# 24.2: Estimating the Perceptual Limits of Mobile Displays

***Fahad Zafar*[†♯], *Mina Choi*[†♮], *Aldo Badano*[†]**

[†]**Center for Devices and Radiological Health, FDA, Silver Spring, MD. ♮Department of Bioengineering, The George Washington University, Washington, DC. ♯Department of Computer Science and Engineering, University of Maryland, Baltimore County, Baltimore, MD.**

## Abstract

*The quality of the displayed image in mobile devices is affected by the varying ambient illumination conditions. Determining appropriate viewing conditions for particular visual tasks requires appropriate instrumentation. However, visual tests can be more practical. We conducted experiments to explore the limitations of mobile displays in terms of the visibility of subtle targets. We explored the effect of background luminance and ambient lighting using two methodologies: a text detection task where the observer identifies characters embedded in noisy backgrounds and a threshold estimation staircase technique. We found that the text detection method holds promise as a surrogate for more complicated tests in the framework of a clinically practical implementation.*

## 1. Introduction

Mobile devices possess extensive portability incentives. Any application ported to such devices can bring maximum availability to users. Physicians could make use of these devices in emergency situations and remote environments. However, the use of these devices to interpret medical images raises many questions that must be answered before their use becomes widespread. It has not yet been demonstrated that mobile displays provide accurate and quality information to the user for correct decision making based on imaging data.

Desktop display systems have been extensively analyzed [1, 2]. Park *et al.* [3] analyzed mobile displays under different viewing conditions to show how the CIECAM02 model can be adopted under different ambient lighting conditions. The authors also showed that the screen luminance differs under different ambient conditions and proposed an enhancement model to compensate for that distortion. Kim *et al.* [4] discussed factors affecting the subjective image quality measures including clearness, naturalness, sharpness and colorfulness for liquid-crystal displays. However, to date, objective assessment of how ambient illumination conditions affect visual task performance has not yet been reported.

In this paper, we study the effects of ambient conditions and image background luminance on detection performance. We introduce a novel methodology based on embedded text detection as a practical methodology which can be applied to multiple devices and any illumination condition. To provide some validation for the novel methodology, we compare data from a staircase sensitivity experiment with results obtained using the faster text detection experiment.

## 2. Methods

All our experimental methodologies present a human subject with a detection task. All these tests were conducted using the Nexus One active-matrix organic light-emitting, mobile display device with a 3.7 in. screen with 480 by 800 pixels.

The observer experiments were conducted for three lighting conditions, bright (∼80,000 lx), average (∼450 lx), and dark (< 1 lx), and five background luminance levels. Each of the five luminance classes represent a subset of the gray levels with a class interval of about 51 gray levels each (0 - 51, 51 - 102, 102 - 153, 153 - 204, 204 - 255). All observers were seated in the particular ambient condition for 10 minutes before conducting the experiment. The set of observers included two females, ages 23 to 25 with perfect vision, and one male, age 25, with corrected vision. These fully informed observers were allowed test trials before conducting the actual data gathering experiments.

| Method | Observers | Illuminance | Trials | Backgrounds | Views |
|---|---|---|---|---|---|
| Staircase | 3 | 3 | 35 | 5 | 1575 |
| DENOTE | 3 | 3 | 10 | 5 | 450 |
| SKEWS | 3 | 3 | 10 | 5 | 450 |

**TABLE I:** Number of observations for all experiments.

Table I presents the number of total observations conducted for all the experiments presented in this paper. Three observers performed the experiments in 3 illumination conditions simulating dark, indoor office ("average") and

From the Division of Imaging and Applied Mathematics, Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, U.S. Food and Drug Administration. 10903 New Hampshire Ave., Building 62, Room 3116, Silver Spring, MD 20993, phone 301-796-2534, email `aldo.badano@fda.hhs.gov`. Fahad Zafar is a Ph.D. student at UMBC. Mina Choi is a M.Sc. student at GWU. The mention of commercial products herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services. 

