Advanced Image Quality Studies of LCTVs

Justin Laird

B.S. Imaging and Photographic Technology
Rochester Institute of Technology (2003)

A thesis submitted in partial fulfillment of the requirements for the degree of Master's of Science in Color Science for the Center of Imaging Science at Rochester Institute of Technology.

September 2005

Signature of the Author

Accepted by Academic Coordinator, Roy S. Berns
M.S. Degree Program

Chester F. Carlson Center for Imaging Science
College of Science
Rochester Institute of Technology
Rochester, New York

Certificate of Approval

M.S. Degree Thesis

The M.S. Degree Thesis of Justin Laird has been examined and approved by three members of the color science faculty as satisfactory for the thesis requirements for the Master of Science Degree

Dr. Jeff B. Pelz, Thesis Advisor

Dr. Mitchell R. Rosen, Thesis Advisor

Dr. Ethan D. Montag, Thesis Advisor

Thesis Release Permission Form

Chester F. Carlson Center for Imaging Science
College of Science
Rochester Institute of Technology
Rochester, New York

Title of Thesis: Advanced Image Quality Studies of LCTVs

I, Justin L. Laird, hereby grant permission to the Wallace Memorial Library of Rochester Institute of Technology to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or for profit.

_____
Signature of the Author

_____
Date

Advanced Image Quality Studies of LCTVs

Justin Laird

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Color Science in the Center of Imaging Science, Rochester Institute of Technology.

Abstract

LCD (liquid crystal display) televisions (LCTVs) are one of the more popular choices for flat panel displays and their popularity will continue to increase as they become more economical. They hold advantages over counterparts including sharper images and higher achievable luminance levels. However, differences between LCDs and the previously dominant CRTs raise important questions. For example, weaknesses inherent in LCDs produce apparent blur in moving objects. Also, recent commercial LCTVs are larger and brighter than traditional televisions. Reported in this thesis is fundamental research targeted at learning about human sensitivity to sharpness for objects in motion. Also experiments testing observer tone scale preference on large, bright displays are described. The research in this thesis was divided into two projects.

Part I of the thesis describes Project 1 where the spatiovelocity contrast sensitivity function (CSF) was investigated. There were a total of five psychophysical experiments. Results included evidence that contrast sensitivity does not change between a constantly drifting grating and a static pattern drifting across the observers' view. Further, it was found that contrast sensitivity to sine wave patterns in motion but stable on the retina and to those stationary on a screen were the same. A 2D spatiovelocity CSF model was modified and tested. The model successfully predicted empirical data.

Part II of the thesis covers Project 2 consisting of a pair of experiments that explored how image quality is impacted by the electro-optical transfer function, maximum screen luminance level and ambient illumination level. Results indicated that, in general, a gamma of 1.6 was most preferred, more so for bright displays. As surround illumination moved from dark to dim, preference to 1.6 gamma was enhanced. Preference for transfer functions was found to be dependent on the display intensity and that this dependence is maintained under natural viewing conditions with a dim surround.

Acknowledgements

I would like to first thank the Lord Jesus Christ who gave me the strength to get through the past two years and blessing me with the following people in my life. Next, I would like to thank Mitchell Rosen for his patience, advice and ceaseless efforts in teaching me what it means to be a scientist. He has been a friend and a great role model for me – thank you Mitch. I would like to thank Jeff Pelz for his advice in matters of school and life and the help he has provided along the way. I have learned a great deal in the past two years and both he and Mitch have had the unfortunate luck to be in the forefront of this.

Thanks to Ethan Montag for his patience and suggestions along the way to keep my research and this thesis as simple and concise as possible.

I would like to thank Lawrence Taplin and Garrett Johnson for their advice on programming and putting up with my stupid questions.

I truly consider myself blessed to have worked for the entire faculty, staff, students and visiting scientist at the Munsell Lab. Thanks to you all for the good times, suggestions and time put in on my behalf.
A special note of thanks to Val Hemink and Colleen Desimone for all the time, advice and help along the way.

I am forever in debt to my friends here in Rochester that have supported me in so many ways. Thanks to Gabe, Dan, Tim, Ferg, and Nate for all the advice, help and support – I could not have made it for 6 years without it and I only hope I can return the favor someday.

I am very grateful to my family for their patience and support. A very special thanks to my Mother for all her support and advice my whole life. I would not be here if it were not for her patient guidance during the last 29 years.

Kathy, Tom and Tawni: I would especially like to thank you for your belief in me, the support and the tireless sacrifice you have made on my behalf. Thank you for everything and I hope I can someday repay you for what you have done.

Thanks to Sharp Labs America who funded this project, in particular thanks to Scott Daly and Xiao-Fan Feng for all their great advice and valuable interaction to make this thesis a success.

Thanks to you all – I could not have done this without a single one of you.

Cheers!

Justin Laird

Table of Contents

List of Figures

List of Tables

Part I – SpatioVelocity CSF

Chapter 1 - Introduction

The motivation for the work described here as Project 1 was to understand how observers' contrast sensitivity changes with retinal velocity. If an object is perceived as having an unacceptable amount of blur when stationary then that object might be deemed acceptable if the image of the object moves across the retina. However, if observers track the object with their eyes making the retinal image stationary then the object can appear unacceptably blurred once again. This is because the perceptual threshold for blur tolerance increases for images in motion across the retina. An understanding of the interaction between image quality and motion blur, and the temporal characteristics of LCTVs may lead to effective algorithms for reducing motion artifacts. Motion artifacts in LCD televisions currently on the market are caused from pixel update rates that are much slower for LCDs than for CRTs. These concerns will be described in more detail later, but the artifacts can cause noticeable blur in moving images.

While there have been different solutions to rectify the motion blur problem[1], one possibility is to use algorithms that can predict where the observer will look and make adjustments to the image based on this knowledge. One requirement for this to occur is to have, a robust model of motion sensitivity. This model could report to the algorithm observer sensitivity based spatial frequency and velocity of image content. Then the algorithm could predict where observers are more likely to look based on this information.

Much of our understanding of the visual system's sensitivity to objects in motion is based on the extensive literature using temporally counterphase flickering gratings.[2] These have been used in building 2D spatiotemporal contrast sensitivity functions (CSF). These are dependent on particular stimulus conditions such as average intensity, stimulus duration, spatial configuration, etc. Flickering is not the best method to test motion sensitivity since most objects in motion do not flicker. However, an object in motion has a temporal and spatial component. If an image of an object is in motion across the retina then each receptor is subject to a temporal frequency that is a function of the image's velocity and spatial content. So it is more important for this application to test the sensitivity to stimuli in motion rather than flickering in space.[3, 4, 5] Eye tracking is an important tool of the experimental design used to determine the actual retinal velocity of a stimulus.

In Part I of this thesis a series of experiments are described that incorporate eye-tracking in the psychophysical determination of spatiovelocity contrast sensitivity. An exploration of whether the velocity of the eye has an impact on its sensitivity is carried out and whether a CSF model can be used to predict sensitivity to moving edges at different velocities and contrasts.

Furthermore, the problem of a natural experimental setup was addressed in this research. This important and practical issue has implications on the interpretation of results. Although the experimental setup, described in later sections, are by no means "natural", the use of non-stabilized retinal images, eye movements, and velocity versus flickering all serve to build a model that incorporates more natural aspects with the hope that this model can be applied to motion imagery viewed on LCD televisions.

Chapter 2 - Background

2.1 – Eye Movements

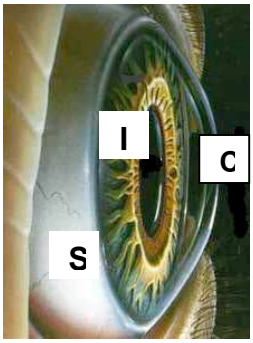


Figure 1. The eye: C - Cornea, I - Iris, S – Sclera.[6]

The eye is an extremely complex set of muscles, tissues and neurons and is essentially an extension of the brain. The image in Figure 1 shows some important features of the outer eye relative to this research. The sclera (S in the image) is the white portion where tendons for moving the eye are attached. It also gives the eye its round shape. The cornea is labeled "C" in Figure 1 and is the outermost surface of the eye. This is where light rays are first refracted in order to bring light from the world into focus on the retina. The iris, labeled "I", controls the amount of light passing into the eye. The iris is a sphincter muscle that contains the pigment responsible for eye color. It controls the size of the pupil, which is the opening inside the iris. Other major sections of the eye important to this research are the retina and the fovea, circled in green in the Figure 2. The retina accommodates two types of light receptors used to transmit information from the world to our brain. The receptors are called rods and cones. The majority of the receptors are rods. These receptors are most sensitive to low light levels and do not allow color discrimination. Cones on the other hand are sensitive to higher light levels and

provide all the color information to the visual system. There are three types of cones: those sensitive to long wavelength light, middle wavelength light and short wavelength light, (roughly equivalent to red, green and blue portions of the energy spectrum). The vast majority of cones are located in a central part of the retina known as the fovea. As seen in the Figure 2, the fovea is a portion of the retina almost directly behind the pupil and is the center of gaze; meaning when a person is staring (or fixating) on an object, then an image of the object is focused onto the fovea. Because of the tightly packed cones in the fovea, humans have the highest spatial acuity and color vision in this region. Spatial resolution of the receptors decreases as distance from the center of the fovea increases. The graph on the right in Figure 2 shows the density of the cones as a function of retinal location. The red portion of the graph on the right is a histogram of cone density and the central spike indicates the majority of cones lie in the fovea.

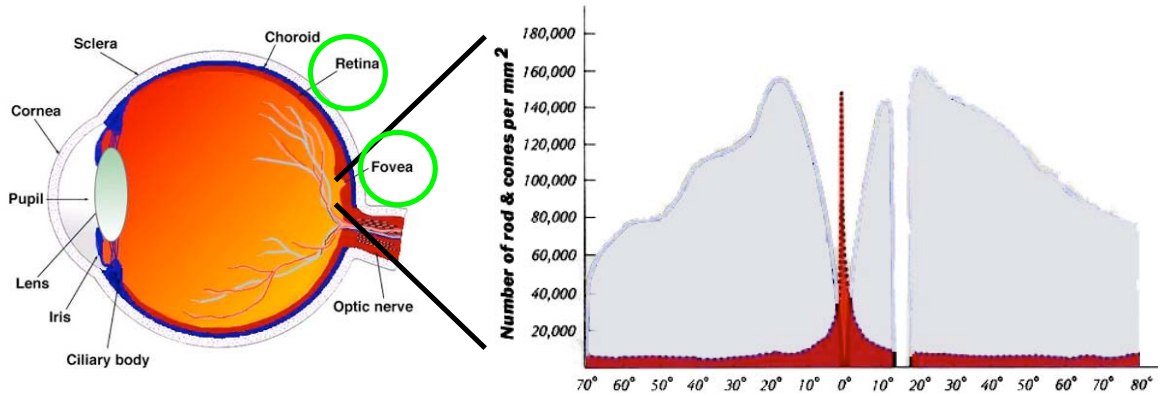


Figure 2. Diagram of eye along with histogram of cone density

In order to properly view the world the eye constantly moves so as to build an internal image of the scene one section at a time. Essentially, the eye moves in order to bring a new portion of the image onto the high-acuity fovea. There are three main types of eye movements that impacted this research. Those are saccades, smooth pursuit and natural drift. Saccades are very fast movements of the eye used to bring a new region of

the scene onto the fovea.[7] Some saccades are voluntary, made in response to flashes or moving targets. Others are guided by higher cognitive processes. Sacacdic eye movements can reach velocities up to 700 degrees per second.[7] Figure 3 is a plot of eye position in visual degrees as a function of time for a portion of one of the experiments in this research. The abscissa is time in seconds and the ordinate is horizontal eye position. Circled in red are three saccade pairs, in which the eye was moved rapidly to one position, and then returned to the starting position. The rightmost circle shows a pair of saccades of approximately 10 degrees.

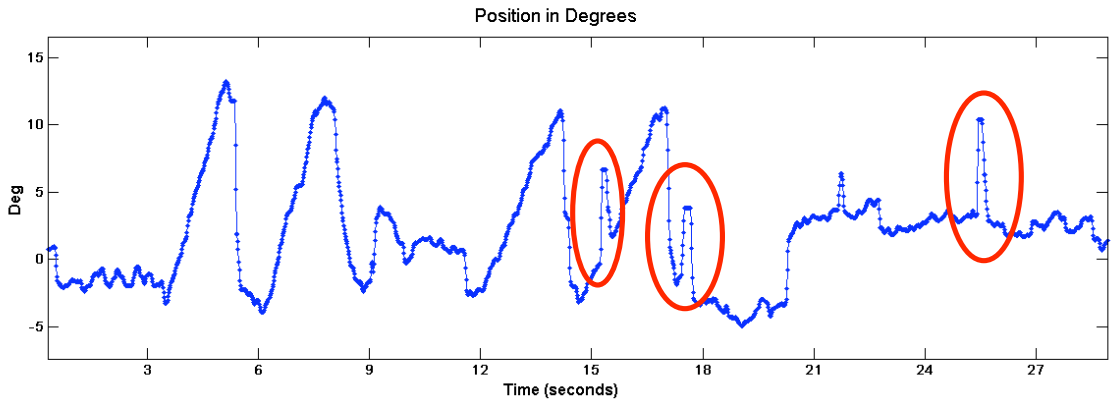


Figure 3. Saccades shown in red circles

Saccades are ballistic movements; once programmed they are executed and cannot be redirected.[7, 8, 9] During a saccade contrast sensitivity falls to nil because of the high retinal velocity[4].

The second type of eye movement is known as smooth pursuit. This occurs when a small target is present and in motion. The observer moves his/her eyes in order to keep the target centrally located on the fovea. Without some object (or target) for the observer to follow, most humans cannot make smooth, uninterrupted eye movements across the

field at relatively slow speeds[10]. Many times the movement of the eye while it is "smoothly" tracking something is actually a combination of saccades and a slower more continuous movement of the eye. The graph in Figure 4 shows a trace of eye movements. It includes a portion where the observer follows an object horizontally across a screen. This plot is the same as above in Figure 3 but reused for a different purpose. The green circle is an example of an observer smoothly tracking a horizontally moving object. There are additional, non-circled examples of smooth pursuit in the plot.

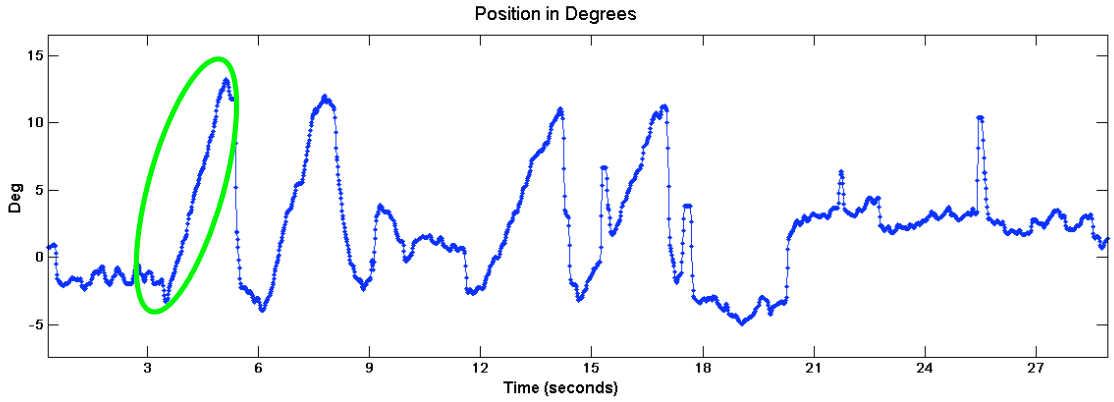


Figure 4. Smooth pursuit of target across screen shown in green circle.

Studies on smooth pursuit[4, 11] show that observers can track targets up to 80 deg/sec before another phenomenon known as "catch-up" saccades occur. This is the so-called "saccadic intrusion" that occurs when targets move fast enough that smooth pursuit eye movements cannot keep up with the target speed. A model can be derived where below a certain target velocity the eye velocity will keep up with it almost perfectly and above that velocity, the eye movement becomes non-constant and the eye velocity either surpasses that of the target or falls behind it. Thus at those high speeds, the observer cannot smoothly track the target. As stated previously, a smoothly tracked object will

have essentially zero retinal velocity, although its shape on the retina will slightly change during the track due to the object angle with respect to the eye. One question that arises is whether it makes any difference to contrast sensitivity when an object is truly stationary or when its image is held stationary on the retina through smooth pursuit. This is one of the questions that this research attempts to answer.

The last type of eye movement relevant to this study is natural drift. This occurs when an observer is fixated on some point. Even when an observer tries to keep the eyes still, there will always be some slight movement. Figure 5 shows two examples of natural drift. The top plot labeled "A" shows a person looking at a marker for one minute and the drift of more than 1/3rd of a degree is clearly visible. The lower plot "B" shows an observer fixated for 1 second and again there is a drift of approximately 1/60th degree. There could be a variety of reasons for this, such as natural body movement due to breathing or other involuntary movements. It could also be the brain's method to keep refreshing the neurons in the retina since it is known that if an image is perfectly stabilized on the retina it quickly fades away.[12] In any case, the result is some very slight movement of the eye which results in non-zero retinal velocity. The velocity of these drift movements is about 0.15 deg/sec.[4]

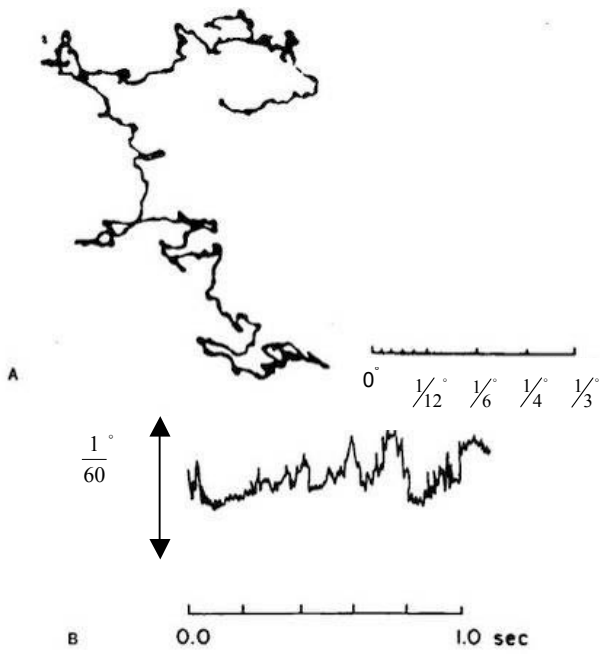


Figure 5. Two examples of natural drift.[13]

The eyes, thus, are never absolutely still and maintain slight movements, or minimal retinal velocity, even when fixated. This is further shown by comparing results from traditional sensitivity experiments, where the pattern is completely stationary to more recent studies where the pattern is moving at the same speed as the hypothesized minimum eye velocity. The results show remarkable similarity to each other.[3,4]

Eye-tracking data were collected for the experiments reported in Part I of this thesis. Analyses show the presence of all three types of eye movements. Observers viewed both objects in motion and stationary objects so smooth pursuit and natural drift eye movements affected the studies. However, none of the objects ever moved so fast that they were beyond the observers' abilities to keep up with them during smooth pursuit, so catch-up saccades did not become an important aspect of the data analysis.

2.2 LCD Motion Artifacts

The basics of an LCD are relatively simple.[14,15,16] Liquid crystals (LCs) are molecules that have a physical state in between that of a solid and a liquid. There is a type of liquid crystal that is naturally twisted, known as twisted nematics (TN). This type of LC has the convenient feature that when an electrical current is applied to it the molecules straighten out. Because of this, liquid crystals can act as an optical light valve. Figure 6 is a schematic of a typical TN LCD pixel that is used in the LCTVs for this research.

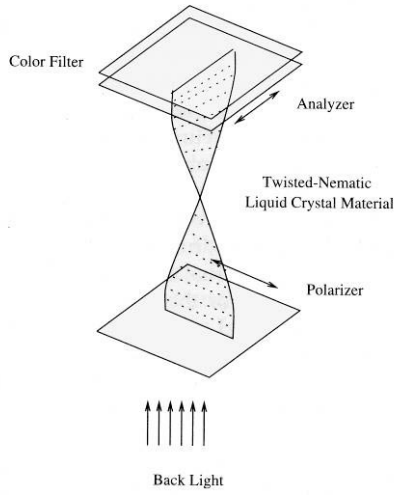


Figure 6. Structure of a LCD pixel[14].

The schematic in Figure 6 is a very basic diagram of a TN LCD pixel. The twisted liquid crystals are sandwiched between two glass plates with a polarizer on the outer side of each plate. The polarizers are placed perpendicular to each other such that light of one polarity coming from the backlight passes through the first polarizer and, if unaffected, would be completely blocked by the second, perpendicular, polarizer. Liquid crystals have the property that when they are twisted, they can change the polarity of light. As

seen in Figure 7, the LCs can be straightened from an electric current, thus preventing the light from rotating and coming out the exit side. By varying the voltage, the LCs can be rotated different amounts and different levels of flux are emitted by the pixel.

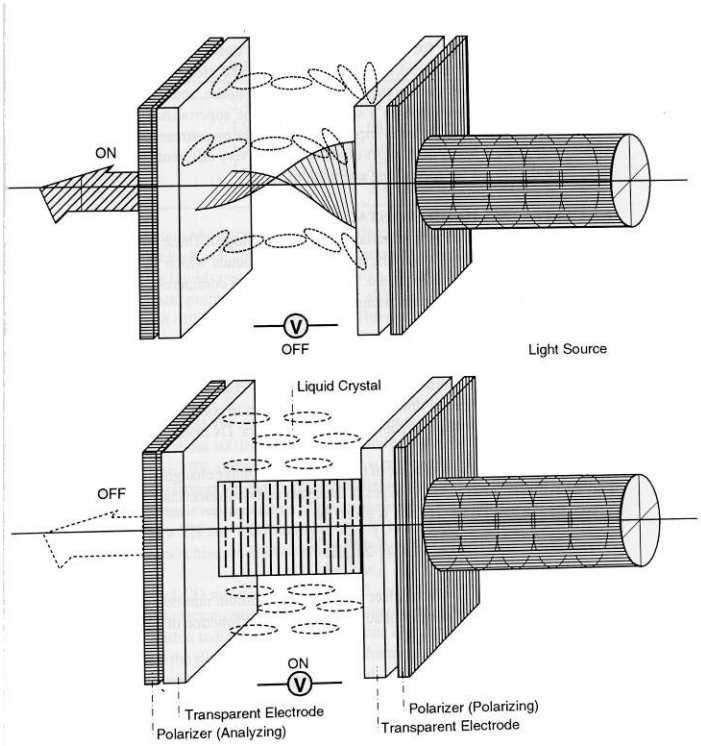


Figure 7. Light flowing from right to left in an LCD system. Above, the LC changes the polarization and allows the light to exit the cell. Below, the LCD does not change the polarization and thus no light exits[15].

The LCTVs used in this research are all active matrix, TFT (thin film transistor) LCD panels. The term active matrix comes from LCDs that have their transistors arranged in a matrix format. Each pixel is addressed by turning on the appropriate row followed by a voltage sent down the appropriate column. Only the intended pixel is addressed since all other rows are turned off at that time. Figure 8 shows this type of arrangement.

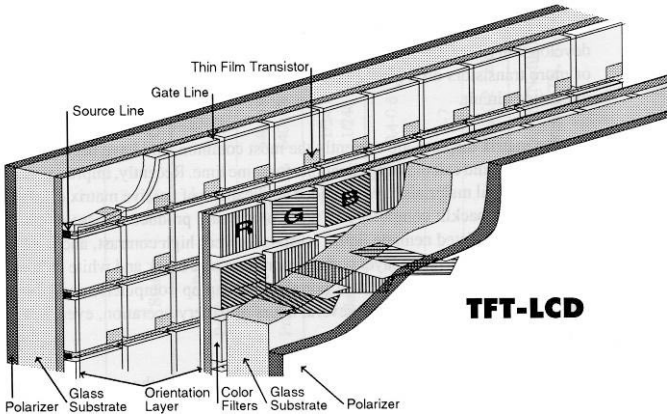


Figure 8. Schematic showing transistors behind each pixel.[15]

This is known as a "sample and hold" procedure[17] since the current pixel remains active until the next refresh cycle. This could be a problem when trying to show an object in motion because each image of the scene is only valid for a single instant in time and not for a complete frame. As a person is tracking an image across the screen his/her eye is in constant motion but if the image is held on the screen too long there is blurring across the person's retina. This is analogous to a person staring at a stationary object; if the person moves his/her eyes suddenly then the image of that object will be blurred because of the motion across the retina.

The speed of change for LCDs is much slower than that for CRTs due to the sample-and-hold procedure and also because of the response time of LCs. Because a CRT must continually excite a phosphor in order to keep it emitting light, the CRT speed is known as its refresh rate. LCDs on the other hand, simply turn a pixel on or off, so its speed is known as response time. LCDs must apply a voltage to either twist or untwist the liquid crystals in each subpixel in order to modulate the light passing from the backlight to the red, green or blue filters. Then the voltage is removed and the liquid crystals (LCs) must revert back to their original state. At best this takes up to 30 milliseconds and

usually it is much longer. A CRT can refresh phosphors in13 milliseconds for a 75Hz display in order to keep the phosphor from darkening.

2.3 Contrast Sensitivity Function

In normal vision, objects are discriminated either by color, luminance or spatial form. Of these, luminance and spatial characteristics of objects are the more prominent aspects that help us recognize objects.[12, 18, 19] The visual system is made up of components that detect and analyze the spatial pattern of light on the retina, see Wandell (ref. 18, chapter 7, page 196). Because of this, the analysis of sensitivity to particular spatial patterns allows a better understanding of the visual process of the human. Sinusoidal patterns (also referred to as sine waves) are a good choice because it is known through linear systems theory, (which is described in section 2.4), that any image can be described through a set of sine wave patterns. If a system is linear then the response to a complex pattern is equal to the sum of responses to the pattern components. Although the visual system is nonlinear with respect to adaptation to different mean light levels, for spatial patterns at a fixed mean level the visual system responds linearly. Therefore, the measurement of contrast sensitivity using sinusoids at a fixed mean level allows evaluation of the visual system as a linear system. By testing the visual system's response to sine wave patterns it is possible to apply the results to more complex stimuli, such as images.

Furthermore, the contrast of these patterns is critical since in natural viewing conditions relative luminances are far more important than the absolute luminances.[19] If an object has little contrast relative to its background then that object will be difficult to recognize. Visual sensitivity to an object increases as less contrast is needed to notice it.

The minimum contrast needed to detect an object is known as its contrast threshold. The reciprocal of contrast threshold is contrast sensitivity.[19] Contrast sensitivity is typically reported on a log scale.

Contrast sensitivity to sine wave patterns is an important characteristic measure of the HVS. Fourier theory shows that all images can be decomposed into basic sine constituents.[20,21] Thus, fully understanding the human response to sine wave patterns can lead to a complete understanding of the human response to any particular input. Sine waves can therefore be applied to other viewing conditions in the natural environment. Practically speaking, sine wave patterns are easy to produce and conveniently easy to flicker or put into motion. Our eyes are quite sensitive to these types of stimuli[18]. The main use of the sine wave pattern in vision experiments is to determine the contrast sensitivity of the HVS, known as the contrast sensitivity function (CSF).

The CSF reports sensitivity against spatial frequency of sine wave patterns. See the graph in Figure 9. The curve in Figure 9 shows peak sensitivity at spatial frequencies of around 4 cycles per degree (CPD). Sensitivity decreases rapidly as the frequency increases and decreases a little as the frequency decreases. Figure 10 is a way to demonstrate how the reader's HVS sensitivity is impacted by spatial frequency. This image consists of increasing spatial frequencies from left-to-right and decreasing contrast from bottom-to-top. For low spatial frequencies the human can only see up to a certain point vertically but as the spatial frequency increases, by moving rightward, more of the pattern is visible at the top of the image and if the spatial frequency is increased further the pattern quickly fades into the background (disregarding any aliasing present due to printing, display, compression, etc. of this document).

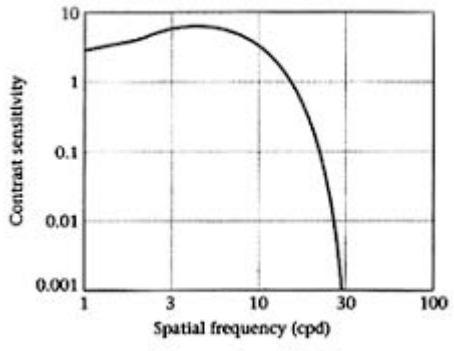


Figure 9. Contrast Sensitivity function modulated by spatial frequency[18].

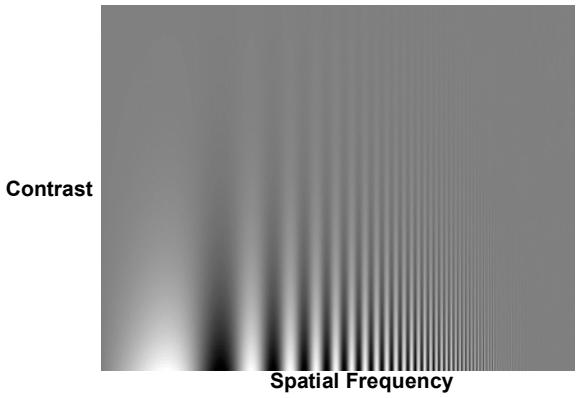


Figure 10. Demonstration of the CSF curve.

Looking at Figure 10, above, there is a spatial frequency with peak contrast sensitivity; where, the pattern is visible to the reader higher in the figure (at lower contrast) than anywhere else. The frequency where this peak contrast sensitivity occurs will be different for each observer and depends upon the physical size of Figure 10 and the reader's distance from it.

