Optimizing Mobile Display Brightness by Leveraging Human Visual Perception

Matthew Schuchhardt
Northwestern University
mjs633@eecs.northwestern.edu

Susmit Jha
Intel Corporation
susmit.jha@intel.com

Raid Ayoub
Intel Corporation
raid.ayoub@intel.com

Michael Kishinevsky
Intel Corporation
michael.kishinevsky@intel.com

Gokhan Memik
Northwestern University
memik@eecs.northwestern.edu

ABSTRACT

Modern smartphones and tablets are battery-constrained by their mobility; this constraint is heavily factored into any design decision made on the device. Furthermore, the display is one of the most power-consuming subsystems. Adaptive display brightness systems attempt to address this high display power consumption by setting the brightness depending on the surrounding ambient light levels.

In this work, we run a series of user studies aimed at gauging the accuracy of these adaptive brightness models. These studies compare subjective satisfaction and readability metrics on a common smartphone display and find that the adaptive brightness system isn't well-tuned to user preferences. We also find that subjective ratings and readability are closely correlated with one another across different display brightness levels, which can be used to give a better understanding of how screen brightness levels impact users.

Additionally, we note that operating systems need to reduce the power envelope of the device for a variety of reasons. This is done without understanding the impact of these decisions on user satisfaction, and can significantly degrade user experience. In this work, we propose a user-aware method of reducing display power consumption, which allows display power to be throttled while understanding the resulting impact on the user. By considering this impact on the user, we devise an optimal dimming scheme which can reduce the time-weighted readability degradation by upwards of 21.5%.

1. INTRODUCTION

In modern mobile devices, the touchscreen display has taken front and center as the most important interface that users interact with. Not only does the touchscreen display act as the primary means of output to the user, but with the absence of hardware keyboards and buttons on many mobile devices, it is now the primary means of input as well. With this elevated station that touchscreen displays have come to enjoy in smartphones and tablets, more and more focus is being placed on the design and features of the display. The trend of mobile displays is increasingly toward larger, brighter, higher-resolution displays [10], and this trend is undoubtedly a sight for sore eyes everywhere.

However, this trend toward larger and brighter displays does not come without cost; although many manufacturers tote the energy efficiency of their displays, Chen et al. [3] find that generational improvements in OLED displays account for minor improvements in power consumption. They find that after two Samsung display design iterations, normalized full-brightness per-pixel power efficiency only improves by 5.7%; this is reduced even further when the current trend toward increasingly larger screens is considered. Furthermore, they find that the most significant contributor to power consumption is the size of the screen itself. With larger devices, although the battery tends to be bigger, the screen is also larger, and considering the lack of significant generational improvements in power consumption, there is no reason to expect that display power will not continue to be a significant consumer of power in mobile devices.

Because the display consumes a significant amount of power, adaptive brightness systems are one way that display manufacturers attempt to reduce the display's power envelope while improving user satisfaction with the display's brightness. Adaptive brightness systems control a display's brightness via ambient light sensors which give an indication of incident environmental light on the screen. These adaptive brightness systems can significantly reduce the power usage compared to a display which is constantly set to a high brightness level. Also, adaptive brightness systems have the advantage of causing less eye strain by avoiding setting the brightness too high when in a dim lighting environment [12].

In this work, we wish to gauge how effective these adaptive brightness systems actually are at satisfying users' brightness requirements. To do this, we design a user study which allows us to gather subjective display brightness satisfaction ratings and readability data for a series of users at a selection of ambient light and screen brightness levels. From these results, we find that the default automatic brightness system that comes shipped with the device isn't well-tailored to the group of users we studied.

Furthermore, we find that subjective user ratings are closely correlated with an objective readability metric, especially in brighter lighting environments. This strong, direct correlation is interesting because it suggests that by using this relationship, it's possible to understand how much of an impact selecting a given brightness level will have on a user, without needing to directly gather subjective ratings from the user.

We also study the impact that various brightness throttling schemes have on users. System power may need to be reduced below the standard power envelope for many reasons. For example, one common reason for reducing display brightness is low-battery situations. When the battery falls under a critical threshold, many smartphones reduce the display's brightness to help extend the device's remaining on-screen time. Additionally, there are systems such as *system-level dynamic thermal management* [9], which reduce system power consumption in an attempt to reduce the system's total heat levels. Any OS subsystem can cause power throttling events, but the underlying issue with these systems is that they do not consider the impact that these power savings decisions have on the user.

To explore this further, we conduct another study which collects data on what lighting environments are commonly encountered throughout the day by a typical user. Using these results along with the previous readability and ratings data, we directly analyze how much power can be saved with a specified degradation in readability. We also compare a user-agnostic system which degrades the display by a constant brightness level or constant brightness percentage, to our proposed user-aware system which degrades the display's brightness more in brighter environments. Any backlight dimming power savings decision will impact users to some degree, but by considering how these decisions affect the user, it's possible to reduce the average impact over a period of time. Using our user-aware brightness throttling scheme can achieve an 8% average system power reduction while only degrading average readability by 3.2%; this is 21.5% less of a degradation than the next best user-agnostic method. Even more importantly, this method allows us to directly understand how the power decisions will affect users, which makes these power decisions more well-informed than arbitrarily throttling system power.

In summary, we make the following primary contributions:

- We find a correlation between subjective user ratings and an objective readability metric
  - We run user studies which gather subjective ratings and visual discernment task data across multiple lighting environments
  - We present analysis which shows a strong correlation between objective visual discernment and subjective ratings
  - We perform the first analysis on the impact of adaptive screen brightness mechanisms on readability
- We create a model for screen brightness which maximizes time-averaged screen readability
  - We compare existing adaptive brightness systems to a properly calibrated one
  - We create a power throttling model which maximizes time-averaged readability for a given power reduction
  - We compare our system's time-averaged readability against other power throttling models

In Section 2, we give some background information regarding the human visual system, mobile display optics, and brightness and power characteristics of our experimental device. Section 3 outlines the user studies that we run to gather data for our analysis. We present our results in Section 4, analyze some related work in the field in Section 5, and conclude in Section 6.