ISSN 0097-966X/11/4201-0305-$1.00 © 2011 SID

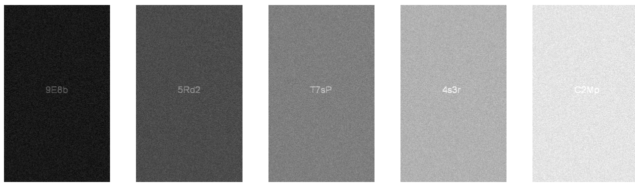



**Fig. 1:** Image samples for the DENOTE method for each of the five luminance classes.

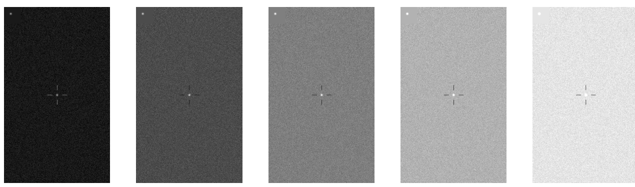


**Fig. 2:** Image samples for the staircase experiment from each of the five luminance classes. There is a reference signal on the top left while the signal in the image center is embedded within the background noise.

outside (“bright”) environment with 5 background luminance classes. The number of trials present in Table I are the repetition of one particular experiment where all other factors are kept constant. Thus the total number of observations conducted for a particular method is a multiple of values in a row of Table I. In addition we developed a Signal Known Exactly (SKE) with search (SKEWS) to test visual search tasks on mobile displays. The significance of this experiment and its results are presented in the discussion section. The techniques were implemented as an application using Java and the Android SDK 2.2 that can be ported to any tablet/mobile/device running the Android OS.

### 2.1. Noise-embedded text detection

The methodology we introduce in this paper is the DEtection of NOisy TExt (DENOTE), a practical way to run a self-contained experiment to judge the current device and ambient conditions for image reading tasks. This method is based on CAPTCHA and the Ishihara vision test [5]. CAPTCHA is a visual test used in computer security while the Ishihara test is used to check for an observer’s color blindness. In this method, a word is distorted by immersing the characters in a noisy background and adjusting the contrast appropriately (see Fig. 1). The observer is asked to identify four alphanumeric characters displayed in the middle of the screen with no time limitations. A percent correct measure is used for the analysis. This proposed method is easy to set up, provides the observer with a simple task with an execution time of less than 2-3 minutes, which is be practical for a clinical reading condition.

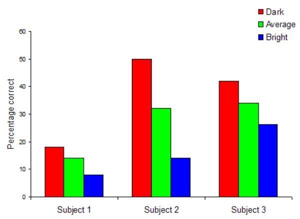
The results from this type of experiment cannot be used for statistical modeling with much confidence but will offer an approximate assessment for any mobile device in different ambient lighting conditions. For this experiment, the mobile device was not stationary; the user was allowed to move the device at will and identify the text on the screen. Four characters were displayed randomly on the screen from a dataset including the characters A-Z, a-z, 1-9. Certain characters that might cause confusion for the observer and present unnatural variability in the results, such as ’0’ and ’O’ or ’I’ and ’l’ were not used. The character amplitude chosen for the embedded text was fixed at 15 gray levels. The user response was recorded and compared to the displayed word. A simple success ratio was generated to gauge user performance. The results are assessed on a binary scale. The observer achieves success only if all four characters in the image are correctly identified. In our analysis, all the other answers are considered incorrect.

### 2.2. Threshold estimation with staircase

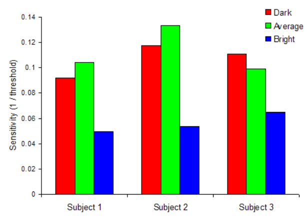
The staircase method is a well understood methodology frequently used for estimating thresholds in sensitivity experiments and detection tasks [6–8]. Using a staircase approach, we determined contrast thresholds for the same experimental conditions. The target was a Gaussian spot 20 pixels in width and height in an ascending staircase paradigm. Noise gray level values clamped to the luminance class limits were displayed using an equal probability occurrence function. Each gray level value in the respective luminance class was equally likely to appear in the background of the signal. The Gaussian signal has the highest amplitude in the center and a variance of 1.5 pixels. A cross hair around the signal and a reference signal is provided at the top left of the screen to aid in detection (Fig. 2). The signal amplitude increases or decreases depending on the response and a threshold value is calculated using the experimental data using the methodology from Ref. [8]. The mobile device was fixed in front of the observer eye for perpendicular viewing to the screen at a distance of about 30 cm. The observer was provided with a chin rest support for fixed eye position in front of the mobile device.

## 3. Results

Fig. 3 shows that user performance deteriorates as ambient illuminance increases. Users of the experiment reported that the reflectivity of the screen was a major issue for the performance drop. The observer was allowed to hold the device while performing the DENOTE experiment. Observers got the best results in dark conditions. Bright conditions which represent an outside environment, presented the highest level of difficulty. A relative performance drop exists for each observer.