A set of experiments that has built upon the CSF have investigated human sensitivity to patterns that change over time. These have tended to rely on sine patterns undergoing counter-phase flicker. Counter-phase flickering has been used as an attempt to understand how sensitive the HVS is to rapid fluctuations from such sources as television, movies and fluorescent lamps.[22] In those studies a sine wave pattern alternated

different intensity stripes at varying speeds, in order to determine the contrast at which perceived flicker is eliminated. The results from these studies are remarkably similar to those where only spatial frequency is modulated for a given temporal frequency, as seen in any one of the four curves on the left graph of Figure 11 below.[23] The plot on the left shows how sensitivity varies with spatial frequency at given temporal frequencies and the plot on the right shows how it varies with temporal frequency for given spatial frequencies. It can be seen there is a drop in sensitivity at low spatial frequencies only at low temporal frequencies and vice-versa. (See 1 Hz curve in left plot and 0.5 CPD curve in right plot).

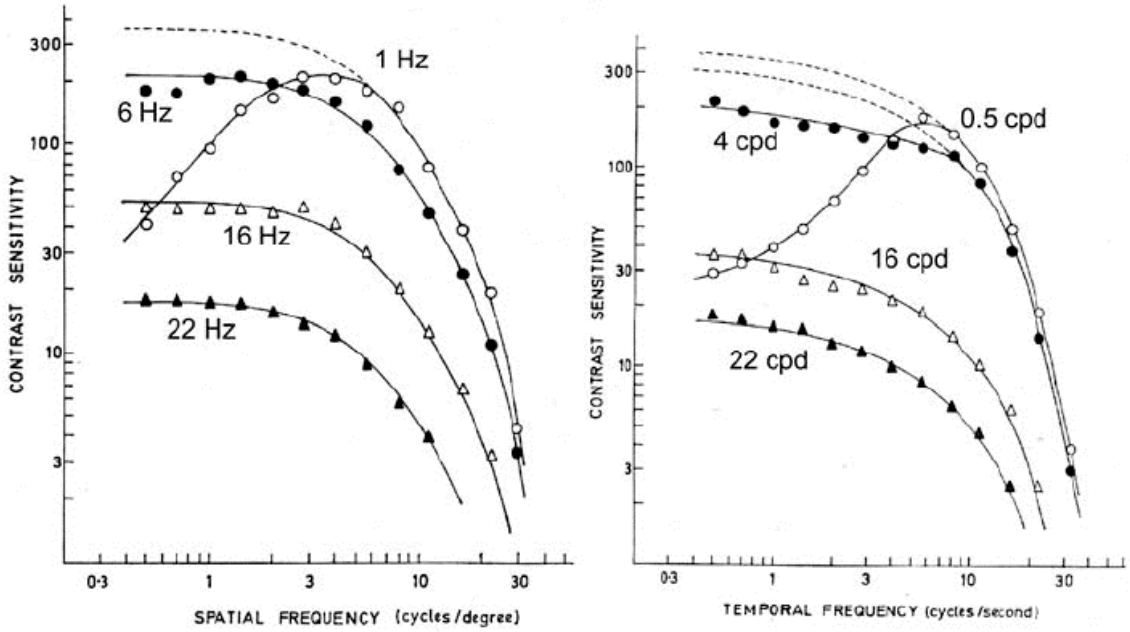


Figure 11. Spatial CSF on left and temporal CSF on right [23]

Historical research has shown that the spatial and temporal contrast sensitivity functions are separable for high spatial and high temporal frequencies above 10 CPD for spatial and 10 Hz for temporal frequencies in Figure 11.[3, 19, 22, 23] In Figure 11, on the

right, the shape of the curves at high temporal frequencies is invariant to spatial frequency. Likewise, on the left, the shape of the curves at high spatial frequencies is invariant to temporal frequency meaning that at these high frequencies the spatiotemporal CSF is simply a product of the spatial CSF and temporal CSF.[22] The interaction at low frequencies is due to the temporal behavior of the photo-receptor signal and the lateral inhibition of the receptors, which is affected by the spatial characteristics of the pattern[19]. Sensitivity can thus be described as an interaction between temporal and spatial frequencies. If a variety of spatial frequencies are shown at a variety of temporal frequencies using counter-phase flicker then there are two dimensions: spatial and temporal frequency.

The early models were based on spatial and temporal frequencies. However, it is more natural to discuss motion in terms of velocity rather than temporal frequency. In nature, targets are less likely to be flickering than in motion, so temporal variations are likely due to spatially varying targets in motion with respect to the retina. So it is spatiovelocity response that should be of interest, rather than the spatiotemporal response to counterphase flickering targets. Additionally, Kelly[3] has shown that contrast sensitivity is different between counterphase flicker stimuli and stimuli in motion.

SpatioTemporal CSF

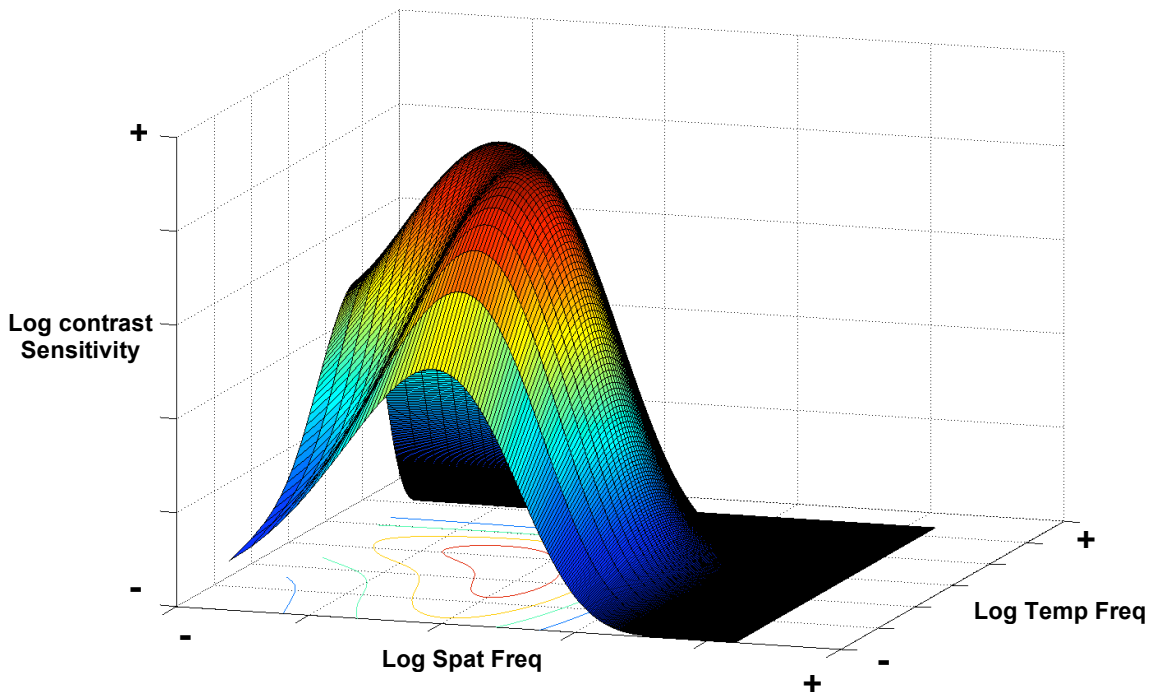


Figure 12. Spatiotemporal CSF

In a break from the traditional spatiotemporal approach, D.H. Kelly investigated contrast sensitivity as a function of velocity across the retina.[3] The results of his studies are more transferable to natural conditions because in general humans do not observe flickering objects, instead objects in motion move at some velocity across the retina or remain fixed on the fovea if a person is tracking an object. Kelly[3] referenced prior work[24] that showed motion sensitivities to be greater than flicker sensitivity. Furthermore, he states, "the fact that motion thresholds are lower than flicker thresholds suggests that moving gratings are somehow better matched to the characteristics of the visual process than are flickering gratings."[3]

Kelly modulated image velocity across the retina through retinal stabilization and could therefore induce movement or keep the stimulus stationary on the observer's retina. He held velocity constant and measured contrast sensitivity to a variety of spatial

frequencies. He built a two-dimensional CSF model through testing contrast sensitivity at a variety of velocities and spatial frequencies. Kelly fit a model to his data using the following formula.

$$CSF(\rho, v) = k \cdot v \cdot \rho^2 \exp\left(-\frac{2\rho}{\rho_{\max}}\right) \tag{2.1}$$

Where $k = s_1 + s_2 \cdot \left|\log\left(v/3\right)\right|^3$ (2.1.1)

Furthermore, $\rho_{\max} = \frac{p_1}{(v+2)}$ (2.1.2)

Where $s_1 = 6.1$, $s_2 = 7.3$ and $p_1 = 45.9$ are the constants provided by Kelly and were determined through optimization. The variable ρ is spatial frequency in cycles per degree and v is the constant-velocity for which the contrast sensitivities were tested at a variety of spatial frequencies. The scale factor k is responsible for the vertical shift of sensitivity and is dependent on velocity, where in general at lower velocities there is simply a vertical offset in sensitivity[3, 4,22, 25], and p_{max} is responsible for the horizontal shift of the peak sensitivity. Both of these scale factors account for the separability of the spatial and temporal components of the CSF for high frequencies.

An example of Kelly's 2D spatiovelocity CSF is shown in Figure 12. This graph shows that as spatial frequency is increased from left to right while keeping temporal frequency constant, sensitivity increases to a peak and then quickly drops off. Likewise, if spatial frequency is held constant and temporal frequency increases from front to back there is also a peak reached followed by a quick falloff.

In 2001 Daly revised Kelly's model to incorporate retinal velocity that took into account smooth pursuit eye movements.[4] Observers can track an object up to a certain

velocity and at faster velocities the person has to make "catch-up" saccades to keep up with the object. The exact velocity at which this happens is dependent on several variables. Daly incorporated retinal velocities with this limitation imposed on sensitivity to Kelly's model, among other changes. This distinction is important because during saccadic eye movements, sensitivity drops to near-zero. Daly also incorporated minimum eye velocity into his revised model since it is known that observers do not keep their eyes still when fixated on a target.[12] The shape of the curve from a traditional, static stimulus CSF experiment is very similar to the shape of the same pattern moving at this minimum velocity. In other words, while sensitivity will drop to zero for fast velocities there will always be sensitivity for minimum velocities.[4]

2.4 Experimental Stimuli

The traditional stimulus used in visual sensitivity experiments is a sine wave pattern because of practical and physiological reasons[18, 19]. When spatial frequency is discussed for this research it refers to the number of cycles subtended at the eye per visual degree as seen in Figure 13. Stimuli used in CSF measurements are often a Gabor pattern. The Gabor pattern is simply a variation of the traditional sine wave pattern, in which the pattern is windowed by a radially symmetric Gaussian function that fades the pattern from its full contrast to the mean level of the background, as seen in Figure 14 below. Fading of the pattern into the background prevents a sharp transition from the pattern edge to the background since observers might respond to such high frequencies along with the frequency of the pattern itself. In other words, the windowing keeps high frequency interference of the pattern edges from influencing observers' judgments. An example of a Gabor is seen in Figure 14 below.

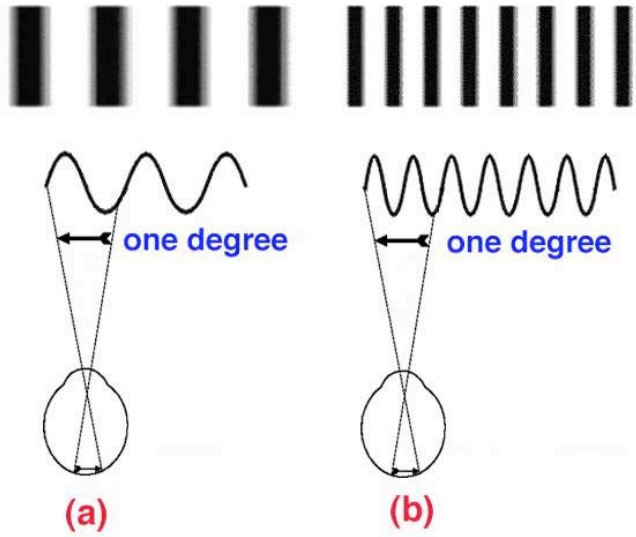


Figure 13. Spatial frequency is a measure of the number of cycles subtended at the eye per degree. (a)one cycle per degree (b) two cycles per degree.[26]

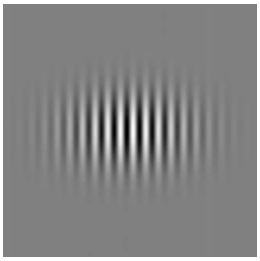


Figure 14. Example of Gabor pattern

2.4.1 Fourier Analysis

Linear systems theory states that any image can be decomposed into a weighted sum of sinusoids. Therefore, it is hypothesized that testing sensitivities to sine wave patterns with a range of spatial frequencies can be directly related to human perception of complex images. More specifically, Fourier analysis is the tool used to compose and decompose a stimulus to a set of sines and cosines. Jean Baptiste Joseph Fourier stated in 1822 (written in 1807, published in 1822 and translated in 1878)[27], that any function which periodically repeats itself can be expressed as the sum of sines and/or cosines of

different frequencies each multiplied by a different coefficient; this is known as a Fourier series. The figure below shows how a set of sine waves can be combined to create a separate function. Note that the sharp edges at the bottom of the image are created using all possible frequencies.

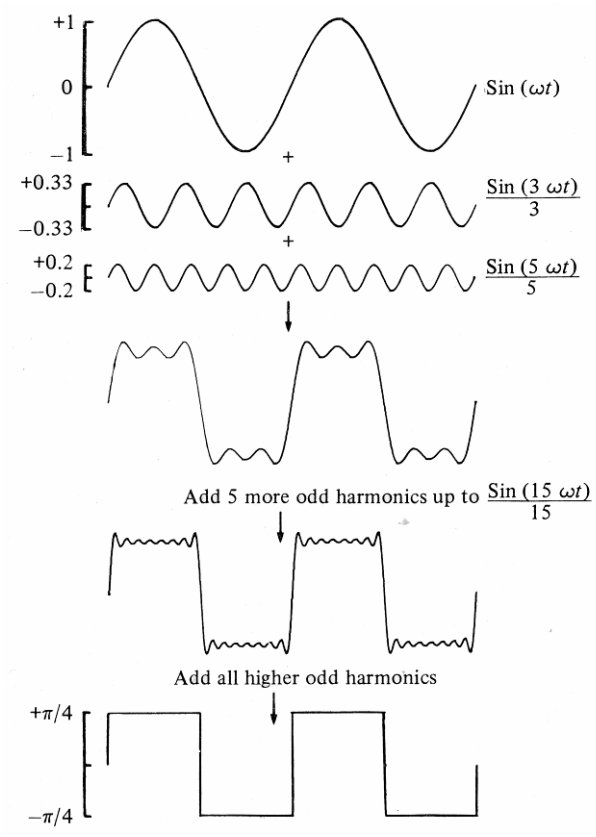


Figure 15. Sine waves combined to form different function.[21]

It is also true that functions that are not periodic but have a finite area under their curve, (images are of this type of function with a finite duration), can be expressed as the integral of sines/cosines multiplied by a weighting function; this is known as a Fourier transform.

The Fourier transform of a continuous function $f(x)$ is defined by

$$F(u) = \int_{-\infty}^{\infty} f(x) e^{-j2\pi ux} dx \tag{2.2}$$

Where $j = \sqrt{-1}$. Likewise, given $F(u)$the original function $f(x)$ can be obtained by

$$f(x) = \int_{-\infty}^{\infty} F(u) e^{j2\pi ux} du \tag{2.3}$$

These are called the Fourier transform pair and show what was stated previously that functions can be transformed forward and back without loss of information. These are the basic functions in Fourier analysis; however since this research deals with images the following functions are of more practical use. The one-dimensional Fourier transform of a discrete function (image) $f(x)$, is given by Eq. 2.4 below.

$$F(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x) e^{-j2\pi ux/M} \quad \text{, for u = 0, 1, 2… M-1} \tag{2.4}$$

Likewise, given $F(u)$ it is still possible to go back to $f(x)$, given Eq. 2.5 below.

$$f(x) = \sum_{u=0}^{M-1} F(u) e^{j2\pi ux/M} \quad \text{, for x = 0, 1, 2… M-1} \tag{2.5}$$

To compute $F(u)$ start at $u = 0$ and sum for all values of x, then set $u = 1$ and sum again for all x, then repeat for all M values of u.

The idea of the frequency domain comes directly from Euler's formula, seen below.

$$e^{j\theta} = \cos\theta + j\sin\theta \tag{2.6}$$

Where j is imaginary and equals $\sqrt{-1}$. Substituting this formula into Eq 2.4 provides

$$F(u) = \frac{1}{M}\sum_{x=0}^{M-1} f(x)\left[\cos 2\pi ux / M - j\sin 2\pi ux / M\right] \tag{2.7}$$

Where $2\pi ux/M = \theta$.

It should be noted that the value of $F(u)$ at each value of u is the sum of all values of $f(x)$ and each value of $f(x)$ is multiplied by sines and cosines of different frequencies. Thus the domain of $F(u)$, which is the values of u, is called the frequency domain since u determines the frequency of the components of the transform. Equations 2.4 through 2.5 will allow the more complex stimuli used in this research to be described by the amount of sine waves contained inside. Furthermore, once the stimulus is decomposed to sine waves then sensitivity to the sine waves at a given velocity can be described.

2.5 Psychophysical Experiments and the QUEST routine

A properly designed psychophysical experiment will answer a question, such as "What is the contrast threshold for a sinusoidal grating?" based on measured task performance of observers. It is an attempt to understand internal human processes through the measurement of observer behavior. There are two types of psychophysical experiments: adjustment and judgment; where adjustment allows the observer to change some stimulus parameter while judgment asks the observer to classify the stimulus. In other words, for judgment experiments the controller makes a change to the stimulus and the observer makes a judgment on the resulting stimulus and for adjustment the controller gives the stimulus to the observer and allows them to makes adjustments to meet some criteria. In general, judgment experiments involve a detection task where the observer might be asked to when a stimulus is present. This is the type of experiment utilized in this research.[28, 29]

Historically, threshold referred to stimulus intensity in which any value below the threshold was invisible to the observer and any value above it was visible. In current methodology threshold refers to the strength of a signal corresponding to a level of performance as defined by the probability of correct response. As stimulus intensity increases, the probability of a correct response also increases (see Figure 16 below).

The experiments used in this research were judgment experiments with a 2-interval forced choice design (2IFC). The observer watched the screen over two timed intervals separated by audible beeps where the stimulus was randomly assigned to be present in one interval and absent in the other. The task of the observer was to determine in which interval the stimulus was present. This allows for a more unbiased observer response, because it reduces internal observer criterion from affecting decisions. In a 2IFC experiment the experimenter begins with a desired criterion for the probability of correct response and then changes a particular stimulus parameter until that criterion has been reached through a predefined psychometric function. This procedure can be done quite effectively through a sequential estimation technique. (See References 30 and 31 for further details.) In the figure below if the desired criterion was a 75% correct response then the corresponding threshold would be S6. Note that each observer would have a slightly different function such that the 75% criterion would relate to a different threshold along the abscissa. The function in the figure below begins at 50% because if the observer is making a guess then there is a 50/50 chance they will get it right (either the first or second interval). Therefore, given enough trials, the observer will have a 50% chance of guessing correctly when they cannot detect the stimulus.

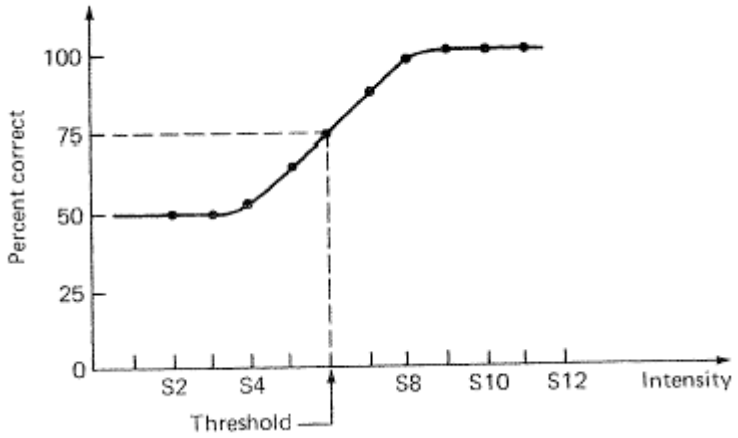


Figure 16. Example of psychometric function.[32]

QUEST[33] is a sequential estimation routine that implements an adaptive procedure to determine threshold of a stimulus. It is an advanced staircase design known as a "maximum likelihood" method, because after each trial the most likely value for threshold is chosen and used as intensity for the next stimulus. In a traditional staircase experiment, the value of the stimulus is modulated based on observer response; for example, if the observer correctly guesses that the stimulus is present in the second interval then for the next trial stimulus intensity is decreased; but, if they incorrectly choose the first interval then intensity is increased for the next trial. QUEST uses a psychometric function, predefined by the controller, and continuously updates the function based on previous trials to make the most efficient estimation of threshold using prior knowledge in the form of a predefined psychometric function and results from all previous trials. The controller decides prior to the experiment what the criterion level should be, such as 75% probability of a correct guess, what psychometric function should be used, such as the Weibull function and finally how many trials are needed for each experimental condition.[32]

Chapter 3 - Experimental

3. 1 Description of Stimuli and Experiments

There were a total of five experiments for the research in this project. The first experiment was a control experiment. Experiments two through four were used to investigate and/or populate the 2D spatiovelocity contrast sensitivity function (CSF). The fifth experiment was used to validate the resulting 2D CSF model. The experiments all used a two interval forced choice (2IFC) method. It should be noted here that for all five experiments observers were instructed to fix their gaze on a circular point, known as a fixation point, which was 4 pixels in diameter. In experiments one, two and three the fixation point was stationary and centered in the display. In experiments four and five the fixation point moved across the screen at the same speed as the stimulus. Note that there was always a fixation point for the observers to track or fixate even during trials where no stimulus was present.

3.1.1 Stimuli Description

The stimuli for experiments one through four were Gabor patterns of 2.46 visual degrees in diameter. As described in the Chapter 2 there are historical and physiological reasons for using a sine wave pattern to test the HVS. For the Gabor pattern, seen in Figure 14, the variables adjusted were the contrast and spatial frequency of the pattern.

The stimulus for Experiment five was a disembodied edge, which is an edge windowed by a Gaussian. Thus there are no sharp edges except for the transition from dark to light in the center of the stimulus as seen in Figure 17 below. The plot on the right in Figure 18 is a cross section of a sharp “disembodied edge” and the plot on the left

shows that it was created from a Gaussian distribution of some width and amplitude. The width of the Gaussian is about 150 pixels, which corresponds to approximately 4 degrees, from mean level to highest intensity and the amplitude is the amount of contrast. The edge stimulus was used in the verification of the CSF model for a couple reasons. One reason is that it is an intermediate stimulus: more complex than a sine wave pattern yet simpler than a typical image. Recall from the Background that all images can be described as a sum of sinosoids. Thus, the edge can be thought of as a sum of sines at multiple frequencies. The edge was blurred by convolving the sharp disembodied edge in Figure 17 with a Gaussian filter. The amount of blur was controlled by changing the size of the Gaussian used by the filter meaning that a Gaussian with a standard deviation of 2 would blur the edge more than a standard deviation of 0.5.

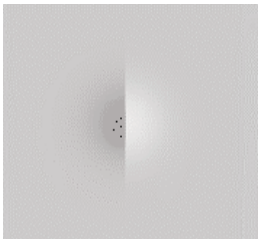


Figure 17. Example of disembodied edge

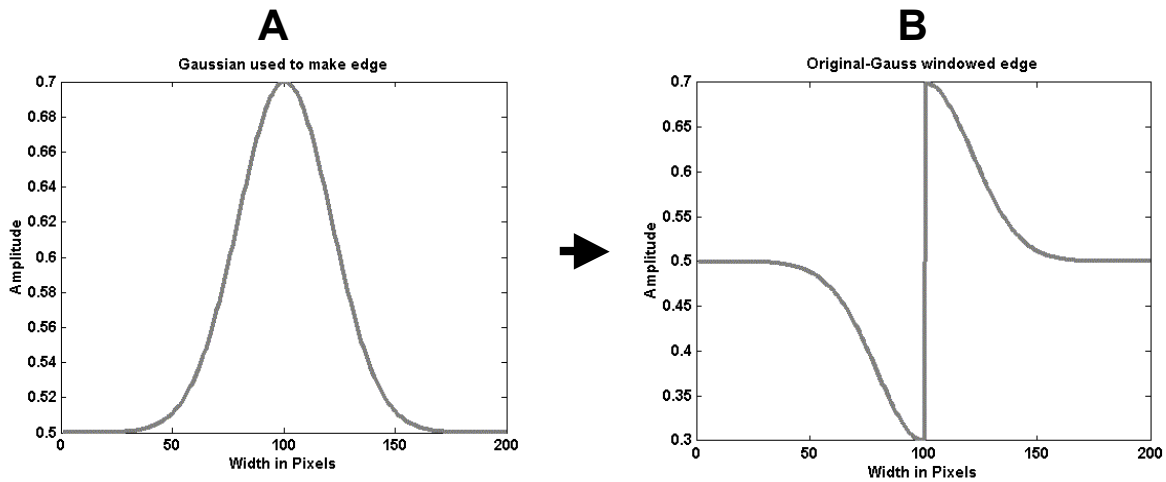


Figure 18. (A) Gaussian blur used to create edge. (B) Cross Section of disembodied edge created from A by splitting Gaussian and inverting first half.

A perfectly sharp edge contains an infinite number of frequencies, as seen in Figure 15. However, as the edge is blurred there is a fall off in the amount of high frequencies in the edge. By testing the threshold of edge detection it is possible to test the minimal amount of high frequency information to be discarded and have observers just barely detect a difference from a sharp edge. Thus the edge stimulus allows many spatial frequencies to be experimented with simultaneously.

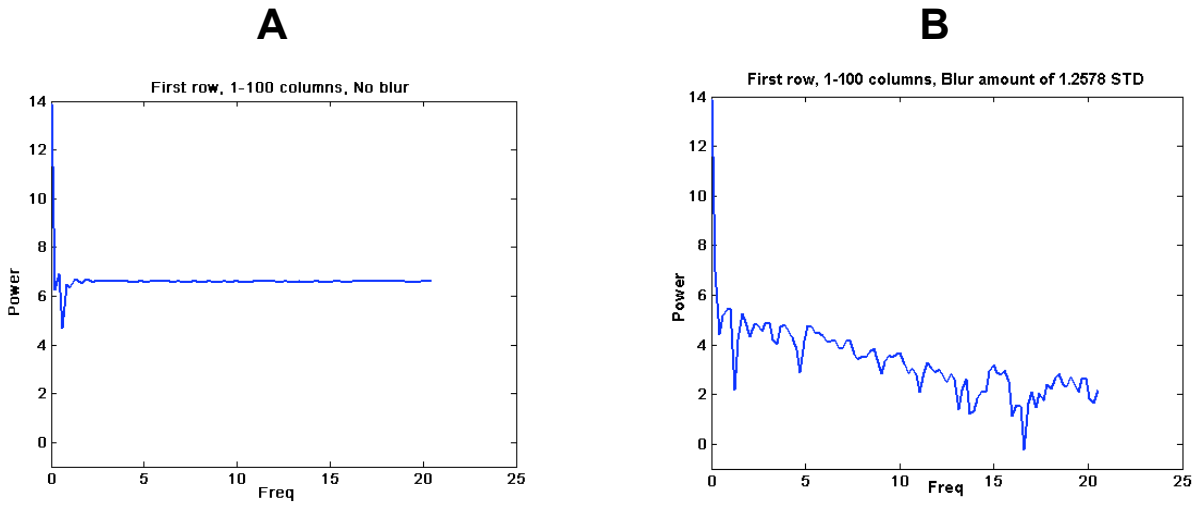


Figure 19. (A) Fourier transform of sharp edge and (B) an edge that has been blurred.

Figure 19 above shows the Fourier transform of an unblurred edge on the left and a blurred edge on the right. The abscissa is spatial frequency and the ordinate is the amount of power at each frequency. It can be seen that the unblurred edge has no decrease in power as the frequency increases. However, the plot on the right shows a falloff in the power as the frequency increases.

3.1.2 Description of Experiments

For experiments two through four, the independent variables were spatial frequency and velocity. There were a total of eight conditions, as shown in the table below. The spatial frequencies of the sine in the Gabor are on the left and the temporal frequencies are along the top. The velocity of the Gabor was changed in order to keep the temporal frequency constant and the particular velocity value, (in deg/sec), is seen in the corresponding cells. There was not a ninth condition, corresponding to 16 CPD and 30 Hz as it was beyond the capability of our system to effectively produce.

Table 1. Gabor pattern data

		Temporal Freq (Hz)		
		10	**20**	**30**
Spat Freq (Cyc/Deg)	**4**	2.5	5.0	7.5
	8	1.25	2.5	3.75
	16	0.625	1.25	

All experiments were controlled using programs written in Matlab. The programs utilize the QUEST routine run within the Psychophysics Toolbox[34] to determine contrast thresholds, which was described in section 2.5 and also in Reference 33. Essentially what Quest does is to make an educated guess for each new trial based on a particular psychometric function and all previous trials. The values it guessed were interpreted as

contrast and used directly to modulate the Gabors for the next trial. Figure 20 below shows QUEST's threshold estimate for all 50 trials of an individual for each condition in one particular experiment. Each subplot is a separate condition in the same order as the elements in Table 1 and is written in the title to each subplot. For example, after the "C" in the title of the upper left subplot there are the numbers "4 0 2.5," which stands for 4 CPD, 0 HZ and 2.5 deg/sec. The linear threshold estimate and standard deviation of that estimate (based on all 50 trials) is also written above each subplot. The curve in each subplot is the log contrast value used for the Gabor in that particular trial, which as stated previously is also QUEST's estimate of threshold. It is seen that the estimates vary greatly and then gradually converge toward a midrange value.