2. BACKGROUND

2.1 Mobile Display Optics

Luminance (L) and *contrast ratio* (C) are two of the most important metrics involved with gauging how easily the human visual system can see details on a mobile display [8, 14]. Luminance is defined as the intensity of light per unit area, and is colloquially referred to as *brightness* in the context of mobile displays. Contrast ratio is related to luminance; generically defined, contrast ratio compares the luminance of the colors *black* and *white* (L_{black} and L_{white}) as they appear on the display. L_{black} is never exactly zero on an LCD display; there is always some amount of light which leaks through the liquid crystal matrix, even on an image which is intended to be fully black. Similarly, L_{white} is limited by how bright the LCD's backlight is. The contrast ratio describes the relative difference between these luminances. However, there is not a single well-defined metric which describes contrast ratio, and there are a number of different equations which can represent this. One common formula for a display's contrast ratio is described in Equation (1):

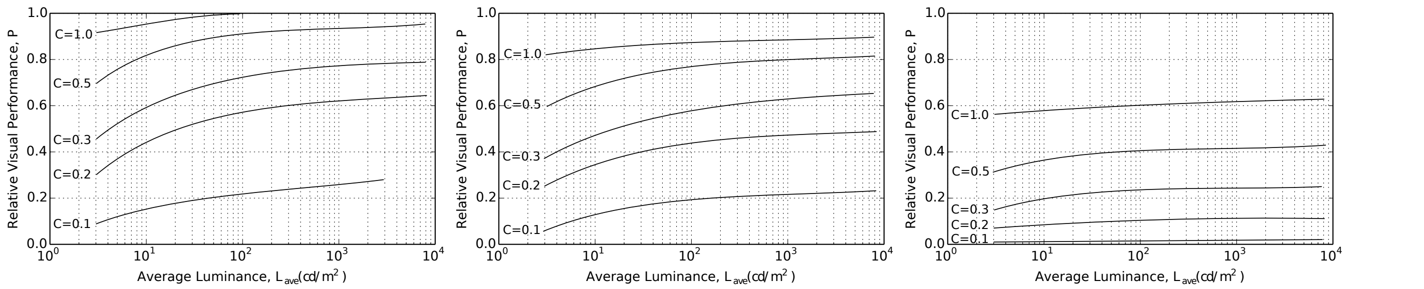
$$C = \frac{L_{white} - L_{black}}{L_{white}} \tag{1}$$

There are some complexities as to how this is measured and reported in consumer devices (such as dynamic contrast ratio, static contrast ratio, etc.), but as far as light perception is considered, this definition is sufficient. It is important to note, L_{black} and L_{white} are not entirely self-contained to the display itself; the device's ambient light can have a significant impact on these values. For instance, in a completely darkened room, the contrast ratio is precisely defined as in Equation (1). However, consider a lighting environment which involves a significant amount of ambient light. Since the display's optical characteristics necessarily don't absorb 100% of incoming ambient light, some of that light is reflected back at the viewer. This reflected light reduces the resulting contrast ratio, as evidenced in Equation (2), with E representing the environmental illuminance and ρ representing the display's reflective coefficient. Contrast ratio has a significant impact on discerning visual detail, and this relationship of contrast ratio with ambient light explains why it is easier to read backlit displays in darker ambient environments. This is also the basis for the inclusion of *adaptive brightness systems* in mobile devices. Adaptive brightness systems use ambient light sensors to help control the display's brightness: in dim environments, the display reduces the display's brightness, and in brighter environments, the display is set to higher brightness levels. The net result of these systems is that the display uses less power and avoids eye strain in dim environments, while maintaining better contrast ratio in bright ones.

$$C = \frac{L_{white} - L_{black}}{L_{white} + \rho E} \tag{2}$$

Now that contrast ratio and its relationship with environmental light has been introduced, how does contrast ratio and luminance actually impact people's ability to view a mobile display? *Readability* is a metric which describes how easily a person can discern visual detail in a given environment. However, similar to contrast ratio, readability doesn't have a single definition, and can be represented in a number of different ways. We will expand further upon one particular readability metric, Relative Visual Performance (RVP), in Section 2.2, but for now, will refer to readability in the generic sense.

Contrast ratio has a heavy influence on readability. This relationship is intuitive; contrast ratio is similar to the concept of signal-to-noise ratio, where the magnitude of the



(a) 20 year old, 1.5' detail (b) 50 year old, 1.5' detail (c) 75 year old, 1.5' detail

Figure 1: Aggregated readability curves [8] for users of varying ages, generated using data from CIE [4]. As users age, their ability to discern detail drops significantly at a given luminance and contrast ratio level. Contrast ratio and luminance are both strong predictors of readability.

signal is compared to the magnitude of the noise on top of the signal. Larger differences between these magnitudes leads to a stronger, more obvious signal. Luminance's relationship with readability, however, is somewhat less direct. One might expect that if the contrast ratio is the same at two different luminance levels, the readability would also be the same. However, this is not the case. The human eye is not a simple sensor; there is a homogeneous mix of types of light-detecting cells, each of which respond to different ranges of luminance levels. Because of this, only some of the light-detecting cells in the eye are active in low-light environments, which reduces the eye's effective resolution [7]. In high-luminance environments, all of the light detecting cells remain active, which improves the eye's visual acuity. The end result of this is that readability is better at higher luminance levels, when the contrast ratio is held constant. This doesn't mean that increased ambient light leads higher to readability, however; increasing ambient light increases the screen's luminance due to reflected light, but it also reduces the contrast ratio, which lowers the effective readability on most LCD displays (as Equation (2) describes). Contrast ratio is a more significant indicator of readability than the overall luminance [7].

The standards organization CIE [4] conducted a study intended to gauge the impact that contrast ratio and luminance have on readability. Readability was measured across a selection of users, at a number of contrast ratios and luminance levels, the results from which are shown in Figure 1. As previously discussed, the study found both contrast ratio and luminance to strongly impact readability. There is also a significant degradation of readability as users age. The curves contain aggregated data from a wide selection of users, and individual users will each have their own personal readability curves, but on average, a 75-year-old requires roughly twice the contrast ratio that a 20-year-old does at the same luminance and contrast ratio to achieve similar readability levels. This trend in readability data is further explored by Kelley et al. [8].

2.2 Readability Metrics

We introduced the notion of readability in Section 2.1, but thus far have only used the term in the generic sense. For the remainder of this work, we use Relative Visual Performance (RVP) as our target readability metric. RVP is an objective metric which directly measures how quickly and accurately a visual discernment task can be completed. A visual discernment task involves asking users to identify textual glyphs of a specified size; the Snellen chart, which contains letters of descending size and is commonly used by optometrists, is one example of this. RVP has been shown to have a strong dependency to contrast ratio and luminance, as described by CIE [4].

