Psychophysical thresholds and digital camera sensitivity: the thousand photon limit

Feng Xiao[ab*], Joyce E. Farrell[b] and Brian A. Wandell[b]

[a] Agilent Technologies, Santa Clara, USA;
[b] Electrical Engineering Department, Stanford University, Stanford, CA, USA

ABSTRACT

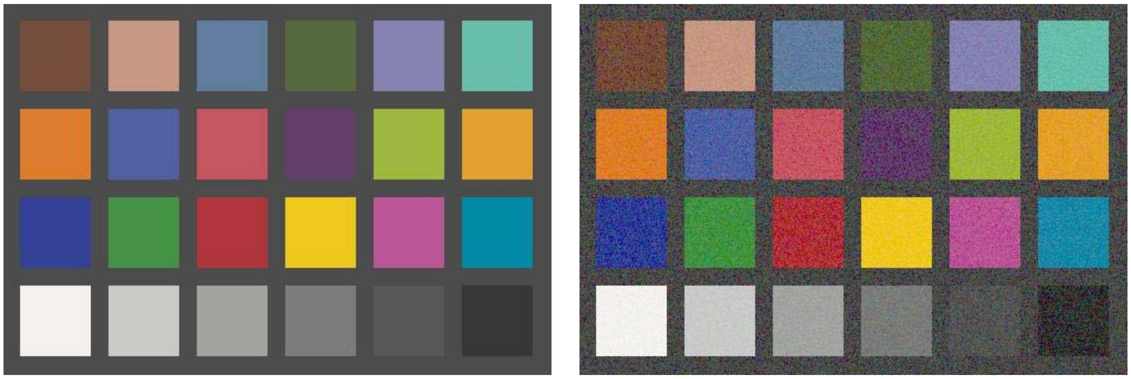
In many imaging applications, there is a tradeoff between sensor spatial resolution and dynamic range. Increasing sampling density by reducing pixel size decreases the number of photons each pixel can capture before saturation. Imagers with small pixels often operate at low irradiance levels where photon and system noise limit image quality. To understand the impact of these noise sources on image quality we conducted a series of psychophysical experiments. The data revealed two general principles. First, the luminance amplitude of the noise standard deviation predicted threshold, independent of color. Second, this threshold was 3-5% of the mean background luminance across a wide range of background luminance levels (ranging from 8 cd/m^2 to 5594 cd/m^2). The relatively constant noise threshold across a wide range of conditions has specific implications for the imaging sensor design and image process pipeline. An ideal image capture device, limited only by photon noise, must capture 1000 photons (1/sqrt(10^3) ~= 3%) to render photon noise invisible. The ideal capture device should also be able to achieve this SNR or higher across the whole captured image range.

Keywords: Psychophysical threshold, color, digital camera, dynamic range, noise sensitivity

1. INTRODUCTION

To increase the spatial resolution of imaging sensors, manufacturers typically decrease pixel size. When chip size is constant, this creates a tradeoff between resolution and dynamic range[Chen, 2000 #71]. Increasing resolution reduces the number of photons each pixel can capture before reaching saturation, thus decreasing sensor dynamic range. Natural images may then exceed the dynamic range of the sensor, resulting in saturated pixels. Further, as we explain here, reducing the dynamic range of a sensor will produce unwanted spatial noise in the dark regions of the rendered image. Figure 1 illustrates the somewhat surprising connection between pixel dynamic range and image noise. Panel (a) shows a synthetic Macbeth color chart illuminated by a D65 light; panel (b) shows a simulation of the image captured by an ideal camera whose dynamic range equals that of the scene (25:1). The ideal camera image contains no saturated components, but the noise is noticeable across the whole image. The noise is visible because at low photon catch levels the photon noise represents significant image contrast. This illustrates the principle that the dynamic range of the capture device must *exceed* the scene dynamic range; otherwise rendered images will include visible spatial noise.