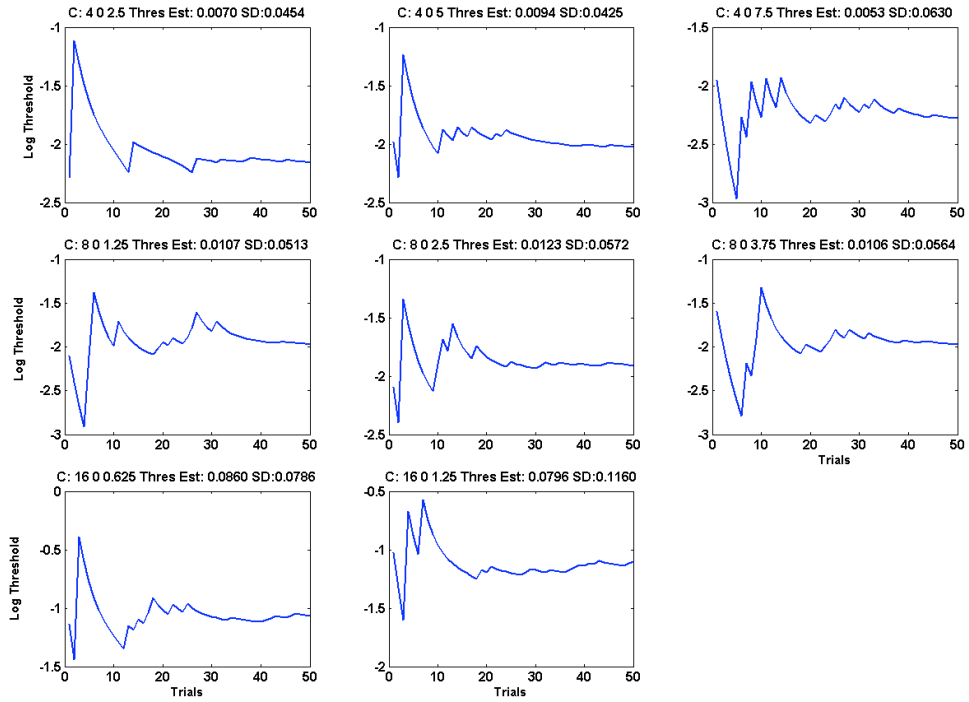


Figure 20. Results from Quest for 1 observer in 1 experiment for all 50 trials.

Experiments one through four were used to build and investigate the 2D spatiovelocity model and because the interest for this research lies in sensitivity as a function of motion, three of the experiments dealt with moving stimuli. There are four figures below that show how each experiment differed and in each figure the arrow represents an increase in time moving back to front. The first experiment was a traditional CSF experiment and is considered a control experiment, which allowed comparisons to be made between results in this research with those in the literature. The Gabors in this experiment were completely stationary and are seen in Figure 21 below. In Experiment Two the sine wave pattern inside the Gabor window was in motion while the window itself remained centered in the monitor, as seen in Figure 22. In Experiment Three the sine wave pattern was stationary relative to the Gabor but the Gabor itself moved left to right across the screen while the observer remained fixated on a stationary point centered on the display, (Figure 23). Experiment four was similar to two except that now the fixation point was in motion along with the Gabor, as seen in Figure 24.

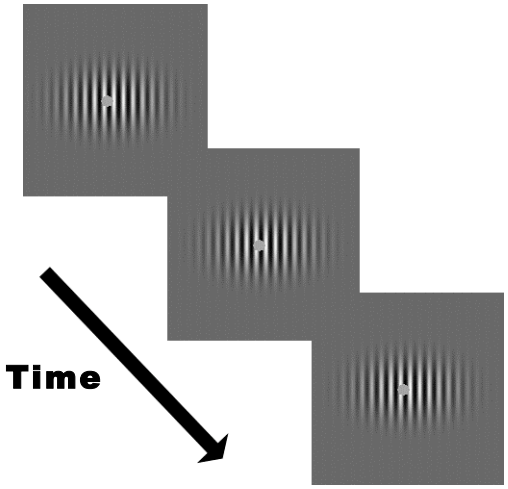


Figure 21. Exp 1, Completely Stationary Gabor: window stationary and sine pattern stationary. Note fixation point remains in center of stimulus (same as center of screen).

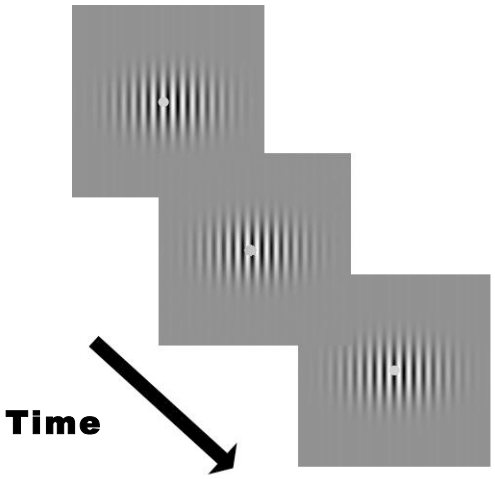


Figure 22. Exp 2, Sine wave pattern in motion within Gabor: window stationary but sine pattern in motion. Note fixation point remains in center of stimulus (same as center of screen).

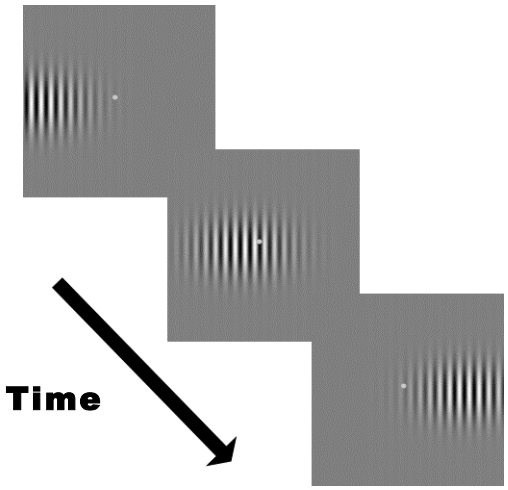


Figure 23. Exp 3, Gabor in motion, eyes fixated: window in motion, sine stationary with respect to window. Note that fixation point remains centered on screen.

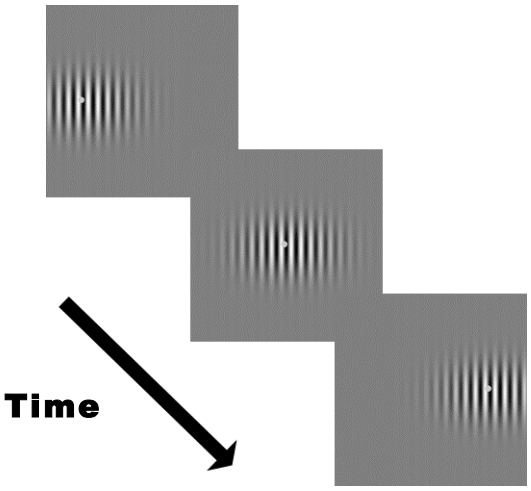


Figure 24. Exp 4, Gabor in motion, eyes tracking: window in motion, sine stationary with respect to window. Note that fixation point remains centered on window as it moves across the screen.

The fifth and final experiment utilized a disembodied edge and was used to verify the model. Although the experiment was a 2IFC as with the other experiments, there were some important differences. The fixed conditions in this experiment were velocity and contrast. There were three contrasts levels used at four velocities, seen in Table 2. There was a stimulus present in both trials; in one interval there was a sharp edge and in the other there was a blurred edge. The experimental design was the same as in Experiment four, shown in Figure 24. The velocities were chosen from those in used in experiments two through four (see Table 2). The question posed after each trial was, "In which interval was the edge sharper?" and the QUEST routine then chose the amount to blur the edge. The answer that QUEST returned is a threshold for edge blurriness. Therefore, the sensitivity described is the reciprocal of threshold for edge detection.

Table 2. Parameters of Experiment Five

Velocity (deg/sec)	Contrast (Michelson)
0	0.05
1.25	0.25
3.75	0.4
7.5	

The contrast values were chosen based on those used by Hamerly & Dvorak[35] in their paper, "Detection and discrimination of blur in edges and lines", with the exception that the maximum contrast valued possible was 0.4 because of reasons described in the next section.

3.2 Experimental Setup

The setup for each experiment consisted of a Sony Trinitron MultiScan G420 CRT monitor, an eyetracker and a chinrest, all of which are seen in Figure 25. Observers were seated a constant distance of 84 cm. from the monitor and placed in the chinrest in order to minimize head movements and maintain the same distance from the monitor throughout the experiment.

Figure 25. Experimental setup, showing monitor, eye tracker and chin bar

Although one of the motivations for this research was to produce results useful for the development of motion algorithms for LCD televisions, the experiment was independent of the medium. The CRT was chosen because of its fast refresh times which reduced motion artifacts. However, in order to relate sensitivities from experiments on the CRT to the much brighter LCD, the brightness of the CRT was set as high possible.

A characterization was performed on the CRT monitor and a TRC-matrix model derived. (See Appendix A for more information.) Measurements were made with an LMT Colorimeter C1210 that measured displayed colors in tristimulus values, (XYZ). It was chosen because of its large dynamic range and for its automated measurement capabilities. The LMT was interfaced with Matlab and all measurements of colors were made with a neutral gray background and a colored patch slightly larger than the diameter of the LMT's measurement head. The model was verified using measurements of 2000 randomly selected RGB colors. The average CIEDE2000 between the measured and calculated values was 0.31. Other tests performed included a channel and spatial independency test. The channel independency test showed the sum of tristimulus values measured for the primaries was greater than the measured full white of the monitor. There was a discrepancy of ΔE_{00} 0.38 between the measured and summed white. The spatial independency test showed an average difference of ΔE_{00} 0.52 between a white on black background and white on other backgrounds (See Appendix A for further details.) Table 3 below shows the settings for the CRT used throughout the experiments followed by Table 4 that shows the results of the characterization.

Table 3. CRT settings for each Experiment

Mean Lum. of screen	60 cd/m^2
Horiz. span of screen	23.95^o
Dist Obs. from Screen	84 cm

Table 4. Verification results using CIEDE2000 between 2000 measured and calculated randomized colors

Verification of model using ΔE_{00}		
Max	**Min**	**Avg**
1.91	0.01	0.31

The mean luminance of the screen was set to 60 cd/m^2, as seen in Table 3, because of the relationship between screen luminance and contrast. The maximum luminance of the screen was 98.5 cd/m^2, therefore there was a range of contrast factors that could be used without any clipping. The maximum contrast factor, which is the same as contrast, is defined by Michelson Contrast[19] to be 0.4. Figure 26 shows the maximum contrast factor possible at each mean luminance level. Michelson contrast is defined as

$$C_M = \frac{L_{Max} - L_{Min}}{L_{Max} + L_{Min}} \tag{3.1}$$

where L_{Max} and L_{Min} is defined as maximum and minimum luminance values respectively. Figure 26 is a plot of the relationship between contrast factor and mean luminance and was used throughout the experiment to keep values in the Gabor from becoming too great and thus being clipped.

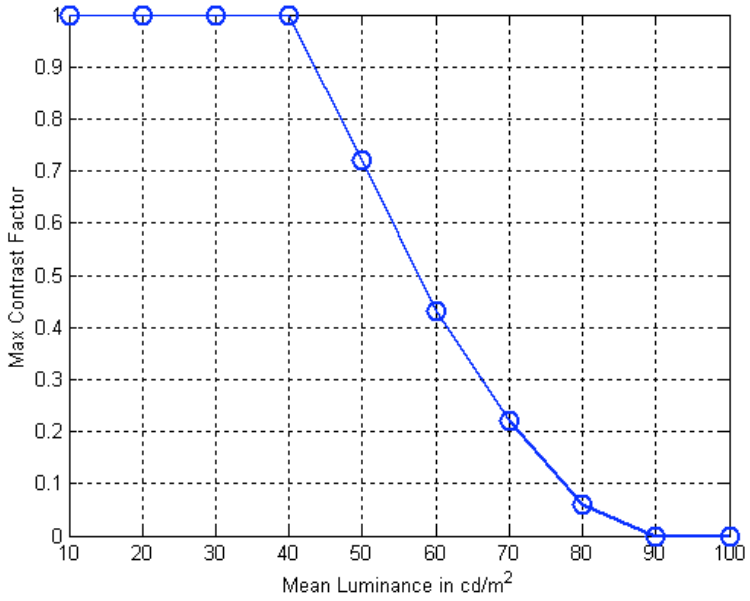


Figure 26. Plot showing the maximum contrast factor possible given mean luminance of CRT.

The ASL Series 504 remote eyetracker[36] was run using ASL's proprietary software package. This system monitors eye position without any contact with the subject by imaging the eye through a camera. The lens is surrounded by infrared light emitting diodes (IRLEDs) providing illumination aligned with the optical axis. This infrared, video-based eyetracker determines the point-of-gaze by using a video camera to extract the center of the subject's pupil and a point of reflection on the cornea. This is known as bright pupil technology because the illumination is coaxial with the axis, resulting in a back-illuminated pupil. The effect is the same as "red eye" and is used by the software to track the eye.

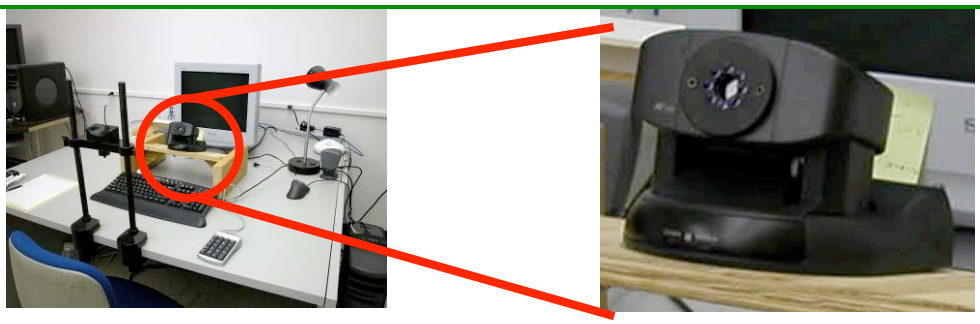


Figure 27. ASL Series 504 Remote eye tracker

Prior to each experiment, the eyetracker was calibrated to the observer using ASL's software and a calibration target seen below in Figure 28. Observers were calibrated by displaying the target and having them look at each of the points while making necessary threshold adjustments in order to get a good image of their pupil and corneal reflection. The next step was to have the observer look at each point and within the software enter where they are looking. The software recorded relative pixel-position, (in its own units), of the point where the observer is looking. Afterwards, the observer looks at each of the points to verify the calibration.

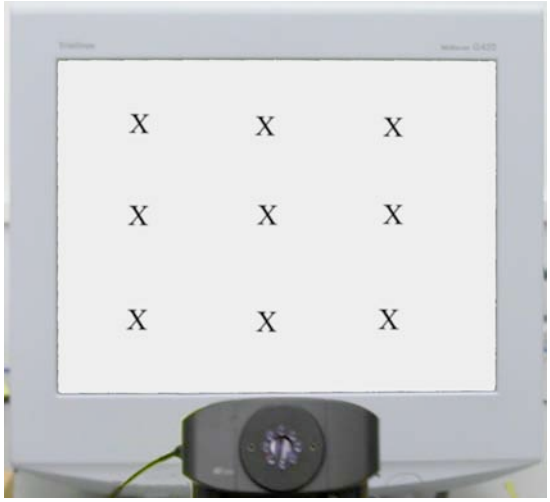


Figure 28. Calibration target for eye tracking

The calibration process sets a relationship between known distances, in centimeters, between points on the monitor to distances the observer's eye moved between those points. The position recorded by ASL can be converted to visual degrees using the formula in equation 3.3, seen in the next section, if the distance between observer and the monitor is known. This calibration provided accurate position of the tracked eye as long as the observer was at the same distance from the monitor throughout the experiment. The accuracy of the eyetrack record is approximately 1° or 16 cm on the display.

3.3 Eyetrack Data

There were two data files collected during the experiments. The contrast sensitivity data from QUEST was stored in one file and in the other the eye movement data from ASL was stored. The data in this second file included horizontal and vertical eye position values in relative ASL units along with codes that represented the condition of the trial and in which interval the stimulus was present or absent. Prior to showing a trial, the Matlab program controlling the experiment sent a list of codes followed by appropriate values to this file, which described pertinent information on that upcoming trial. For this system, these user-definable codes are called Xdat codes. An example of Xdat codes generated during an eye scan is shown graphically below in Figure 29. The interval where the stimulus is present is the horizontal line to the right and has a code value of 20. This value represents the 20th trial of this particular spatial frequency and velocity condition. The horizontal line on the left has values of 120. For this dataset, any value 100 and greater represented the stimulus-absent trial. Therefore, in the first interval of this trial the stimulus was absent and in the second interval the stimulus was present. It should be noted that although it is possible to know in which interval the stimulus was

present, it was not possible to know if it was visible. While the contrast of the Gabor may be known it does not necessarily mean it is visible.

The actual Xdat codes used to record trial information and their definitions are seen in Table 5. Studying the Figure 29 example trial, the value after 95 is 4 which according to Table 5 means the spatial frequency is 4 CPD and the value after 96 is 0 which means a temporal frequency of 0 Hz, (which was the case for all conditions in Experiment Three). In Figure 29, the value after 97 is 7, which corresponds to a velocity of 7.5 deg/sec, and the value after 98 is 120, (which means the stimulus is absent and it is the 20th trial for this condition). The value after 201 is 1 and the value after 202 is 18, which is interpreted according to Table 5 as meaning this trial is the 118th overall trial of this experiment.

Table 5. Xdat codes used and what they represent

Code value	Represents
95	Spatial frequency
96	Temporal frequency
97	Velocity
98	Interval present/absent
201	Overall trial 1-400 in tens
202	Overall trial 1-400 in hundreds

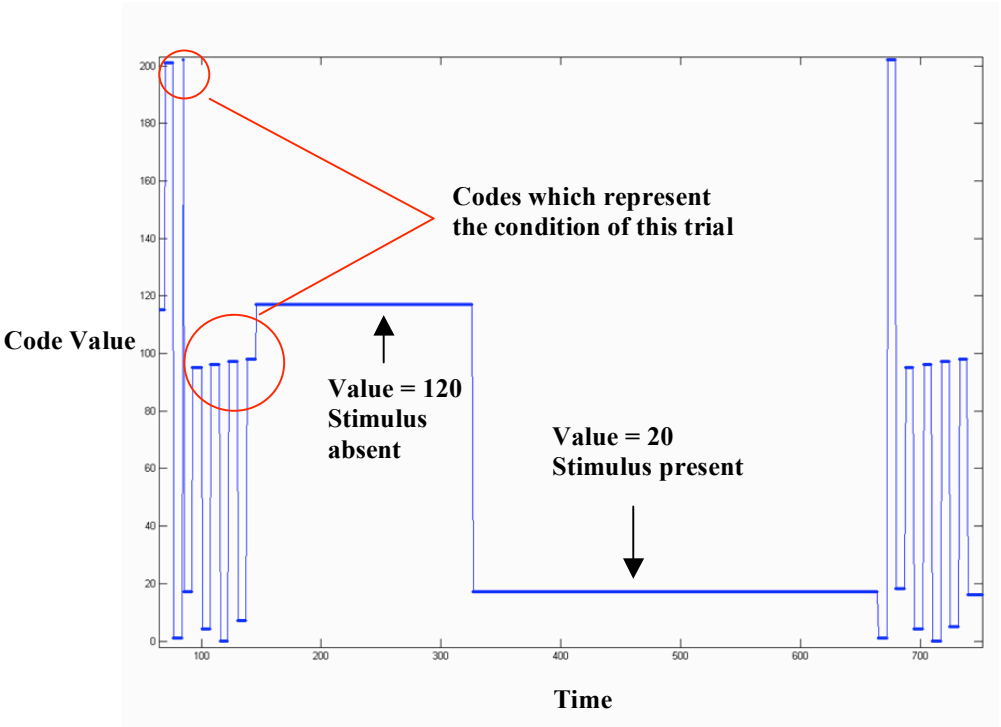


Figure 29. Example of Xdat data.

The list of code values in this dataset was useful for two main reasons: it provided a backup for trial information throughout the experiment and allowed easy parsing of the eye position data.

The first step in converting the positional data from relative ASL units to visual degrees was accomplished through a conversion factor derived from calibration for translating ASL units to position in degrees. The distance between the corner points on the calibration target were measured and divided by the difference in ASL units of the same points. Each of three points in each row and column were equidistant from each other on the screen so only the distance in one row and column was needed, as seen in the figure below.

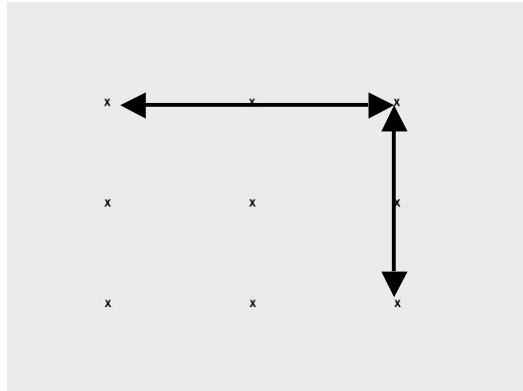


Figure 30. Calibrating eyetracker

The equation to get eye position in terms of visual degrees on the screen from relative units is

$$Deg = \left(\frac{CF}{A}\right) * ASL \tag{3.2}$$

where, *ASL* is the eye position data in relative ASL units, A is the distance between points in relative ASL units and CF is defined as

$$CF = \left(\tan^{-1}\left(\frac{d_{cm}/2}{d_{obs}}\right) * 2\right) \tag{3.3}$$

where d_{obs} is the distance in centimeters the observer is from the screen and d_{cm} is the difference in centimeters between the corresponding points in Figure 30.

Figure 31 below shows an example of eye position data after the conversion to visual degrees for one observer in Experiment Four. The saw-tooth patterns are representative of when the observer tracks from left to right, following the Gabor across the screen and then returns rapidly to the starting point. Each Gabor stayed on the screen for the same length of time. Thus, faster moving Gabors moved further distances than the slower moving ones. Some of the saw-tooth shapes are longer with greater slope than

others. This is indicative of faster, longer moving Gabors. Indeed, the slowest moving Gabor barely had any rise in the position data since it moved very slowly and over a small distance.

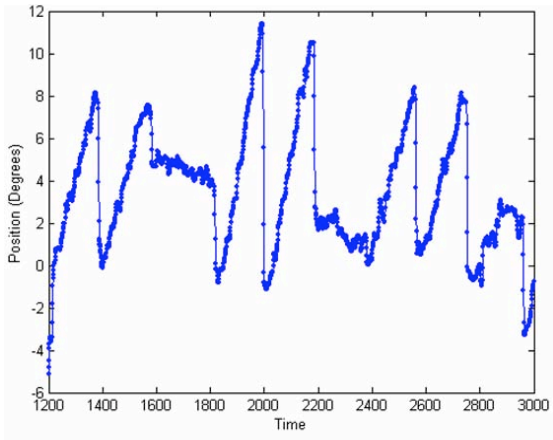


Figure 31. Example graph showing an observer's eye position data for Experiment Four.

There were 50 trials per condition and 8 conditions, which meant there were 400 trials per experiment. The resulting data file contained a long list of numbers after the experiment was over and consequently a robust method of pulling out only the pertinent information was needed and this is where the Xdat codes became useful. The figure below shows how the two datasets from the eyetracking record were used together. The codes not only described the condition of the particular trial but also signaled the beginning of the intervals. A program was written to look for the codes and record the corresponding index values in order to get the eye position data for each interval of each trial.

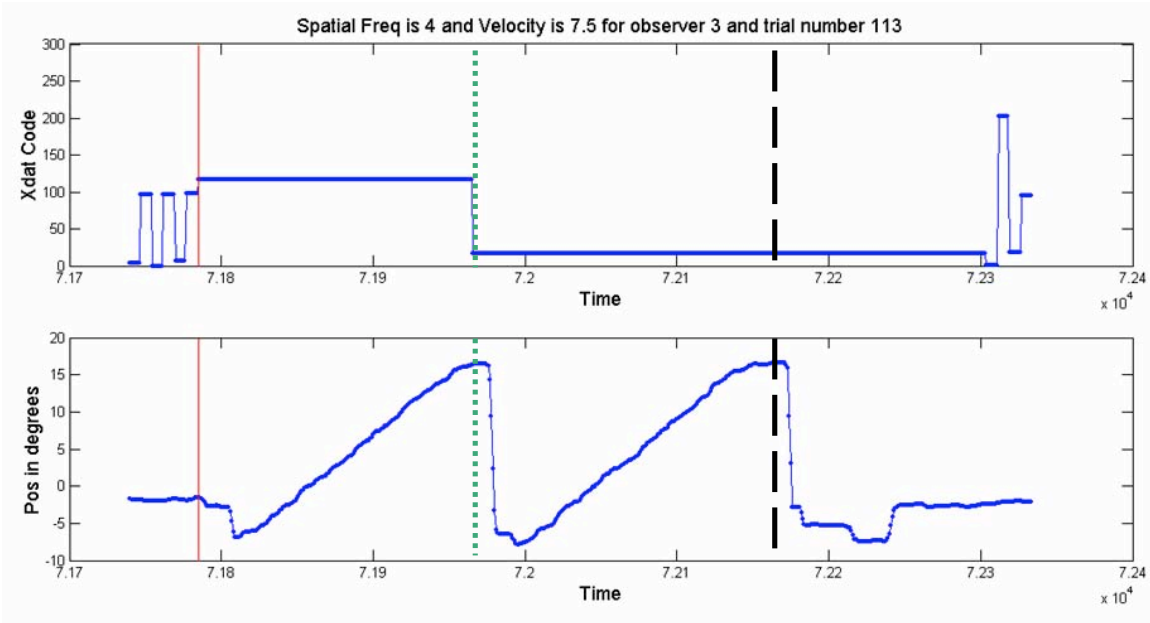


Figure 32. Plot showing how Xdat could be used to parse eye position

In Figure 32 above, the solid red line at far left signals the beginning of the trial, the green dotted line signals the beginning of the second interval and the black dashed line at far right signals the end of the trial. The Xdat values continue to run after the trial is over, as seen to the right of the black line in the top plot of Figure 32, because the program is waiting for user response after the second interval. The Xdat codes enable the parsing software to align the eye position data with the start and end of each interval per trial. The Xdat values also encoded the spatial frequency of the Gabor patterns, the velocity of the Gabor across the screen and the trial numbers making it a very robust data set for experimental analysis.

Once the data was in angular units it was possible to convert the values to eye velocity. There were two methods employed. One was to take the average of the instantaneous velocity computed at the ASL rate of every 1/60th of a second over some time range. Another was to calculate the average slope of the eye position data over a

time range. The figure below shows an example of both methods over the same time period. The black lines over the red region in the top plot is the range along each interval that the slope was calculated and the blue region in the bottom plot is the range that was used to average the velocity. Both methods provided similar results per interval and were therefore averaged together.

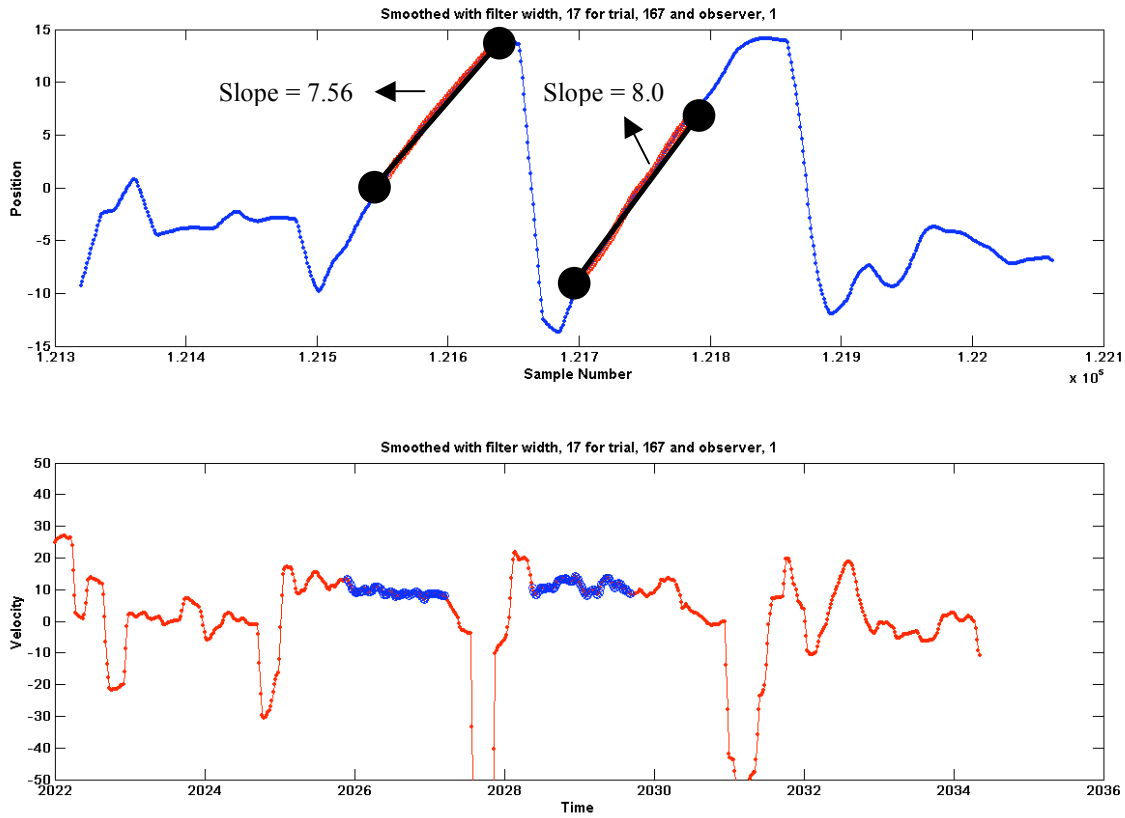


Figure 33. Two methods to calculate velocity. The top plot shows the ranges used to calculate slope. The bottom plot shows the ranges used for the average velocity.

