Influence of Screen Size and Field of View on Perceived Brightness

ALEXANDRE CHAPIRO, TIMO KUNKEL, ROBIN ATKINS, and SCOTT DALY,
Dolby Laboratories

We present a study into the perception of display brightness as related to the physical size and distance of the screen from the observer. Brightness perception is a complex topic, which is influenced by a number of lower and higher order factors - with empirical evidence from the cinema industry suggesting that display size may play a significant role. To test this hypothesis, we conducted a series of user studies exploring brightness perception for a range of displays and distances from the observer that span representative use scenarios. Our results suggest that retinal size is not sufficient to explain the range of discovered brightness variations, but is sufficient in combination with physical distance from the observer. The resulting model can be used as a step towards perceptually correcting image brightness perception based on target display parameters. This can be leveraged for energy management and the preservation of artistic intent. A pilot study suggests that adaptation luminance is an additional factor for the magnitude of the effect.

CCS Concepts: • **Human-centered computing → Displays and imagers**;

Additional Key Words and Phrases: perception, brightness, size effects, computational display, viewing distance

ACM Reference format:
Alexandre Chapiro, Timo Kunkel, Robin Atkins, and Scott Daly. 2018. Influence of Screen Size and Field of View on Perceived Brightness. *ACM Transactions on Applied Perception* 0, 0, Article 0 (2018), 12 pages.
https://doi.org/0000001.0000001

1 INTRODUCTION

Brightness perception is an important topic for computational display. If a content creator masters image assets on a certain screen, it is desirable to preserve the artistic integrity of the result, avoiding the introduction of variations in terms of brightness, contrast or other factors on the consumer side when viewed on a different display. In addition, knowledge of potential brightness offsets would help to plan for power consumption and maximum brightness capabilities for a given display to present an equivalent image at a different size.

Previous experience leads us to consider that physical screen size may influence the perceived qualities of displayed imagery. Image sizes may vary significantly between different types of displays, with commercially available technologies spanning from 70 ft. large-scale cinema screens, traditional cinema screens, various home theater displays, televisions and computer monitors to small form factor mobile displays and even fully immersive virtual reality headsets. Conversations with content creators, colorists and cinema professionals led us to believe that this effect is taken into account when mastering content in practice. We would like to understand and leverage the perceptual effects of this variation and how it affects image qualities such as brightness, contrast and chroma. In this work, we target brightness perception, as previous art indicates this is the most affected quality.

Author's addresses: Dolby Laboratories, 432 Lakeside Dr., Sunnyvale, CA, 94085, USA..

XXXX-XXXX/2018/0-ART0 $15.00
https://doi.org/0000001.0000001

Apart from obvious influences such as properties of the content being displayed, existing works suggest that various further factors may impact brightness, such as the frequency of the pattern being viewed, simultaneous contrast, various grouping and gestalt effects and the area effect, among others (see section 2 for an in-depth overview). Many of these may be contingent on retinal size and distance from the observer.

In this work we explored the effect of display size and field of view on perceived stimulus brightness in a series of subjective studies. To obtain a comprehensive view of the magnitude of the effect in practice, we explored screen sizes ranging from small handheld devices to full-size cinema screens. To help interpret the cause behind this effect, our studies employed a variety of stimuli of increasing complexity. Our results indicate that both screen size and distance from the observer influence brightness perception, with significant differences emerging within the scope of commonly used display technologies.

2 RELATED WORK

2.1 Brightness Perception

The first psychophysical question on this topic is whether to query the subjects on lightness or brightness. Most people do not have a clear understanding of the difference, as these are strictly perceptual terms. Even in the display and optics industries, experts often use the perceptual term brightness when they actually mean physically measurable luminance. *Lightness* is relative to a maximum diffuse white and does not apply to the emissive regions (e.g., light sources or specular highlights) of the imagery. On the other hand, *brightness* is the perception of the absolute luminance [13], but is not linear to it [26] and is not relative. We want to apply the results of this experiment to overall imagery including optically-captured natural and civilized imagery, computer-generated imagery, art, and possibly test targets. All of these categories may have emissive regions, so the lightness term would not be applicable. Further, lightness tends to be bounded by anchors within framed regions [8], and involves estimations of a scene geometry and illumination [25]. However, important applications include side-by-side viewing of displays where comparisons would be made across anchored frame regions. Additional applications may involve very simple or extremely abstract imagery where lightness relationships may not be constructed by the visual system. Because of these factors we decided to query on brightness.

It is unknown what aspects of the imagery subjects might use to assess the overall brightness of an image. Obvious candidates include the average luminance or the maximum luminance, but there are certainly less expected possibilities such as the image black level, and histogram criteria derived from nonlinear functions of luminance. A recent study [21] to find the best correlates to overall perceived brightness of natural and civilized imagery found the average luminance of the image was the best predictor of those tested, but others that performed nearly as well included the 96th percentile of luminance, and the mean of the luminance after being raised to the power of 0.82. Since the mean luminance was the best and simplest predictor, it suggests that a simple flat image of constant luminance would thus be the best predictor of perceived brightness. However, the tested imagery did not include a simple flat field of constant luminance, so it is unknown if that simple stimulus would give the same results. In addition, while they tested 15 images, there is always the question of the sufficiency of the image statistics of the test set. That is, whether their test set included cases that may be outliers or ‘challenge cases’. Further, there were no synthetic images or those that would be in the category of ‘art’. While that study gave very useful results to the field of image perception, we did not want to fully rely on it for our stimulus choices.