There are two main challenges in determining the relationship between the design parameters of an image capture device and the visibility of noise. First, since noise is introduced at various stages of the imaging pipeline, it is important to develop simulation technologies that predict how the noise or algorithms in one part of the imaging pipeline will influence noise in the final rendered image. To understand this whole process we have developed a simulation system for the digital imaging pipeline[Farrell, 2004 #49]. This simulation tool not only models the multiple noise sources introduced at the sensor stage (photon shot noise, reset noise, readout noise, fix pattern noise and so on) but also covers the effects of most image processing functions and algorithms (such as white balance, color correction, tone-mapping and preference enhancement).



(a) (b)

Figure 1: Illustration of spatial noise caused by limiting camera dynamic range. The target is a synthetic Macbeth color chart under D65 light with scene dynamic range of 25:1. (a) Simulated image capture by an ideal camera (shot noise only) with a high dynamic range imager, and (b) an ideal camera with whose dynamic range is matched to the scene. The image captured by the low dynamic range camera never saturates, but the shot noise is visible across much of the image.

Second, because human sensitivity to visual noise depends on the spatial, temporal and chromatic properties of the spatial noise, as well as the background image that the noise is viewed against, it is important to carry out experiments to measure the visibility of noise pattern as a function of these properties. Previous psychophysical studies mainly used noise patterns as masks for the detection/discrimination of regular patterns, such as sinusoid patterns[Meeteren, 1988 #75; Blackwell, 1998 #74; Gegenfurtner, 1992 #76; Klein, 1997 #77; Eckstein, 1996 #78]. The importance of the noise visibility on overall digital image quality warrants studies in which the noise pattern is the target instead of the mask. In one of the few studies to measure the visibility of noise patterns, Winkler et.al[Winkler, 2004 #73] used a set of grayscale natural images as the background and investigated the effects of noise filtering and envelope on the noise threshold. The noise threshold shows a similar trend for varies types of spatial noise filters as the contrast sensitivity measurements using sinusoid patterns. It also shows that image structure increases the noise threshold significantly and the noise threshold for a uniform background is the lowest among all backgrounds. These results are a good first step towards the understanding of human visual sensitivity to noise. However, there are several important issues that need to be addressed before we can build a visual noise sensitivity model to guide the design and optimization of digital imaging devices. First, we need a device-independent noise threshold definition that links the human visual system with digital imaging devices. Second, we need to measure the chromatic effect on noise threshold (individual color channels and interactions between color channels). Third, we need a systematic study of the effect of absolute background luminance level on the noise threshold. Most commercial monitors available today have a limited output luminance range and do not have enough resolution to measure noise thresholds across a wide dynamic range of backgrounds (on very dark backgrounds, for example).

In this paper, we build a link between the human visual sensitivity to noise and the noise property of digital imaging devices. First, we describe the construction of a high dynamic range display device with consistent high accuracy for noise threshold measurements across a wide range of luminance output levels. Second, we design psychophysics experiments to measure how the luminance level of background, color and surroundings affect the noise threshold. Third, we show how these results can be used in the design and optimization of digital imaging devices.

2. HIGH DYNAMIC RANGE DISPLAY DEVICE

One challenge we faced in measuring noise visibility accurately is that most commercially available display devices, such as CRT and LCD monitors, do not permit fine control of the noise contrast across a wide range of mean levels. For example, to measure the noise sensitivity at very dark levels (mean digital output at 1), the minimum step is far too coarse. In addition, the maximum luminance output level is quite limited (80 cd/m^2 for typical CRT display and 250 cd/m^2 for typical LCD display). To overcome these problems, we built a high dynamic range display device. The device we built* uses a DLP device to project an image onto the back of a LCD panel. The DLP projector has a maximum luminance output of 2000 lumens and a native resolution of 1024 by 768. The LCD panel, which was stripped from a 12-inch LCD monitor, also has a native resolution of 1024 by 768. A Fresnel lens was placed between the DLP and LCD panel to improve the spatial uniformity of the output and to direct the light over a narrower viewing angle. Several narrow angle holographic diffusers were attached to the back of the LCD panel to increase the maximum luminance output and also improve the spatial uniformity of the output. The color wheel was removed from the DLP projector to improve the maximum luminance. Hence, the DLP generates grayscale images $I_{DLP}(x,y)$ while the LCD generates full color image $I_{LCD}(x,y)$. A simplified form of the final output image I(x,y) can be expressed as:

$$I(x,y) = I_{DLP}(x,y) \times I_{LCD}(x,y) = I_{DLP}(x,y) \times \begin{bmatrix} R_{LCD}(x,y) \\ G_{LCD}(x,y) \\ B_{LCD}(x,y) \end{bmatrix} \qquad (1)$$