Chapter 4 – Results & Discussion

4.1 Experiments One through Four

4.1.1 Overview of Results

Experiments one through four were conducted in order to parameterize and test a 2D spatiovelocity CSF model. The output from these experiments was contrast sensitivity to sine wave gratings. The graphs in Figures 34 and 35 are examples of how all the experimental results will be shown. The error bars represent two Standard Error of the Mean (±2SEM) for all graphs in this section. In the graph below, (Figure 34), the abscissa represents temporal frequency and the ordinate represents log contrast sensitivity. As stated in the Experimental section, the Gabor patterns were created in order to keep spatial and temporal frequency constant for each condition. Through the relationship between spatial and temporal frequency a range of velocities were used for each spatial frequency (see Table 1) and because of this, the motion is described in terms of temporal frequency to show the results in a more consistent manner. Figure 34 below represents a spatial frequency of 4 cycles per degree (CPD) for Experiment Two. So, for this spatial frequency the sensitivity decreased as the sine wave pattern increased in temporal frequency, (moving rightward along the abscissa). In other words, for a Gabor pattern with spatial frequency of 4 CPD, as the sine wave image moved across the retina faster, contrast sensitivity decreased.

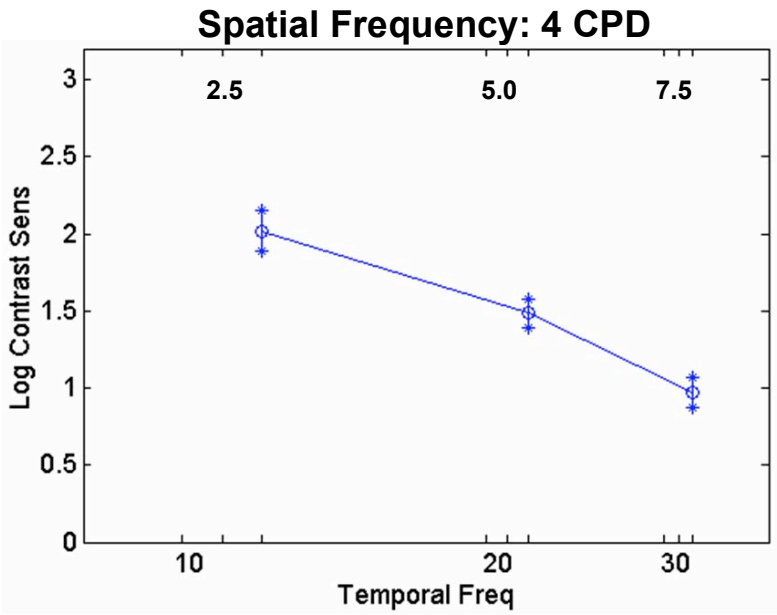


Figure 34. Example plot of contrast sensitivity for 4 CPD as a function of temporal frequency with velocity in deg/sec along top.

The next graph, Figure 35, also shows sensitivity data but plotted a different way. The abscissa now represents spatial frequency and the graph is for a temporal frequency of 10 Hz. This results in a different velocity at each spatial frequency along the abscissa, which is seen at the top of the graph. This graph shows that as spatial frequency increases for a temporal frequency of 10 Hz, the sensitivity decreases. The results in this section are shown in one of these two graph formats. The graphs for experiments two through four, beginning with Figure 37, show three curves, the blue curves represent all observers, while the red and green curves represent two observers who ran the experiments four times each. They are plotted together in order to show intra and inter-observer variability.

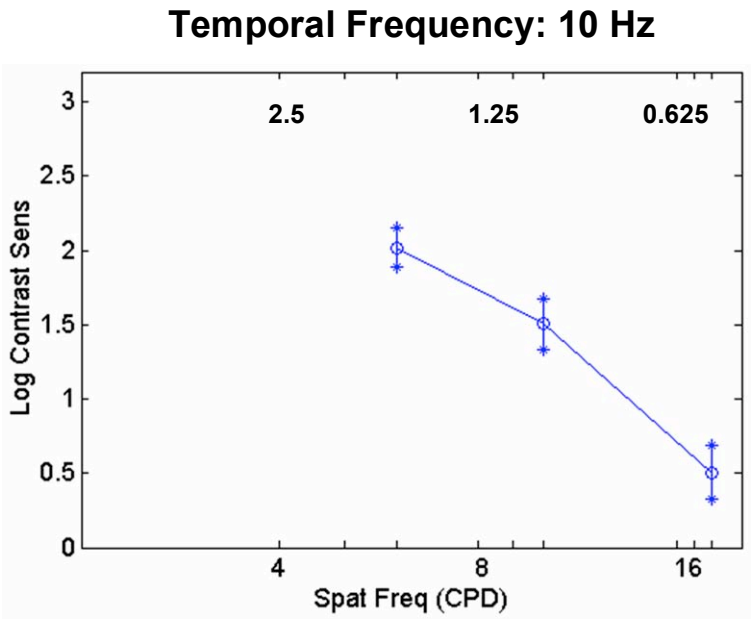


Figure 35. Example plot showing contrast sensitivity for 10 Hz as a function of spatial frequency with velocity in deg/sec along top.

It should be noted that there were four major experimental situations described above in the Experimental section. In the first situation, which was the control experiment and seen in Fig 22 above, there was no movement at all. It was expected that as spatial frequency increased sensitivity would decrease. Indeed this is the case as seen in Figure. 36 below.

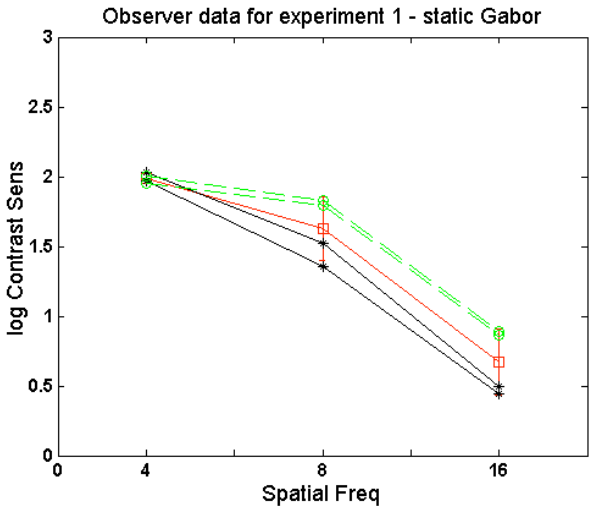


Figure 36. Results for Experiment One for 2 observers. The top dotted, green lines correspond to observer 1 and the bottom solid, black lines correspond to observer 2. The middle line, with errorbars, is the average of the 2 observers.

In the second situation, which was Experiment Two and seen in Figure 22 above, the sine wave pattern moved within the Gaussian window, which was stationary and a fixation point centered over the stimulus. In this case it was expected that sensitivity would decrease as the sine wave moved faster, as well as a decrease when spatial frequency increased. Results, as seen in Figure 37 and 38, showed that indeed sensitivity decreased as spatial and temporal frequency increased.

In the third situation, which was Experiment Three and shown in Fig 23 above, the movement was the entire Gabor along the screen but the sine wave pattern inside remained stationary relative to the Gabor and observers kept their eyes fixated on a centered, non-moving point. Likewise, it was expected that sensitivity would decrease with increasing velocity and spatial frequency. Results were as expected and very similar to those in Experiment Two. Figure 40 and 41 shows that as the Gabor increased in speed across the screen, sensitivity decreased and as the spatial frequency increased at a particular velocity, sensitivity decreased.

The fourth situation, which was Experiment Four and shown in Fig 24 above, had the same movement as the second experiment except that now observers' eyes tracked a fixation point that was centered over the stimulus while it was in motion. If the observer tracked the stimulus perfectly then the stimulus would remain stationary on the retina. If that were the case, sensitivity would be similar to results from Experiment One because the stimulus would have 0 retinal velocity. As seen in Figure 44 sensitivity dropped as spatial frequency increased. However, there was no drop in sensitivity as velocity increased, as seen in Figure 43. Since sensitivity only changed with spatial frequency, the question was whether the results would be similar to the traditional CSF experiment where the Gabors were not in motion. In the discussion below the experimental results will show that indeed contrast sensitivity of a tracked stimulus is the same as that of a stationary stimulus.

4.1.2 Results with Discussion

The results from Experiment Two are shown in Figure 37, where the sine wave pattern is in motion relative to the Gabor, and that as temporal frequency increased for a given spatial frequency, sensitivity decreased. Likewise, Figure 38 shows that as spatial frequency is increased for a given temporal frequency sensitivity decreased. Again, this was the expectation based on previous research[23] and what was already known about the HVS. Wandell states (pp. 201) that the contrast and thus sensitivity of high spatial frequency targets is reduced due to neural and optical effects.[18] It is also known that center-surround retinal ganglion cells in the retina are less sensitive to high spatial frequencies.[22] Watson has discussed in more detail physiological reasons for lower sensitivity at higher spatial and/or temporal frequencies.[22]

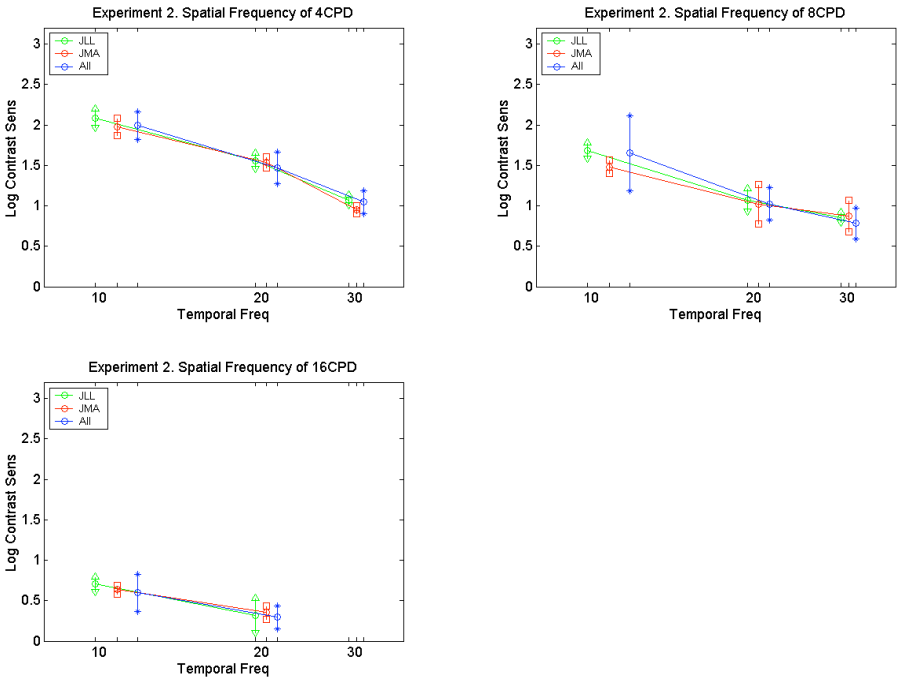


Figure 37. Experiment Two; plot of each contrast sensitivity for each spatial frequency as a function of temporal frequency.

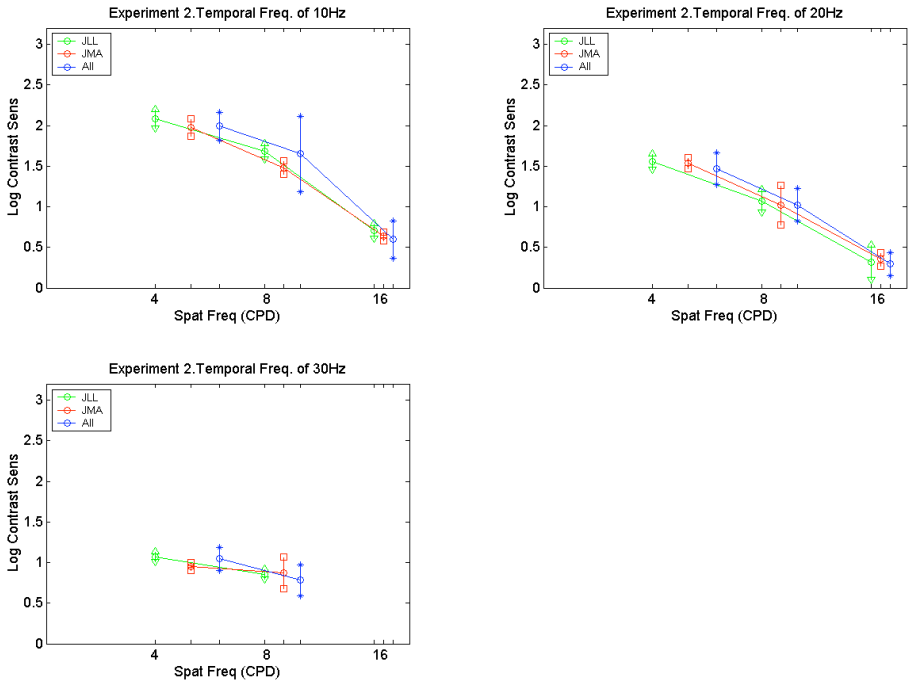
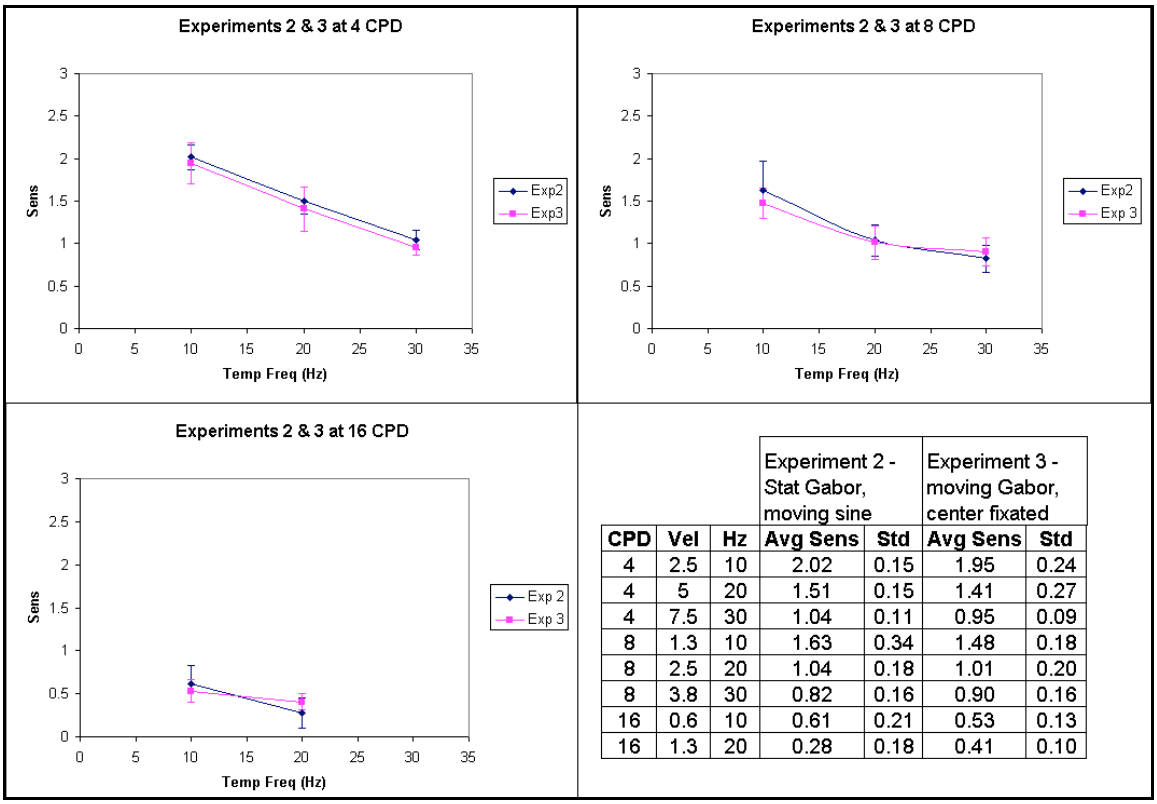


Figure 38. Experiment Two; plot of contrast sensitivity for each temporal frequency as a function of spatial frequency.

The motion in Experiment Three was different than Experiment Two in terms of the movement along the screen. Unlike Experiment Two, temporal frequency for this was the actual movement of the Gabor across the screen. The difference is that the Gabor is moving across a larger section of the retina in Experiment Three and for higher temporal frequencies the Gabor moves into and out of the field of view of the observer while in Experiment Two the motion was always in the central fovea of the observer. This could affect the results because the peripheral retina, with lower spatial resolution, may be contributing. However the contrast of the Gabor was gradually increased from the background level to full contrast. This effectively put the Gabor at full contrast within the same region on the screen as when the Gabor was stationary. If the motion on the retina

was the same then there should be no difference between the results of Experiment Two and three in terms of sensitivity. Indeed, the results are statistically the same between both experiments as seen from in figure below. Each plot in the figure is a different spatial frequency and within each graph the results for Experiments 2 and 3 are plotted together for comparison. Additionally, the table in the lower right of the figure has the sensitivity values with the standard deviation for both experiments.



			Experiment 2 - Stat Gabor, moving sine		Experiment 3 - moving Gabor, center fixated	
CPD	**Vel**	**Hz**	**Avg Sens**	**Std**	**Avg Sens**	**Std**
4	2.5	10	2.02	0.15	1.95	0.24
4	5	20	1.51	0.15	1.41	0.27
4	7.5	30	1.04	0.11	0.95	0.09
8	1.3	10	1.63	0.34	1.48	0.18
8	2.5	20	1.04	0.18	1.01	0.20
8	3.8	30	0.82	0.16	0.90	0.16
16	0.6	10	0.61	0.21	0.53	0.13
16	1.3	20	0.28	0.18	0.41	0.10

Figure 39. Comparison of results for Experiments 2 & 3

This indicates that processes in the visual system respond to sine wave gratings in a similar manner regardless of how it is placed in motion. This will allow placement of the Gabors in motion across the screen while keeping the sine wave inside the Gabor stationary. The graphs in Figure 40 show the results from Experiment Three as the temporal frequency increased. In these graphs the velocity (in deg/sec) are along the top

of each plot and the curves are similar to those in Figure 37. Figure 41 again shows similar results to Figure 38 in Experiment Two.

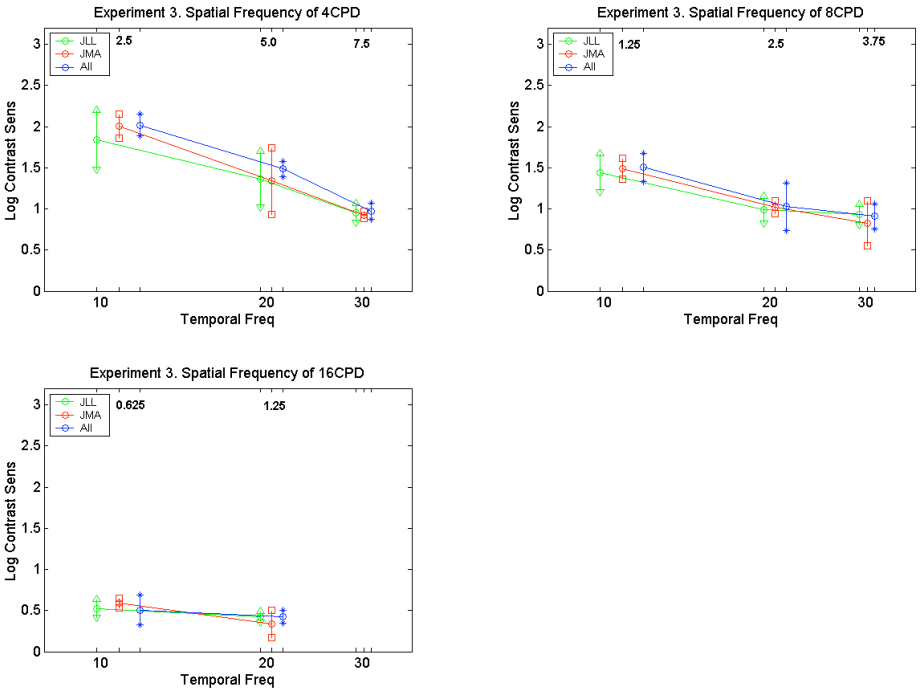


Figure 40. Experiment Three; plot of contrast sensitivity for each spatial frequency as a function of temporal frequency

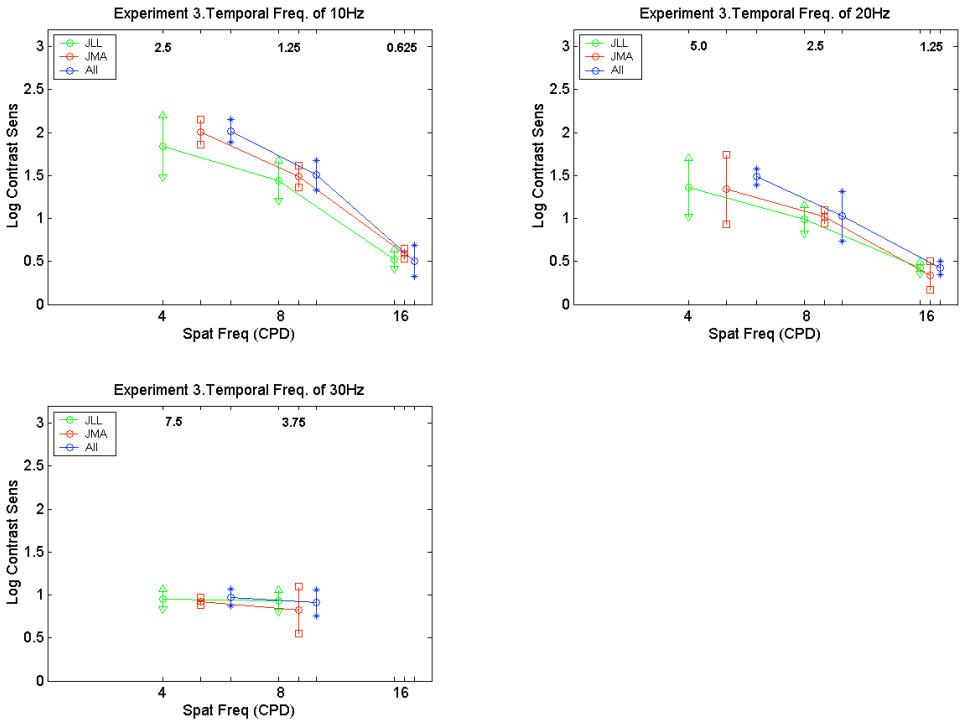


Figure 41. Experiment Three; plot of contrast sensitivity as a function of spatial frequency at each temporal frequency.

As seen by the curves in the graphs of Figure 42 below, there was a difference in results for experiments two, moving sinusoids within a stationary Gaussian envelop, and three, moving Gabor/center fixated, and the results in four, tracked/moving Gabor. While the stimulus motion was the same in Experiment Four as in three, the difference is that observers were fixated on the screen center in Experiment Three and tracking the Gabor in Experiment Four. As the temporal frequency increases, the sensitivity does not decrease for Experiment Four as it does for Experiments 2 and 3.

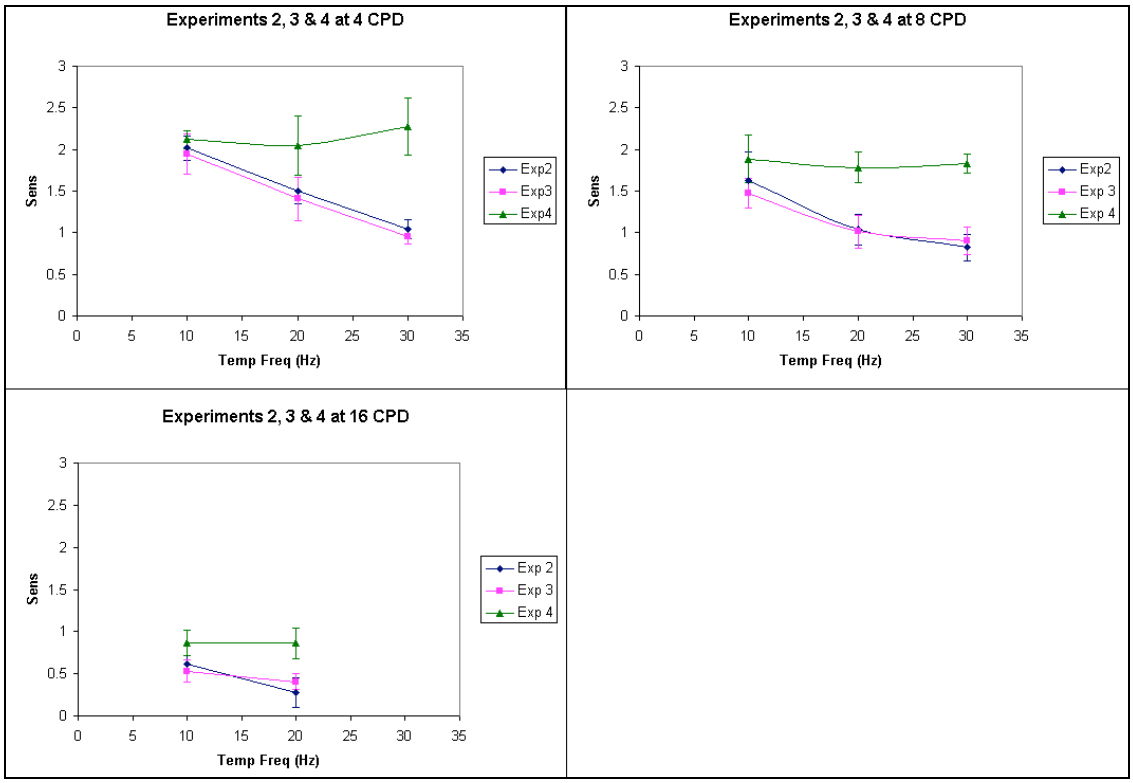


Figure 42. Comparison between results for Experiments 2 & 3 and those in Experiment Four.

Furthermore, the graphs in Figure 44 show that as spatial frequency increases for a given temporal frequency, sensitivity decreases. However, graphs in Figure 43 show there is no decrease in sensitivity as temporal frequency is increased. This indicates that observers were doing a good job tracking the Gabor and that there is no substantial retinal velocity. Furthermore, it will be shown in section 4.3 that eye movements did not affect contrast sensitivity in these experiments. This will be further supported in section 4.3 EyeTrack Analysis when the results of observer eyetrack are discussed. Furthermore, Figures 45 and 46 show that contrast sensitivity was the same in this experiment as in the control experiment, (stationary Gabor).

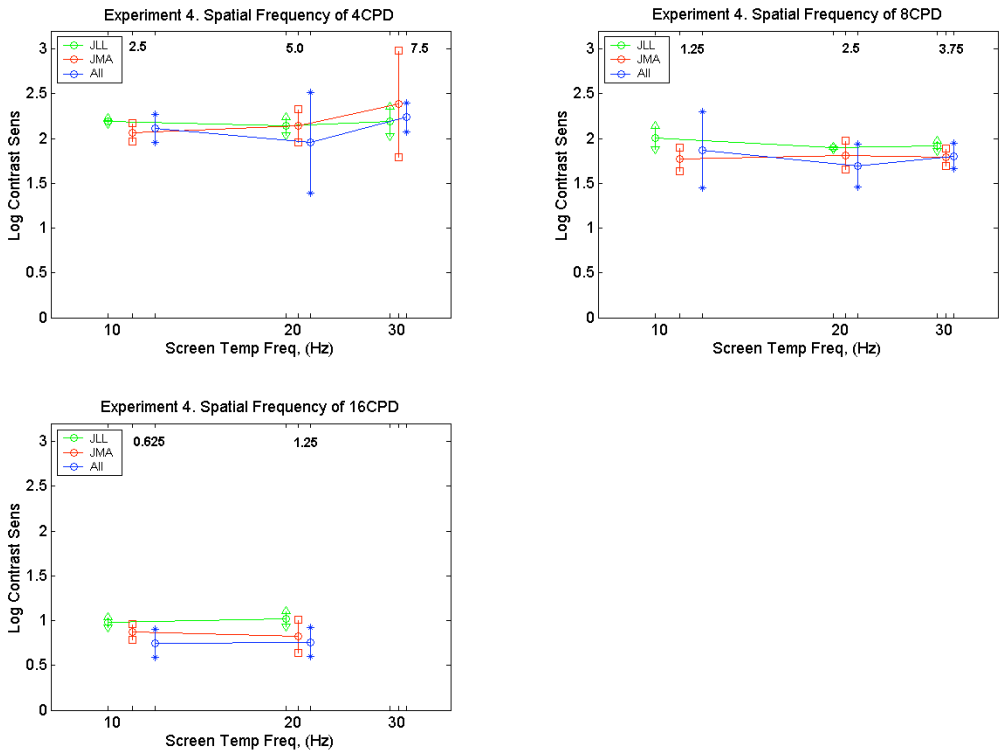


Figure 43. Experiment Four; plot of contrast sensitivity for each spatial frequency as a function of temporal frequency

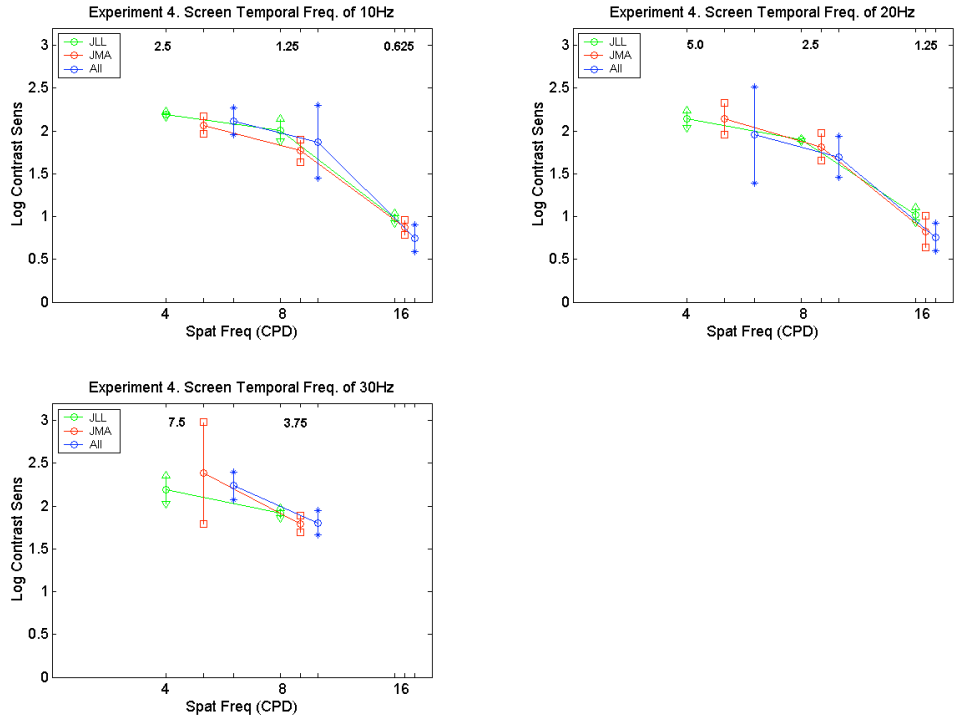


Figure 44. Experiment four; plot contrast sensitivity for each temporal frequency as a function of spatial frequency.