2.2 Image Surround and Content

Gilchrist [8] and colleagues investigate the area effect, i.e. how the appearance of a patch is affected by its surroundings [1]. Many aspects of the area effect are explored, with an overall focus on relative sizing of stimuli. The general conclusion is that patches surrounded predominantly by brighter regions appear dimmer and vice versa. This effect is discussed in a Gestalt framework and ample evidence is provided through careful psychophysical experimentation. Although this work is related to the studies presented here, it does not contain a numerical model that can be applied to the practical scenarios of display-related corrections for brightness perception that we target in our work.

Image properties have been investigated from a perceptual viewpoint [9, 28], and some algorithms already use this type of information to predict perceived color changes. Kim et al. [14] conducted experiments that showed that a pattern with a blurred edges appear lighter than sharp ones under certain circumstances. This could be relevant to screen size, as a pattern will exhibit lower spatial frequencies when seen with a larger field of view by an observer. Fridell [7] writes about the perception of color for building façades, focusing on size, as well as outdoor conditions and atmospheric influence. Finally, Tan and colleagues [27] show that the perception of brightness can be significantly influenced by illumination factors such as direction and elevation.

2.3 Size Effects

Note that in vision science the size of an object as seen by an observer is commonly discussed in terms of visual angles. In this work we refer to the visual angle of an object as the *field of view* (FOV) subtended by the stimulus, measured in degrees of arc and computed through simple geometry based on linear size and distance. A number of relevant studies exist that specifically compare brightness and lightness perception for stimuli with different sizes. Burnham [5, 6] produced studies where patches are compared based on their angular and physical size and surroundings. A small effect is found, with significant variation between participants. A similar study done by Burgh and Grindley [4] found no effect for the size or duration of presentation. Kutas et al [15, 16] produced a subjective study where participants compared a large LCD panel inside a mirrored booth that filled the participants' field of view displaying a flat pattern to a small patch displayed on a CRT monitor. They conclude that larger stimuli will appear lighter. Nezamabadi and colleagues [18–20] produced a series of user studies focused on the perception of artwork and noise patterns when shown in different sizes. They used a projector and compared the resulting image with a smaller display, also concluding that larger stimuli will appear lighter. Xiao et al [30–32] present studies where larger and smaller color patches are compared using a variety of technologies such as viewing booths, CRT displays, paint patches on a wall and large form factor LCD displays. Although their results vary in magnitude between trials, the general conclusion is that larger stimuli appear lighter and more colorful. Han and colleagues [10] conducted a subjective study on three differently sized televisions placed side by side and found that a 75 in. screen will appear almost 80 nits brighter than a 55 in. screen when both are set to display a 400 nit flat field. Their experimental procedure consisted of putting the screens right next to each other, leaving the process open to Cornsweet and crispening effects.

Table 1. Summary of perceptual effects found in previous size-related studies.

	Lightness	Chroma	Saturation
Burnham [5, 6]	small	yes	small
Burgh [4]	no	n/a	n/a
Gilchrist [8]	small	n/a	n/a
Kutas [15, 16]	yes	no	no
Xiao [30–32]	yes	yes	no
Han [10]	yes	n/a	n/a
This study	yes	n/a	n/a

A summary of this information can be seen in Table 1. All these studies have some common weaknesses that we address in the current body of work. Firstly, none of these focus on screens that differ in physical size by a large amount. Our work spans most of the application range by going from mobile handheld displays to cinema screens. Secondly, all studies mentioned here use only one or two different types of stimuli, and most only employ flat fields. This makes it difficult to interpret the origins of the results effectively or infer the behavior that would occur in practical scenarios. To counter this, our work is done on a variety of stimuli meant to discriminate between different parts of the visual pathway.

3 HARDWARE

For this project, our goal was to test a broad set of screen sizes that would cover most common use case scenarios. In order to do this, we selected three general cases to explore: handheld devices, home theater and cinema. We picked three displays with representative sizes:

- An LCD pannel used in LG E980 mobile displays with a 5.5 inch (14 cm) diagonal, subsequently referred to as *mobile*;
- A Dolby Professional Reference Monitor 4220 (PRM), with a 42-inch (106 cm) diagonal, subsequently referred to as *home theater*;
- A Christie CP4220 projector, displaying on a cinema screen with a 14 foot 2 inch (431 cm) diagonal, subsequently referred to as *cinema*.

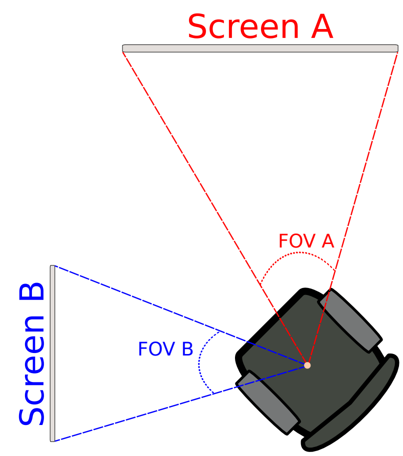

Fig. 1. Top view of the experimental setup.

All displays operated at 1920x1080 resolution. They were characterized and calibrated using a Photo Research SpectraScan PR740 spectroradiometer targeting the center point of the screen as seen from the subjects' position. The white point was set to the standard CIE D65 illuminant ($x = 0.3127$, $y = 0.329$) using hardware controls when available and, if not, a primary correction matrix when generating stimuli. A luminance look up table was generated with values measured across the feasible range. All stimuli were presented in grayscale only. Prior to each experimental session, displays were allowed a one hour warm-up period where a flat gray field was shown at half of the maximum supported intensity to stabilize temperature. This was followed by an additional calibration step if drift was present. All experiments were performed inside of a dark cinema room with all light sources and reflective surfaces in the room covered with dark material to avoid light contamination of the participant's visual field. Displays were placed at approximately a right angle from the viewers' position as shown in Fig. 1 so that only one screen could be seen simultaneously. Brightness settings on the displays were adjusted so that minimum and maximum luminance values matched the most limited display in the set as best possible.