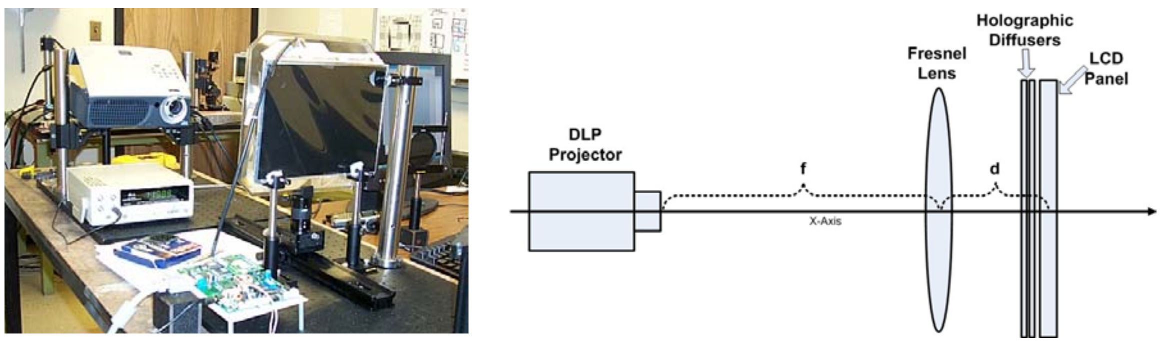


Figure 2: The prototype high dynamic range display device consisted of a DLP projector and an LCD panel. Together these devices create images with a mean intensity ranging from 1 cd/m^2 to 10,000 cd/m^2. The quantization steps are smaller and more uniform across this range compared to commercial displays.

The combined output from these devices forms an image with a large dynamic range (1 cd/m^2 to 10,000 cd/m^2) and fine control of quantization across a wide range of output levels. The colorimetric values of the display white point and primaries are listed in Table 1. The LCD panel has a near linear gamma curve across most of its range and this simplifies the noise generation for later experiments.

Table 1: CIE Yxy values of the display white point and primaries. Y (luminance) in cd/m^2

White	Red	Green	Blue
9962, 0.32, 0.35	2343, 0.55, 0.36	6026, 0.35, 0.55	1593, 0.15, 0.15

3. GENERAL METHODS

Sensitivity to visual noise depends on the nature of the noise pattern as well as on the background image that the noise is viewed against. It is important to design experiments to measure the visibility of noise patterns as a function of these properties. In a practical imaging device, noise is generated by different sources, each of which can be approximated by spatially independent white Gaussian noise[Tian, 2001 #65]. Noise thresholds measured for spatially independent white Gaussian noise patterns can, therefore, predict the visibility of these noise sources. To simplify the experimental design and data analysis, we used a uniform gray image as the background for a Gaussian noise target. In section 5, we showed how to apply these results for images with more complicated backgrounds

5.1 Stimulus Generation

We used the DLP image to change the overall background luminance level (either uniform or varying slowly) and the LCD image to generate noise pattern and color formation. This arrangement minimizes the error incurred by the misalignment between the DLP and LCD images. Equation 1 can be further simplified as:

$$I(x,y) = I_{DLP} \times \begin{bmatrix} 128 + N_r(0,\sigma_r) \\ 128 + N_g(0,\sigma_g) \\ 128 + N_b(0,\sigma_b) \end{bmatrix} \tag{2}$$

Where N(0,σ) is the spatially independent white Gaussian distribution. The mean level of the LCD image was set to 128 in order to maximize the accuracy of threshold measurement. To generate stimuli with different background luminance levels, only the DLP image was changed. To generate stimuli with different color mixings and correlation, we only changed the LCD image. The relatively linear gamma curve of the LCD device across its whole range simplified the noise generation.