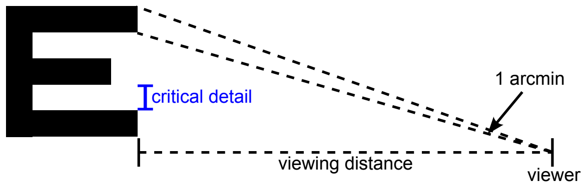


Figure 2: Diagram of critical detail size of the letter 'E'. Critical detail is 1/5th of overall height of letter. 20/20 vision requires ability to resolve 1 arcmin of detail.

Figure 3: A series of tumbling E glyphs.

To make an RVP measurement, a user is asked to visually identify a series of characters with a *critical detail* of a known size. A critical detail is a feature in a textual glyph which is required to be properly resolved to correctly identify the glyph (see Figure 2) [7]. One test, commonly used with young or otherwise illiterate test subjects, consists of a series of E characters which are tumbled in one of four orientations, as in Figure 3. The subjects are asked to identify the orientation of each of the letters in this test. The accuracy and speed at which users can identify these is measured, and RVP is calculated via Equation (3). Because RVP is an experimentally-derived metric, it will vary between iterations and users to some degree.

$$speed = \frac{correctly\ identified\ E's}{time}$$

$$accuracy = \frac{correctly\ identified\ E's}{total\ E's}$$

$$RVP = accuracy \cdot speed \tag{3}$$

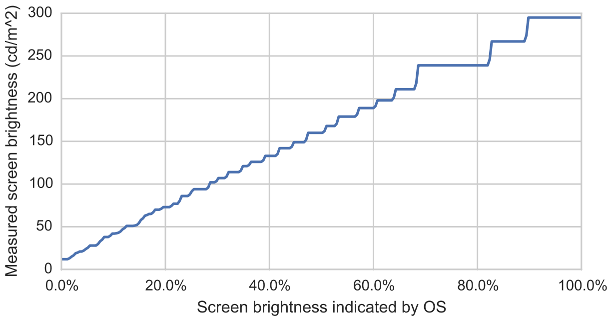


Figure 4: Comparison of brightness levels as indicated by the OS to actual brightness levels, as measured by an external light meter. Screen brightness is linear with the OS's indicated levels, although the granularity of brightness drops at higher screen brightness levels.

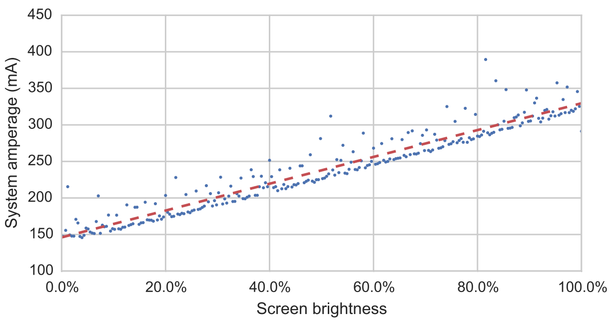


Figure 5: Comparison of Nexus 4 current draw at various screen brightnesses with idle system. Blue dots represent individual amperage readings, and the dotted red line represents the linear best fit.

2.3 Display Characteristics

In this section, we examine the power and brightness characteristics of our targeted smartphone device. We use the Google® Nexus™ 4 smartphone, which contains a 4.7", 320ppi IPS LCD display [5]. LCD displays are more common and more mature than the newer OLED displays, but the two behave similarly from an optical standpoint.

Most modern mobile displays have brightness controls which can be accessed programmatically by the operating system. However, these brightness values as reported by the OS aren't guaranteed to be directly related to the display's actual brightness. Because of this, we use an external light sensor from a second Nexus 4 device to verify that the display brightness levels indicated by the operating system are linear with the screen's actual brightness (Figure 4), and find that they are linearly correlated. This is conducted in a darkened room to remove the effect of environmental light. It is interesting to note that the granularity of the measured screen brightness is lower at the higher screen brightness levels. This may be because the human visual system's sensitivity to change in stimuli is reduced at higher brightnesses, making it unnecessary to implement all of the brightness levels on the display.

We similarly analyze the relationship between brightness

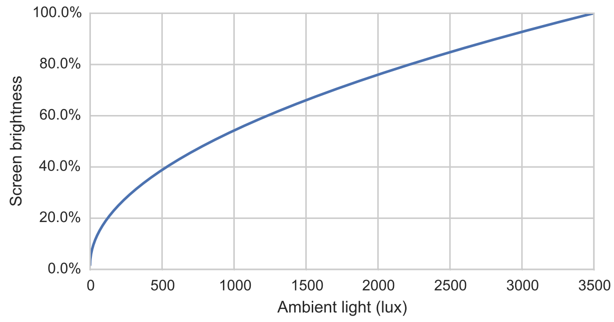


Figure 6: Brightness curve of adaptive brightness model on the Nexus 4 device. Screen brightness is linear with the square root of ambient light.

and power by comparing the device's power consumption across multiple screen brightnesses via the Nexus 4's internal current sensor. The results are presented in Figure 5, and suggest a linear trend between power and display brightness. These amperage numbers are only valid for a idle system; system power will be higher when actively running workloads. The idle system power at minimum brightness is 562.4mW. From Figure 4 and Figure 5, we can treat the OS-indicated brightness levels as linear with measured screen brightness and power consumption.

Finally, we wish to note what adaptive brightness system the device uses by default. In the case of the Nexus 4, the brightness levels are adjusted via a static model which translates the ambient light levels to a corresponding screen brightness. This data was retrieved from the operating system, and the continuous model of this data is presented in Figure 6. According to this model, the screen brightness is linear with the square root of the ambient light levels, reflecting Steven's Law [19].

3. EXPERIMENTAL SETUP

The goal of this work is to develop a better understanding of how a wide variety of users interact with and respond to mobile displays, and to explore potential power and user experience optimizations in light of this new data. The first two experiments we perform are conducted in a lab environment which allows us to directly control ambient light and screen brightness levels, giving us further insight into the impact that screen brightness and ambient light has on users. We also conduct a third experiment which allows us to analyze how users interact with their devices in typical day-to-day use. The details of these experiments are described below.

3.1 Screen Readability Study

As described in Section 2, readability is important to consider in any mobile display. RVP, which relates the speed and accuracy of a visual discernment task, has the advantages of being a bounded and normalized metric, and also uses easily gathered data which doesn't require sophisticated eye tracking apparatus. Furthermore, RVP also is significantly dependent on contrast ratio and luminance.

We wish to understand how RVP varies with screen brightness and ambient light levels across a variety of users by directly comparing RVP to subjective user ratings in identical ambient lighting and display environments. To accomplish this, we analyze readability in this study, and then perform the subjective rating study described in Section 3.2. This study was set up as follows.