4 STIMULUS OVERVIEW

Our study was conducted using nine stimuli with increasingly complex structure to help understand the origin of the changes in brightness. The stimuli were designed per current knowledge of the physiological processing of human vision, including both optical and retino-cortical physiology.

Since in most applications viewing distance is a key aspect involved in the comparison of images of different physical sizes, we had to consider that accommodation may play a role in any observed effect. The consequence of this is that the test stimuli must provide a sufficient features to allow the visual system to accurately focus on the display surface. A recent study of refractive compensation due to focus distance as a function of spatial frequency found very complete compensation for high frequencies and civilized imagery containing many sharp edges, but a significant reduction as frequency is lowered [2]. Based on this, we suspected a simple flat field may not be representative of imagery containing focus cues. Therefore, we decided to include imagery containing high frequencies. We included both civilized imagery, which contains the high frequencies of sharp edges and textures non-uniformly distributed across the image, as well as imagery containing uniformly distributed high frequencies without structural scene context.

There has been advanced study of the mechanisms of vision that contribute to the perception of contrast in complex imagery (natural and civilized) [11, 23], and the results point to specific cortical processing in V1 of the striate cortex, generally characterized as having a Gabor receptive field. These processes have been modelled by radial and oriented bandpass filters, such as Laplacian Pyramids, the Cortex transform, Steerable filters, and to a lesser degree by Wavelets. Phase is also known to be an important element of visual processing, as elucidated by psychophysics where the phase uncertainty within a frequency band has been found to be 90 deg [12]. That phase carries the basic sematic aspects of imagery [22], and that it is important for image quality [17]. Perceived contrast in complex imagery is dominated by the frequency bands between 1 and 6 cycles/deg, contrast constancy was found not to occur in natural images, and darker regions were judged to have higher overall contrast [11]. While we were not explicitly testing for perceived contrast or these more complex perceptual aspects, we knew there would be perceived contrast and phase relationships occurring in the relevant imagery, with possible effects of contrast on perceived overall brightness. Based on this knowledge, we designed our stimuli to build up from simpler stimulation of these cortical processes to the more complex.

4.1 Stimuli

Our stimuli choices are listed below and can be seen in Fig. 3. A detailed explanation of each stimulus and associated psychophysical aspects can be found in the supplementary material.

(1) Flat field;
(2) Single Gabor pattern;
(3) Field of smaller Gabors with the same orientation (45°);
(4) Field of smaller Gabors with random orientations;
(5) Pink noise ($1/f$);
(6) Low-pass of a natural image;
(7) Bandpass favoring low frequencies of the same image;
(8) Bandpass favoring the highest response of the CSF of the same image;
(9) Unprocessed grayscale natural image.

For the main part of the experiment, the average luminance of the stimuli was 10 cd/m^2 with an amplitude of 10 cd/m^2, resulting in images ranging from 0-20 cd/m^2, which is then offset by the highest minimum luminance of the displays in the set for feasibility. The stimuli were generated with the desired per pixel luminance values, followed by the application of the appropriate electro-optical transfer function (e.g. gamma) and corrections for the target display.

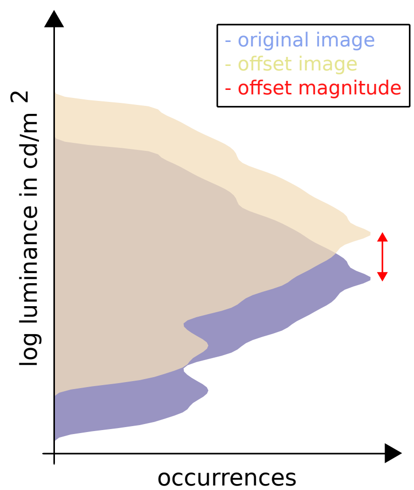

Fig. 2. Histogram showing the frequency of occurences in the original image and its shifted version. Logarithmic luminance shifts are used to change image brightness without significantly affecting contrast.

4.2 Displaying Stimuli

We will call M the lowest maximum luminance in the considered set and m the highest minimum luminance. The stimuli I_1 and I_2 are linear luminance images used in the experiment, and have amplitude A and mean luminance La as described in the previous section. Participants adjusted the stimuli by controlling parameter $o \in [-R, R]$, as shown in Fig. 2, within its range such that $I_2 = 10^{log_{10}(I_1)+o}$, where R is calculated as