In each trial, two stimuli (one was a uniform disk and the other one with noise superimposed) were displayed side by side (10 degrees apart) on the front panel of the high dynamic range display. Each stimulus spanned a spatial angle of 10 degrees (or 250 pixels by 250 pixels) To minimize the border effect of sharp edges on the noise detection, the noise amplitude was modulated by a Gaussian envelope with kernel size (or standard deviation) of 2 degrees (or 50 pixels) A cross pattern subtending half a degree of visual angle was superimposed at the center of each stimulus to help observers locate the stimuli (Figure 2).

Figure 3: A pair of stimuli was displayed side by side: one was a uniform disk and the other one with noise superimposed.

5.2 Experimental Procedure

The experiments were controlled in Matlab using extensions of the Psychophysics Toolbox[Brainard, 1997 #56]. The room light was off and stray light reflection was minimized throughout the experiments. The stimuli were displayed on the front panel of the high dynamic range display. Three unpaid observers (one female and two males, ages from 22 to 50 years old) with normal color vision participated in the experiments. Two of the subjects had extensive experience in psychophysical experiments. The viewing distance was 40 cm and a chin-bar was used to support subjects' heads. During each trial, observers pressed a button to indicate which location contained the Gaussian noise. Subjects were given feedback about the correctness of their response after each trial. A two-up-one-down staircase[Watson, 1983 #67] was used to control the noise level σ for each trial and the psychometric (Weibull) function determined the noise threshold level required for observers to identify 82 percent of the trials correctly:

$$p = 0.5 + 0.5 \times (1 - e^{-\frac{\delta}{\alpha}})^{\beta} \tag{3}$$

Subjects were given sufficient time to fully adapt to the background luminance level before each session and they were also allowed to take as many breaks as they want during each session. A total of 160 trials were run for each session and the entire experiment took two to three hours to complete.

4. EXPERIMENTS AND RESULTS

4.1 Experiment one: effect of background luminance level

For a typical imaging capture device, the amount of noise typically depends on number of photons it received or the local luminance level in the captured image. So the first question is how the background luminance level affects the noise threshold. The high dynamic range display device enabled us to measure noise thresholds across an enormous range of background luminance levels. In the first experiment, the noise was achromatic ($N_r(0,\sigma_r)$= $N_g(0,\sigma_g)$= $N_b(0,\sigma_b)$ in Equation 2) and the DLP image was varied to achieve nine different background luminance levels from 8 cd/m^2 to 5594 cd/m^2. There are many ways to define the noise threshold. Ideally, the noise threshold should be simple and independent of the device specific values (such as the device RGB values). In line with the definition of signal noise ratio (SNR) in engineering practice, we picked a candidate noise threshold as:

$$T = \frac{I_{DLP} \times \sigma_L}{I_{DLP} \times Y} = \frac{\sigma_L}{Y} \tag{4}$$

Where σ_L is the threshold (Equation 3) of the standard deviation of the luminance channel of the noise image and Y is the average luminance level of the image. For achromatic noise, the standard deviation of luminance channel can be derived from the standard deviation of the three color channels as:

$$\sigma_L = w_r \times \sigma_r + w_g \times \sigma_g + w_b \times \sigma_b \tag{5}$$

where [w_r=0.2352, w_g=0.6049, w_b=0.1599] are the luminance weighting factors for each color channel.

Figure 4 plots the relative noise threshold as a function of background luminance levels. The relative noise threshold is roughly constant (3-5%) across a wide range of background luminance levels, ranging from 8 cd/m^2 to 5594 cd/m^2. This is consistent with other detection/discrimination experiments that find that contrast (defined as the ratio of the difference between brightest and darkest parts over the mean background level) stays relatively constant over a smaller range of background luminance levels. Our results extend these findings to a much larger range of background luminance. As expected, the two experienced subjects have lower thresholds than the inexperienced subject.

4.2 Experiment two: effect of color

The results of the first experiment show that detection thresholds for achromatic noise patterns are relatively constant for a wide range of background luminance. Since most image capture devices have multiple color channels, in the second experiment we investigated the effects of color on the noise threshold. Luminance noise threshold for three types of color noise (red, green and blue) at two background luminance levels (51 cd/m^2 and 5594 cd/m^2) were measured for the same subjects. For comparison, we plot noise thresholds defined in the CIE X and Z values in Figure 5. It is very clear that target noise threshold measured in luminance (CIE Y value) is nearly constant while the CIE X and Z values vary. The results are similar at two other background luminance levels.