The results of Experiment One, the control experiment, compare favorably to results in the literature where increasing spatial frequency decreases sensitivity in the region tested, (as seen in Figure 9). For the two top graphs in Figure 43 the CSF is statistically indistinguishable from those in Experiment Four, as seen in the figure below from the overlap in errorbars at each point and similarity in curve shape. Comparison of the top two graphs in Figures 43 and the curves in Figures 45 and 46 below shows that the CSF shape is quite similar to that from Experiment One, the control experiment.

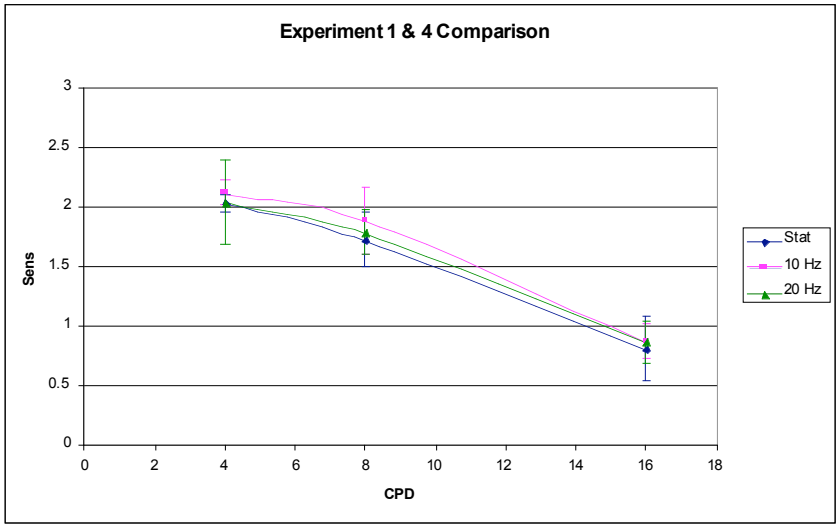


Figure 45. Experiment One, Static Gabor, plotted with Experiment Four, tracked Gabor; showing statistically similar results

In order to test the hypothesis that sensitivity results would be the same for stationary Gabors as for tracked Gabors the results from Experiment One, the control experiment, were compared to those from Experiment Four. Experiment One is considered a traditional CSF experiment in the sense that there is no motion involved, so the independent variable is spatial frequency. In the figure below, the blue line represents average sensitivity from all observers for Experiment One and the red, green and black lines are from Experiment Four and represent average contrast sensitivity for observers tracking a Gabor moving along the screen at corresponding temporal frequencies of 10, 20 and 30 Hz respectively. It is evident from Figure 46 there was not significant differences.

There does appear to be a slight trend, although not statistically significant, for higher sensitivities for eyes in motion, especially at lower spatial frequencies. It is plausible that eye movements slightly increase sensitivity for low spatial frequencies but as spatial frequency increases these slight eye movements, which result in retinal

velocity, become less important. The idea that eye movements actually increase contrast sensitivity have been shown by Kelly[12] and in this research the fact that retinal velocity decreases sensitivity will be discussed in more detail later. Nevertheless, Robson shows that sensitivity to gratings at low temporal frequencies is modulated by spatial frequency and also that low sensitivity to low spatial frequency is modulated by temporal frequency. Therefore, it is plausible that contrast sensitivity increases slightly for a low spatial frequency when the eyes are in motion because there could be small eye movements resulting in small retinal velocity[4, 12].

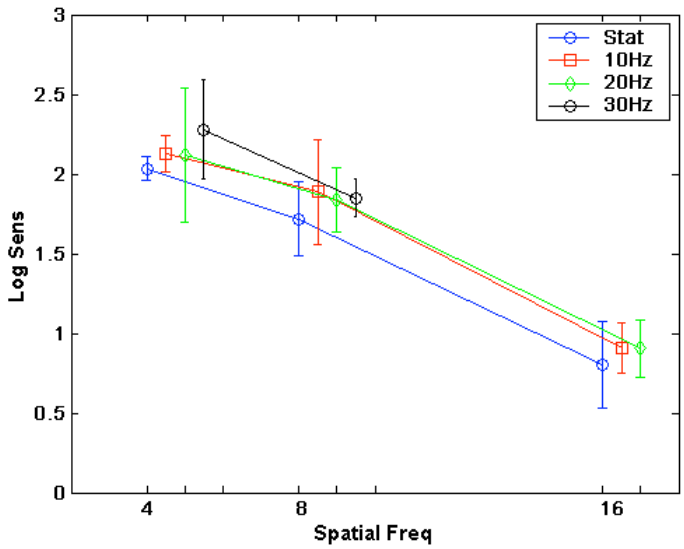


Figure 46. Contrast sensitivity data for static Gabor compared to Gabor tracked across screen. The offset of each data point is for comparison only. In actuality these points do overlap each other.

The results from Experiment Three, (Gabor in motion with observers screen-center-fixated), will be used in the parameterization of the 2D spatiovelocity CSF model. The results from this experiment are useful because they indicate how sensitivity changes with retinal velocity. It was important to determine if there was a difference between the two types of retinal motion between Experiments two and three. Because the statistics,

seen in Figure 46 above, show that the results were the same, those from Experiment three will be used since they better relate to natural eye movements; where different spatial frequencies are in motion at the same velocities across the retina. The results from Experiment one was also referred to in the 2D CSF development.

4.2 Parameterizing and Verifying the Spatiovelocity CSF Model

4.2.1 Parameterizing CSF Model

As stated previously the model from D.H. Kelly, which was further updated by S. Daly, is used as the 2D spatiovelocity model. Daly's modification adds constants c_0, c_1 and c_2 to Kelly's model (found in Equation 2.1). The Kelly-Daly equation is found in Equation 4.1:

$$CSF(\rho, v_R) = k \cdot c_0 \cdot c_1 \cdot c_2 \cdot v_R \cdot (c_1 2\pi\rho)^2 \exp\left(-\frac{c_1 4\pi\rho}{\rho_{\max}}\right) \tag{4.1}$$

Where $k = s_1 + s_2 \cdot \left|\log\left(c_2 v_R / 3\right)\right|^3$ (4.1.2)

Furthermore, $\rho_{\max} = \dfrac{p_1}{(c_2 v_R + 2)}$ (4.1.3)

s_1, s_2 and p_1 are the original constants provided by Kelly. The variable v_R is velocity which for purposes in this chapter is actually retinal velocity. Daly's constants allow the model to be adjusted based on a particular experimental setup. The constant c_0 can be adjusted for peak sensitivity, c_1 for maximum spatial frequency cutoff and c_2 for the maximum critical flicker frequency[4]. If these values are set to 1 then Kelly's model is produced. The other variables, ρ, k, and ρ_{max} are defined as in section 2.3, page 20.

The constants c_0, c_1, c_2 were optimized in the following way. A nonlinear least-square routine was run in MATLAB[37] in which the sensitivity values from the model was fit to the experimental results corresponding to the spatial frequency and velocity data points. Values of the constants were modified until the differences between the model and the experimental results were minimized. These values found by fitting the model to our experimental results are as follows:

$c_0 = 0.6329$

$c_1 = 0.8404$

$c_2 = 0.7986$

In Figure 47 below the 2D CSF model was built from the optimized constants, c_0, c_1, c_2.

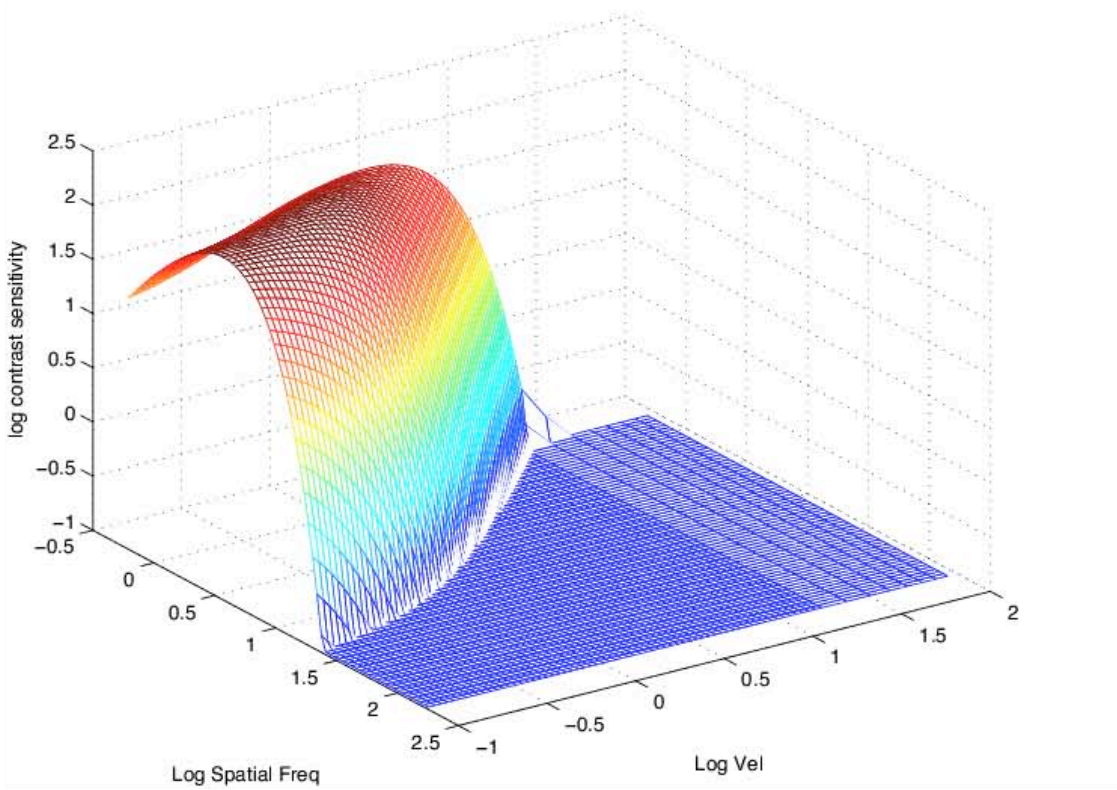


Figure 47. Spatial-velocity CSF model found after optimization of Daly-constants

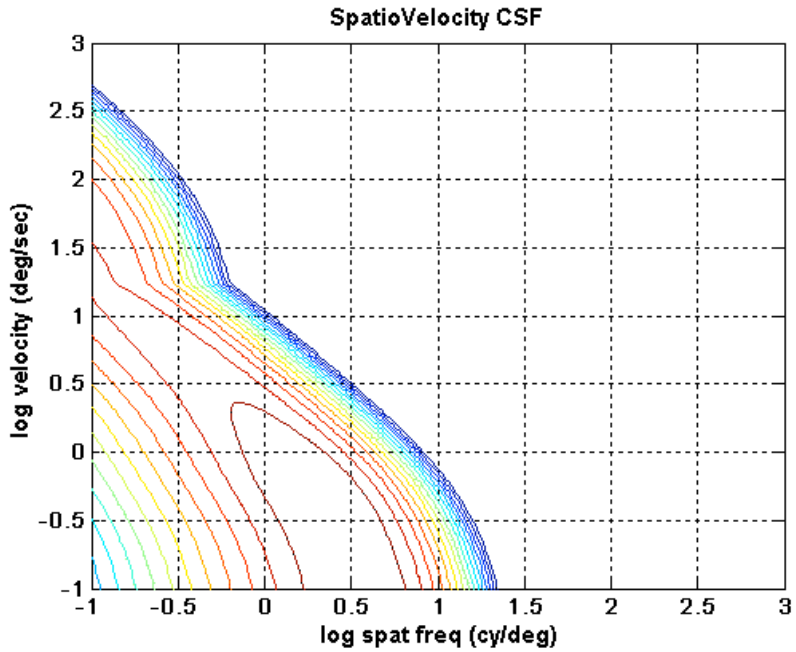


Figure 48. Contour plot of 2D spatioVelocity CSF

The contour plot in Figure 48 above shows a maximum cut-off spatial frequency of approximately 20 CPD. Because of viewing distance and parameters of the CRT there were 41 pixels-per-degree and because there is a minimum of 2 pixels for 1 cycle there was a maximum of 20.5 CPD available in these experiments. The discontinuity seen between 1 and 1.5 deg/sec and between -0.5 and 0 CPD is an artifact of plotting resolution.[4]

4.2.2 Verifying the model

The final experiment, which was a moving disembodied edge, was used to verify the prediction results from the model. In this case the edge was shown at three contrast levels and four velocities. Results showed no change in sensitivity to the edge as velocity increased. However, the minimum contrast level had the lowest sensitivity while the middle and high contrast level had the same sensitivity, which was higher than the lowest contrast level. The results from this experiment are seen in Figure 49 below and show

how sensitivity to edge sharpness changes with different contrast and velocities. The abscissa represents the three contrast levels tested and the ordinate is edge-blur sensitivity. Each line in the graph represents a different velocity that the edge moved. The results, seen in Figure 49 below show no difference in the results between the velocities. Sensitivities to the two higher contrast levels are statistically identical. The error bars for each velocity overlap each other, indicating zero retinal velocity at each contrast level. These results will be compared to the model predictions.

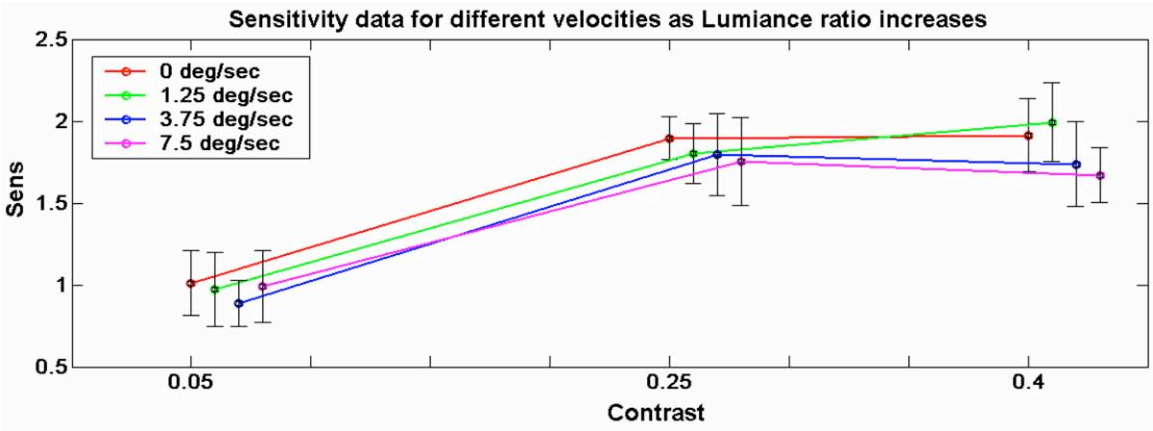


Figure 49. Results of Experiment Five where each line represents a different velocity.

As stated previously, a blurred edge can be expressed in the frequency domain, which allows the edge to be described by its spatial frequency content. See Figure 19 above for a graph of a sharp and blurred edge in the frequency domain. Due to good tracking of the moving edges indicated by overlapping error bars in Figure 49 and eye tracking records (discussed below), the minimum velocity data from the 2D CSF model was used for verification.

For verification, the CSF was normalized to between 0 and 1. Then, the blurred edge 1D Fourier transform was multiplied by the normalized CSF data for zero velocity followed by integration, see Figure 50.

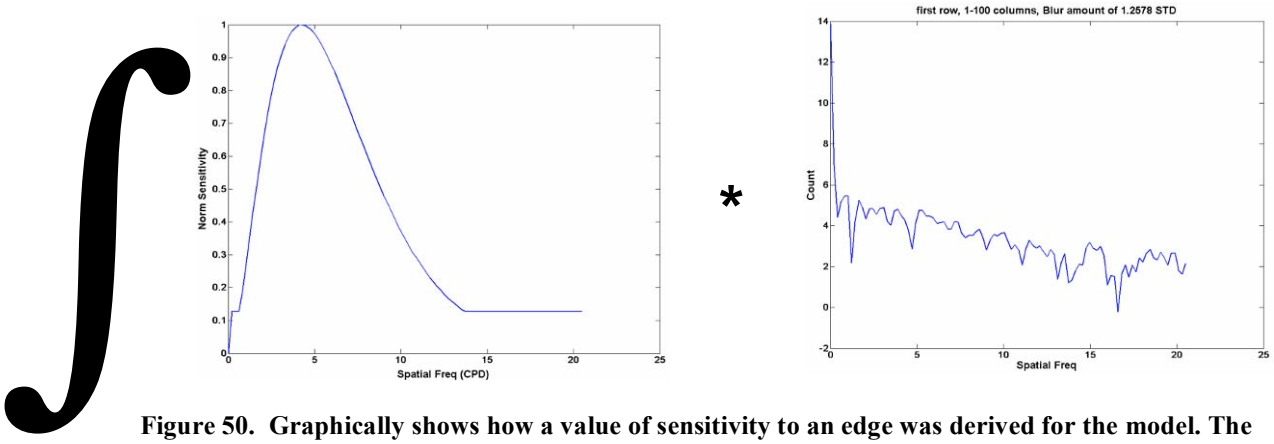


Figure 50. Graphically shows how a value of sensitivity to an edge was derived for the model. The normalized CSF is shown on the left and the FFT of a blurred edge on the right.

Because the results from the experiment are in terms of threshold, the values are different than the values from the above process. Therefore, both the experimental and model results are normalized and plotted together. Figure 51 below shows the similarities.

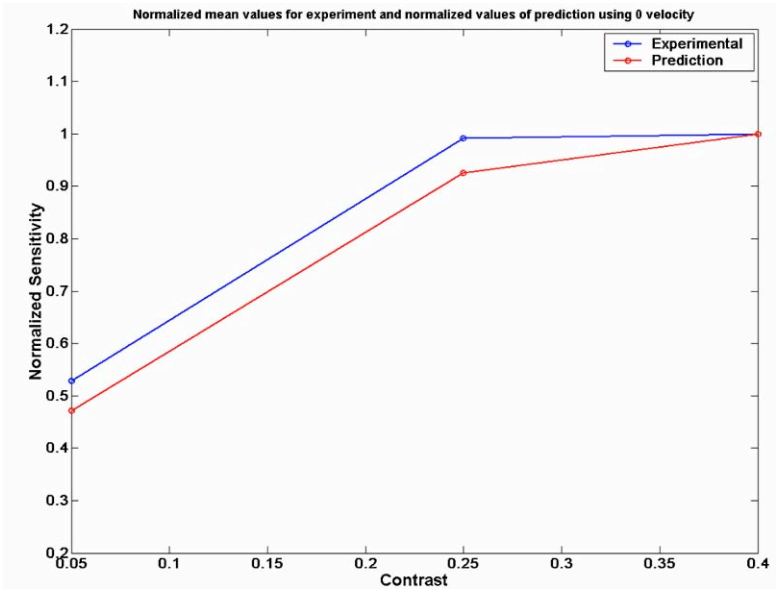


Figure 51. Model results compared to experimental results

4.3 EyeTrack Analysis

The metric of how well a person tracks a stimulus in this research is referred to as gain. If an observer tracks something perfectly then his/her gain would be 1 and if they did not track at all it would be 0. If the observer moves his/her eye faster than the object, the gain is greater than 1, and if eye movements are slower than the object, the gain is less than 1.[38] The method for calculating gain was to first calculate the eye velocity and then divide the calculated velocity by the target velocity. For example, in Figure 52 below the observer's calculated velocity is 11.41 and 11.35 for interval 1 and 2 respectively. Therefore, the average calculated gain between the two intervals would be 1.07 if the target velocity is 10.59 cm/sec. Note that this gain value is for a single trial out of a total of 50 for this condition. To analyze the eyetrack results an analysis of variance (ANOVA) was run to determine what variables in Experiment Three affected gains. The ANOVA results will be shown in terms of an average gain across 50 trials since there is only a single sensitivity value that is calculated from all 50 trials. There are several possible reasons that the gain metric would not accurately report tracking: there could be noise in the eyetracker, the person could be looking away from the screen, they could be lagging behind the Gabor and then making "catch up" saccades. All these would have to be accounted for prior to calculating velocity and gain values. However, it was expected from results thus far that most observers would have gains near 1 since it was already shown that the sensitivities did not change with different velocities.

It is hypothesized that if an observer has a gain value different than 1 for a particular trial then his/her contrast sensitivity will be different. In other words, as gain approaches 1, sensitivity approaches a maximum for that particular condition. In this

research observers' eye movements were tracked for the main purpose of calculating their gains and actual velocities.

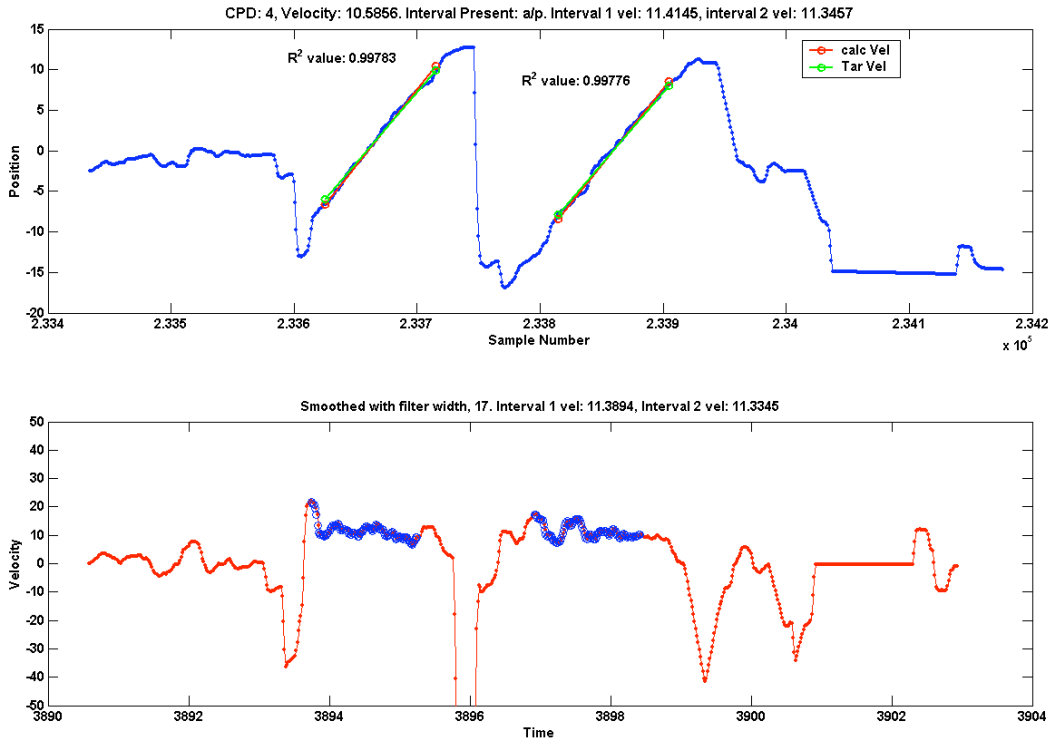


Figure 52. Two different ways of calculating velocity. Top graph uses slope and the bottom graph uses average of calculated velocity.

As stated previously it was found that observers' sensitivity did not change as they tracked Gabors but now it can be explicitly stated why this is the case. Correlations could then be made between observers' eye movements and sensitivity data.

Through Experiment Three, where the Gabor pattern moved across the observer's retina while his/her eyes were fixated, it was shown that sensitivity decreased with increasing retinal velocity. Through Experiment Four, where the observer tracks the Gabor across the screen, it was shown that contrast sensitivity did not change. This result shows that eye movements did not affect contrast sensitivity. Therefore, it can be inferred that gain was not affected and that observers tracked very well. It can be inferred that if

the gain value deviates from 1 then there will be retinal velocity and possibly a change in sensitivity.

The following paragraph will discuss the results of calculated gains. It should be remembered that there was not a change in sensitivity as a function of eye movements. However, there were some noticeable trends and there were some observers with lower or higher sensitivity. It was hoped that an analysis of their eyetrack records relative to all other observers would explain these different values. An analysis of variance was performed on the eyetrack data and the calculated gains for each observer. In order for an appropriate analysis of gains versus velocities using an ANOVA three velocity levels were used so that the minimum temporal frequency per spatial frequency is 1 and the medium temporal frequency is 2 and the highest frequency is 3, as shown in Table 6.

Table 6. Table on left is actual velocities which correspond to spatial frequencies. Table on right is relative velocities used in ANOVA.

Spat Freq (Cyc/Deg)	Temporal Freq (Hz)		
	10	**20**	**30**
4	2.5	5.0	7.5
8	1.25	2.5	3.75
16	0.625	1.25	

Spat Freq (Cyc/Deg)	Temporal Freq (Hz)		
	10	**20**	**30**
4	1	2	3
8	1	2	3
16	1	2	

Some interesting results of the ANOVA are examined. First, the average gain over all observers was 0.956 +/- 0.017. Gains for a spatial frequency of 16 CPD were higher and statistically different than those for 4 and 8 CPD. The table below shows results from the ANOVA and there is a significant difference for 16 CPD compared to both 4 and 8 CPD.

Table 7. ANOVA results showing differences in gain for the different spatial frequencies.

(I) SF	(J) SF	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
4.00	**8.00**	.0005	.01972	.980	-.0386	.0396
	16.00	-.0717(*)	.02205	**.002**	-.1154	-.0281
8.00	**4.00**	-.0005	.01972	.980	-.0396	.0386
	16.00	-.0722(*)	.02205	**.001**	-.1159	-.0286
16.00	**4.00**	.0717(*)	.02205	**.002**	.0281	.1154
	8.00	.0722(*)	.02205	**.001**	.0286	.1159
Based on observed means.						
* The mean difference is significant at the .05 level.						

Likewise the gains for the slowest velocities (0.625, 1.25 & 2.5 deg/sec) were statistically different than those at the medium and fast velocities. In this case as velocity increased the gains decreased. The table below also shows that for a relative velocity of 1, which corresponds to the slowest velocity for each spatial frequency, there is a significant difference in gains from relative velocities of 2 and 3.

Table 8. ANOVA results showing differences in gain for the different relative velocities.

(I) RELV	(J) RELV	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	**2.00**	.0421(*)	.01972	**.035**	.0030	.0812
	3.00	.0533(*)	.02205	**.017**	.0096	.0970
2.00	**1.00**	-.0421(*)	.01972	**.035**	-.0812	-.0030
	3.00	.0112	.02205	.613	-.0325	.0549
3.00	**1.00**	-.0533(*)	.02205	**.017**	-.0970	-.0096
	2.00	-.0112	.02205	.613	-.0549	.0325
Based on observed means.						
* The mean difference is significant at the .05 level.						

This result is counterintuitive given the overall result of no change in sensitivity for the different spatial frequency and velocity conditions. However, it is reasoned that although there is a difference in terms of eye movements for particular spatial frequencies and velocities, this difference did not lead to a statistically significant change in sensitivity for the region tested. Also, the Gabors physically moved over a shorter distance around the center of the screen for the slowest velocities and it is possible that observers made anticipatory eye movements that were masked by the noise in the eye tracker. These type of eye movements would result in higher gain values at the lowest velocities since observers would be jumping ahead of the fixation point they were tracking, which is what is observed through the ANOVA calculation. Although not a statistical difference, there is the trend that as velocity increases for each spatial frequency the gain decreases, as seen in the table below in the column labeled "Mean".

Table 9. ANOVA results showing interaction between relative velocities and spatial frequencies.

RELV	SF	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1.00	4.00	.948	.024	.900	.995
	8.00	.945	.024	.897	.993
	16.00	1.061	.024	1.014	1.109
2.00	4.00	.932	.024	.884	.980
	8.00	.938	.024	.890	.986
	16.00	.958	.024	.910	1.006
3.00	4.00	.934	.024	.886	.982
	8.00	.929	.024	.881	.977

This indicates that there is an interaction between spatial frequency and velocity, where overall there was a decrease in gain for increasing velocity but as the spatial frequencies increased the gain increased. This may indicate that for the regions tested observer's track better for higher spatial frequencies at higher velocities and therefore were more sensitive to these patterns. However, these trends are small compared to the overall results.

Additionally, there was no difference between the Stimulus Present or Stimulus Absent intervals for each trial. This makes sense because observers were told to track the fixation point and it was always present and at the same contrast level. However, there was a *trend* for higher gain values for the Present interval. This can be explained because if the stimulus was present and visible then observers would be more attentive and therefore track the stimulus better. It could also be that when the stimulus was present, visible and below a maximum pattern velocity of some amount (Daly used 80 deg/sec[4]) observers would move their eyes around a little more instead of remaining fixated on the stimulus if it were only a point and not a more interesting stimulus. This might result in higher gain values although it is recognized that the gain value could average out if the observer moves their eye back and forth around the fixation point. In any case this would result in higher sensitivity due to additional retinal velocity as the eyes scanned across the Gabor pattern. The sensitivity would increase at the contrast threshold level because of this motion. Note that this is not counter to the previous results of sensitivity decreasing for increasing velocity across the retina since there is an increase in sensitivity as velocity increases slightly from a minimum velocity followed by a decrease in sensitivity as the velocity continues to increase.

It is seen that the majority of gains are close to 1. However, there were some observers with a low average gain. There could be many reasons for this such as inattentiveness, a bad eyetrack record, or an actual inability to track the fixation point. Analysis of the video taken during each experiment would answer this question.

Since the majority of observers tracked the stimuli acceptably their data were included in the development of the 2D spatial-velocity CSF model, which, because of the use of non-stabilized retinal images is a more natural model for human contrast sensitivity than previous models. Furthermore, it can now be said that observers' contrast sensitivity is the same for tracked targets as when the target is stationary.