$$R = log_{10}(M) - log_{10}(m + La + A)$$

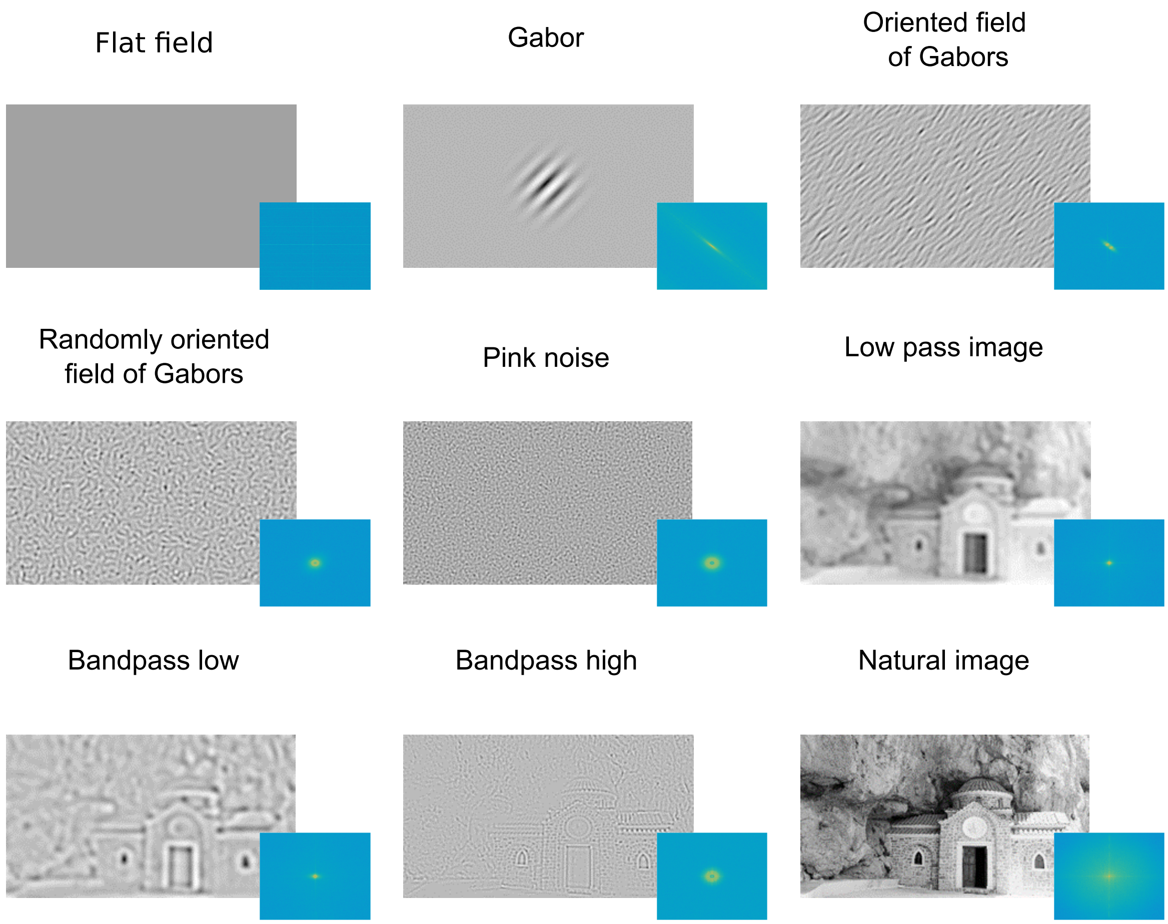

Fig. 3. Stimuli used in our experiment. Insets show the log absolute value of the Fourier transform for each stimulus.

This keeps both images within the feasible range of both displays. In practice this range far exceeded the corrections required by all users in the experiment. In addition, this luminance change does not affect the Michelson contrast of the content.

It is important to note that displaying the same image with a different FOV will alter the frequency spectrum perceived by the observer. Although it is possible to reduce this effect by rescaling the smaller image until it matches the larger FOV, we chose not to perform this correction in our main experiment because in a realistic use case the images will not be scaled this way. We keep in mind that visible frequencies may have an effect on the measured responses of our participants, which in this case will be tied in to the retinal size of the stimulus for participants.

In addition to this, we noticed that screens that are physically smaller are generally perceived to take up a lesser field of view than large screens, even when visual angles are matched. This illusion can be partially countered by closing one eye, suggesting that higher-order depth perception effects are present.

5 EXPERIMENTAL PROCEDURE

Participants were comfortably seated in a dark cinema room. The experiment was performed with binocular viewing with natural pupil size. Two displays were placed at approximately a right angle following the work of Braun et al. [3] (see Fig. 1) such that both could not be observed simultaneously to avoid framework effects as discussed in Sec. 2.2. The task at hand was to adjust imagery on a test display in order to perceptually match stimuli in terms of brightness to a reference display.

The home theater display was used as the test display in all cases as it exhibits stable behavior over the range of brightness and contrast settings variations present in this work. The second display was either of the three options described in Sec. 3, mobile, home or cinema, placed at a distance that would either match the visual angle seen by the observer or bring it to half the value. See Table 2 for a summary of all the experimental conditions tested in this work. A distance of approximately 3 picture heights [24, p.9] is considered standard: we tested distances such that the mean of the conditions fell within this standard: for instance the home display has a height of 20.5 inches (52.5 cm), translating to a recommended distance of 62 inches (157.5 cm). Since we tested two conditions with distance varying by a factor of two, we selected 47 inches (120 cm) and 94 inches (240 cm) as a reasonable compromise. Please note that for the *home | mobile* case, larger relative distances were selected for the displays than in other trials. This is necessary because we found that it was difficult to focus on a mobile-sized display if it was placed close enough to non-miopic observers to mimic the 47 inches (120 cm) condition used for other screens.

The experiment followed a method-of-adjustment (MOA) procedure. Participants were tasked with adjusting the image on the test display by using a sliding hardware controller to tune a log-luminance offset as explained in Sec. 4. Each stimulus appeared twice, totaling 18 trials per run, with all presentation being done in random order. Stimuli appeared with a mean luminance of 10 cd/m^2 on the reference screen at all times. The stimulus started out randomly at either the largest positive or the largest negative offset on the test screen to make the subjects' initial adjustments easier and to avoid frustration.

In pilot experiments we attempted to employ pairwise comparisons using the QUEST procedure [29] instead of MOA, but found that binary choice pushed users towards a strategy where one small region was used for the comparison (e.g. comparing highlights). Subjective feedback from post-experiment questionnaires suggested that the MOA procedure made observers more conscious of the effects of the offset on the image overall.