Users are asked to sit down in front of a Google Nexus 4 smartphone device which is fastened in front of the user. A high-brightness halogen lamp stand with a variable brightness slider is used to control the ambient light levels to simulate the full range of brightnesses that a user may encounter in day-to-day life. This lamp allows us to create ambient lighting environments ranging from a dark room (0 lux) to a sunlit day ($\geq$ 4999 lux). The lamp is placed behind the user's shoulder and aimed toward the fastened Nexus 4, causing reflections which reduce the effective contrast ratio.

To conduct each iteration of the study, users begin with a blank screen which contains a *start* button. Upon pressing that button, users are presented with a grid of 20 E's which are tumbled in either an up, down, left, or right orientation (as in Figure 3). Users are asked to touch the E's which are in either an *up* or *down* orientation, while ignoring those in a *left* or *right* orientation. Upon completing this visual discernment task, the user presses the *stop* button, and our system logs the screen brightness, ambient light level, and the speed and accuracy at which the user completed the task. This same task is repeated 60 times (5 discrete ambient light levels, and 12 screen brightness levels for each), with a 15-second gap in-between tasks, each at a different screen brightness and ambient light combination. The ordering of these ambient light and screen brightness combinations is randomized to reduce order-effect bias.

1 arcmin is a generally accepted threshold of detail that a healthy human eye can typically resolve, and is the visual discernment resolution required for 20/20 vision [7]. For this reason, we use 1 arcmin as our target critical detail size, as shown in Figure 2. We set this gap size to exactly 2 pixels on our study, which allows us to avoid dealing with subpixel display issues like anti-aliasing. Since the Nexus 4 has a 320ppi screen, we situate users 21.6 inches from the display to achieve the 1 arcmin critical detail gap. Users are given the option of using a larger E if they couldn't discern 1 arcmin of detail, but no users required this.

We ran this experiment on a set of 30 undergraduate and graduate students. These users are primarily in the 20–30 year age range, and so is not representative of the entire population. This does, however, allow us to analyze the variability in that subgroup. Older users are expected to have a similar reaction to changes in brightness, but with generally brighter display requirements.

3.2 Subjective Brightness Satisfaction Study

In addition to readability analysis, we are also interested in relating user satisfaction to ambient and brightness levels. We use the same lighting setup and the same 30 users as in Section 3.1. We again fasten the phone at a constant distance of 21.6 inches from the users.

This time, however, instead of having the users perform a synthetic visual discernment task, users are asked to read and interact with the BBC News Android application [1]. While they are browsing the articles on the application, a dialog box pops up every 30 seconds, and users are asked to rate how satisfied the current ambient light and screen brightness levels are. The options are *much too dim*, *slightly too dim*, *perfect*, *slightly too bright*, and *much too bright*. Users are instructed to interpret the modifier *much* as a lighting situation where the user would likely go out of their way to adjust the screen's brightness. *Slightly* is defined as a situation where users likely wouldn't adjust the device, but it isn't set exactly where they'd prefer. *Perfect* is just as it sounds; users are completely satisfied in this situation.

As in the previous study, after each of these subjective ratings, the ambient light and screen brightness levels are randomly adjusted to one of the 60 discrete ambient light/screen brightness combinations. These ambient light and screen brightness combinations are identical to the previous experiment to allow for direct comparisons between the two studies. This is repeated 60 times, and the ambient and screen brightness levels and the subjective ratings are recorded after each user indication.

3.3 Device Usage Study

As a final data collection component to this study, we also want to understand how real users interact with their devices over the course of a day. Specifically, the data we wish to collect includes the average amount of time per day that users spend on their phones, as well as the ambient light levels that they encounter throughout the day. With this data, we can then integrate any power savings techniques we come up with over time to determine the daily battery impact that these techniques have.

We collected data from 5 individual users over the course of 14 days. We log the times that the users turn on/off their smartphone screens, as well as the ambient light levels that they are in during this time. We poll this data at a rate of once every 5 seconds. This collection system was devised as an Android application which was distributed via the Google Play Store [6] to users from a wide range of geographies, which helps account for bias due to latitudinal location. No data regarding readability or subjective ratings is collected in this experiment.

4. RESULTS

In this section, we describe the analyses of the experiments described in Section 3. We analyze the correlation between screen readability and subjective ratings at various ambient light and screen brightness levels in Section 4.1. We investigate how well-suited existing default adaptive brightness system are for our sampling of users, and compare this to a calibrated model in Section 4.2. We then analyze the potential power savings that we can achieve by degrading the readability and ratings metrics by known quantities in Section 4.3. In Section 4.4, we present our analysis of the ambient light and device usage study. We then present our readability-aware brightness throttling model in Section 4.5. Finally, we give a comparison of the impact that power saving decisions made by brightness throttling systems have on readability in Section 4.6.

4.1 Screen Readability and Subjective Ratings

In Sections 3.1 and 3.2 we described the studies that were conducted on 30 users to gather data on users' readability and subjective satisfaction in these lighting scenarios. Now, we average the 30 users' data for each of the 60 ambient lighting and screen brightness combinations. We aggregate the user data study rather than deal with individuals because we want to understand how the average behaves, although individual tendencies are important in final user-facing applications [15,