Chapter 5 - Conclusion

Results from all experiments show that observers did a good job tracking the stimuli and therefore it was possible to use their results to populate a spatiovelocity CSF model. Proper eyetracking was important because, from Experiment Three, it was shown that sensitivity is determined by retinal velocity and the eyetrack records proved that eye movements did not affect contrast sensitivity. Because the majority of observer gains were close to 1 for their eye movements it was determined that they were tracking the stimuli quite well. This inference was supported by the analysis of eye movement records. It was further reinforced from the sensitivity data from Experiment four, where observers were tracking moving Gabors.

A spatiovelocity CSF model was built from Experiments One, where Gabors were static, and three, where observers were fixated at the center of the screen and the stimuli were moving Gabor patterns. The model was verified with results from Experiment Five in which an edge moved at several different velocities. The predictions from the model agreed favorably with results from Experiment Five. The experiments only measured 11 points in the spatiovelocity space, so further experiments should be used to better refine the model.

Additionally, these experiments have shown that contrast sensitivity is the same whether a sine wave is in motion inside a Gaussian window or stationary with the Gabor pattern in motion across the field of view. This agrees well with the idea that at the lowest level of visual processing, retinal ganglion cells, are responding to both temporal variation and spatial pattern of the light without regard to the type of pattern motion. The

practical implication is that the model can be built using natural stimuli without having to consider the type of motion of the stimulus.

From the eyetrack records it was shown that observers did an acceptable job of tracking the stimuli based on the fact that the average gain for all observers was very close to one. This allowed the use of calculated retinal velocities in the 2D spatiovelocity CSF model from actual observers taking part in the experiment. Although it was good for building the CSF model, using velocities that allowed eyes to track the targets with no difficulty did not probe the limits of smooth pursuit. Observers had no trouble tracking a Gabor moving at 7.5 deg/sec, (which was the fastest speed in these experiments). Future experiments should test the limit of smooth pursuit. The results of that experiment could be used to make the CSF model more accurate.

Part II – Gamma Preference for LCTVs

Chapter 6 - Introduction

LCD televisions have staked their claim as one of the more popular choices for flat panel displays and their popularity will continue to increase as they become more economical. They hold several advantages over their counterparts including being larger, sharper and having higher luminance levels than other flat panels and traditional displays. Because they are larger and brighter, the viewing conditions associated with these TVs are not the same as with CRTs.

The goal of the research associated with Part II of this thesis is to determine how image quality is affected by the change in viewing condition brought on by these displays, (as compared to the traditional viewing condition with CRTs) and whether changes should be made to account for the particular viewing condition brought on by these displays. To explore this, two electro-optical transfer functions, (EOTFs), were simulated: a traditional gamma-function and an exponential modification to the inherent LUTs used to drive the red, green, and blue channels of the display.

In the first experiment a set of images modified to simulate how they would appear on displays with different transfer functions were presented in a dark surround at both the default luminance of the LCD television and at a reduced luminance level. The preferences for the various transfer functions were measured at both screen intensity levels using a paired-comparison experiment. By changing the luminance of the display, the results of the experiment on a more traditional, lower intensity, display were simulated.

In the second experiment, a surround illumination of 10% of the TV's white point, with chromaticities closely matched to the white point of the display, was introduced and the procedure from the first experiment was repeated in order to examine the effect of a dim surround on preferences to images shown on displays with different transfer functions.

Chapter 7 – Background

7.1 Viewing Condition and Simultaneous Contrast

Typical flat panel displays have sizes that range from 20” up to 60” and the average size is around 35”-40”. Additionally, these televisions have maximum luminance’s approaching 600 cd/m^2.[39, 40] These differences mean that the viewing conditions associated with new LCTVs are different from those of previous generations of televisions. In these experiments the change in luminance of the displays and how different surround conditions may affect the perceived contrast of displayed scenes are examined. The change in the perceived contrast may necessitate a change in the EOTF, from the standard used in traditional CRT televisions, in order to maintain or improve image quality under these new viewing conditions.

Changes in the surround conditions of a display can lead to changes in the perceived contrast of images on the display, which is related to the effect of simultaneous contrast.[41] For example, given a particular color, as the surround becomes lighter the color will appear darker. This can be seen in the image in Figure 53. In this image, each of the inner squares is the same gray value. From left to right, as the surround becomes darker and the inner square appears lighter. The phenomenon of simultaneous contrast has been attributed to lateral inhibition in the retina although higher mechanisms likely play a larger role[42]. As with simultaneous contrast, changes in the luminance of a display and its surround can change the perceived contrast in a scene.

Figure 53. Example of simultaneous contrast

Two historical studies on perceived contrast are Bartelson & Breneman[43] and the Stevens Effect.[44] Bartelson & Breneman showed that as the surround illumination increases, without substantial flare on the screen, the dark regions of the scene appear darker while the light regions remain white. This change in apparent lightness will therefore increase the perceived contrast by increasing the dynamic range in the dark regions of a scene. Likewise, the Stevens Effect shows that as luminance level increases, the whites appear whiter and the darks can appear darker. Note that the Stevens Effect may not cause darkening of the darks if the physical characteristics of the display actually increase the luminance of the blacks, as with a LCD. Again it is seen that scene contrast can increase due to increased dynamic range in both the dark and light scene regions. Depending on the display conditions, the increase in contrast may be due to changes in the white point only if the changes in luminance do not affect the black point.

7.2 Psychophysical Experiments

In order to examine the effect of different simulated EOTFs and backgrounds on image quality, the Method of Paired Comparison was used in these experiments and the results were analyzed using Thurston's Law of Comparative Judgment, Case V.[45] For each image, the analysis gives an interval scale of image quality based on the observers'

judgments of preference. The goal of any psychophysical experiment is to create a relationship between some physical attribute of a stimulus and the perceptual responses to that stimulus. In this case the stimulus was an image and the attribute judged was image quality.

The analysis entails creating a scale of the judged attribute. There are different scales created depending on the type of experiment.[28] For a paired comparison experiment an interval scale is created. Interval scale values have equal intervals between each sample on the scale. Zero on the scale is not defined so the scale is relative rather than absolute.

In a paired comparison experiment where the judgment is that of preference, the scale will indicate relative amount of preference between stimuli. The first step is to create a frequency matrix that counts the number of times stimulus X was preferred over each other stimulus. The next step is to create from the frequencies the probability of stimulus X being chosen over another stimulus. These probabilities are used to create Z-scores. The formula for calculating z-scores is seen below.

$$z = \frac{x_i - \bar{x}}{s} \tag{7.1}$$

Where s is the standard deviation, x_i is the particular stimulus and $\bar{x}$ is the average probability. A Z-score is a number that indicates how far away from the mean a value is. The average of Z-scores for a particular stimulus is the interval scale rating for that stimulus. A Z-score is the distance a particular value falls away from the standard deviation for a particular distribution. Thurston's Law of Comparative Judgment[28, 45] was used in the analysis to create an interval scale from the Z-scores.

Chapter 8 - Experimental

8.1 Experimental Setup

The experiments were performed using a 30” Sharp AQUOS LC-30HV6U LCD television, with observers seated 3 image heights from the display, approximately 33 inches. Figure 54 below shows a schematic of the experimental layout. The display subtended approximately 42° of visual angle horizontally and 26° vertically. The experiments took place in a specially built room that allowed the surround luminance to change both in lightness and color. The surround was lit behind the LCTV using 12 uniformly distributed high power LED lights.[46] The LED lights illuminated a white semicircle shaped diffusely reflecting screen. The surround filled more than the complete field of view when the observer faced the display. The walls and ceiling of the room were covered in a black material in order to keep flare off the display screen. The LEDs were all situated behind the display so that no direct illumination reached the LCD

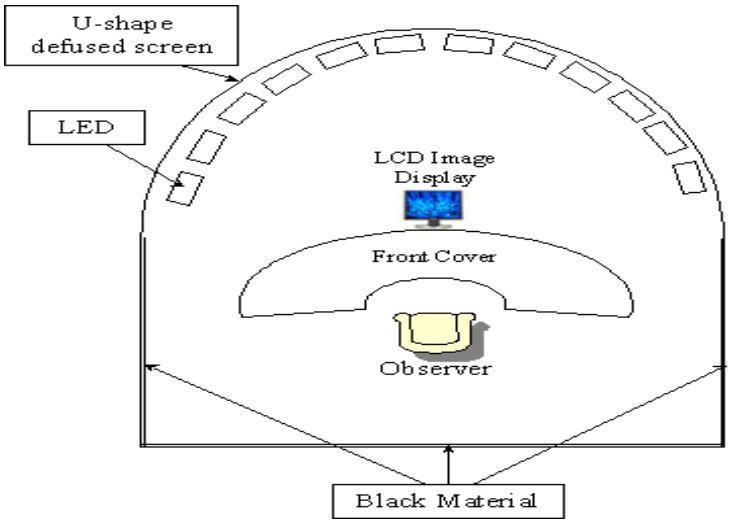


Figure 54. Experimental setup showing diagram of room with display.

The LED lights were previously characterized by C. Lui of the Munsell Lab using a 3x3 matrix followed by three 1D LUTs.[47] The figures below show the 1D LUTs used in the characterization and the gamut of the LED lights.

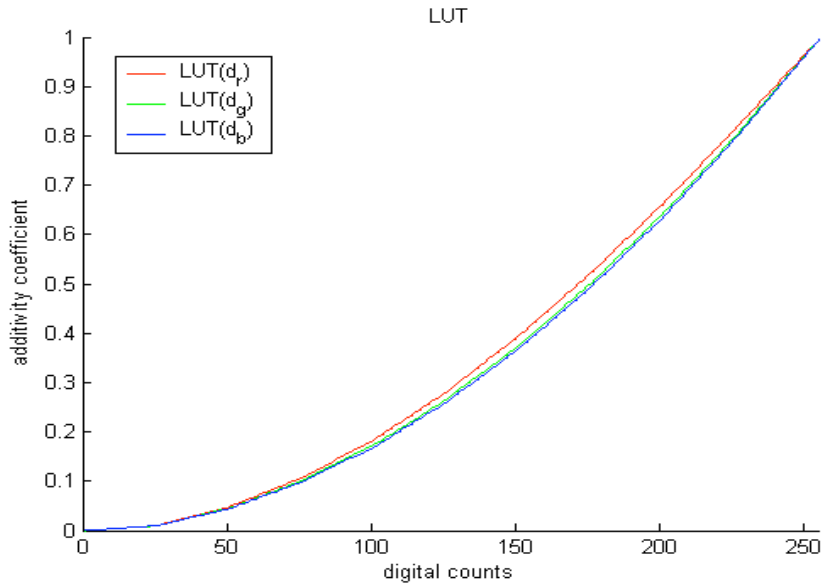


Figure 55. LUTs for each primary of the LED lights (reproduced from C. Lui 2004)

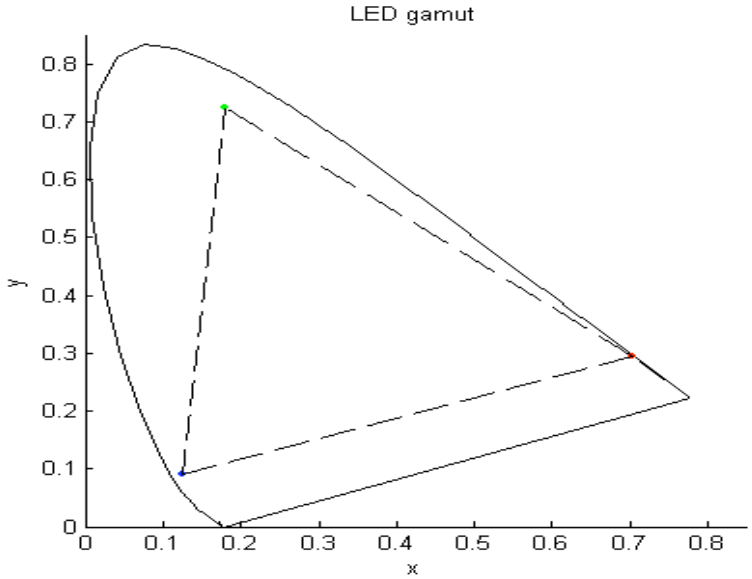


Figure 56. Gamut of the LED lights (reproduced from C. Lui 2004)

The LCTV was characterized using a 3D-LUT with an average error of 0.36 ΔE_{00} for 2000 random RGB colors, seen in Table 10 below. The characterization was performed using the LMT C1210 colorimeter and interfaced with Matlab.

Table 10. Characterization results

Max	1.77
Min	0.02
Avg	0.36
Std Dev	0.20

For the characterization of the LCTVs, a more simple characterization approach could not be accomplished because of observed cross-talk between the three primary channels.[48] (See Appendix D for details). Figure 57 shows the native EOTFs for a red, green, and blue ramp.

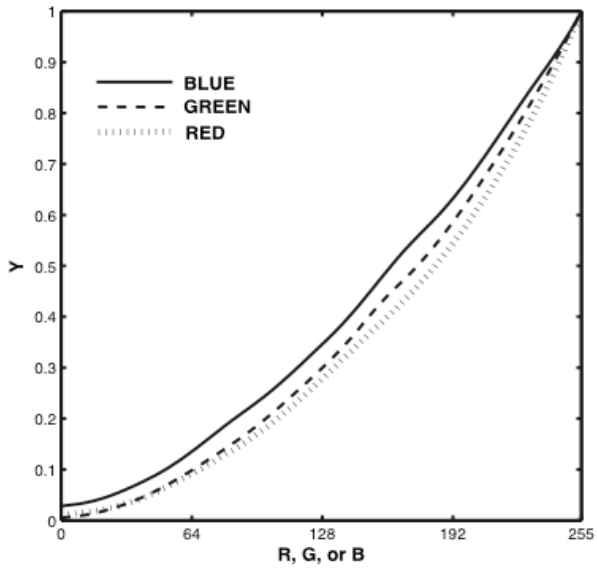


Figure 57. EOTFs of primary channels

The ten images used for the first experiment are shown in Figure 58. These images were chosen from a much larger dataset based on their content and characteristics so that they would cover a wide range of images seen on LCD televisions. This set of images

included DVD frame grabs, nature scenes, skin tones, and high and low key images. Based on the results from the first experiment, only four of the images from the original set were used in the second experiment and are shown in Figure 59.

Experiment one was performed in a dark surround. In Experiment two the LEDs were directed at the U-shaped diffuse screen to raise the surround illumination to the SMPTE's recommended luminance level of 10% of the display's white point, while closely matching the chromaticities. The chromaticites of the ambient illumination and display white for the dim surround are seen in the Table 11 below. Appropriate experimental settings are seen further below in Table 12. The labels across the top of the table describe the particular experimental design. Condition 1 is the default luminance of the display and Condition 2 is the lowered luminance of the display.

Table 11. Chromaticity of room at 10% surround and display's white point

	x	y
Ambient Illumination	0.308	0.307
LCD White	0.30	0.26

Table 12. Experimental settings

	Exp 1 Cond. 1	Exp 1 Cond. 2	Exp 2 Cond. 1	Exp 2 Cond. 2
Surround (cd/m^2)	**0**	**0**	**40**	**40**
Disp. Luminance (cd/m$^{2)}$	**400**	**170**	**400**	**170**
Number Images	**10**	**10**	**4**	**4**
Gamma Values (Meth. 1)	**1.3, 1.6 1.9, 2.2**	**1.3, 1.6 1.9, 2.2**	**1.3, 1.45, 1.6, 1.75, 1.9, 2.20**	**1.3, 1.45, 1.6, 1.75, 1.9, 2.20**
Exponential Values (Meth. 2)	**0.75, 0.875, 1.125, 1.25**	**0.75, 0.875, 1.125, 1.25**	--	--

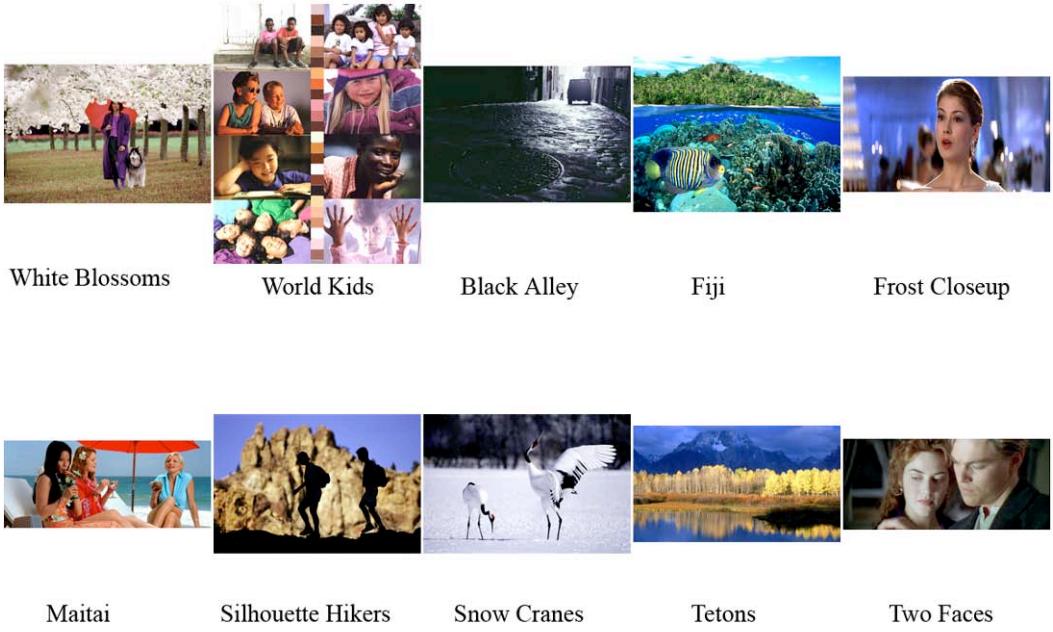


Figure 58. Images used in Experiment One

Figure 59. Four images used in Experiment Two

Figure 60. LCTV setup for all experiments

The background of the LCTV, seen around the centered image in Figure 60 above, for all conditions in both experiments was set to 20% of the white point of the display. In Condition 2, lower display luminance, of both experiments the overall luminance of the display was reduced by placing a neutral filter over the entire display. While not a perfect neutral-density filter, it effectively reduced the luminance with a minimal shift in color. The average ΔE_{00} shift in color was 0.98 for white, black and the three primaries. Figure 61 shows this shift for each primary and it can be seen how small the shift really is. Figure 62 shows a spectral plot of radiance for the unfiltered LCD and a spectral plot of transmittance for the filter. This plot shows the unevenness of the transmittance as a function of wavelength. Figure 63 shows plots of the spectral radiance for white, black and the primaries. Spectral transmittance for the filter was highly nonselective so that the

curve shapes of the spectral radiances were only slightly affected. With the filter in place the luminance of white was reduced approximately 57% from 400 cd/m^2 to 170 cd/m^2. It is interesting to note that when it was in place over the display observers were unaware that a filter was covering the screen.

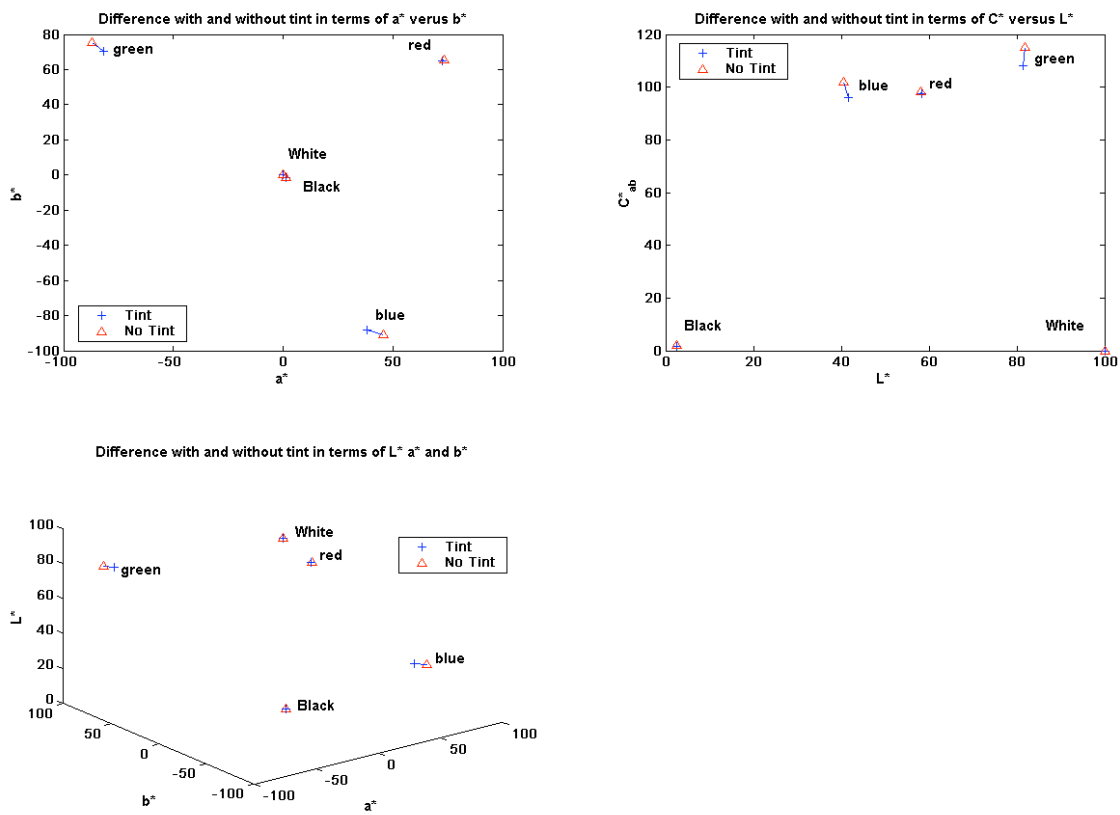


Figure 61. Color shift in CIELAB with and without filter over display

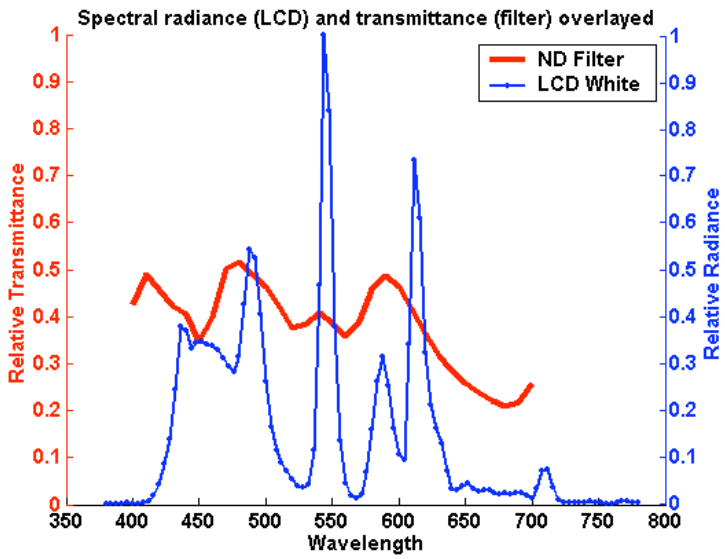


Figure 62. Spectral transmittance of filter over spectral radiance of LCD white

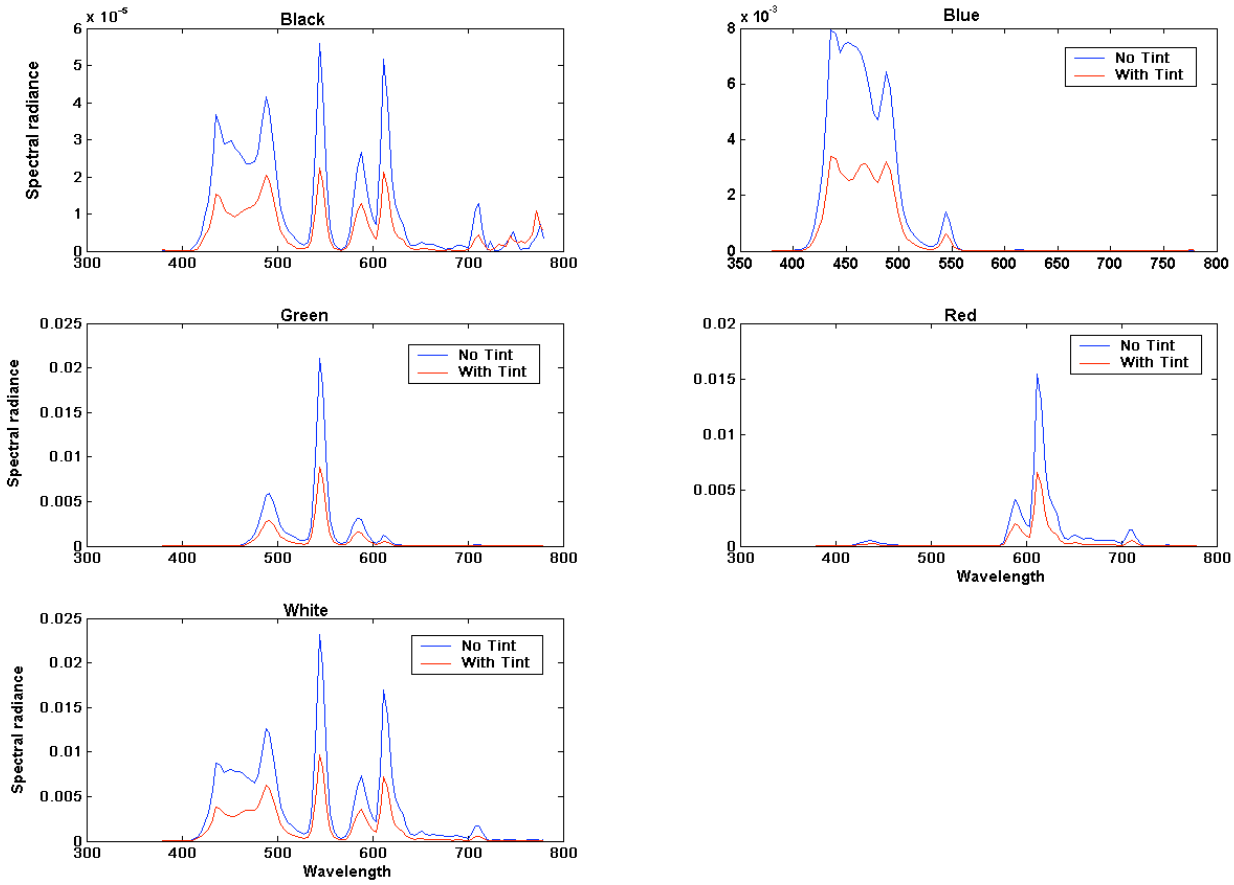


Figure 63. Spectral plots of primary channels with and without filter

8.2 Experimental Design

8.2.1 Gamma Method

Two methods were used to compute images that simulated their appearance as if shown on a display with a different tone-mapping function. In Method 1, which is referred to as the Gamma method, the raw RGB values were translated to RGB image values that correspond to a particular display. The raw RGB values are translated to scalars through the gain, offset, and gamma (GOG) formula[49] in Equations 8.1 to 8.3 using a gamma value that corresponds to the simulated display. See Table 12 for the gamma values used in the experiment.

$$R_{Scalar} = \left(gain_r \left(\frac{dc_r}{dc_{r,\max}} \right) + offset_r \right)^{gamma_r} \tag{8.1}$$

$$G_{Scalar} = \left(gain_g \left(\frac{dc_g}{dc_{g,\max}} \right) + offset_g \right)^{gamma_g} \tag{8.2}$$

$$B_{Scalar} = \left(gain_b \left(\frac{dc_b}{dc_{b,\max}} \right) + offset_b \right)^{gamma_b} \tag{8.3}$$

The *gain* and the *offset* values were based on best fits to the native EOTFs of the display (see Figure 57). The "*dc*" in each equation above is the input digital count, or raw RGB values, between 0 and 255. For a well behaved, additive display, the normal process when using the GOG method is to then multiply the R,G, or B scalars by a 3x3 matrix, which has elements corresponding to the maximum values of the primary channels, seen in Equation 8.4.

$$\begin{bmatrix} X_{pixel} \\ Y_{pixel} \\ Z_{pixel} \end{bmatrix} = \begin{bmatrix} X_{r,\max} & X_{g,\max} & X_{b,\max} \\ Y_{r,\max} & Y_{g,\max} & Y_{b,\max} \\ Z_{r,\max} & Z_{g,\max} & Z_{b,\max} \end{bmatrix} \begin{bmatrix} R_{scalar} \\ G_{scalar} \\ B_{scalar} \end{bmatrix} \tag{8.4}$$

where $X,Y,Z_{r,max}$ refers to tristimulus values for the maximum red and the same applies for green and blue primary channels.

However, since there was significant cross-talk between the primary channels in the display used here, a more complicated method, which took into account this interdependence of the channels, was used to estimate the tristimulus values for the modulated images. A 3x11 matrix multiplied by an 11x *n* vector transform was used to convert the scalars for each image of the *n* image pixels to tristimulus values. The vector accounted for channel interdependence by modeling some of the possible interactions, as seen in Equation 8.5. The 3x11 matrix is a transform matrix whose elements were found through optimization and converts the RGB scalars to XYZ.

$$\begin{bmatrix} R_{scalar} \\ G_{scalar} \\ B_{scalar} \\ R_{scalar}{}^2 \\ G_{scalar}{}^2 \\ B_{scalar}{}^2 \\ R_{scalar} * G_{scalar} \\ R_{scalar} * B_{scalar} \\ G_{scalar} * B_{scalar} \\ R_{scalar} * G_{scalar} * B_{scalar} \\ 1 \end{bmatrix} \tag{8.5}$$

Therefore, the method used to obtain XYZ for each pixel in the newly modulated image was

$$\begin{bmatrix} X_{pixel} \\ Y_{pixel} \\ Z_{pixel} \end{bmatrix} = [3x11][11xn] \tag{8.6}$$

The use of this method resulted in an average error of approximately 1.4 ΔE_{00} for a uniform sampling of 3375 colors in the display gamut for an independent verification routine. The inverse 3D-LUT was then used to get the RGB values for the new image based on these recomputed tristimulus values. The inverse 3D LUT was derived from the characterization performed on the LCTV (see Appendix E for further details).