Table 2. Distances and horizontal screen sizes used in our experiments. The comparison is always done between the *home* display and the display shown in the table. The final column shows the measured log10 luminance offsets and 95% confidence intervals, which are the main result of this work. These will be discussed in more detail in Sec 6.

Test | Reference	Condition	Size (cm)	Distance (cm)	FOV (deg)	Diopters	Offset | 95% conf.
Home | Mobile	Same	93.5 | 12	240 | 31	11.0 | 11.0	0.42 | 3.23	0.064 | 0.014
Home | Home	Same	93.5 | 93.5	120 | 120	21.2 | 21.2	0.83 | 0.83	-0.002 | 0.007
Home | Cinema	Same	93.5 | 377	120 | 490	21.2 | 21.0	0.83 | 0.20	-0.030 | 0.008
Home | Mobile	Half	93.5 | 12	240 | 62	11.0 | 5.5	0.42 | 1.61	-0.011 | 0.013
Home | Home	Half	93.5 | 93.5	240 | 120	11.0 | 21.2	0.83 | 0.42	0.038 | 0.014
Home | Cinema	Half	93.5 | 377	240 | 490	11.0 | 21.0	0.42 | 0.20	0.002 | 0.013

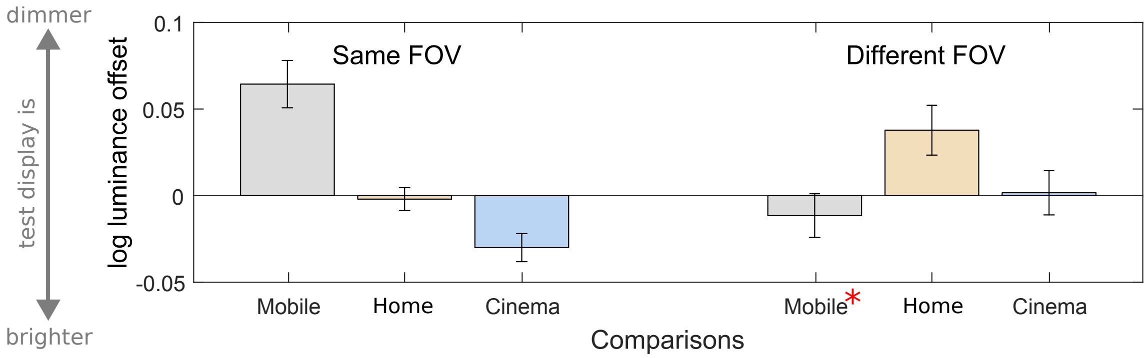

Fig. 4. Values shown are the mean *log*10 offsets set by participants for each reference and error bars show 95% confidence intervals. Offsets were added to images displayed on the test screen: a 42” monitor (*home*) in a matching task. Negative offsets mean the reference screen was perceived as dimmer than the test and a positive offset means the reference screen was perceived as brighter.
*Note that in this case the reference screen occupies a smaller field of view than the test.

6 RESULTS

The results of our experiment, averaged over all stimuli and participants, can be seen in Figure 4. Each condition had the following number of participants, with *same* and *different* referring to FOV matching: *home same* 5 (0F5M) , *home different* 5 (0F5M), *mobile same* 11 (2F9M), *mobile different* 9 (1F8M), *cinema same* 11 (3F8M), *cinema different* 11 (2F9M). Some participants took part in the experiment twice to provide additional data. N-way ANOVA analysis was performed, with results for each relevant independant variable reported below.

Note that the values shown represent a *log*10 offset in luminance added to the home display to match the screen in question. The first group of columns showing the *Same FOV* condition shows a clear pattern, contrary to anecdotal evidence, where larger offsets are required for smaller screens. Notably, the home display needed to be offset by about 0.065 log10 units to match a 5.5 inch mobile display, meaning it was perceived to be about 16% dimmer than the mobile screen. The second condition shows a comparison between two identical 42” screens shown with the same FOV, and the required offset is not significantly different from zero, with a far smaller 95% confidence interval than all other results, indicating participants had no problems matching two identical displays. Finally, the larger cinema screen led participants to adjust the home display by -0.03 log10 units, that is, it was perceived to be 7% dimmer than the test when FOV was matched. Our conclusion is that in our setup, when FOV is matched, larger screens are placed further away from the observer and appear dimmer. This was consistent across all display sizes from small form-factor mobile displays to full-size cinema screens. ANOVA analysis showed screen size to be a significant factor ($F = 52.03$, $p << 0.001$).

The results shown in the second group of columns, indicating the *different FOV* condition, change significantly. For these results, the smaller screen is always presented with half the FOV. As indicated by previous work discussed in Sec. 2.3, we expect the larger FOV to appear brighter. This counteracts the effect seen in the data of the *Same FOV* condition, leading to different offsets. At half FOV, the mobile display appeared slightly dimmer than the home display. A 42” screen at half the FOV appears approximately 9% dimmer than another identical screen. Finally, participants saw the cinema display as similar in brightness to a home display that occupied half the FOV. ANOVA analysis showed FOV (either same or different) to be a significant factor for the measured offsets ($F = 12.81$, $p << 0.001$). This result is in agreement with previous studies discussed in Sec. 2.3 and empirical experience from industry.

6.1 Higher Mean Luminance Levels

The experiments presented in this work were done at a mean luminance level of 10 cd/m^2 and a maximum luminance of 20 cd/m^2. These values were selected based on the mean luminance of cinematic content and were considered to be the most relevant in practice for current technologies. However, because informal experience points towards the screen size effect being larger for higher luminance levels, we ran an additional trial study at a higher luminance level.