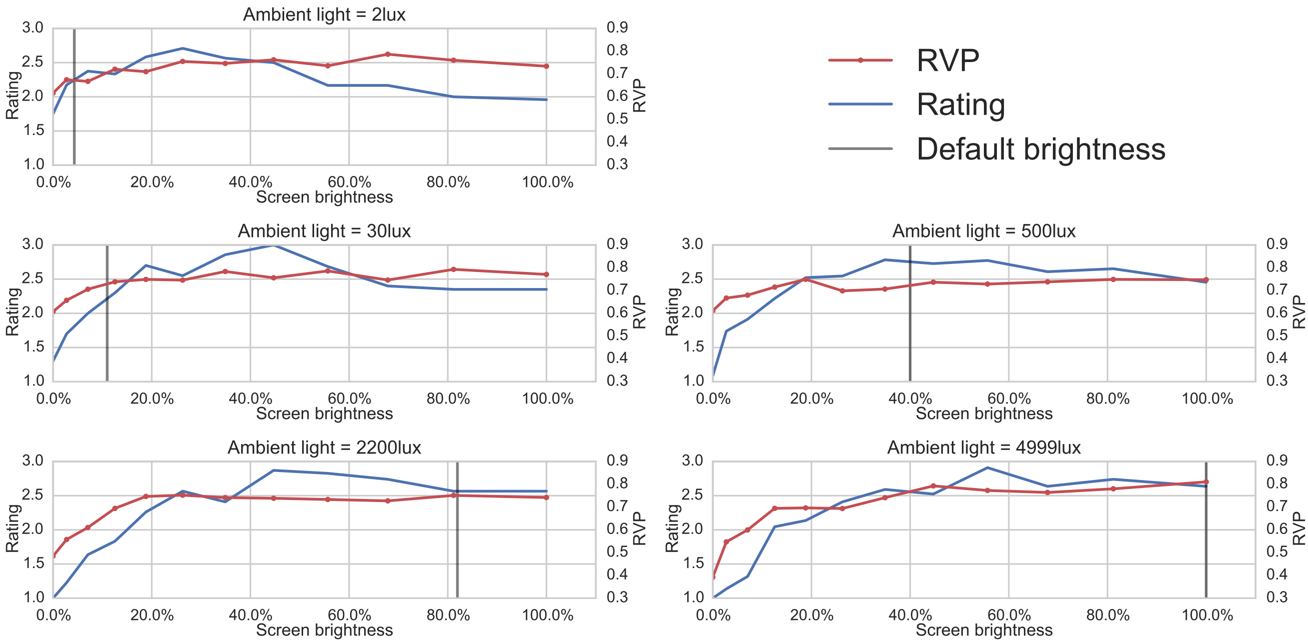


Figure 7: Comparison of RVP and subjective ratings at various ambient light and screen brightness levels.

17, 18].

The aggregated data is presented in Figure 7. The readability and ratings data are plotted on separate axis, but the axis are scaled to make them directly comparable. The readability data is presented as raw data; it is a continuous metric, and allows for presentation in this manner. The rating data, however, is nominal; to make it possible to graphically present this data, we convert the nominal ratings to numeric data before averaging each user's data. Any rating of *perfect* is given a value of 3, any rating of *slightly too dim* or *slightly too bright* is given a value of 2, and any rating of *much too dim* or *much too bright* is given a rating of 1. This allows us to not only see the relative satisfaction of the users with the screen's brightness, but also see the point at which the ratings peak; either side of this peak would be relatively too bright or too dim.

Figure 7 shows that the default adaptive brightness system doesn't align well with either RVP or the ratings peaks in these graphs. This suggests that the adaptive brightness system is not calibrated well on this device. We explore this further in Section 4.2. RVP doesn't seem to decline after a certain point like the subjective ratings do, suggesting that although users can read the detail on the device well, it is still not set to an optimal level. This is important to consider as we begin to draw similarities between RVP and the subjective user ratings. Finally, the display's brightness seems to have diminishing returns as it is increased; after a certain point, the screen has to be significantly brightened to achieve just a small improvement in ratings or RVP. This is a potential point for optimization, which we investigate further.

In addition to the overlaid RVP and ratings plots, we analyze the correlation between RVP and the subjective ratings for each ambient level. To perform this analysis, we consider each ambient lighting scenario separately. We then calculate each user's individual correlation coefficient by comparing

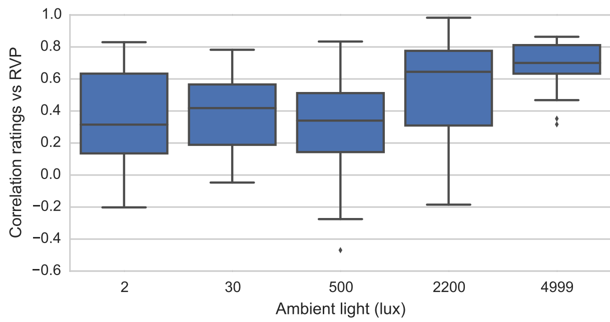


Figure 8: Correlation between RVP and subjective ratings at various ambient light levels. Correlation is computed using Pearson's correlation factor for RVP and subjective ratings at selected screen brightness levels across 30 users. Correlation is more significant in higher ambient environments

their RVP and subjective ratings at each screen brightness level and present the aggregated correlation coefficients in Figure 8. Readability tends not to worsen when the display is too bright, although the subjective ratings do degrade; because of this, we also create Figure 9, which only calculates the correlation between the lowest brightness level and the subjective ratings peak for each ambient level. As these graphs show, there is a strong correlation between the two metrics. However, the correlation is more significant at higher brightness levels. Additionally, the correlation between readability and RVP is even higher when you only consider this up to the maximal ratings point for each ambient level. Because RVP is so closely mapped to contrast ratio and luminance [8], this data suggests that it is possible to use objective task performance metrics (such as RVP, in the

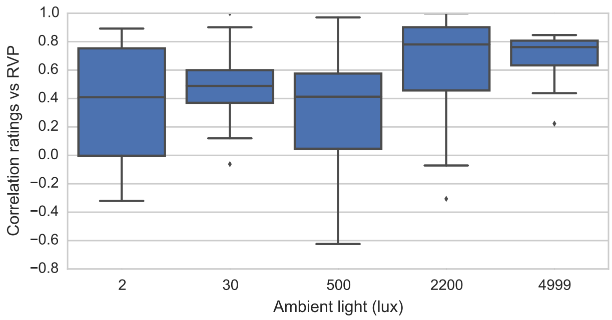


Figure 9: Correlation between RVP and subjective ratings. Instead of correlating the full range of brightness levels, each boxplot only contains data between brightness level 0 and each ambient level's ratings peak. This second analysis demonstrates stronger correlation because ratings drop when the screen is deemed to be too bright; overly bright screens do not reduce readability in acclimated eyes, however.

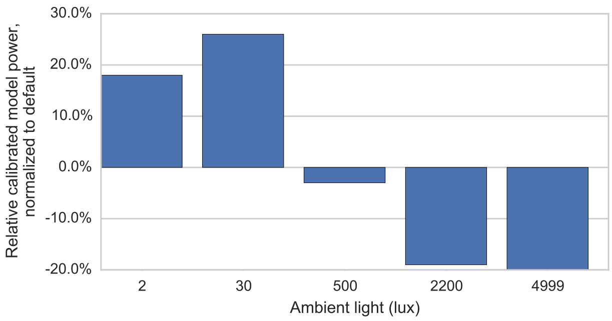


Figure 10: Relative power consumption of a perfectly calibrated adaptive screen brightness model (screen brightness at the highest average rating), normalized to the default adaptive brightness system's power consumption.

case of display brightness) as an indication of a system's effectiveness, rather than needing to rely on subjective metrics, which require direct user interaction.