The above process is shown graphically in the flow diagram below. It shows the steps to translate raw RGB image values to RGB image values that correspond to a particular display. The raw RGB values are translated to scalars through the GOG formula in Equations 8.1 to 8.3 using a gamma value that corresponds to the simulated display. The scalars are then transformed to XYZ values using Equation 8.4. These XYZ values are converted to CIELAB units and sent through the inverse 3D LUT. The output of these steps is then converted RGB values that correspond to a simulated display.

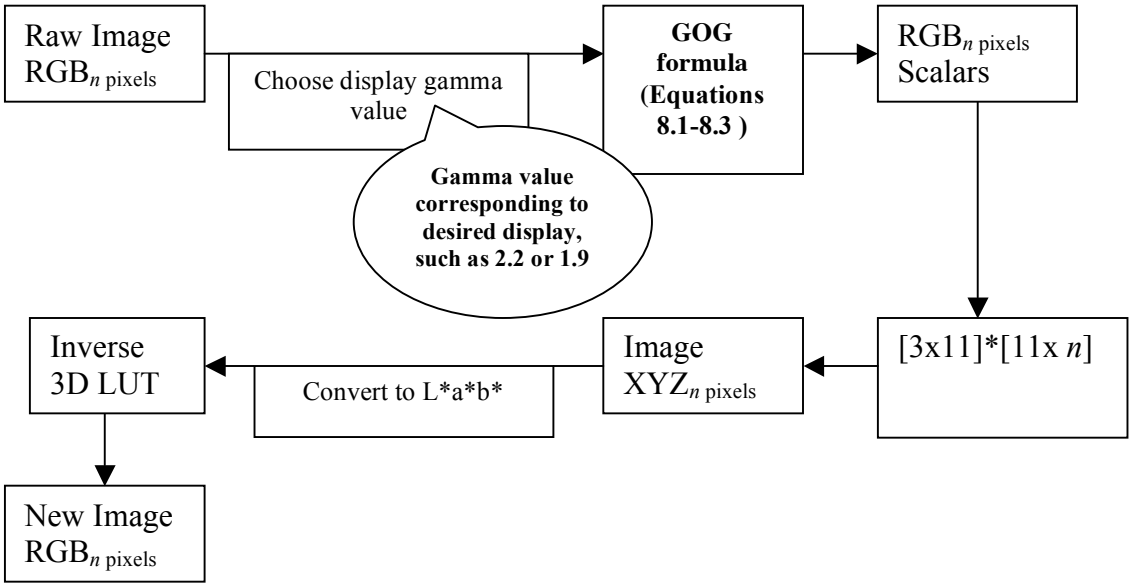


Figure 64. Flowchart for Gamma Method

Figure 65 below shows an example of the three primary ramps and a neutral ramp for the display using the two extreme gamma values (1.3 and 2.2) in the experiment. The ordinate for this graph is absolute Y value and because of this the curves do not fall on top of each other. The intrinsic functions of the display for each primary had a different gamma value, as seen in Figure 57, but for the process used in Method 1, each primary channel was forced to have the same gamma value. This could have implications on how images will be shown on the display. Although not researched in this thesis, this could affect results and should be considered for future experiments. Figure 65 shows the resulting transfer function for each primary and a neutral ramp for two gamma values. Table 12 shows the gamma values used in the experiment.

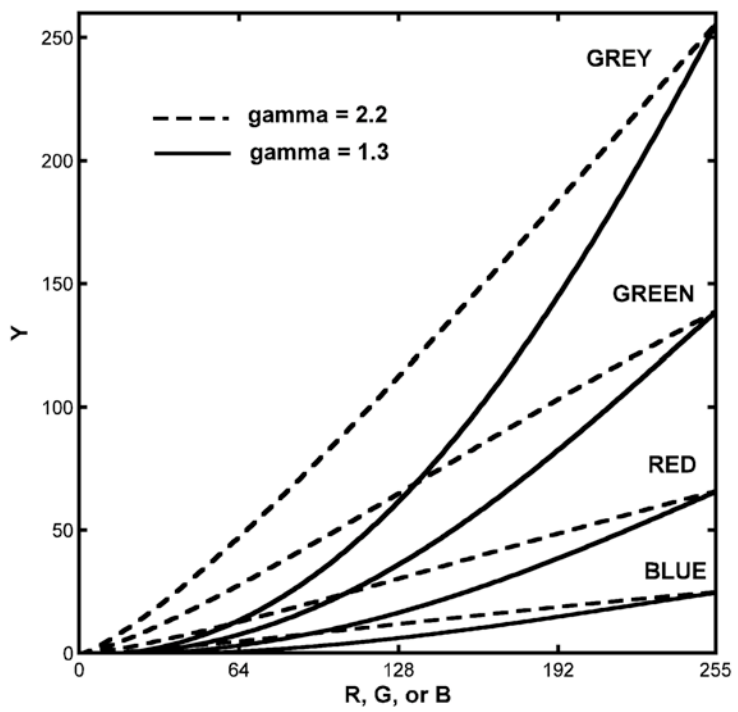


Figure 65. Primary and neutral ramps displayed using 2 gammas

8.2.2 Exponential Method

The second method was entitled the Exponential method and was a much simpler approach for simulating different displays. The native transfer functions, seen in Figure 57, were raised to values seen in Table 12. This method tested whether keeping the same relation and shape of each primary channel, yet changing gamma values would impact image preference. Figure 67 shows an example of the primary ramps and a neutral ramp using this method for the extreme values of 0.75 and 1.25.

This process is seen in the flow diagram below. It simply scaled the digital RGB values in an image between 0 and 1 and then raised these values to an exponent. These values were then rescaled between 0 and 255 to create the new image RGB values. This approach uses the intrinsic gamma of the image and serves only as a "gamma-boost" or "gamma-reduction".

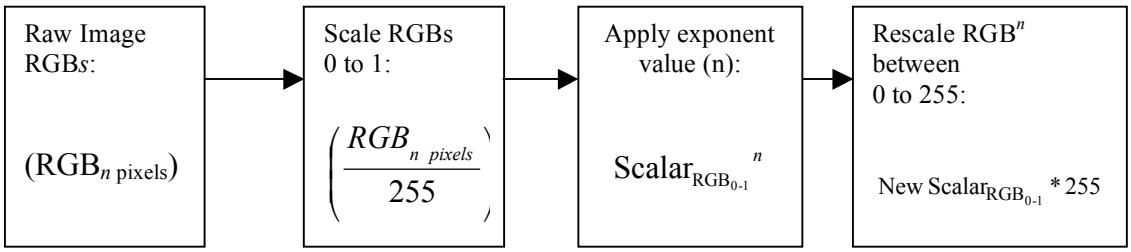


Figure 66. Flowchart for Exponential Method with equations at the bottom of each box representing each step.

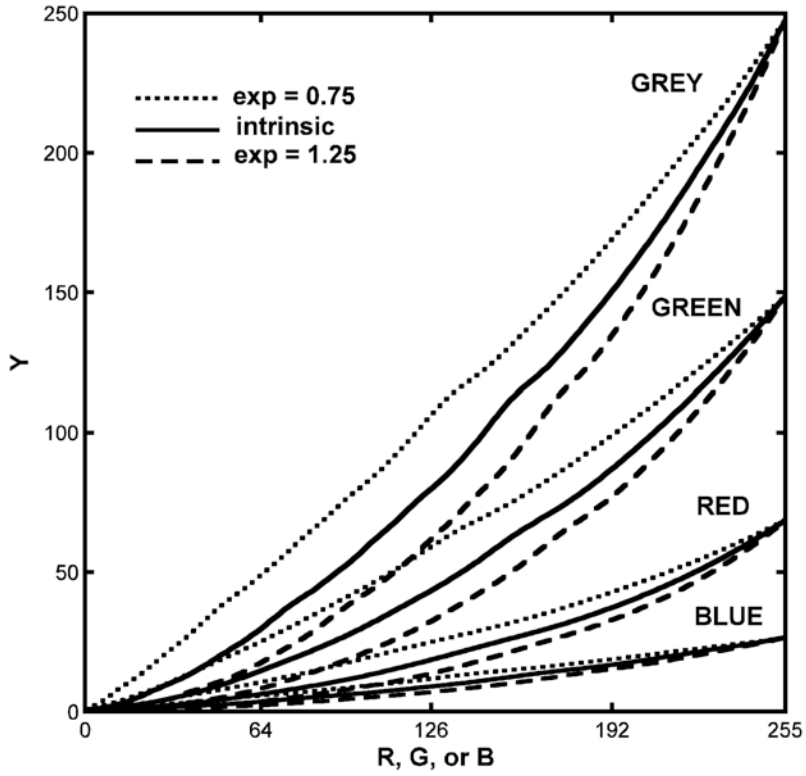


Figure 67. Primary and neutral ramps displayed using 2 exponential values

8.2.3 Psychophysical Experiment

In both conditions for Experiments 1 and 2, a paired-comparison experiment was performed in which each image, including the original unprocessed image, was presented with all the others. The observers' task was to choose which of the two images in the pair they preferred based on overall image quality. The observers could toggle between the two images and select their choice by hitting the "Return" key while the preferred image was displayed. There were a total of 360 trials in Experiment One, (10 images with 9 variations each) and in Experiment Two there were 84 trials, (4 images with 7 variations each). The data was analyzed using Thurstone's Law[45], which produces interval scale values of image preference. Additionally, the 95% error bars for curves in all the below

graphs were created using a method based on Monte Carlo simulation.[50] There were 26 observers for Experiment One, ranging in age from 22-49. For Experiment Two there were 20 observers in the same age range.

Chapter 9 - Results

9.1 Experiment One: Dark Surround

Figure 68 shows the individual image results for all 10 images for Experiment One. The interval scale values from Condition 1 (default display luminance – solid line) and Condition 2 (lower luminance display – dashed line) were shifted by an additive constant so that the scale value of the original image (far left value) was 1. This allows the trends for both sets of data to be compared relative to the original. As a result, the scale values between conditions cannot be compared but the trends relative to the original can be.

The curves on the left of the vertical line represent Method 1, the Gamma method, and there are several noticeable trends. Overall, a gamma of 1.6 produced images that are preferred over the original. At the lower luminance level (dashed line) the curve appears wider, thus suggesting less selectivity for different gamma values. Conversely the higher luminance levels (solid line) result in a narrower curve, suggesting that gamma values are more critical and that higher values are more objectionable. In other words, for brighter displays the choice for gamma becomes more constrained. The images from Method 2, the Exponential Method, to the right of the vertical line, on average, are not preferred to the original tone functions. (An exponent of 1 would produce images identical to the original).

Figure 70 shows results for four images that are representative of the trends observed in the other 6 images. These four images are used later in Experiment Two. The different curve shapes for the images, as seen in the subplots of Figure 68, show a great

deal of image dependence in terms of the absolute preference relative to the original. Nevertheless, a gamma value of 1.6 is clearly preferred over the original for the majority of images. However, the trends within each type of image manipulation are quite similar for both luminance levels. This indicates that display luminance is not impacting observer preference as much as the change in gamma.

It is seen that there is a small effect of display luminance on the preferred tone function. Furthermore, the preferred tone mapping of 1.6 gamma value does not change for the two screen intensities tested. However, at the higher luminance, it appears that obtaining this optimal gamma is more critical. Additionally, it is clear that the intrinsic tone-mapping of the display can be improved.

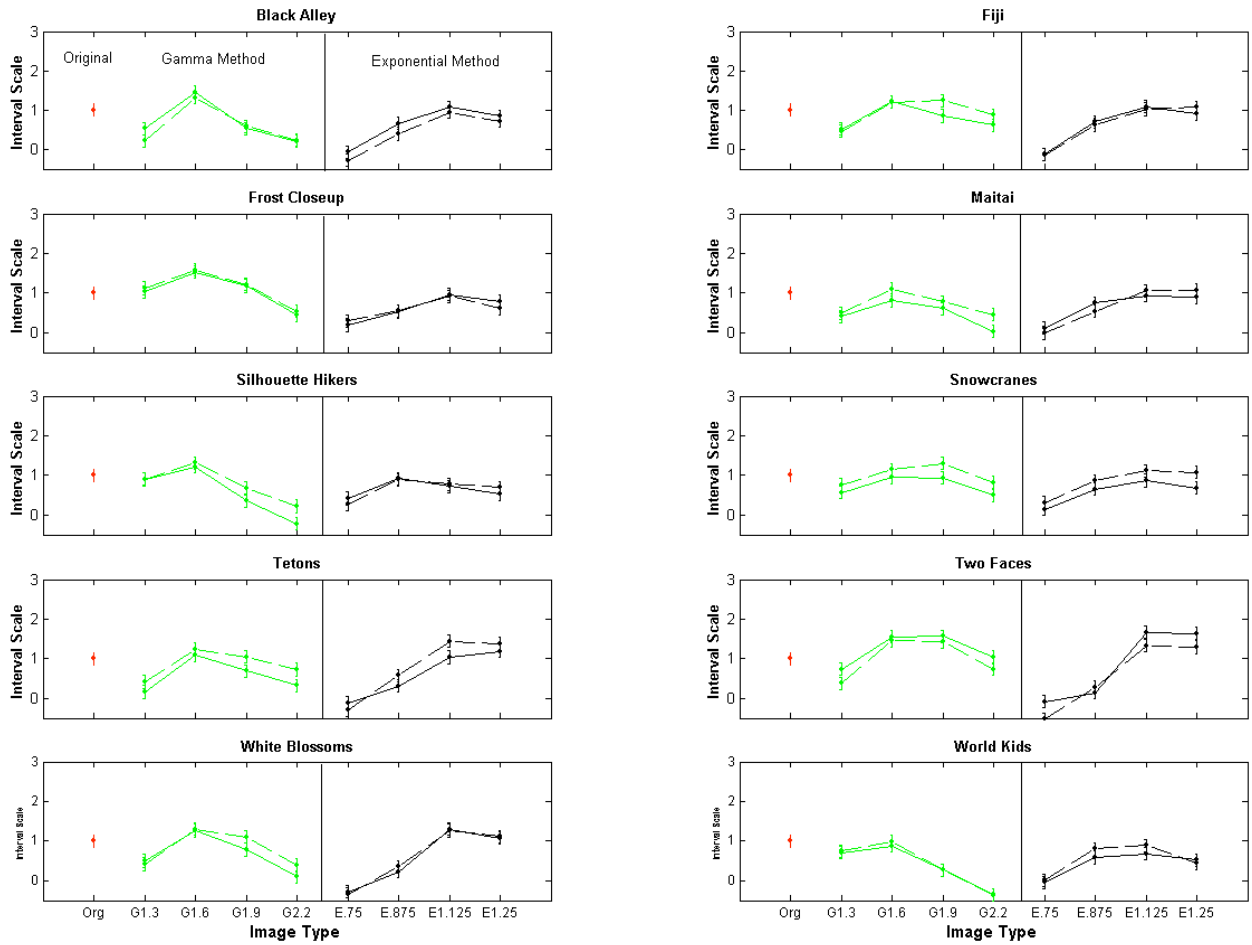


Figure 68. Results of Experiment One per image. Solid line is default luminance and dashed line is lowered luminance.

Figure 69 shows that the preference scale increased for the lower luminance display relative to the default luminance. In other words, for higher gamma values at the lowered luminance level, observers' preference increased but at the default luminance their preference decreased for the higher gamma values. In general this was the case for the individual images, seen in Figure 68. On average, an image shown on a display with a gamma value of 1.6 had a higher preference relative to the original. This can be seen in Figure 69.

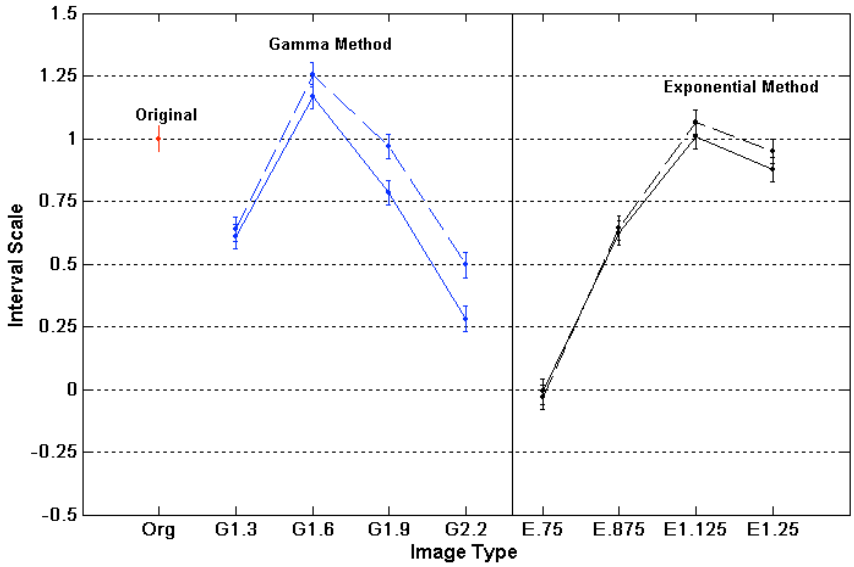


Figure 69. Average of all images for dark surround. Solid line is default luminance, dashed is lowered

Figure 70 below shows experimental results from Experiment One for the four images that were used in Experiment Two. This figure shows results of Experiment One for both display luminances. From this figure an easy comparison can be made between the two experiments, which will be described in the next section.

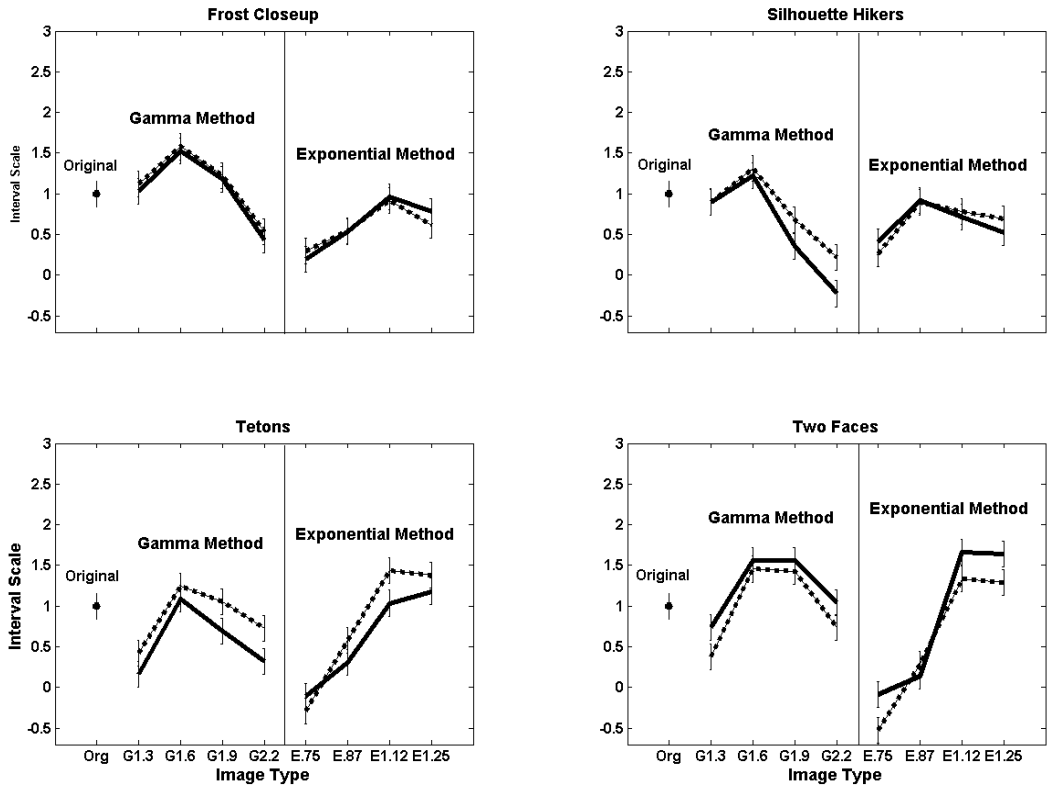


Figure 70. Results of individual images for dark surround. Solid line is default luminance, dashed is lowered

9.2 Experiment Two: Dim Surround

The surround luminance level was raised for Experiment Two. The SMPTE's[51] guideline of 10% of the display's white point was followed and consequently the luminance level of the room was raised to 40 cd/m^2 with chromaticities closely matched. Note that in this experiment the surround luminance was the same even when the display luminance was decreased. The images in Figure 59 were used in Experiment Two with two additional gamma values added in order to better determine the optimal gamma. In Figures 68 to 70 above there is a peak at a gamma value of 1.6; however the spacing of the gamma values was too far apart to determine whether a lower or higher value would lead to optimal performance.

Figure 71 below shows the average results from all four images, seen in Figure 59. At the higher display luminance level, (solid line), the gamma value of 1.6 is clearly preferred but at the lower display luminance it is either 1.6 or 1.75. Figures 72 and 73 below show the differences between the dim and dark surround, comparing results from Experiments One, the dim surround, and Experiment Two, the dark surround. Figure 72 shows the differences between the experiments at the default luminance of the display and Figure 73 shows the differences at the lowered luminance. In Experiment Two, the dim surround, for both display luminance levels there was a greater range between least and most preferred compared to the original. This signifies that the display luminance had less of an effect when compared to the surround condition. As seen in Figure 69, which shows results at the dark surround for both display luminance levels, there was significant overlap for observer preference at a gamma of 1.6 (and indeed for other gamma values as well). This indicates observer preference was statistically similar for the two display luminance levels. However, Figures 72 and 73 show results for observer preference were statistically different for most images. Furthermore, within each graph for Figures 72 and 73 there was a greater difference between the surround conditions, (the two curves in each graph). While the above results are generally true, it is recognized that there is image dependence for these results.

In general, the optimal gamma value of 1.6 had a higher preference scale value with a dim surround than for a dark surround. Therefore, under more natural viewing conditions with a surround recommended by SMPTE[51], greater improvement in image preference is seen when using a gamma value of 1.6 in each channel as compared to the intrinsic tone curves and other gamma values.

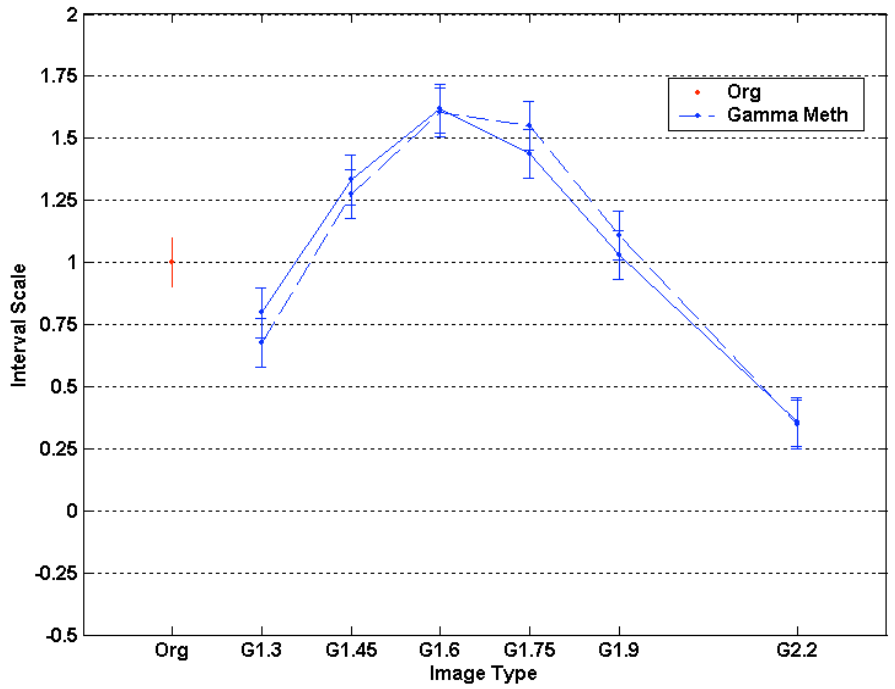


Figure 71. Average of four images for dim surround. Solid line is default (bright) screen luminance and dashed is lowered.

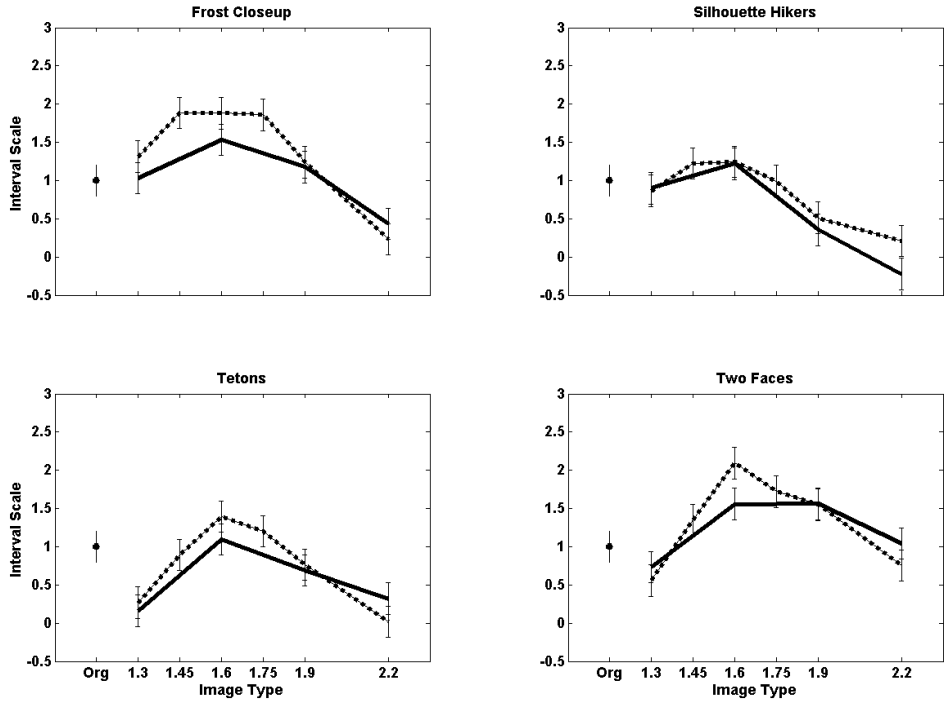


Figure 72. Comparing surround at default (bright) screen luminance. Dashed line is dim surround and solid is dark surround.

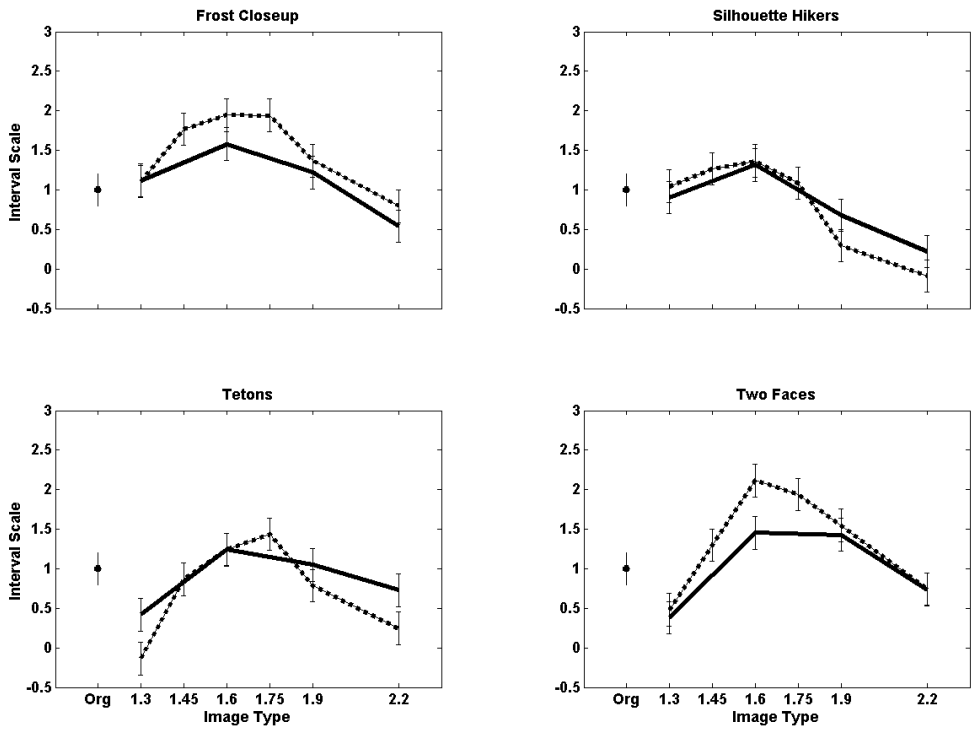


Figure 73. Comparing surround for lowered screen luminance. Dashed line is dim surround and solid is dark surround.

Chapter 10 - Discussion

Based on the Steven's Effect[44], at the default luminance of the display there is a higher perceived contrast than at the at the lower luminance level with the overlayed filter. Based on the Bartelson & Breneman[43] effect, the dark surround will have a lower perceived contrast than the dim surround. With newer, brighter TVs the perceived contrast of the display is higher; therefore adjustments may need to be made in the tone mapping of images in order to compensate for these changes. The prediction is that when the luminance of the display is increased, a lower gamma value is preferred due to the increased contrast produced by the brighter display.

The surround and the intensity of the display can have profound effects on the appearance of a scene (see Fairchild[52] for a discussion of these effects). Bartleson and Breneman found that if the surround illumination around a display is increased, without causing substantial flare, the white point remains stable but the blacks appear darker thus increasing the perceived contrast. This effect generally enhances the appearance of an image but it is noted that the effect depends on the image whether it enhances or makes the appearance worse.

Conversely, by reducing the default LCD luminance through a filter both the white point and absolute black were reduced, resulting in a change in preference with the lower luminance. The Stevens Effect describes the decrease in perceived contrast with reduced luminance. As the luminance level decreases the bright areas of an image do not appear very white and the dark areas do not appear very dark, thus the perceived contrast has decreased. This would account for the fact that as the overall luminance decreases

thus decreasing contrast and causing observers to choose images with higher gamma values.