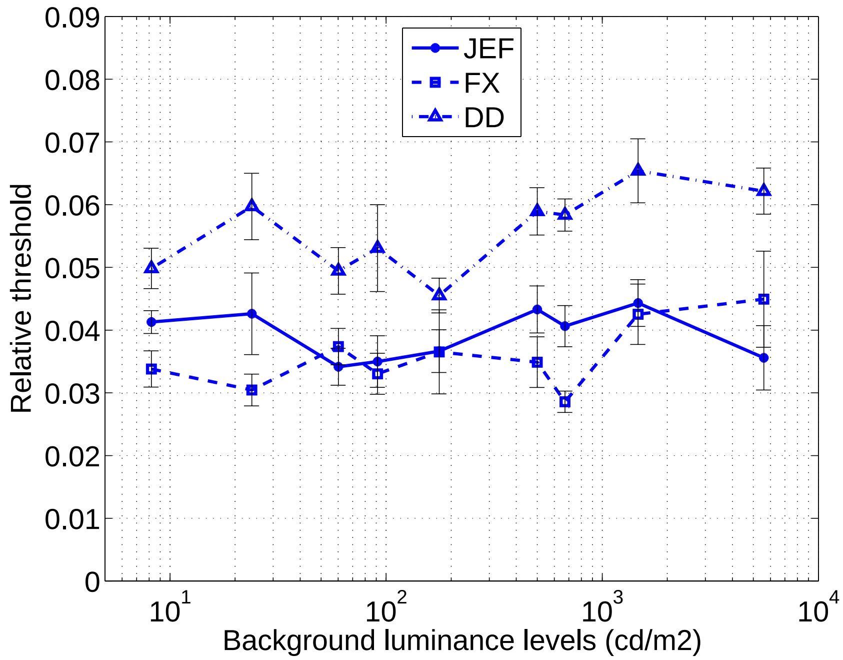
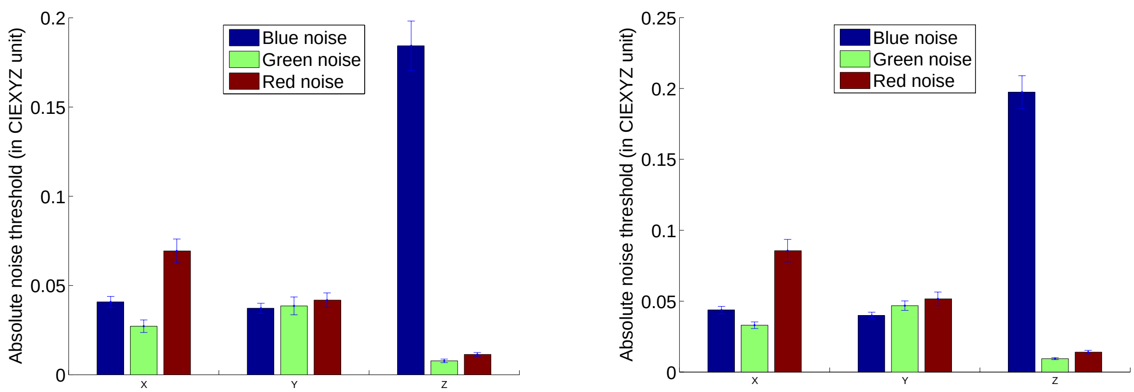


Figure 4: The effect of background luminance on the relative noise threshold.



(a) (b)

Figure 5: Average noise thresholds defined in CIE X, Y, Z values across all subjects: (a) background luminance at 51 cd/m^2; (b) background luminance at 5594 cd/m^2.