The hardware setup, experimental procedure and stimuli were identical to those presented in Sections 3, 4 and 5, but mean luminance was set to 40 cd/m^2, with maximum luminance reaching 80 cd/m^2. We compared the 42" home monitor with the 14 ft. cinema screen seen at twice the FOV using the same setup described in Table 2, *Home | Cinema | Half*. 7 participants (2F5M) took part in this experiment. The mean offsets found can be seen in Fig. 5. When using this higher luminance level, the resulting offset for this condition changed from 0.0017 to 0.0595. This result was analyzed through ANOVA, with luminance found to be a significant factor ($F = 16.42$, $p << 0.001$). We believe this means the effect is likely to scale up with higher luminance levels, a result that is especially relevant in the emerging high dynamic range (HDR) display landscape that is quickly becoming mainstream. We plan to further explore size effects for brightness perception at different adaptation luminance levels in the future.

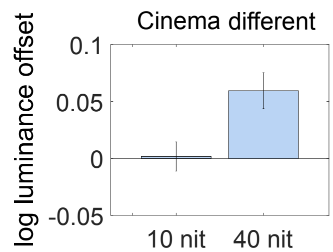

Fig. 5. Results for original experiment with mean luminance set to 10 nits compared to trial at 40 nits for the *cinema, different FOV* condition.

6.2 Results over Stimuli

Different stimuli were also found to be a factor in participant responses ($F = 6.99$, $p << 0.001$). Figure 6 shows results for the *same FOV* condition separated by stimulus type, fitted with the best linear fit for each stimulus across all display sizes on a log scale. Participants tended to require larger offsets to match more complex images: the average y-intercept (shown in the legend, the y-intercept is the point at which this line would cross the y-axis

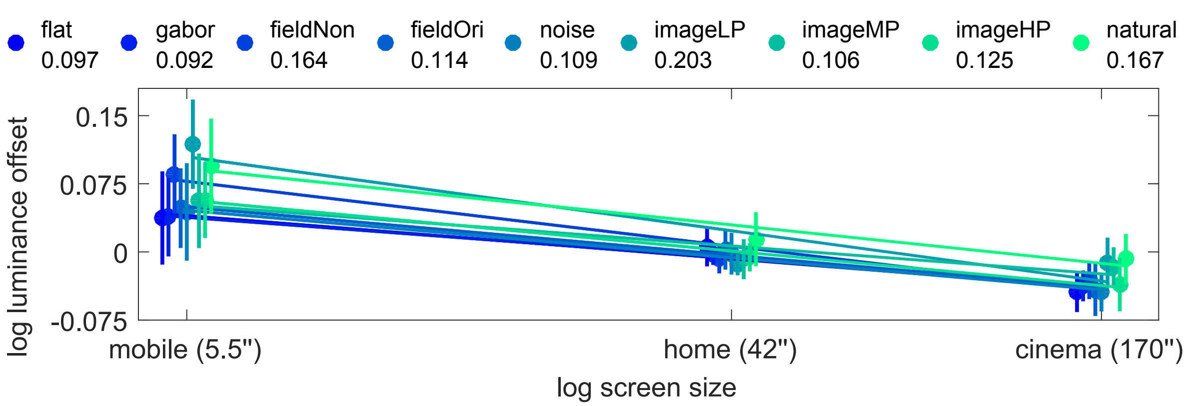

Fig. 6. Mean responses per stimulus for the *same* FOV condition, shifted horizontally for clarity. Bluer colors represent simpler stimuli and greener colors show more complex stimuli. Vertical bars show 95% confidence intervals. Lines show best linear fit for each stimulus across display sizes on a log scale with the corresponding y-intercept values shown in the legend.

for an argument value of zero) for natural image derivatives is 0.15 while the average for non-natural imagery is smaller, 0.115.

6.3 Results over Participant Strategies

Finally, we collected subjective feedback for this task from all participants. The general consensus was that the matching task was considered challenging, and even after a nearest match was achieved there was still a qualitative difference between conditions (i.e. it was impossible to find a "perfect" match). Participants gave significantly different responses based on ANOVA analysis ($F = 29.18$, $p << 0.001$).

In addition, we asked participants to explain their matching strategies and classified their responses in one of three groups as *judged based on average*, *judged based on feature* and *other*. Participants in the first group judged the images based on their overall brightness perception; the second group matched based on specific features in the image (such as the disc above the door in the natural image). The final group used other strategies, such as discomfort due to brightness when raising it to higher values or were unable to explain their strategy clearly. The results can be seen in Fig. 7. Note that while there is significant variations of the effect between individuals, participants tended to be internally consistent in their answers. We consider this to be a good indication that our model can be effectively applied as a personalized tool.

