test person is requested to watch a real-world headlight from an Audi A4 (2008). The headlight is places at a distance of 2m to the viewer and at the same height to insure direct visibility. The daytime LED lights within the headlight system are connected to a custom-made controllable LED driver, which allows the setting of pulse width modulation frequency and duty cycle. Starting from a high PWM frequency (240 Hz), the frequency is lowered, till the test person perceives flicker without a display. The test is conducted for direct and peripheral view (with a fixed spot left to the headlamp) and 5% and 50% duty cycle.

2. Subjective Grading of Live Flicker

Subsequently, the same headlight is perceived through the left monitor of the car environment. The headlight is captured with the same global shutter camera, used for the video acquisition, but at 60 frames per second. The following four different pulse width modulation parameters of the LEDs are then evaluated by test person according to the grading scale.

Table 2: Flicker settings of the headlights.

ID	PWM Frequency	PWM Duty Cycle	Flicker
I	$120Hz$	5%	-
II	$76Hz$	5%	Fast
III	$65Hz$	5%	Slow
IV	$65Hz$	50%	Slow

3. Subjective Grading of Driving Sequences

In order to evaluate the effects of flicker on different viewing areas (central or peripheral view), the driver is requested to focus on one monitor. For one video sequence and flicker setting, central, right and left perception are graded successively. However, while looking on one screen, noticeable flickering from a neighboring screen is considered for the grade of the screen under evaluation. The video is repeated for each screen.

Given the original real-world driving sequence and pixel-wise annotated light source mask, synthetic flicker is rendered on the original sequence according to a fixed flicker sequence or a physical model. Latter allows to set the frequency and duty cycle of a pulse-width modulated light, as well as the exposure time and capture frequency of the camera. Additional rolling shutter artefacts can be simulated relating to the adjustable rolling shutter speed. These parameters allow the simulation of a wide range of flicker artefacts, including low frequency amplitude variation close to beat frequencies and partially visible light sources due to rolling shutter. However, for the user study, the eight fixed flicker patterns in Table 3 (global shutter) have been selected to represent a wide range of different influences on the human visual perception.

Table 3: Flicker sequences of the masked regions.

ID	Sequence	Frequency	Amplitude
0	1.0	25.0 Hz	0.00
1	1.0, 0.5	12.5 Hz	0.25
2	1.0, 0.0	12.5 Hz	0.50
3	1.0, 0.9	12.5 Hz	0.05
4	0.5, 0.1	12.5 Hz	0.20
5	1.0, 0.9, …, 0.0 …, 0.9	1.25 Hz	0.50
6	1.0, 0.9, …, 0.5 …, 0.9	2.5 Hz	0.25
7	random	div	0.5

Figure 4: Critical LED pulse-width modulation frequency and duty cycle for flicker perception.

Results and Evaluation

This chapter is structured according to the three different tasks of the task person.

1. Critical flicker frequency assessment

In Figure 4 a box plot of the critical pulse width modulation frequency, below which the test person perceives flicker, is illustrated. For both LED illumination intensities, flicker is perceived more widely perceived from the peripheral view. The mean critical flicker frequency increases from $55.3Hz$ to $59.5Hz$ for 5% duty cycle, and from $58.3Hz$ to $66.4Hz$ for 50% duty cycle respectively. Flicker is perceived $3Hz$ or $7Hz$ earlier for brighter LEDs (50% duty cycle) in the central and peripheral view in comparison to the 5% duty cycle. Additionally, a lower critical flicker frequency is observed for older test persons.

2. Subjective Grading of Live Flicker

Beside the introduction of the test person to the monitor setup and grading scale, the flicker introduced to the side mirrors by the LED headlight is subjectively evaluated. Across all assessments from 20 test persons and four flicker settings each, an overall mean of grade 3 is reached.

According to Table 2, the non-flickering sequence I is rated with 5 by all candidates and proves that there are no other flicker sources in the image. Higher frequency flicker (sequence II) has the highest mean grade of 2.5. Comparing sequence III and IV at a fixed pulse width modulation frequency of $65Hz$ and consequently resulting in lower frequency flicker in the side mirror, the brighter LEDs (50% duty cycle) are again more disturbing (mean grade of 2.35 to 2.15). However, for flicker sequence III and IV, test persons report blinking or headlight horn behavior, which can lead to potentially dangerous driving decisions in road traffic.

