DTU Library

Subjective quality of video sequences rendered on LCD with local backlight dimming at different lighting conditions

Mantel, Claire; Korhonen, Jari; Pedersen, Jesper Mørkhøj; Bech, Søren; Andersen, Jakob Dahl; Forchhammer, Søren

[Link back to DTU Orbit](#)

Subjective quality of video sequences rendered on LCD with local backlight dimming at different lighting conditions

Claire Mantel[a], Jari Korhonen[a], Jesper Melgaard Pedersen[b], Søren Bech[b,c], Jakob Dahl Andersen [a], Søren Forchhammer[a]

[a]Department of Photonics Engineering, Technical University of Denmark, 2800 Kongens Lyngby, Denmark
[b]Bang & Olufsen A/S, Peter Bangs Vej 15, 7600 Struer, Denmark
[c]Section of Signal and Information Processing, Department of Electronic systems, Aalborg University, 9100 Aalborg, Denmark

ABSTRACT

This paper focuses on the influence of ambient light on the perceived quality of videos displayed on Liquid Crystal Display (LCD) with local backlight dimming. A subjective test assessing the quality of videos with two backlight dimming methods and three lighting conditions, i.e. no light, low light level (5 lux) and higher light level (60 lux) was organized to collect subjective data. Results show that participants prefer the method exploiting local dimming possibilities to the conventional full backlight but that this preference varies depending on the ambient light level. The clear preference for one method at the low light conditions decreases at the high ambient light, confirming that the ambient light significantly attenuates the perception of the leakage defect (light leaking through dark pixels). Results are also highly dependent on the content of the sequence, which can modulate the effect of the ambient light from having an important influence on the quality grades to no influence at all.

1. INTRODUCTION

In a Liquid Crystal Display (LCD), the image is rendered by the combination of the light emitted by the backlight unit (nowadays usually composed of Light Emitting Diodes (LEDs)) and the Liquid Crystal cell (LC) grid that forms the image. Local backlight dimming is a technology that aims at both saving energy and improving the quality of images rendered on LCDs. The backlight of a display with local backlight dimming capability is composed of various segments which intensity can be set independently from one another. Local backlight dimming consists in reducing the luminance of the display backlight in areas where the local image content is dark and does not require full intensity.

Two types of defects can emerge from the intensity of the backlight: *leakage*, when the LCs fail to block light completely, producing grayish black pixels; and *clipping*, when not enough light is provided to the LC cells to reach the intended luminance. This paper investigates the role of the ambient light level in a viewing room on the perception of videos rendered with local backlight dimming. As leakage can appear only in the dark areas, the increase of the minimum black level due to reflections and the decrease in contrast due to ambient light adaptation particularly influence the perception.

The term ambient light usually encompasses both the light from the room reflected on the screen and the light included in the field of view besides that of the screen (called surrounding light)[1,2] . Although it is possible to model the ambient light in a perceptually accurate way,[3] characterizing it[4] and computing it is still quite complex and subjective testing are mostly set up in actual light laboratories (by opposition to modeled environments).

The influence of ambient light on watching images on a screen can be approached in two different ways: it can influence the visual performance[5] and it can influence the visual preference[6–8] . Lin and Huang showed for example that an ambient light of 500 lux worked best, among the 250-2000 lux range they tried, for a performance task (character identification and text comprehension). Hanhart et al.[9] have shown in a side-by-side test of HDR

Further author information: (Send correspondence to C.M.)
C.M.: E-mail: clma@fotonik.dtu.dk

displays using peak whites of 100-4000 cd/m2 that the brighter of the two display settings was preferred. This paper focuses on the preference aspect. For preference, previous studies report little effect of the ambient light on the preference of the image[6] or of the display brightness[6,7].

Liu and Fairchild have investigated[8] the relationship between surround illumination and preferred contrast. They show that, as also reported in previous studies, observers preferred to increase the contrast as ambient light decreased. However they show that the results highly depend on the content, i.e. the complexity and dynamic range of the image, and the viewers, i.e. the reported tendency is unclear for non-experts viewers.

In this paper the relationship between the ambient light level and the preferred quality of videos displayed with local backlight dimming is investigated. Local backlight dimming modifies the contrast of a rendered image by influencing both the minimum black level and maximum values. Indeed, leakage amounts to locally increasing the minimum black level and clipping to reducing the brightness. The idea is here to evaluate whether results obtained in prior literature for images displayed without any processings[6] or by modifying the gamma[8] are transposable to sequences displayed with local backlight dimming.