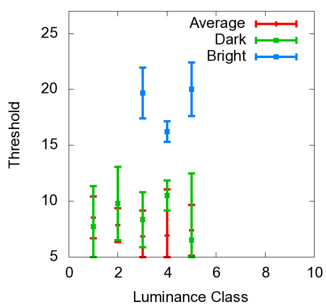



(a) DENOTE

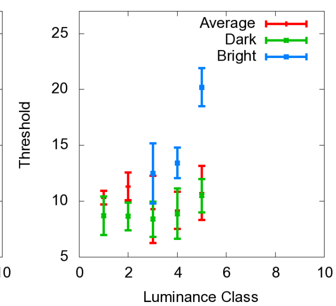



(b) Staircase

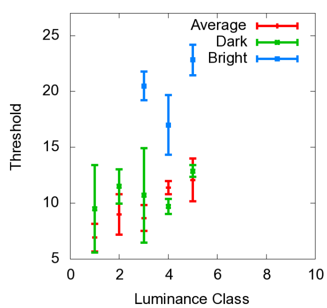
**Fig. 3:** User performance for embedded text detection and average threshold estimates conducted for 3 different ambient lighting conditions using the Nexus One.




(a) Observer 1.




(b) Observer 2.




(c) Observer 3.

**Fig. 4:** Threshold results for the staircase experiment in 3 ambient conditions for 3 observers. The error bars represent +/- a standard deviation.

The staircase results in Fig. 3 suggest that the signal detection threshold increases in a highly lit ambient environment. The mobile screen reflects (mostly specularly for this technology) the environmental lighting into the user's eye which makes it difficult to spot the signal. The subject settles on a higher threshold compared to when he/she performs the same experiment in the dark. The staircase method is time consuming and requires test trials to determine starting thresholds for different conditions. The observers were unable to locate the signal within a reasonable gray level distance for the first and second luminance class in bright conditions.

Our results indicate that 450 lx as an indoor condition does have some measurable effect on performance. Results

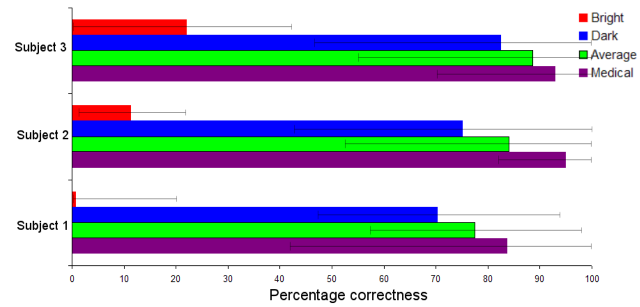



**Fig. 5:** Results from the SKEWS experiment.

for average and dark condition for observer 2 (see Fig. 4) in the second luminance class and observer 3 in the 4th luminance class suggest that image background luminance affects the visual perception at 450 lx. The trends show that the increase in ambient illumination will only increase this difference of perceptual thresholds. The observer data do not follow a particular trend per luminance class. However, the compiled results shown in Fig. 3 suggest that 2 observers have smaller thresholds in average compared to that of dark conditions (more work is needed to understand the contributions to the uncertainty estimates in this method).

## 4. Discussion

We have explored some factors that can hinder the image quality for medical imaging uses of mobile displays. We also introduced a practical test that can be used to detect the perceptual limit of a human observer under varying ambient conditions. The results unequivocally show that detection tasks are hindered to some extent by the ambient lighting conditions in mobile display devices. Each of the two experiments conducted indicate that recognizing letters or identifying signals have a similar dependence on the mobile display's response to environmental lighting. However, more evidence is needed to make a claim about the usability of the DENOTE method as a surrogate for the more accurate but time consuming methodology. This change in response can alter the decision making for an observer when performing sensitive detection tasks.

The luminance values for average and dark conditions might be too similar to extract clear trends and relationships between contributing factors in the experiments. Results for bright conditions stand out indicating that the effect of an increase in luminance to 80,000 lx causes the performance to significantly decrease.

The two methods we describe have advantages and shortcomings. The staircase experiment does not test the observer for a search task. The signal location and amplitude are known to the user through the cross hair indicator and the reference image. The embedded text experiment is a simple task to assess the use of any mobile display which come in vastly different specifications (resolution, luminance). The test is simple and can be performed quickly. Results from such a study can help identify the limitations of a certain mobile display in any condition.