4.3 Experiment three: interaction between color channels

In the second experiment, noise thresholds were measured for noise patterns presented in individual color channels. In practice, noise includes contributions from multiple color channels. For example, color noise is introduced by demosaicing and color correction. In this experiment, we determined whether the correlation among color channels affects the noise threshold. For stimuli with correlated color noise, noise patterns from three color channels were identical and thus appeared achromatic. For stimuli with uncorrelated color noise, noise patterns from three color channels were statistically independent and thus appeared chromatic. Note that for uncorrelated color noise case, the calculation of noise threshold is now different from Equation 5:

$$\sigma_L = \sqrt{w_r^2 \times \sigma_r^2 + w_g^2 \times \sigma_g^2 + w_b^2 \times \sigma_b^2} \tag{6}$$

Figure 6 shows the average noise thresholds across all subjects for two different background luminance levels. The noise thresholds were very similar for both correlated and uncorrelated color noise although the noise standard deviation [σ_r,σ_g,σ_b] is higher for the uncorrelated color noise. This can be predicted by the difference between Equation 5 and Equation 6.

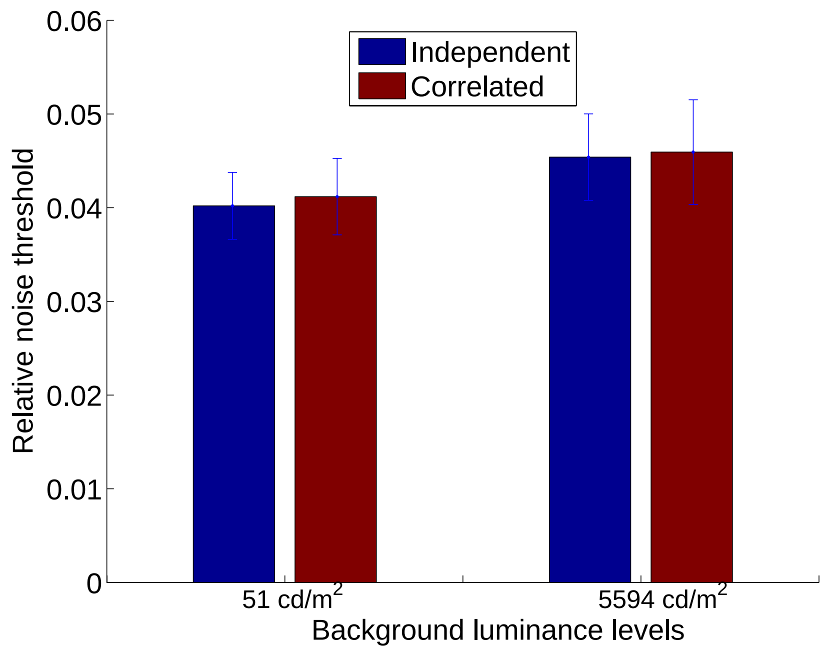
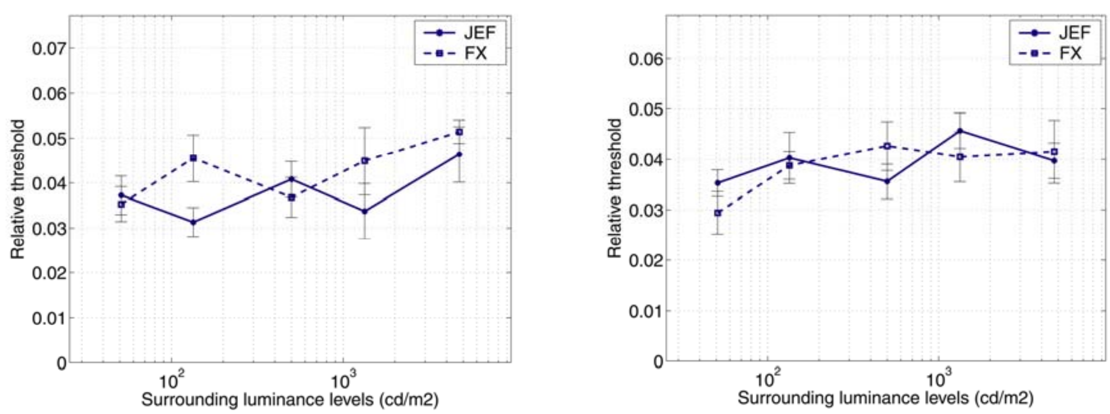


Figure 6: Noise thresholds were similar for uncorrelated noise (appears achromatic) and correlated color noise (appears achromatic).



(a) (b)

Figure 7: Effect of surrounding luminance level on the noise threshold for center target area luminance level at (a) 51 cd/m^2 and (b) 500 cd/m^2.