The paper is organized as follows: firstly the subjective test is detailed in Sec. 2 then the analysis of the results is presented in Sec. 3.

2. EXPERIMENT

We set-up a subjective test in which participants rated the quality of video sequences presented on an LCD platform with a controllable backlight at various ambient light levels.

Two local backlight dimming methods were applied: the first method represents a conventional LCD (Full backlight), i.e. all LED segments are set uniformly to their full intensity. The second method (the Gradient Descent algorithm[10,11]), aims at achieving best quality by adapting the backlight to the content according to a display model. Using a gradient descent algorithm, it reaches the best possible quality as evaluated by the display model in terms of MSE. As varying the intensity of a LED can produce a flashing defect called *flicker*, the Gradient Descent algorithm includes a flicker control mechanism to ensure that no such defect appears by keeping the LED variations below a chosen threshold. The Gradient Descent algorithm does not contain any spatio-temporal filtering therefore its spatio-temporal variations can be seen as the major difference between the two methods.

Sequences were shown in a room with dark walls at three different ambient light conditions: no light (approximately 0 lux), low ambient light (approximately 5 lux, 2400K) and higher ambient light (60 lux, 2800K). The three ambient light levels were obtained by varying the power supplied to three halogen lamps located symmetrically on each side of the display. This light was diffused with translucent paper. The ambient light values were measured on the display. Participants were given time to adapt to each light condition before performing the required task.

For the test, five sequences all containing dark areas that could show leakage were used. The selected sequences (duration 5-8s) are: *'Stars'*,[12] *'Titles'*,[12] *'Volcano'*,[13] *'TchDwn'*[13] and *'Uboat'*; frames of the first four are depicted in Fig. 1, whereas the fifth cannot be shown due to copyright reasons. All sequences have Full HD (1920x1080) resolution and are available in original or extremely good quality (i.e. blu-ray).

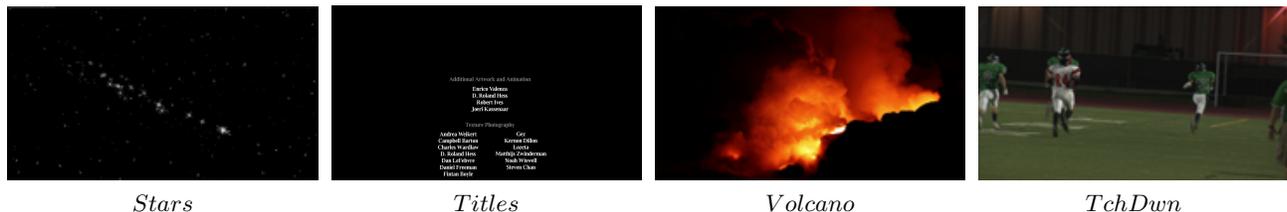| *Stars* | *Titles* | *Volcano* | *TchDwn* |

Figure 1: A frame extracted from three of the four sequences used in the experiment

A total of 10 stimuli were shown at each ambient light level. The participants were instructed to evaluate the quality of each stimuli by placing a cursor on a continuous scale without label. Each participant repeated the

experiment twice. The method order was randomized and the ambient light level and sequence order followed Latin Square designs. The ambient light level was a block variable, meaning that the participants were rating all stimuli at one lighting condition at one go, then they would get adapted to the next lighting condition and rate all the stimuli again and so on. Therefore the results from the experiment do not consist in a direct comparison of the quality at different ambient light levels. As the instructions to the participants did not mention this comparison either, their grades are not calibrated across the different ambient light levels.

Participants were located at a distance of three times the display height and the display used is a 46" LCD which was rotated by 15 degrees. The peak white of the display was kept at the constant value of 490 Cd/m^2 for all ambient light levels. For each method and each sequence, the LED intensities and corresponding LC values were computed off line prior to the experiment and then displayed using a special playback mode that allows to control them on the platform.

Twenty subjects participated in the experiment, all naive regarding its purpose but as they are all employees of Bang & Olufsen they can be considered as experienced in video processing.