7 PROPOSED MODEL

We propose that a log_{10} offset for perceptual matching of brightness between displays can be calculated using our data as a look-up table. Our data can be described by a function $o = f(x, y)$ where o is a log_{10} offset and with x and y being the *reference screen size* and *distance to the screen in picture heights*. Consider a colorist that graded an image I expressed here in cd/m^2 on a display with screen size x_1 at a distance y_1. Let's think of a virtual comparison between our display and a hypothetical Professional Reference Monitor (home), which has a known constant size x_p and which we assume is located at 3 picture heights. We find the offset $o_1 = f(x_1, y_1)$ necessary to match the colorists' display to it. We then consider the offset required to match the consumer side display to

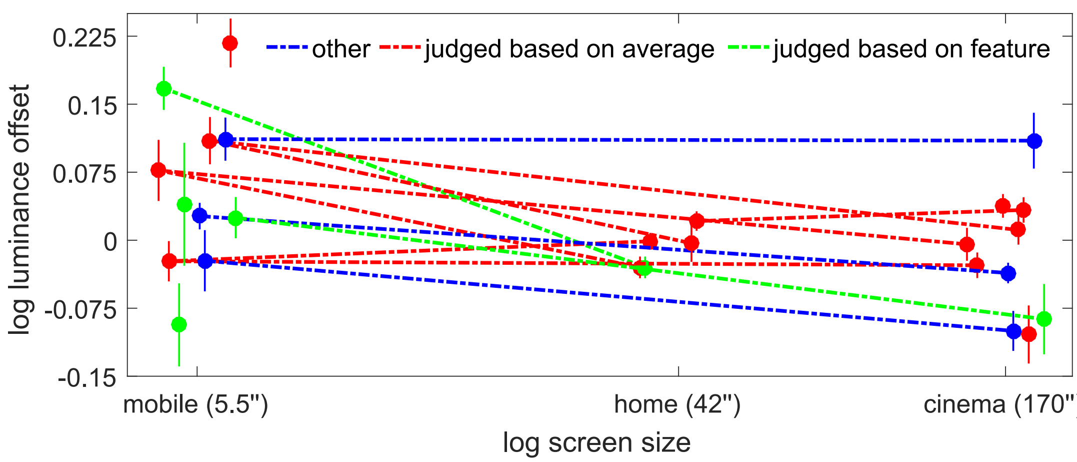

Fig. 7. Subjective strategies for the matching task were collected from participants and classified into three rough groups. Lines join data from the same participant, averaged over all responses for the *same FOV* condition. Note that not all participants participated in all studies.

the same hypothetical reference $o_2 = f(x_2, y_2)$. The final offset is the difference of these offsets $o = o_1 - o_2$ which represents the offset from display 1 to 2.

By way of example, if our goal is to perceptually match the brightness of a scene mastered on our cinema screen seen with a FOV of approximately 21 degrees to a mobile display seen with a FOV of 11 degrees we would use the data in Table 2. We calculate an offset of 0.002 for the cinema to a home display at 11 degrees, and the offset from a home display to a mobile display both at 11 degress as 0.064 . The offset required to match these two displays according to our experiment would be $0.002 - 0.064 = -0.062$ log10 luminance units. Offsets for displays that were not studied in this work can be obtained by interpolating the existing data.

In practical systems, it may be desirable to maintain maximum contrast between certain parts of the displayed imagery. In this case, a different editing method that preserves contrasting regions while changing the mean of the content may prove more desirable. In this work we chose to use a simple log offset in order to avoid the contamination of our results by changes in contrast [14].

8 DISCUSSION AND FUTURE WORK

In this work we have studied the effect of screen size and distance from the observer on perceived brightness. We found this to be a complex topic with a number of relevant factors. We found significant effects across the range of explored screens and distances, with perceived brightness scaling positively with FOV, adaptation luminance, and image complexity, and negatively with distance. Although the effect is relatively small, generating a perceptual increase of at most 15% in our trials (Table 2, first line *home|mobile|half*), this difference is noticeable and may be effected for practical applications such as appearance matching and perception modeling for displayed imagery.

Although different display technologies were used, we can rule out that display properties such as differing emissive spectra or environmental contaminants are the cause of the effect as it was present when comparing two identical Professional Reference Monitors at different FOV but absent when the displays were shown at the same FOV. We avoided introducing perceptual biases by positioning displays in a way that only one display can be seen at a time by participants. Specifically, placing displays side-by-side was avoided as luminance offsets and falloffs towards the displays' boundaries could induce a luminance edge, inducing a Cornsweet illusion. Given that users were generally expected to be adapted to the background gray for several of the stimuli, this illusion could be exacerbated by Whittle's Crispening effect if local gradients occur in opposing luminance directions.

Although our work explores a wide range of screen sizes commonly used in practice, it would be interesting to explore additional screen sizes such as very large cinema screens. In addition, our work could be extended to consider virtual reality displays, which present us with a qualitatively differing viewing mode where users are fully immersed in a wearable display.

Our pilot study at higher luminance levels presented in Sec. 6.1 suggests that the effect of screen size on brightness may scale nonlinearly with higher luminance levels. This could become very relevant with the advent of high dynamic range displays and cinema projectors currently entering the market which are capable of higher output luminances. Additional exploration into higher luminance is necessary to quantify this effect.

ACKNOWLEDGMENTS

The authors would like to thank all members of the Applied Vision Science Group at Dolby Laboratories for the insightful discussions. We are grateful to the user study participants for their efforts. We thank Shane Ruggieri for assistance in using the cinema and Thomas Wan for providing and maintaining the mobile-sized display.

REFERENCES

[1] Sarah R Allred, Ana Radonjić, Alan L Gilchrist, and David H Brainard. 2012. Lightness perception in high dynamic range images: Local and remote luminance effects. *Journal of vision* 12, 2 (2012), 7–7.