4.2 Default vs. Perfectly Calibrated Power Consumption

As shown in Figure 7, there is a particular brightness level where increasing the brightness of the screen any further stops causing increased user satisfaction, and actually begins to degrade the user's subjective satisfaction with the system. This phenomenon occurs because after some optimal brightness, the screen is excessively bright for the user and actually begins to degrade their experience [12]. Furthermore, it shows that these ratings "peaks" don't actually always align with the default brightness model's brightness at that ambient light level. This disconnect between the ratings peaks and the adaptive brightness system suggests that the adaptive brightness system isn't well-tuned. This also suggests that if we were to use this "perfect" brightness model instead of the default adaptive brightness model, the display would consume different amounts of power as a result.

To explore this further, we first begin by finding the highest-rated screen brightness level for each of the ambient levels

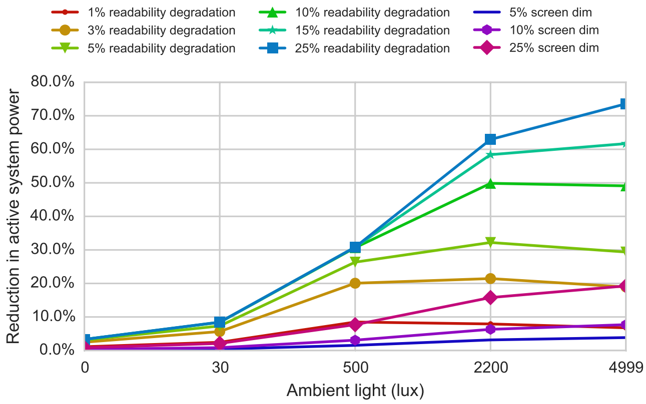


Figure 11: Amount of power that can be saved at a specific amount of readability degradation, across selected ambient light levels.

that we explore in our study. Additionally, we find the screen brightness that the default adaptive brightness system would use for each of these ambient lighting environments. Finally, we calculate the difference in screen brightness between the two and apply the power model from Figure 5, which gives us how much more or less power is consumed by the display at the various ambient light levels. This data is presented in Figure 10.

As the figure shows, users actually preferred the display's brightness to be set higher than the default model's brightness at lower ambient light levels. Additionally, users preferred the display to be dimmer than the default model at higher ambient levels.

4.3 Power Savings and Readability Degradation

In Section 2.3, we presented a series of graphs which relate OS-specified screen brightness values, measured screen brightness, and screen power consumption. These three sets of data together allow us to determine the impact on system power consumption that any change in screen brightness has. Furthermore, Section 4.1 shows us exactly how much of a change in screen brightness impacts users' RVP at a given ambient level. These individual pieces of data can be used in tandem to gauge the impact that improving or degrading screen readability has on screen power consumption.

From Figure 7, we analyze the change in screen brightness that results from degrading the 30 aggregated users' readability at a given ambient level by a known percentage. The results of this analysis are presented in Figure 11. As this figure shows, a significant amount of power can be saved at the higher ambient light levels with a relatively small readability or ratings degradation. Also, the amount of power that can be saved at the lower ambient light levels is much smaller. This trend is reflected in the work done by Stevens [19], which suggests that human perception generally does not follow a linear trend. For instance, the human eye can detect a tiny flicker of light in a dark room, but that same flicker would be imperceptible in a brighter environment. We use this result in our further analysis, which allows us to consider the human visual system in any power savings decisions we make. There is slightly less potential power savings for a given readability level at 4999 lux than there is at 2200 lux. This is likely due to the initial default adaptive brightness

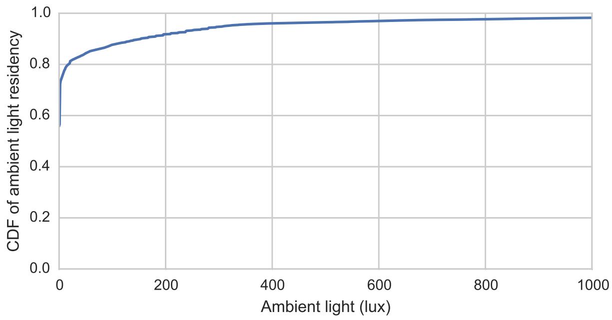


Figure 12: Amount of time that users spend in various ambient light levels with their screens active.

levels; this suggests that the 4999 lux screen brightness levels better matched user requirements than at 2200 lux, and is reflected in the slightly lower power savings.

4.4 Ambient Light Residency Analysis

In Section 3.3, we outlined a study in which we gather ambient light data from a sampling users to get a better idea of how users interact with their devices. Specifically, we wanted to discover the daily averages of how long users' screens were active and what ambient light levels the users encountered. For this analysis, we combined all of the users' ambient residency durations. We then calculated the CDF of these ambient lighting environments over time, as presented in Figure 12. We note that the daily on-screen duration average is 3.87 hours.

Dim environments ($\leq$ 30 lux) are shown to make up about 80% of the total on-screen time. Typically, only interior and nighttime environments produce this range of ambient light levels, and so this suggests that users spend most of their on-phone time indoors. Because most of the time is spent at lower ambient light levels, if you were to use the perfectly tuned adaptive brightness model as discussed in Section 4.2, the system would consume more power since this model consumes more power than the default at lower ambient light levels (and vice-versa). However, this would improve the overall satisfaction levels, effectively making this a design decision.

4.5 Readability-Aware Brightness Model

Thus far, we have presented results which characterize the relationship between user satisfaction, readability, and power consumption. In the remainder of this work, we focus on the actual design and characteristics of a dimming algorithm which is better tuned to reflect actual user readability requirements. Our end goal is to be able to maximize the time-weighted readability average over a period of time. We learned in Section 4.3 that the brightest ambient regions provide the most significant power savings per readability degradation level. Hence, in order to maximize the time-weighted readability, we will throttle the screen's brightness more significantly in higher ambient-lit environments.

We begin with our calibrated readability curve, which is a static curve of screen brightness vs ambient light. For each amount of desired power saved (we generate data between 1% and 8%), we then generate a readability-aware brightness curve. For each of the discrete ambient levels that we studied, we iteratively dim the brightest ambient regions (while enforcing monotonicity of the screen brightness) until our time-averaged power savings are met. This shape is most closely matched to a root function (and is supported by the evidence from Stevens [19]), so we generate the line of best fit through these discrete points. We present these generated curves in Figure 13.