The settings of the experiment were chosen in order to emphasize leakage perception: using dark contrasted videos, using methods with very different leakage levels (they were rated as the most and less contrasted method in a previous experiment[14]), rotating the display by 15 degrees and using a high peak white for the display.

3. RESULTS ANALYSIS

As stated in the description of the experiment, the observers did not directly compare the stimuli at the various ambient lights (which was impossible due to the required adaptation delay). Therefore the grades from one ambient light level to another are not really calibrated on a similar scale but the analysis will focus mainly on the differences between the two methods.

An ANalysis Of VAriance (ANOVA) of the data was applied with fixed factors *Method*, *Sequence* and *Ambient light* and random factors *Repetition* and *Subjects*. The results, available in table 1, show that the used dimming method (Method) and content (Sequence) have a significant effect on the ratings. The most influential significant factor is the interaction of Sequence and Method as illustrated in Fig. 2. The Gradient Descent algorithm is significantly ($p<0.05$) preferred to full backlight under all light conditions, as can be seen in Fig. 3. The ambient light (Light) does not have a significant influence alone, but the interaction between Light and Sequence, as well as Light, Sequence and Method, are significant.

Factor	p-value	F
Method	0.018	7.2
Sequence	0.003	4.1
Subject	0.003	2.9
Method $\times$ *Sequence*	0.000	53.2
Method $\times$ *Subject*	0.000	9.4
Sequence $\times$ *Subject*	0.000	4.2
Ambient light $\times$ *Sequence*	0.025	4.1
Method $\times$ *Sequence* $\times$ *Ambient light*	0.000	4.7

Table 1: Significant factors for the ANOVA of the quality grades

Post hoc tests (Tukey $p < 0.05$) show that there is a significant difference between the ratings attributed at the two lower ambient light levels and those attributed at the higher ambient light level. On the other hand, the two lower light levels are not statistically different.

The ambient light by itself is not a significant factor but its interaction with the sequence is significant, as illustrated in Fig. 4. However this figure also shows that when averaging over the two methods, the effect of ambient light is quite light and mainly depends on two of the sequences used: *Titles* and *Volcano*.
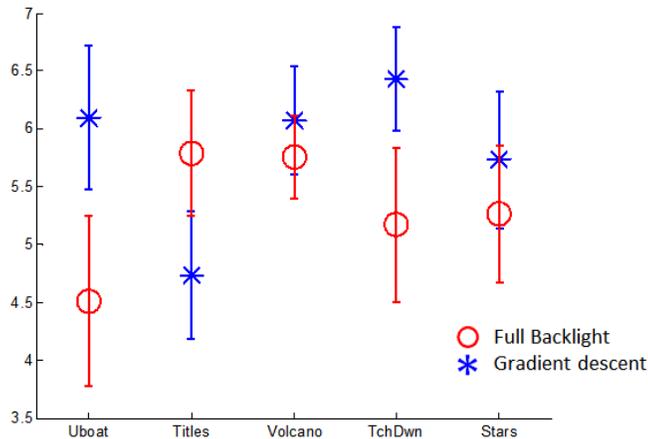
Figure 2: Mean Opinion Score as a function of the sequence for each method (averaged over all ambient light levels)
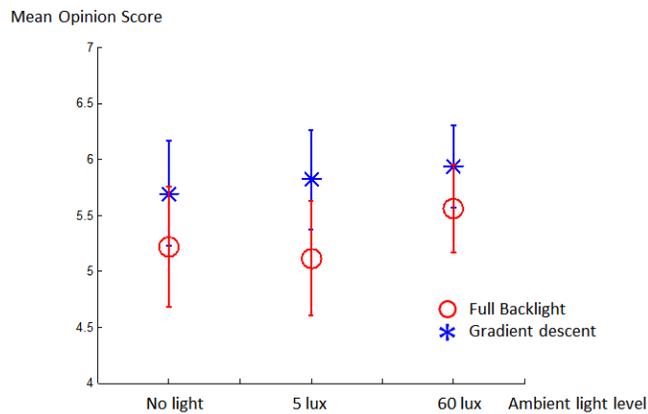

Figure 3: Mean Opinion Score as a function of the ambient light level for each method

Our work hypothesis was that the subjective quality differences between full backlight and Gradient Descent algorithms would diminish when the ambient light level becomes higher. The results support this hypothesis as the differences between the two methods tested are less visible (i.e. the grades obtained are closer) at the higher ambient light. This is shown in Fig. 5.