[2] Arthur Bradley, Renfeng Xu, Larry Thibos, Gildas Marin, and Martha Hernandez. 2014. Influence of spherical aberration, stimulus spatial frequency, and pupil apodisation on subjective refractions. *Ophthalmic and Physiological Optics* 34, 3 (2014), 309–320.
[3] Karen M Braun, Mark D Fairchild, and Paula J Alessi. 1996. Viewing techniques for cross-media image comparisons. *Color Research & Application* 21, 1 (1996), 6–17.
[4] Peter Burgh and GC Grindley. 1962. Size of test patch and simultaneous contrast. *Quarterly Journal of Experimental Psychology* 14, 2 (1962), 89–93.
[5] Robert W Burnham. 1951. The dependence of color upon area. *The American journal of psychology* 64, 4 (1951), 521–533.
[6] Robert W Burnham. 1952. Comparative effects of area and luminance on color. *The American journal of psychology* 65, 1 (1952), 27–38.
[7] Karin Fridell Anter. 2000. *What colour is the red house? Perceived colour of painted facades.* Ph.D. Dissertation. Institutionen för arkitektur.
[8] Alan Gilchrist, Christos Kossyfidis, Frederick Bonato, Tiziano Agostini, Joseph Cataliotti, Xiaojun Li, Branka Spehar, Vidal Annan, and Elias Economou. 1999. An anchoring theory of lightness perception. *Psychological review* 106, 4 (1999), 795.
[9] Daniel J Graham and David J Field. 2007. Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial vision* 21, 1 (2007), 149–164.
[10] Taeseong Han, Sungjin Kim, and Dongwoo Kang. 2016. Perceptually Equivalent Luminance Level of Large-screen TVs. *Color and Imaging Conference* 24 (2016).
[11] Andrew M Haun and Eli Peli. 2013. Perceived contrast in complex images. *Journal of vision* 13, 13 (2013), 3–3.
[12] Martin Hübner, Terry Caelli, and Ingo Rentschler. 1988. Visual phase resolution for gray-scale textures. *Perception & psychophysics* 43, 4 (1988), 319–325.
[13] Robert William Gainer Hunt. 1967. The reproduction of colour. (1967).
[14] Min H Kim, Tobias Ritschel, and Jan Kautz. 2011. Edge-aware color appearance. *ACM Transactions on Graphics (TOG)* 30, 2 (2011), 13.
[15] G Kutas and P Bodrogi. 2008. Color appearance of a large homogenous visual field. *Color Research & Application* 33, 1 (2008), 45–54.
[16] Gábor Kutas, Katalin Gócza, Peter Bodrogi, and János Schanda. 2004. Colour size effect. In *Conference on Colour in Graphics, Imaging, and Vision*, Vol. 2004. Society for Imaging Science and Technology, 70–73.
[17] Sean T. McCarthy. 2016. Setting Video Quality & Performance Targets for HDR and WCG Video Services. In *Proceedings of Cable-Tec Expo*. Society of Cable Telecommunications Engineers and International Society of Broadband Experts.
[18] Mahdi Nezamabadi and Roy S Berns. 2005. The effect of image size on the color appearance of image reproductions. In *Color and Imaging Conference*, Vol. 2005. Society for Imaging Science and Technology, 79–84.
[19] Mahdi Nezamabadi and Roy S Berns. 2006. Effect of image size on the color appearance of image reproductions using colorimetrically calibrated LCD and DLP displays. *Journal of the Society for Information Display* 14, 9 (2006), 773–783.
[20] Mahdi Nezamabadi, Ethan D Montag, and Roy S Berns. 2007. An investigation of the effect of image size on the color appearance of softcopy reproductions using a contrast matching technique. In *Electronic Imaging 2007*. International Society for Optics and Photonics, 649309–649309.
[21] Katy C Noland, Manish Pindoria, and Andrew Cotton. 2017. Modelling brightness perception for high dynamic range television. In *Quality of Multimedia Experience (QoMEX), 2017 Ninth International Conference on*. IEEE, 1–6.
[22] Alan V Oppenheim and Jae S Lim. 1981. The importance of phase in signals. *Proc. IEEE* 69, 5 (1981), 529–541.
[23] Eli Peli. 1990. Contrast in complex images. *JOSA A* 7, 10 (1990), 2032–2040.
[24] Charles Poynton. 2012. *Digital video and HD: Algorithms and Interfaces.* Elsevier.
[25] Dale Purves, S Mark Williams, Surajit Nundy, and R Beau Lotto. 2004. Perceiving the intensity of light. *Psychological review* 111, 1 (2004), 142.
[26] JC Stevens and Stanley S Stevens. 1963. Brightness function: Effects of adaptation. *JOSA* 53, 3 (1963), 375–385.
[27] Minghui Tan, Jean-François Lalonde, Lavanya Sharan, Holly Rushmeier, and Carol O'Sullivan. 2015. The perception of lighting inconsistencies in composite outdoor scenes. *ACM Transactions on Applied Perception (TAP)* 12, 4 (2015), 18.
[28] Andrew B Watson, HB Barlow, John G Robson, et al. 1983. What does the eye see best? *Nature* 302, 5907 (1983), 419–422.
[29] Andrew B Watson and Denis G Pelli. 1983. QUEST: A Bayesian adaptive psychometric method. *Attention, Perception, & Psychophysics* 33, 2 (1983), 113–120.
[30] Kaida Xiao, CJ Li, M Ronnier Luo, and C Taylor. 2004. Colour Appearance for Dissimilar Sizes. In *Conference on Colour in Graphics, Imaging, and Vision*, Vol. 2004. Society for Imaging Science and Technology, 12–16.
[31] Kaida Xiao, M Ronnier Luo, Changjun Li, Guihua Cui, and Dusik Park. 2011. Investigation of colour size effect for colour appearance assessment. *Color Research & Application* 36, 3 (2011), 201–209.
[32] Kaida Xiao, M Ronnier Luo, CJ Li, Peter A Rhodes, and C Taylor. 2003. Specifying the Colour Appearance of a Real Room. In *Color and Imaging Conference*, Vol. 2003. Society for Imaging Science and Technology, 308–312.

Received December 2017; accepted January 2018