4.4 Experiment 4: effect of surroundings

In this experiment we measured the effect of a non-uniform background on noise thresholds. The background images were divided into a center area (20 by 10 degree) and a surrounding area. The noise target was presented in the center area. Figure 7 shows the noise threshold for two center luminance levels

(51 and 500 cd/m^2) over a range of surrounding luminance levels. The luminance levels of surrounding area do not affect noise threshold substantially.

4.5 Summary

These experiments have shown that contrast threshold (standard deviation of noise in luminance channel divided by the mean background luminance level) is a robust measure of noise sensitivity. The noise contrast threshold is relatively constant across subjects, color content, wide range of background luminance levels (8 cd/m^2 to 5594 cd/m^2) and surroundings luminance levels. The interaction of noise from different color channels is predicted by the luminance noise component. In summary, noise is barely visible at contrast levels below 3% for all subjects.

5. APPLICATIONS AND DISCUSSION

We measured noise thresholds on uniform gray backgrounds instead of complex images to simplify the experimental design and data analysis. A natural question is how to apply these results to real images with much more complicated backgrounds. First of all, most complex images can be segmented into local regions with relative uniform background. Then for each local region, the ratio of luminance noise over background can be derived from the color data and compared with the 3% noise threshold. Even for regions with significant activity, the noise threshold for uniform backgrounds can be used as a lower bound of noise visibility since image activity would only increase the noise threshold as reported in other studies [Winkler, 2004 #73]. In engineering terms, the corresponding SNR threshold for 3% noise threshold will be around 30 db (~20 $\log_{10}$ (1/0.03)). The relatively constant noise threshold across a wide range of conditions has specific implications for the imaging sensor design and image process pipeline.

5.1 The thousand-photon limit and dynamic range requirement for single capture device

Even for an ideal capture device with only photon shot noise, the noise threshold of 3 percent implies that the minimum number of photon absorptions per pixel should exceed 1000 (i.e.,$1/\sqrt{1000} \approx 0.03$). In terms of SNR, this is equivalent to requiring that image captured in the dark portion of a scene should exceed 33 db (~20 $\log_{10}$ (1/0.03)). The psychophysical results show that when the mean number of photon absorptions in a region of the image is below this level, the spatial pattern of photon noise may be visible. As pixel size shrinks, designers need to account for this thousand-photon limit at the low end.

At the high end, a traditional single capture device must be able to capture the full scene dynamic range without saturation. For typical natural images, we have measured this range to be on the order of 1000:1, excluding specularities[Xiao, 2002 #43]. Hence, the total well capacity of an ideal system should be on the order of 10^6 electrons.

In summary, an ideal single capture camera must be designed to capture 10^3 photons in the dark part of an image to avoid visible photon noise. The pixel must be able to capture 10^6 photons to encode the dynamic range of natural images. These are the basic constraints for an ideal camera that can render the vast majority of natural images with no visible noise and no saturated pixels. For real cameras, the requirement can be even higher due to the addition of electronic noise and color process functions (such as color correction). However, aggressive noise reduction algorithms might be able to suppress certain amount of the noise.

5.2 Optimal exposure scheme for multiple capture device

These dynamic range requirements for single capture device to record typical natural scenes without visible noise are very demanding. The relatively constant noise threshold across a wide range of conditions indicates that the optimal solution would be a capture system being able to hold SNR below 33dB across the entire captured image. In practice, there are many different ways to achieve this goal. The logarithmic ADC scheme has the potential to achieve this goal in a single capture[Ricquier, 1995 #69] and the self-reset pixel technology[Liu, 2002 #84] is another promising technology. However, the multiple-capture-single-image (MCSI) technique is arguably the most widely used method[Yang, 1999 #10; Wandell, 2002 #18; Chen, 2002 #72; Liu, 2003 #80]. How to find the optimal capture time scheduling (number of exposures and length of each exposure) for a specific scene using MCSI technique is still an open question. The capture time scheduling algorithm proposed by Chen and colleagues[Chen, 2002 #72] tries to maximize the average SNR of the whole image instead of achieving the target local SNR suggested by our experiment results. A better MCSI capture time scheduling for a scene with N objects at different luminance levels would be to minimize the following cost function so that the local SNR is always below 33dB:

$$C = \sum_{n=1}^{N} f(33 - SNR_n) \tag{7}$$

where

$$f(x) = \begin{cases} x & if\ x > 0 \\ 0 & otherwise \end{cases} \tag{8}$$

Figure 8 illustrates an example of the proposed capture time scheduling method. For each input luminance level, the broken line represents the maximum exposure time before saturation and the solid line represents the minimum exposure time to reach SNR of 33dB. Therefore, the exposure time for any input luminance level to achieve SNR of 33dB lies between these two lines. The solid line can also be seen as the most efficient exposure time scheduling if self-reset pixel technology is used. For MCSI, the exact exposure durations are the starting points of the horizontal line (three captures needed in this case). This method can easily be adapted to SNR levels other than 33 dB.

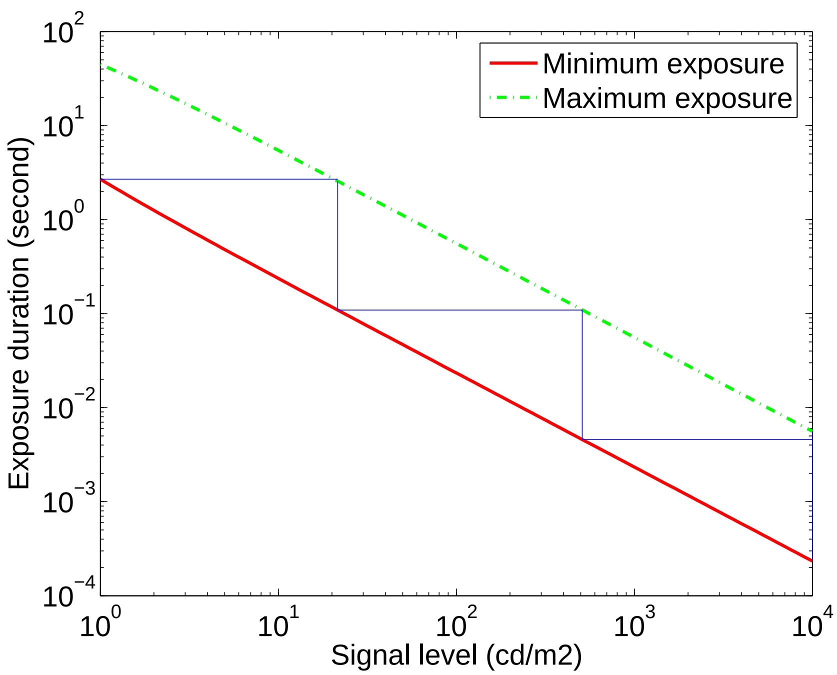


Figure 8: An example of the proposed capture time scheduling method for multiple-capture-single-image (MCSI) systems. It tries to achieve the target local SNR of 33:1 instead of maximizing the average SNR of the whole image. Please see text for detailed descriptions.

5.3 Discussion and future directions

There are a few things that need to be addressed in the future. First, the noise threshold for uniform backgrounds only provides a lower bound on the visibility of noise. A thorough investigation of how more complex backgrounds increase the noise threshold will further help the design of digital imaging device. Second, a spatially independent Gaussian noise target can provide a first-order approximation of other noise sources. More accurate predictions will require the investigation of spatially correlated noise (such as the structured noise introduced by noise reduction algorithms or compression). Third, the overall image quality not only depends on the visibility of noise but also on other image attributes such as sharpness and color saturation. In many cases, tradeoff has to be made to achieve optimal image quality[Barnhofer, 2003 #85]. Lastly, the 3 percent threshold might be too demanding for practical applications (digital photography for example) where a higher level of noise can be tolerated. It is worthwhile to investigate the relationship between the amount of noise and the perception of image quality, where 3 percent noise contrast defines the ideal or perfect image.

ACKNOWLEDGEMENTS

We thank Dr. Peter Catrysse and Professor Abbas El Gamal for useful discussions and Angela Chau, Ulrich Barnhoefer and Ian McDowall for help on the high dynamic range display device.

6. REFERENCES