The aim of the Gradient Descent algorithm is to achieve the lowest possible error as evaluated by the display model while at the same time saving energy. The main difference between the two methods is that the Gradient Descent algorithm shows globally less leakage and varies over time as opposed to the Full backlight. It can be measured through the normalized energy consumption and the error on leakage of each local backlight dimming, as presented in Table 2. Therefore change in the method ratings implies that the perception of leakage is less significant at higher ambient light.

As stated earlier, the subjective preferences are highly dependent on the different contents. Indeed, depending on the sequence, the ambient light levels can cause differences in the subjective grades ranging from more than 15% of the whole scale to no difference at all. Therefore we performed t-test on each sequence at each ambient light separately to investigate the details of the third order interaction $Method \times Sequence \times Ambient\ light$ (as this test has only 20 observations, the power is smaller than that of the global ANOVA). The results, visible in Table 3, show that for all sequences the obtained p value is higher at the low ambient light levels. Only for the sequence $TchDwn$ the difference of ratings between the two methods is statistically significant at the high ambient light level (60 lux).

The content dependency can be explained by different temporal and spatial characteristics affecting the visibility of leakage, i.e. for some contents the leakage and its temporal variations have higher impact on
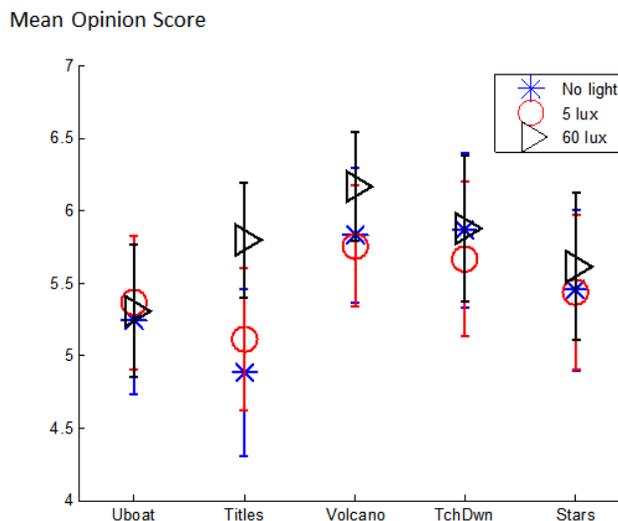
Figure 4: Mean Opinion Score as a function of the sequence for each ambient light level
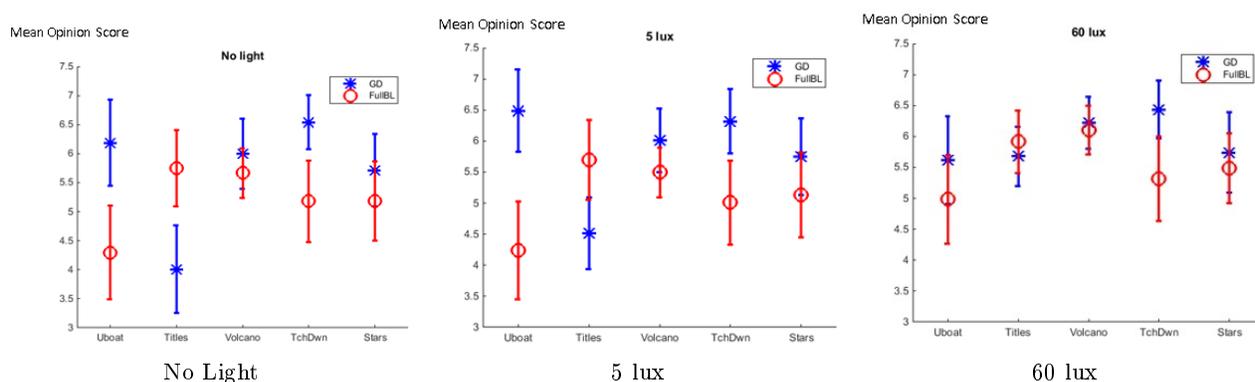

| No Light | 5 lux | 60 lux |

Figure 5: Mean Opinion Score as a function of the sequence for the two methods

perceived quality than for other contents. The two sequences for which the impact of the ambient light is the most important, i.e. *Uboat* and *Titles* are the ones presenting the most spatio-temporal variation of the backlight.