Testing methodologies for mobile displays have one added complexity. If you restrict the mobile device in front of the user, it denies the user the advantage of adjusting the viewing. On the other hand, if free movement of the device is allowed for an experiment, some observers might move the device into a position where experimental factors might begin to change. Using one hand to shield light or moving the device in a position where the ambient illumination changes significantly are factors hard to control during experimentation. We did however notice that fixing the mobile device restricts many viewing angles which the observer can use to their advantage and that this might be the reason behind the clear trend present from the results of the embedded text experiment, while the staircase experiments have significant variations for dark and average conditions.

In addition to the two experiments described in this paper, we conducted a Signal Known Exactly with Search (SKEWS) task using the mobile display. The results suffered from high variability perhaps due to satisfaction of search [9]. If the attention level of an observer decreases as he or she scans the whole image, high variability is added to the experimental results. This experimental layout also suffers from a false positive and a false negative error. The user might get the number of signals right but there is no guarantee he saw the right signals to make that decision. He might be considering noise clumps as signals while actually missing the actual signal. In one of the variations of the SKE with search experiment, we allowed the user to touch the Gaussian signal on the screen for identification. This test sounds intuitive as most mobile devices have touch sensitive screens. This method added two additional variables to the experiment. First, the observers were unable to click a 10x10 signal on the screen due to parallax error while holding the device. Multiple attempts were needed for some observers who identified the spot early in the trial but took 10 seconds to touch the screen for identification affecting observer attention. The second problem was related to the highly specular screen present in the device. One touch can leave marks on the screen that need to be wiped repeatedly for every trial. We noticed that the condition of the screen has a significant role in the observer performance. Scratches and finger prints present added difficulty to the user especially when environmental lighting is present. This is one of the areas of this study that needs further research.

The results for SKEWS on a mobile device were compared to a 5-million-pixel liquid-crystal medical display. We used an EIZO RadiForce G51 with a 10-bit monochrome look-up table, a maximum luminance of 450 lx and a contrast ratio of 800. It is important to understand the performance of mobile devices in comparison to medical devices currently in use. Fig. 5 shows that user performance for all conditions was lower when compared to that of a medical device. The worst performance occurs when using the mobile device in average conditions specifically with the fifth background luminance class. The results show a substantial error for each data point (only 10 trials were conducted for each single data point). This added variability in the results might present a challenge for search tasks on mobile devices.

## 5. Conclusion

We have analyzed some of the factors that affect observer performance when using mobile displays for image interpretation. The ambient conditions clearly alter the perceptual limit and affect reader performance. Some device specific features such as the touch screen and the adjustable device brightness can also add difficulty when performing visual detection tasks. Experiments that involve signal search assessments produce results with high variability due to satisfaction of search problems [9]. Mobile displays bring added complexity to medical imaging along with their portability benefits. Our method DENOTE, a noise-embedded text detection task, shows promise as a practical image quality assessment tools for mobile displays in varying ambient illumination conditions. More research is needed to validate the methodology for different mobile display technologies and illumination conditions and to understand the additional aspects associated with search tasks.

## References


[1] A Badano, D Sharma, G Guarnieri, G Ramponi, and L Albani. Effect of case variability on the measurement of reader performance when comparing medical display technologies: application to the study of high-luminance-range displays. *Proceedings of IDRC*, 2009.

[2] MJ Flynn and A Badano. Image quality degradation by light scattering in display devices. *Journal of Digital Imaging*, 12:50–59, 1999.

[3] Y Park, C Li, Y Kwak, D Park, and C Kim. Applying CIECAM02 for mobile display viewing conditions. *15th Color Imaging Conference Final Program and Proceedings*, 2007.

[4] Y Jin Kim, M. Ronnier Luo, W Choe, H Suk Kim, S Ok Park, Y Baek, P Rhodes, S Lee, and C Yeong Kim. Factors affecting the psychophysical image quality evaluation of mobile phone displays: the case of transmissive liquid-crystal displays. *J. Opt. Soc A.*, 25:2215–2222, 2008.

[5] S. Ishihara. Tests for colour-blindness. *Handaya, Tokyo, Hongo Harukicho*, 1917.

[6] TN Cornsweet. The staircase-method in psychophysics. *The American Journal of Psychology*, 75:485–491, 1962.

[7] J Nachmias and RM Steinman. An experimental comparison of the method of limits and the double staircase-method. *The American Journal of Psychology*, 78:112–115, 1965.

[8] MA García-Perez. Yes-no staircases with fixed step sizes: Psychometric properties and optimal setup. *Optometry and Vision Science*, 78:1, 2001.

[9] S Samuel, H L Kundel, CF Nodine, and L C Toto. Mechanism of satisfaction of search: eye position recordings in the reading of chest radiographs. *RadioGraphics*, pages 895–902, 1995.