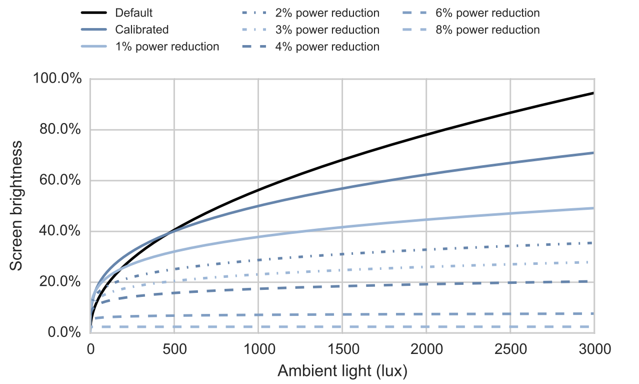


Figure 13: Comparison of adaptive brightness curves. The depicted curves include the default adaptive brightness algorithm, as well as the readability-optimized curves for selected power savings levels.

These resulting curves have a number of interesting characteristics. First of all, the dimmest ambient levels tend to actually be slightly brighter than the default algorithm, suggestion that the default brightness algorithm isn't properly calibrated. The highest power reduction curves are nearly flat; these approach the maximum power that can be saved on the screen, and so the brightness is aggressively low at all ambient levels. The less significant power throttling curves look more like a traditional root function.

These graphs represent how the proposed brightness algorithm will behave in terms of screen brightness and power savings; in the following section, we examine how they impact the time-weighted readability levels.

4.6 Power Savings and Readability Comparison

We now analyze how much readability is reduced for a given amount of power savings for a selection of brightness dimming schemes. These dimming schemes all reduce the brightness of the display in order to achieve a power reduction target. How these schemes differ, however, is how much they dim the display when they are in various ambient lighting regions. Because we are evaluating these schemes in a time-weighted fashion, a scheme which dims the display more aggressively in dim environments will have to dim the display less in brighter environments to compensate for this, and vice-versa. These schemes have been observed in various subsystems which throttle display power consumption.

- **Constant:** throttles the screen's brightness by a specified amount, regardless of the current ambient light or display brightness levels. The resulting effect is that the power is reduced by the same amount at any ambient light level. For example, this scheme might reduce a display at 100% brightness to 90% brightness, and a display at 20% brightness by the same absolute amount, to 10%. This scheme is effective when there is a constant required amount of

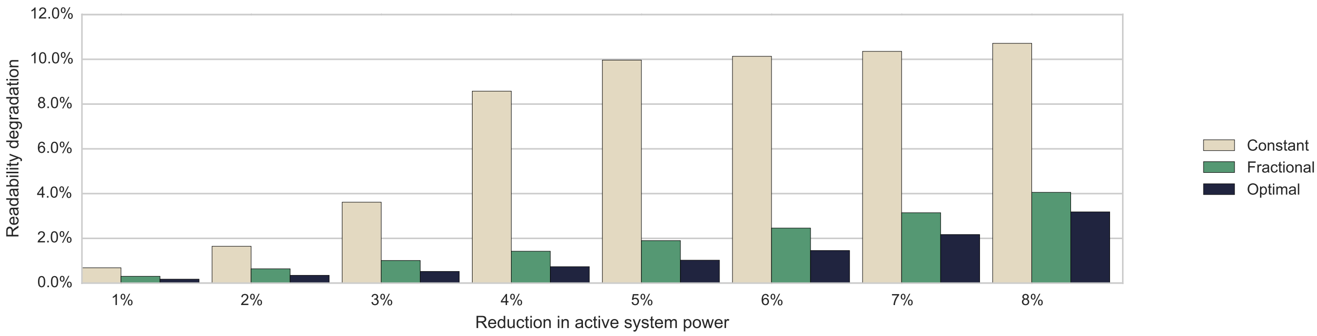


Figure 14: Comparison of the impact on readability that various dimming schemes have at selected power savings thresholds.

power savings.

- **Fractional:** instead of reducing the brightness by a constant amount, the brightness is reduced depending on the screen's starting brightness. If this scheme reduces a display at 100% brightness to 90%, it would reduce a display at 20% by the same relative percentage, to 18%. This system will reduce the power more in brighter environments, typically leading to higher power savings.

- **Optimal:** considers the user's readability curve, and first dims the regions which give the best power vs readability trade-off, as introduced in Section 4.5. This scheme maximizes the time-averaged readability at a given amount of power savings.

To perform this analysis, we determine how much the average readability is degraded at various power savings levels. We use the data from Figure 12 to apply the ambient light residencies data which gives us a time-weighted average impact on readability.

As Figure 14 shows, these schemes have significantly different time-weighted average readabilities. The constant scheme has the worse time-weighted readability average. This makes intuitive sense; since the lower ambient level regions are much more sensitive to brightness changes, a large brightness decrease in these has a huge negative impact on readability. The fractional and optimal models have better power reduction vs readability reduction performance, as they dim the displays more severely in the brighter ambient lighting environments, and vice-versa. At an 8% power reduction target, the time-weighted average readability drops by 10.7% for the constant model, 4.1% for the fractional model, and 3.1% for the optimal model, a 21.5% improvement over the fractional model. This result shows that an optimized power model is able to deliver a significantly better average user experience at a given power savings level.

Being able to directly understand the impact that a display brightness degradation decision has on the user is useful. Instead of arbitrarily making system-level decisions which directly affect the end user, these curves allow the system to first consider the user impact when throttling the device's power. Secondly, it allows for the creation of models (like our optimal model), which are optimized to maximize the user's satisfaction with the system for a given set of constraints.

It is important to note, these methods are only applicable when saving power over longer periods of time. If the system decides that it needs to reduce the amount of power being consumed immediately (e.g., a high-priority thermal emergency), it may have to do this without being able to consider the impact on readability. However, even in this situation, the system is aware exactly how much of an impact such a decision has on the user, which enables better-informed power saving decisions.

5. RELATED WORK

Finding methods of altering the display's image to allow the backlight to be dimmed with minimal impact on the user's perception is one common method of reducing LCD backlight power consumption. Chang et al. [2] present a work which allows manipulation of the on-screen image to allow the LCD backlight to be reduced while minimally impacting the user's perceived image. Ranganathan et al. [13] explore the feasibility of using *energy-aware user interfaces*, which are designed to use high-contrast colors schemes and UI elements, which allow the screen to consume less power without sacrificing readability. Shin et al. [16] explore a method of power-saving image compensation for OLED displays, rather than backlit LCD displays.

Another way of reducing display power is by improving the display's optical characteristics, allowing the display to reduce the brightness of the backlight while maintaining readability. Zhu et al. [20] present a survey of the design characteristics of transflective LCD displays, which use ambient light to increase the screen's effective brightness without using as much additional backlight power. Lee et al. [11] describe an LCD display which uses either an OLED display or a transflective LCD display, depending on how much ambient light is currently available.