3. Subjective Grading of Driving Sequences

In this study, video sequences with overlaid flicker sequences are evaluated in a randomized order for each test candidate. It is ensured that the same video is not evaluated several times in a row. The first six evaluations per screen are not used for evaluation, as the driver calibrates itself to the grading scale.

These subjective assessments show an overall mean grading of 3.5 and 3.3 without the original videos, which proves a good selection of various data points for the study. In one video sequence, flicker from the tail lights of a truck in front has already

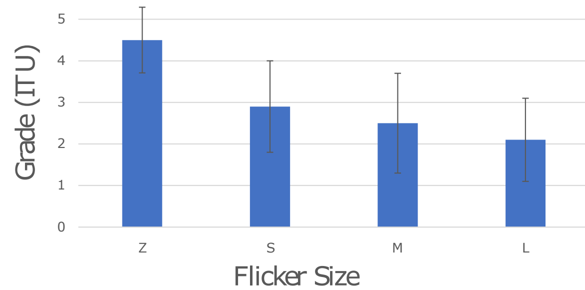


Figure 5: Mean subjective grade and standard deviation for real-world flicker sequences for different affected light source sizes: small (S), medium (M), large (L) and zero flicker (Z).

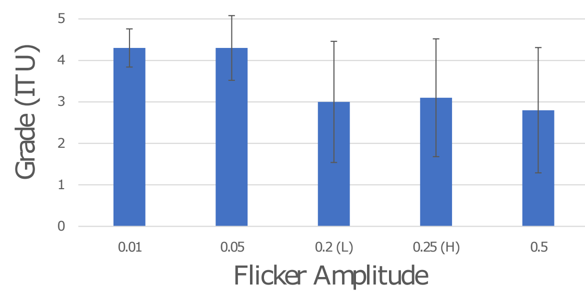


Figure 6: Mean subjective grade and standard deviation for real-world flicker sequences for different flicker amplitudes at 12.5 Hz. In this case, (L) is related to low intensity flickering (sequence 4), wheres (H) relates to flicker sequence 1.

been recorded by the front camera and results in a mean grade of 3.9 for the playback of the original video on that screen. The remaining videos show a mean grade of 4.9 for the original playback and thereby prove the quality of the source videos.

During the studies, the highest sensitivity can be seen for the larger front screen with an overall mean grade of 3.2, whereas the average grade for the left and right mirror display are 3.6 and 3.7.

In Figure 5 the mean subjective grade and standard deviation of all sequences structured for different light source sizes are depicted. For zero flickering (Z), a mean grade of 4.5 is reached. With increasing size of the flickering region, the mean grade drops from 2.9 for small (S) to 2.1 for large (L) flickering regions. However, already a small flickering area is perceptible to the test persons.

For all flicker sequences of 12.5 Hz, the mean grade and standard deviation for different flicker amplitudes is depicted in Figure 6. Up to a certain flicker amplitude threshold (≈ 0.1), no flicker is perceived by the test persons. Flicker with larger amplitudes, however, is roughly equally annoying (mean grade 3). Additionally, it can be perceived, that the flicker mean value has no influence on human perception in contrast to the flicker amplitude.

Conclusion

In this paper we conducted a novel psychophysics study on the humanoid visual perception on flicker artefacts caused by pulse width modulated LED light sources captured with digital image sensors, with a focus on automotive mirror replacement systems. In a first part, the critical flicker detection frequency is searched by lowering the PWM frequency of real daytime headlights and found to be higher for brighter, as well as for peripherally perceived lights. In the second part, different flicker situations have been found to distract the driver by the mistaken perception of a direction indicator or headlight horn. Through the subjective assessment of real-world driving sequences with synthetically overlaid flicker, correlations between different flicker sizes and amplitudes have been derived. The disturbance level increases with larger areas. However, flicker perception is not linearly dependent on the flicker amplitude, as all flicker above a certain amplitude is equally disturbant.

References

[1] N. Behmann, G. Schewior, S. Hesselbarth, and H. Blume, "Selective LED Flicker Detection and Mitigation Algorithm for Non-HDR Video Sequences," in *IEEE Intl. Conf. on Consumer Electronics*, vol. 8, 2018.