The second most influential significant effect, as can be seen from the F value in Table 1, is the interaction between the Subject and the Method. In practice it means that the grades show a 'bimodal' distribution relatively to the Method preferred. Among the 20 participants, 6 test subjects tend to prefer full backlight even in cases when the majority (14) of test subjects show a clear preference towards the Gradient Descent algorithm. We have therefore separated the observers in two groups and redone the ANOVA on each of them but the significant factors remain the same although the F values of the factors are modified.

Different individual preferences can be explained by personal tendencies of paying major attention on different aspects in visual content.

4. OBJECTIVE METRICS PERFORMANCE

To the best of our knowledge, no published metric takes the ambient light into account while evaluating the quality of a stimuli. In order to apply objective metrics to the data, the ambient light level needed to be included in the modeled stimuli.

To do so two different methods were selected. The light coming into the eye of the observer from any part of the display is composed of the light emitted by the display plus the light reflected by the display. Therefore

Sequence	Power consumtion		MSE leakage	
	Gradient Descent	Full Backlight	Gradient Descent	Full Backlight
Uboat	0.045	1	1.2E-05	0.000311
Titles	0.436	1	0.000186	0.000424
Volcano	0.591	1	0.000128	0.000237
TchDwn	0.409	1	3.8E-05	0.000112
Stars	0.300	1	0.000131	0.000418

Table 2: Power consumption (normalized relatively to the maximum power consumption) and MSE of the leakage for each sequence.

Sequence	Ambient Light		
	No light	5 lux	60 lux
Uboat	0.004	0.001	0.269
Titles	0.0006	0.005	0.446
Volcano	0.202	0.021	0.477
TchDwn	0.0001	0.0001	0.002
Stars	0.133	0.121	0.488

Table 3: Significance value p of Student's t-test of equality between the two methods for each Sequence at each Ambient light level (degree of freedom = 19)

the first method consists in adding the reflected ambient light to the emitted light in the physical domain. The modeled luminance y_i^{light} for pixel i is then:

$$y_i^{light} = y_i + \frac{k}{\pi} E_{amb}, \tag{1}$$

where y_i is the modeled rendered luminance, k is the reflection coefficient of the display (3% in this case) and E_{amb} is the illuminance in *lux*. Then the second chosen approach is the one developped by Devlin et al. in[15] where they use the following transform to account for luminance modification due to ambient light:

$$y_i^{light} = \begin{cases} \frac{y_i L^2}{(L-L_R)^2 + y_i L_R} \\ \frac{y_i (m-L)^2 + m L_R (y_i + m + L_R - 2L)}{(m-L)^2 + L_R (y_i + m + L_R - 2L)} \end{cases}, \tag{2}$$

where L_R is the reflected light, m is the maximum value taken by the image and L a pedestal value of 20% of m. The modeled rendered frames were therefore modified with those two methods and the result was provided to two quality metrics (SSIM[16] and HDR-VDP-2[17]). As image metrics are computed separately for each frame, 11 temporal pooling methods were applied to turn each set of instantaneous measures into single values representing the whole sequence:

- 5 types of average (over all frames, the worst 10%, the best 10%, the first 2 seconds and the last 3 seconds),

- Minkowski summation with powers 2, 3, 4 and 5,[18]

- the low pass FIR described by Hamberg and DeRidder[19] and

- the asymmetrical pooling introduced by Ninassi et al. in.[20]

The Pearson linear correlation coefficient and Spearman rank order correlation coefficients are available in Table 4. The matter here is to evaluate the effect of the transforms applied to the images and the temporal

pooling has no influence on the relative performance of the two methods so only the results with the best temporal pooling (the one from Hamberg and DeRidder) are shown.

Metric	Additive		Devlin et al.	
	CC	SROCC	CC	SROCC
SSIM	0.420	0.319	0,488	0,3956
HDR-VDP-2	0.589	0.581	0.760	0.753

Table 4: Correlation coefficients for the two tested metrics

Results obtained with the transform from[15] are statistically better (as verified by Williams test[21]) than those obtained with the Additive approach.