Schuchhardt et al. [15] present a system which uses an online learning model with multiple contextual sources to control system brightness. The end result is a system which uses an individualized adaptive brightness model that is better tuned to user brightness requirements. However, this system does not create a readability-aware method of throttling display power, nor does it draw a correlation between subjective ratings and objective readability metrics.

Although these studies are all concerned with improving the power consumption for a given readability level, there are a number of ways that our work differs from these prior works. First of all, [2, 13, 16] involve methods of reducing the display's power consumption without the user perceiving the change; power throttling schemes which reduce screen brightness at the cost of reduced readability can be applied on top of these methods to further reduce power consumption. Also, although these studies examine either readability or subjective satisfaction, none of them attempt to draw a correlation between the two. [20, 11] are concerned with improving the display's optics in ambient light; however, any system with a dimmable backlight will still be able to apply our findings on power consumption and readability.

6. CONCLUSION

In this work, we ran a series of user studies to analyze subjective user satisfaction and objective readability metrics on mobile LCD displays in a variety of ambient lighting environments. From these studies, we were able to find a strong correlation between our subjective satisfaction and objective readability metrics; this is important when attempting to design a system which is aware of the user's satisfaction with a system, but doesn't interrupt the user for subjective ratings. We also found that the adaptive brightness systems currently in use are not necessarily well-tuned. Furthermore, from this data, we found that many brightness throttling systems do not consider the impact on the user when they are active. By taking this impact into consideration, we were able to create a display power throttling system which achieves a 21.5% higher average readability and user satisfaction at a given level of power reduction over a period of time compared to a naive solution.

Having an understanding of the impact that a user-facing system decision has on the user is important, as this makes it possible for the system to optimize these power decisions along with the user's satisfaction level. However, the best way to get this information is by directly asking the user. Unfortunately, this is an invasive process which doesn't fit well in the user's work flow. In addition, subjective ratings can be noisy. Because of this, devising objective models which are closely correlated to subjective satisfaction metrics is a promising way of including the user in the system optimization loop, allowing for the system to achieve its own power and performance requirements while maximizing the user's satisfaction in the process.

References

[1] BBC News Android application. `https://play.google.com/store/apps/details?id=bbc.mobile.news.ww`, March 2015.

[2] N. Chang, I. Choi, and H. Shim. DLS: dynamic backlight luminance scaling of liquid crystal display. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 12(8):837–846, 2004.

[3] X. Chen, Y. Chen, Z. Ma, and F. C. Fernandes. How is energy consumed in smartphone display applications? In *Proceedings of the 14th Workshop on Mobile Computing Systems and Applications (HotMobile)*, pages 21–26. ACM, 2013.

[4] CIE 145. The correlation of models for vision and visual performance. Technical Report 145:2002, Commission Internationale de l'Eclairage (International Commission on Illumination), Vienna, Austria, 2002.

[5] Google Inc. Nexus 4 tech specs, November 2014. URL `https://support.google.com/nexus/answer/2840740?hl=en&ref_topic=3415523`.

[6] Google Play Store. `https://play.google.com/store/apps/`, March 2015.

[7] M. Kalloniatis and C. Luu. Visual acuity, 2007. URL `http://webvision.med.utah.edu/book/part-viii-gabac-receptors/visual-acuity/`.

[8] E. F. Kelley, M. Lindfors, and J. Penczek. Display daylight ambient contrast measurement methods and daylight readability. *Journal of the Society for Information Display*, 14(11):1019–1030, 2006.

[9] A. Kumar, L. Shang, L.-S. Peh, and N. K. Jha. Hybdtm: a coordinated hardware-software approach for dynamic thermal management. In *Proceedings of the 43rd annual Design Automation Conference (DAC)*, pages 548–553. ACM, 2006.

[10] Larger Screens and Improved Resolution Drive Growth in Smartphone Displays, According to NPD DisplaySearch. `http://www.prweb.com/releases/2013/6/prweb10850494.htm`, June 2013.

[11] J.-H. Lee, X. Zhu, Y.-H. Lin, W. K. Choi, T.-C. Lin, S.-C. Hsu, H.-Y. Lin, and S.-T. Wu. High ambient-contrast-ratio display using tandem reflective liquid crystal display and organic light-emitting device. *Optics Express*, 13(23):9431–9438, 2005.

[12] N. Na, J. Jang, and H.-J. Suk. Dynamics of backlight luminance for using smartphone in dark environment. In *Proceedings of the International Conference on Perception and Cognition in Electronic Media (HVEI)*. SPIE, 2014.

[13] P. Ranganathan, E. Geelhoed, M. Manahan, and K. Nicholas. Energy-aware user interfaces and energy-adaptive displays. *Computer*, 39(3):31–38, 2006.

[14] M. S. Rea and M. J. Ouellette. Relative visual performance: A basis for application. *Lighting Research and Technology*, 23(3):135–144, 1991.

[15] M. Schuchhardt, S. Jha, R. Ayoub, M. Kishinevsky, and G. Memik. CAPED: Context-aware personalized display brightness for mobile devices. In *Proceedings of the 2014 International Conference on Compilers, Architecture and Synthesis for Embedded Systems (CASES)*, pages 19:1–19:10. ACM, 2014.

[16] D. Shin, Y. Kim, N. Chang, and M. Pedram. Dynamic voltage scaling of oled displays. In *Proceedings of the 48th Annual Design Automation Conference (DAC)*, pages 53–58. ACM, 2011.

[17] A. Shye, Y. Pan, B. Scholbrock, J. S. Miller, G. Memik, P. A. Dinda, and R. P. Dick. Power to the people: Leveraging human physiological traits to control microprocessor frequency. In *Proceedings of the 41st Annual International Symposium on Microarchitecture (MICRO)*, pages 188–199. IEEE/ACM, 2008.

[18] A. Shye, B. Scholbrock, and G. Memik. Into the wild: studying real user activity patterns to guide power optimizations for mobile architectures. In *Proceedings of the 42nd Annual International Symposium on Microarchitecture (MICRO)*, pages 168–178. IEEE/ACM, 2009.

[19] S. S. Stevens. To honor Fechner and repeal his law. *Science*, 133(3446):80–86, 1961.

[20] X. Zhu, Z. Ge, T. X. Wu, and S.-T. Wu. Transflective liquid crystal displays. *Journal of Display Technology*, 1 (1):15–29, 2005.