[2] IEEE P2020 Working Group, "IEEE P2020 Automotive Imaging: White Paper," in *IEEE P2020 Automotive Imaging*. IEEE-SA, 2018. [Online]. Available: https://www.image-engineering.de/content/library/white_paper/P2020_white_paper.pdf

[3] J. L. Brown, "Flicker and intermittent stimulation," *Vision and visual perception*, vol. 1, pp. 251–320, 1965.

[4] D. H. Kelly, *Flicker*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1972, pp. 273–302. [Online]. Available: https://doi.org/10.1007/978-3-642-88658-4_11

[5] H. E. Ives, "Critical frequency relations in scotopic vision," *JOSA*, vol. 6, no. 3, pp. 254–268, 1922.

[6] D. Kelly, "Flicker fusion and harmonic analysis," *JOSA*, vol. 51, no. 8, pp. 917–918, 1961.

[7] H. De Lange Dzn, "Research into the dynamic nature of the human fovea - cortex systems with intermittent and modulated light. i. attenuation characteristics with white and colored light," *Josa*, vol. 48, no. 11, pp. 777–784, 1958.

[8] G. Hauske, *Systemtheorie der visuellen Wahrnehmung*. Springer-Verlag, 2013.

[9] R. I.-R. BT, "Methodology for the subjective assessment of the quality of television pictures," 2019.

Author Biography

Nicolai Behmann received his B.Sc. (2014) and M.Sc. (2015) in electrical engineering from the Leibniz University of Hanover. Since then he has worked in the Research on algorithms and architectures for camera-based advanced driver assistance systems at the Institute of Microelectronic Systems, Leibniz University Hanover.

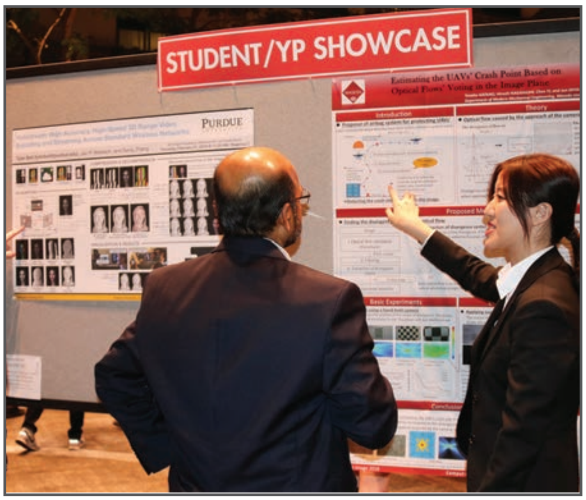
JOIN US AT THE NEXT EI!

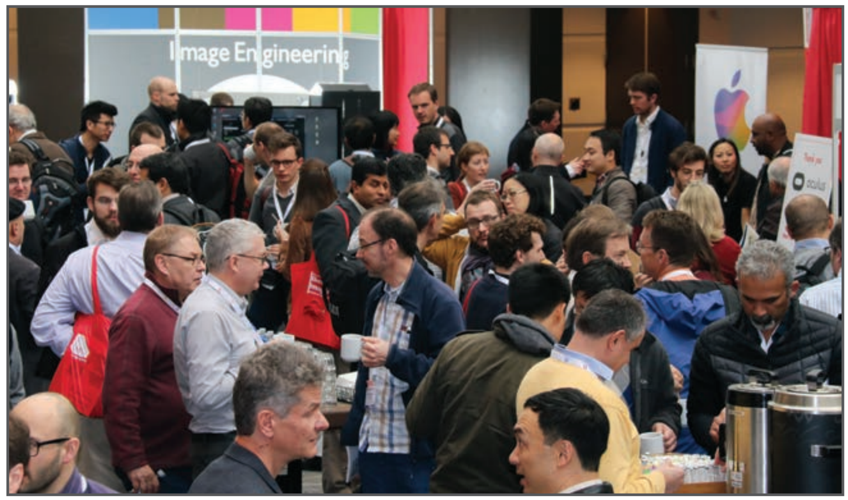
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org