5. CONCLUSION

This paper presents a subjective test investigating how the ambient light affects the quality of videos displayed with local backlight dimming. As also observed in the literature on preference studies under various ambient light, the effect of the ambient light factor in our subjective test does not manifest itself clearly, e.g. as a significant single factor. Indeed the factor *Ambient light* is not significant directly but only via second order interaction with the *Sequence* and third order interaction with *Sequence* and *Method*.

Test results show that observers can better perceive the difference between the two local backlight dimming methods when the light is low. The results of the test highly depend on the content considered and sequences presenting a higher spatio-temporal backlight variation are the ones for which the effect of the ambient light is the most important.

Objective metrics were also tested with two transforms applied on the images to take into account the ambient light. Among the two transforms, the one from Devlin et al.[15] seems to provide encouraging results.

REFERENCES

[1] ITU-R, *Recommandation BT.500-13 Methodology for the subjective assessment of the quality of television pictures* (2012).

[2] Mantiuk, R., Daly, S., and Kerofsky, L., "Display adaptive tone mapping," *ACM Trans. Graph.* **27**, 68:1–68:10 (Aug. 2008).

[3] Engelke, U., Stokkermans, M. G. M., and Murdoch, M. J., "Visualizing lighting with images: converging between the predictive value of renderings and photographs," (2013).

[4] Murdoch, M. J. and Stokkermans, M. G. M., "Effects of image size and interactivity in lighting visualization," (2014).

[5] Lin, C.-C. and Huang, K.-C., "Effects of ambient illumination conditions and background color on visual performance with tft-lcd screens," *Displays* **34**(4), 276 – 282 (2013).

[6] Guterman, P. S., Fukuda, K., Wilcox, L. M., and Allison, R. S., "75.3: Is brighter always better? the effects of display and ambient luminance on preferences for digital signage," *SID Symposium Digest of Technical Papers* **41**(1), 1116–1119 (2010).

[7] Rempel, A. G., Heidrich, W., Li, H., and Mantiuk, R., "Video viewing preferences for hdr displays under varying ambient illumination," in [*Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*], *APGV '09*, 45–52, ACM, New York, NY, USA (2009).

[8] Liu, C. and Fairchild, M. D., "Measuring the relationship between perceived image contrast and surround illumination," *Color and Imaging Conference* **2004**(1), 282–288 (2004).

[9] Hanhart, P., Korshunov, P., and Ebrahimi, T., "Subjective evaluation of higher dynamic range video," *Proc. SPIE* **9217**, 92170L–92170L–13 (2014).

[10] Burini, N., Nadernejad, E., Korhonen, J., Forchhammer, S., and Wu, X., "Modeling power-constrained optimal backlight dimming for color displays," **9**(8), 656–665 (2013).

[11] Burini, N., Mantel, C., Nadernejad, E., Korhonen, J., Forchhammer, S., and Pedersen, J. M., "Block-based gradient descent for local backlight dimming and flicker reduction," *J. Display Technol.* **10**, 71–79 (Jan 2014).

[12] "Sita Sings the Blues." *http : //archive.org/details/Sita_Sings_the_Blues*.

[13] *http : //www.cdvl.org*, *Consumer Digital Video Library* (2010).

[14] Mantel, C., Bech, S., Forchhammer, S., Korhonen, J., and Pedersen, J., "Investigating subjective attributes of quality for videos displayed with local backlight dimming.," *Proceedings of QoMEX 2014* (2014).

[15] Devlin, K., Chalmers, A., and Reinhard, E., "Visual calibration and correction for ambient illumination," *ACM Trans. Appl. Percept.* **3**, 429–452 (Oct. 2006).

[16] Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E., "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing* **13**, 600 –612 (april 2004).

[17] Mantiuk, R., Kim, K. J., Rempel, A. G., and Heidrich, W., "HDR-VDP-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graph.* (2011).

[18] Watson, A. B., [*Handbook of Perception and Human Performance*], ch. Temporal Sensitivity, New York: Wiley (1986).

[19] Hamberg, R. and de Ridder, H., "Continuous assessment of perceptual image quality," *Journal of the Optical Society of America A* **12**, 2573–2577 (1995).

[20] Ninassi, A., Le Meur, O., Le Callet, P., and Barba, D., "Considering temporal variations of spatial visual distortions in video quality assessment," **3**(2), 253–265 (2009).

[21] Howell, D. C., [*Statistical methods for psychology, 7th Edition*], Wadsworth (2010).