It is noted that the Bartleson & Breneman effect explains the shift in preference between the two surround conditions while the Stevens Effect explains the change in preference between the display's luminance levels. While it might seem that these two effects might balance each other the results indicate that surround conditions chosen here had a larger impact on observer preference than display luminance at the levels tested. In other words, changing display luminance had less of an effect on preference than changing surround. Figures 72 and 73 show that as surround illumination is increased observer preference decreased for higher gamma value images. This result is directly in line with the Bartleson & Breneman effect because observers already perceived a higher contrast so therefore objected more persistently to images with high gamma values.

Chapter 11 - Conclusion

The results indicate that at the default luminance of the 30" LCD television observers preferred a gamma of 1.6 in each channel over the intrinsic EOTFs of the display. At the lower luminance there was more tolerance in the absolute choice for preferred gamma but on average, observers still preferred a gamma of 1.6 for the dark surround and either 1.6 or 1.75 for the dim surround. The end result is that for higher luminance displays viewed in a dim surround the choice of gamma in a display becomes critical.

It is also clear that these effects are dependent on image content. As noted in section 8.2.1 ,*Gamma Method*, the change in color due to the uniform gamma curves relative to the intrinsic curves, which are different in each channel, had unknown impact on these results. Further experimentation is needed to explore both the effect of these color changes on preference and determine whether a completely different tone-mapping function would produce even better results. From the results found in this study what can be said is that manufactures of these display types need to choose the particular gamma value carefully and using similar values for each channel should be considered. It is clear that using the traditional value of 2.2 would not be optimal.

Further studies are required in order to quantify the effect of different luminance levels and backgrounds on image preference. In addition, these experiments only tested two forms of the EOTF: a gamma function and the intrinsic LUTs built into the display hardware raised to an exponent. The results showed that modifying the shape of the intrinsic EOTFs by an exponent did not improve image quality. However, the use of a more traditional gamma function did improve image quality. The change in the images was not confined to changes in lightness and contrast with the use of the gamma function

because the intrinsic LUTs of each of the display primaries are slightly different. Therefore, the use of the gamma functions also leads to color shifts. It is possible that other shaped curves, such as sigmoids, may produce even better results.

The method of simulating different gamma functions is susceptible to image artifacts due to quantization and loss of image detail in the dark regions due to the implementation of the algorithm to recompute the images. Despite these artifacts, the image manipulations did lead to improvement in image quality.

Beside changes in display intensity and surround conditions, the next generation of televisions are also larger and sharper than older CRTs. Both of these factors may also produce changes in image quality. As display technology improves and displays become brighter and larger, it is possible that the effect of the display luminance may necessitate different tone mapping due to local adaptation to the display. Further study is needed to determine how these factors should be accounted for in image processing with these displays.

References

1. H. Pan, X. eng and S. Daly. Quantitative Analysis of LCD Motion Blur and Performance of Existing Approaches. *Proceedings of the SID Symposiums,* **36**, 1590-1593 (2005).
2. D.H. Kelly, "Flicker", Ch. 11 in *Handbook in Sensory Physiology, Vol. VII/4, Visual Psychophysics*, D. Jameson & L.M. Hurvich, (Eds.), Springer Verlag, New York, (1972).
3. D. Kelly, Motion and vision. II. Stabilized spatio-temporal threshold surface, *JOSA*, **69**, pp. 1340-1349 (1979).
4. S. Daly, Engineering observations from spatiovelocity and spatiotemporal visual models, in *Vision Models and Applications to Image and Video Processing.* C.J. van den Branden Lambrecht, (Ed.), Kluwer Academic Publishers (2001).
5. M.P. Eckert and G. Buchsbaum, The Significance of Eye Movements and Image Acceleration for Coding Television Image Sequences, Ch. 8 in *Digital Images and Human Vision,* A.B. Watson, (Ed.), MIT Press, Cambridge, MA (1993).
6. http:\\www.ctlens.co.uk/problems.htm
7. R.H. Carpenter, "Saccades", Ch. 4 in *Movement of the Eyes*, 2nd edition, Pion: London (1988).
8. W. Becker, Metrics, *The Neurobiology of Saccadic Eye Movements*, R.H. Wurtz & M.E.Goldberg (Eds.), Elsevier Science Publishers (1989).
9. E. Kowler, Cogito Ergo Moveo: Cognitive Control of Eye Movements, Ch 2 in *Exploratory Vision: The Active Eye*. M.S. Landy, L.T. Maloney M. Pavel (Eds.), Springer Verlag (1996).
10. R.H. Carpenter, "Optokinesis and Smooth Pursuit", Ch. 3 in *Movement of the Eyes*, 2nd edition, Pion: London (1988).
11. C. H. Meyer, A. G. Lasker, and D. A. Robinson, The upper limit of human smooth pursuit velocity, *Vision. Res*., **25(4)**, pp. 561-563 (1985).
12. D.H. Kelly, Eye Movements and Contrast Sensitivity, Ch. 3 in *Visual Science and Engineering: Models and Applications,* D.H. Kelly, (Ed.), Marcel Dekker, New York (1994).

13. A.L. Yarbus (1956), Haigh B. (translator) (1967), *Eye Movements and Vision*, Plenum Press: New York.
14. G. Sharma, Color fundamentals for digital imaging, Ch. 1 in *Digital Color Imaging Handbook*, G. Sharma, (Ed.), CRC Press, Boca Raton (2003).
15. W. O'Mara, *Liquid Crystal Flat Panel Display*. Chapter 1. Chapman & Hall, New York (1993).
16. P. Yeh and C. Gu, *Optics of Liquid Crystal Displays*. John Wiley and Sons, New York (1999).
17. M.A. Klompenhouwer, Temporal Impulse Response and Bandwidth of Displays in Relation to Motion Blur, *Proceedings of the SID Symposiums,* **36**, 1578-1581 (2005).
18. B.A. Wandell, *Foundations of Vision,* Sinauer Associates, Sunderland, MA (1995).
19. P. G. J. Barten, *Contrast Sensitivity of the human eye and its effects on image quality*, Uitgeverij HV Press, Knegsel, Netherlands (1999).
20. R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, *2nd Ed*., Prentice Hall, New Jersey (2002).
21. H. B. Barlow, *The Senses*, Cambridge University, London (1982).
22. A. B. Watson, "Temporal Sensitivity", Ch.6 in *Handbook of Perception and Human Performance*. John Wiley and Sons, New York (1986).
23. J. G. Robson, Spatial and Temporal contrast sensitivity functions of the visual system, *JOSA* **56,** pp. 1141-1142 (1966).
24. E. Levinson and R. Sekular, The Independence of channels in human vision selective for direction of movement, *Journal of Physiology,* **250**, 347-366 (1975).
25. D. Kelly, Motion and vision. I. Stabilized images of stationary gratings, *JOSA* **69**, pp. 1266-1274 (1979).
26. http:\\webvision\med\utah.edu\images\KallSpat22\mod
27. J. Fourier (1822), Freeman, A. (translator). *The Analytical Theory of Heat*, Cambridge: University Press, London (1878).

28. R. Boynton, "Psychophysics", Ch. 6 in *Optical Radiation Measurements, Vol.5: Visual Measurements*, F. Grum & C.J. Bartleson (Eds.), Academic Press, Orlando, Florida (1984).
29. J.G. Snodgrass, Psychophysics, Ch. 2 in *Experimental Sensory Psychology*, B Scharf (Ed.), Scott, Foresman and Company, CA. (1975).
30. D.G. Pelli and B. Farell, Psychophysical Methods, Ch. 29 in *Handbook of Optics and Psychophysical Methods,* M. Bass, E. W. Van Stryland, D. R. Williams, and W. L. Wolfe (Eds.), McGraw-Hill, New York (1994).
31. D.G. Pelli and B. Farell, Psychophysical Methods or How to Measure Threshold and Why, in *A Practical Guide to Vision Research*, J.G. Robson and R.H.S. Carpenter (Eds.), Oxford University Press, New York (1996).
32. M.W. Levine, Psychophysics, Ch. 2 in *Fundamentals of Sensation and Perception 3rd Ed.*, M.W. Levine and J.G. Shefner (Eds.), Oxford University Press, New York (2000).
33. A.B. Watson and D.G. Pelli, QUEST: A Bayesian adaptive psychometric method, *Perception & Psychophysics* **33(2)**, pp. 113-120 (1983).
34. D. H. Brainard, The Psychophysics Toolbox, *Spatial Vision* **10**, pp. 433-436 (1997).
35. J.R. Hamerly and C.A.Dvorak, Detection and discrimination of blur in edges and lines, *JOSA*, **71**, pp. 448-452 (1981).
36. Applied Sciences Laboratories (2001). Eye Tracker Manual (Model 504), Applied Sciences Group, Inc.: MA.
37. http://www.mathworks.com
38. M. Spering, D. Kerzel, D.I. Braun, M.J. Hawken, & K.R. Gegenfurtner, Effects of contrast on smooth pursuit eye movements. *Journal of Vision*, **5(5)**, 455-465 (2005).
39. N. Kimura et al., New Technologies for Large-sized High-Quality LCD TV. *Proceedings of the SID Symposiums*, **36**, 1734-1737 (2005).
40. S.D. Yeo, C.H. Oh, H.W.Lee, M.H. Park, LCD Technologies for the TV application, *Proceedings of the SID Symposiums*, **36**, 1738-1741 (2005).

41. M.D. Fairchild, "Considering the surround in Device Independent Color Imaging", *Color Res. and Appl*., **20**, 352-363 (1995).
42. S.E. Palmer, *Vision Science Photons to Phenomenology*, MIT Press, Cambridge MA (1999).
43. C.J. Bartleson and E.J. Breneman, Brightness Perception in Complex Fields, *JOSA*, **57**, 953-957 (1967).
44. J.C. Stevens and S.S. Stevens, Brightness Functions: Effects of Adaptation, *JOSA*, **53**, 375-385 (1963).
45. L.L. Thurstone. A Law of Comparative Judgment, *Psych. Review* **34**, 273-286, (1927).
46. Color Kinetics Incorporated, ColorBlast 12 PowerCore System. http://www.colorkinetics.com
47. C. Lui and M.D. Fairchild, Measuring the Relationship between Perceived Image contrast and Surround Illumination, *Proc. IS&T/SID 12th Color Imaging Conf.,* 282-288 (2004).
48. J.L. Laird and M.R. Rosen, Characterization Methods for LCD Devices with Crosstalk Issues, *Proceedings of the SID Symposiums*, **36**, 550-554 (2005).
49. R. S. Berns, Methods for Characterizing CRT Displays, *Displays* **16**, 173-182 (1996).
50. E.D. Montag, Louis Leon Thurstone in Monte Carlo: Creating Error Bars for the Method of Paired Comparison, *Proc. IS&T/SPIE Symposium on Electronic Imaging Science & Technology*, **5294**, 222-230 (2004).
51. SMPTE RP 166-1995: Critical Viewing Conditions for Evaluation of Color Television Pictures.
52. M. D. Fairchild, *Color Appearance Models, 2nd Ed*., John Wiley & Sons, New York (2005).
53. E. A. Day, L. Taplin & R. S. Berns, Colorimetric Characterization of a Computer-Controlled Liquid Crystal Display, Munsell Color Science Laboratory Technical Report. RIT Rochester, NY
54. M.D. Stokes, M.D. Fairchild, and R.S. Berns, Precision requirements for digital color reproduction, *ACM Transactions on Graphics* **11**, 406-422 (1992).

55. G. Marcu, W. Chen, K. Chen, P. Graffagnino, O. Andrade. Color Characterization issues for TFTLCD displays, *SPIE Proc*. **4663**, 187-198 (2002).
56. Y. Kwak, L. W. MacDonald. Accurate Prediction of Colours on Liquid Crystal Displays, *Proc. 9th IS&T/SID Color Imaging Conf*., pg. 355-359 (2001).
57. D.H. Brainard, D.G Pelli, and T. Robson, Display Characterization, in J. Hornak (Ed.), *Encyclopedia of Imaging Science and Technology,* pp. 172-186, Wiley, New York, (2002).

Appendix

Appendix A

CRT Characterization

Summary

A characterization was performed on a Sony Trinitron CRT monitor. Subsequently, a TRC-matrix model of the CRT system was derived. The model was verified using measurements from 2000 randomly selected colors. The average ΔE_{00} between the measured and calculated values was 0.31. Other tests performed included a channel and spatial independency test. The channel independency test showed the sum of tristimulus values measured for the primaries was greater than the measured full white of the monitor. There was a discrepancy of ΔE_{00} 0.38 between the measured and summed white. The spatial independency test showed an average difference of ΔE_{00} 0.52 between a white on black background and white on other backgrounds. Interestingly when the background was blue the measured white was shown to actually increase its luminance beyond that of white on black. The gamma for each channel was optimized to 2.4 for the red channel and 2.3 for green and blue channels.

CRT Characterization

The monitor described in this report is a Sony Trinitron Multiscan G420. The white point, or color temperature, was set to a factory default of 9300K. The brightness/contrast settings were adjusted so that a JND was detected between two patches at digital count of [0,0,0] and [1,1,1] and also at [255,255,255] and [254,254,254].

The measurement device used was an LMT Colorimeter C1210 that measured displayed colors in tristimulus values, (XYZ), in units of cd/m^2. It was chosen because of its large dynamic range and for its automated measurement capabilities. The LMT was interfaced with Matlab and all measurements of colors were made using a GUI that had a neutral gray background, (except where noted in the spatial independency test), and a colored patch slightly larger than the diameter of the LMT's measurement head. Figure A.1 shows the GUI used to take all measurements throughout the characterization process. This GUI was developed using the Psychophysics Toolbox plug-in, which allowed the user to take over the screen and completely fill the background with any desired color, as well as the relative ease of changing the size and color of the center patch.

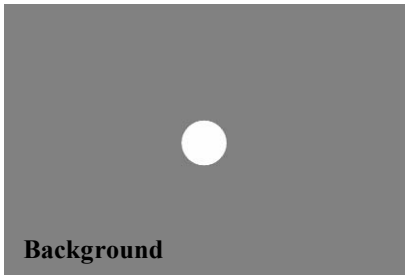


Figure A.1 – Matlab GUI used to make measurements

There were three main sections of the characterization process: channel independency test, spatial independency test and measurements appropriate to building the CRT model. The channel independency test analyzes whether the sum of the primaries, (R+G+B), equals the measured white point of the monitor. Berns[49] states that a lack of channel independence is due to overdriving the gun amplifiers and can be remedied by decreasing the contrast of each channel. The spatial independency test analyzes how the background of the display affects the center color. A lack of

independency suggests a power supply limitation[1] available to the guns. After the monitor has passed these two tests, primary ramps from 0-255 each were measured along with black and white. Results of the tests are seen below.

First is the channel independency test where each channel was summed and compared to the measured white point of the display, seen in Tables A.1 and A.2 (13 & 14)

Table A.1 - Channel independency test, differences in tristimulus measurements

Channel	X	Y	Z
Red	24.53	13.74	1.54
Green	18.73	41.36	7.87
Blue	15.35	6.65	80.38
**********			
R+G+B	58.61	61.75	89.78
Meas White	58.08	61.12	88.91
Perc Diff	0.91%	1.03%	0.98%
**********			
Average Percent Difference			0.97%

Table A.2- Channel independency test, differences in CIELAB units

Channel	L	a	b	ΔE_{00}
Meas. White	100	0	0	0
R+G+B	100.4	-0.21	0.03	0.38

Second is the Spatial Independency test where the same white patch was measured with 6 different backgrounds: gray, and red, green, blue, followed by white and black, seen in Table A.3 (15).

Table A.3-Spatial independency test, differences based on white patch with black background

Differences between measurements of white on different backgrounds.

Background	X	Y	Z	L	a	b
Black	57.9	60.93	88.76	100	0	0
Red	57.55	60.67	88.86	99.83	-0.29	-0.36
White	57.32	60.3	87.8	99.6	0.04	0.04
Green	57.75	60.47	88.48	99.71	0.82	-0.29
Gray	58.04	61.07	88.83	100.09	0.01	0.11
Blue	57.9	61.09	88.03	100.1	-0.44	0.73

Note that the white patch measured on the black background was the default value. This was decided because it would exhibit the least amount of flare and the guns should be at their maximum potential at this setting. However, Tables 3 and 4 shows that this was not always the case. Note that in table 4, the differences are signed, for example a negative a* represents a color with green content, and for L* any difference value that is positive represents an L* value less than 100 and any difference that is negative represents an L* value that is higher than 100. All differences are taken from the white on black measurement.

Table A.4- CIELAB differences

Differences from white patch on black background			
	L	**a**	**b**
Black	0	0	0
Red	0.17	-0.29	-0.36
White	0.4	0.04	0.04
Green	0.29	0.82	-0.29
Gray	-0.09	0.01	0.11
Blue	-0.1	-0.44	0.73
Average	0.13	0.03	0.05

Table A.4 (16) shows the ΔE_{00} values between the white patches with different backgrounds and the white patch on the black background. It shows that white on green had the largest shift in color, which Table A.3 (15) shows resulting from a change in a*. It should be noted that the a* change is towards red, which indicates a lack of available voltage from the green gun to the center white patch.

Table A.5 - CIEDE2000 between white on black background and all others

Background	**ΔE_{00}**
Black	--
Red	0.56
White	0.24
Green	1.24
Gray	0.12
Blue	0.96
Average	**0.52**

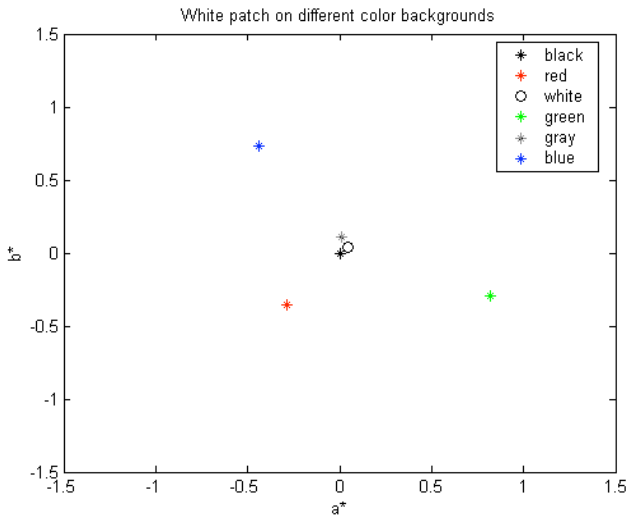


Figure A.2 - a* b* diagram of the white patch on different backgrounds. Each marker represents the white patch on the particular color of the background.

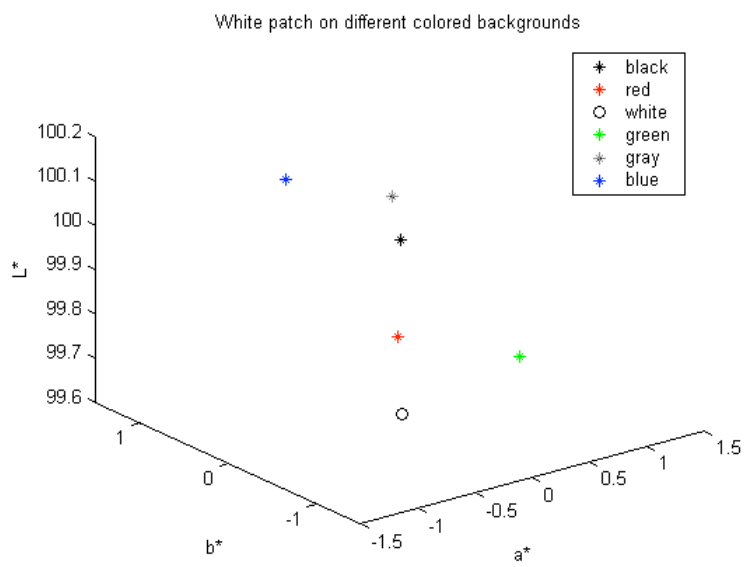


Figure A.3 - White patch on different backgrounds in 3 dimensions

The above two graphs show the problem the monitor has with voltage. It is seen in Figure A.2 that the colored backgrounds cause the white patches to take on the opposite hue. For example the red background causes the white to become cyan-ish, blue causes the white to become more yellow and green causes the white to become magenta-ish. The reason is

due to insufficient voltage of the guns to keep all three at maximum power when they need to produce color across the screen. When only the white patch is displayed and no other color is produced on the background, (i.e. a black background), the guns are able to produce maximum voltage at the small patch in the center of the screen, however when the background is changed from black to some other primary color there is less voltage available to that particular gun and this causes the white to shift to that primary's opposite color, ("less red" for example equals "more cyan"). Figure A.3 (71) shows the same as Figure A.2 (70) but in 3 dimensions. The major contribution of this figure is that it shows the L* component of the backgrounds. It is seen that white background has the lowest value, corresponding to the voltage problem discussed previously. However, it is interesting that flare is less of an issue than voltage. This is different for the gray background. It is seen that the white patch on this background has a higher luminance than white or black. It is reasoned that the voltage is not enough to significantly lower the luminance of the white patch but that flare is now the main issue. This leads one to the conclusion that flare at first is the main problem but at some point voltage begins to decrease and overpowers any effect of flare. Another interesting phenomena seen in Figure A.3 (71) is the blue background creating a higher luminance than the white on black condition. One possibility for this stems from the relatively low luminance (Y value) of the blue primary, (see Table A.1 (13)). When blue becomes the background there is less voltage needed for the blue, which then allows more power to the red and green guns. Since these primaries have higher luminance values it is hypothesized that the instantaneous increase in voltage sent to the red and green guns allow them to reach a higher luminance, thus allowing the luminance of the white to become higher than for the

white on black where all three guns have approximately the same voltage supplied to them.

It should be noted that although there were trends in the above data, the overall error was low. So, this data shows why the channel and spatial independency tests had some error – albeit a small error.

Building the TRC/Matrix Model

The three, (R, G, B), TRCs were populated using the normalized Y tristimulus value of each ramp. Normalization was done by dividing all values in the primary ramp by its particular maximum. Finally verification of the model was completed using the TRC and a 3x3 matrix. This equation is seen below, where XYZ_{min} is subtracted to account for flare and it is the measurement of the tristimulus values of black.

$$\begin{bmatrix} X_{\min} \\ Y_{\min} \\ Z_{\min} \end{bmatrix} + \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_{r\max} - X_{\min} & X_{g\max} - X_{\min} & X_{b\max} - X_{\min} \\ Y_{r\max} - Y_{\min} & Y_{g\max} - Y_{\min} & Y_{b\max} - Y_{\min} \\ Z_{r\max} - Z_{\min} & Z_{g\max} - Z_{\min} & Z_{b\max} - Z_{\min} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Essentially, digital counts of the verification colors are first run through the appropriate TRC, (red through R-TRC, green through G-TRC, et cetera), then the values returned from them are assigned to the $\begin{bmatrix} R \\ G \\ B \end{bmatrix}$ vector in Equation A.1. These values are considered linearized RGB scalars that are between 0-1 and are not RGB digital counts. Equation A.1 provides the tristimulus values of the verification colors, which are then compared to the measured tristimulus values in CIELAB space.

The TRCs are also known as LUTs and are plotted below in Figure A.4 (73). They represent the luminance portion of each primary channel, (The Y tristimulus value of the red, green and blue channels). Note that the LUTs are a little noisy but this is only

because all 256 values from 0 – 255 for each primary ramp were measured. Next the gammas of each LUT were determined and are shown in Figure A.5. It can be seen that the gamma values for blue and green are essentially the same, with a value of 2.3 and the red LUT has a little higher gamma at 2.4. The gamma values were found by running an optimization routine that minimized RMS between the LUT values and another set of values between 0 and 1 raised to a gamma value. The gamma was changed on the second set of values until the RMS was minimized.

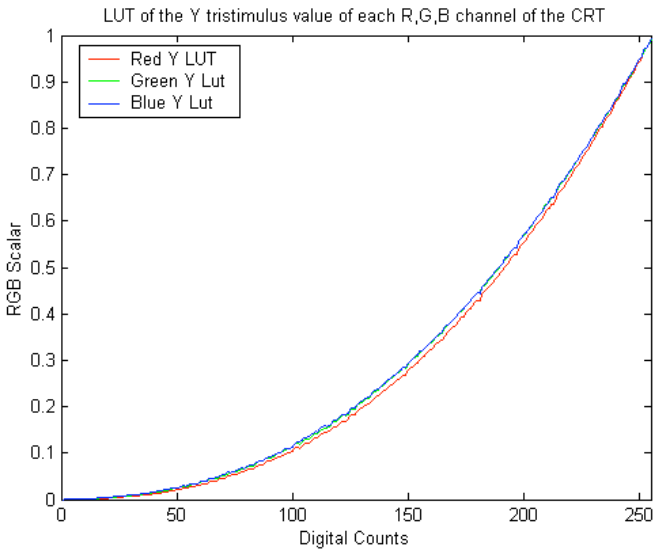


Figure A.4 - TRC for each primary channel

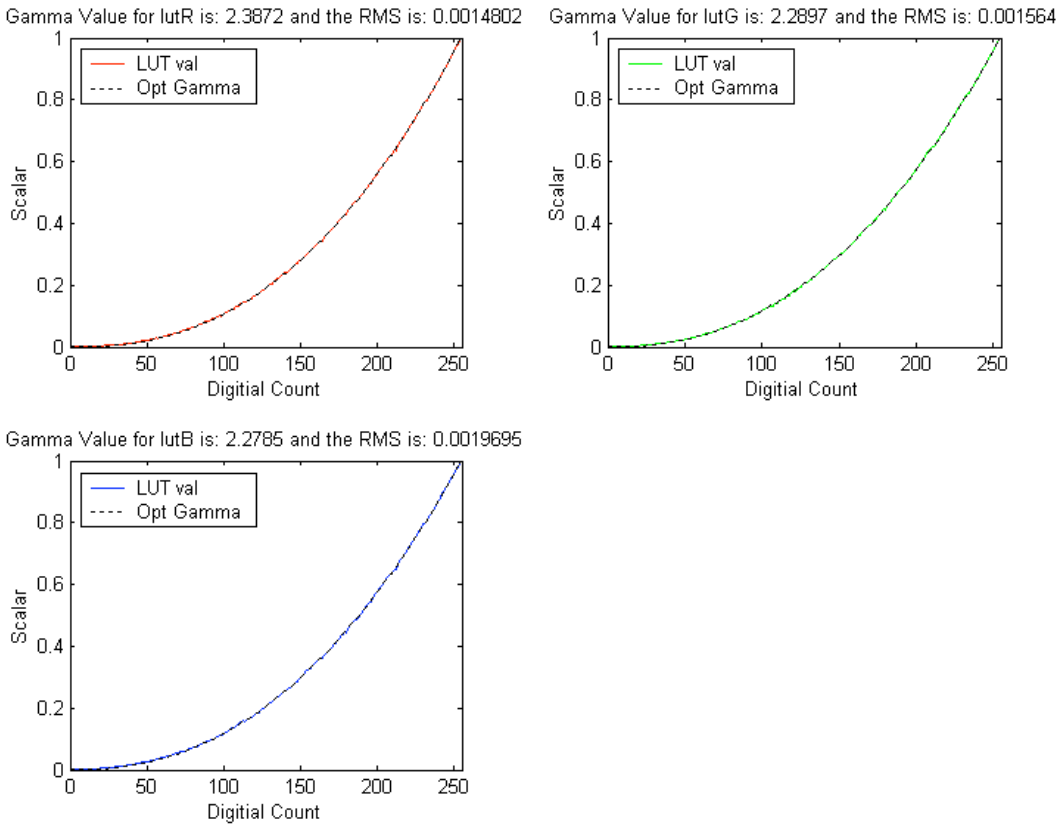


Figure A.5 - Gamma determination of each TRC

Verification of the TRC/Matrix Model

Table A.6 (18) is the result of 2000 randomly selected colors uniformly distributed in RGB space. The RGB values were transformed to XYZ via the TRC and matrix and then they were converted to L*a*b* using white on a black background as the nominal white. The histogram of the L* values and a plot of the data in CIELAB space is shown below in Figures A.6 (74) and A.7 (75).

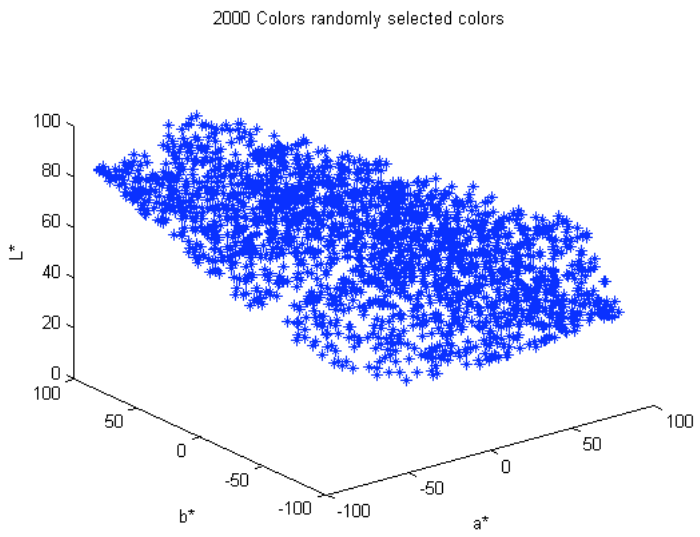


Figure A.6 – 2000 colors selected in RGB space from a uniformly distributed population

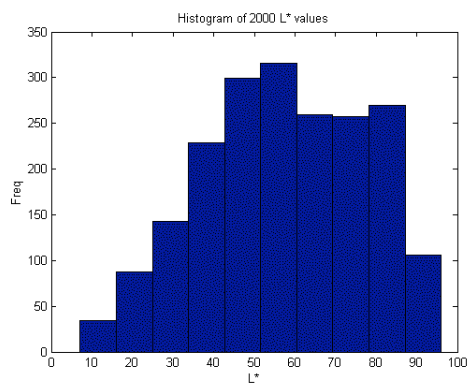


Figure A.7 – Histogram of 2000 L* values

The measured tristimulus values are compared to the same digital values run through the LUTs and tristimulus values calculated from the resulting values. It can be seen that the model provides very good results.

Table A.6 - Verification results using CIEDE2000 between 2000 measured and calculated randomized colors

Verification of model using ΔE_{00}		
Max	**Min**	**Avg**
1.91	0.01	0.31

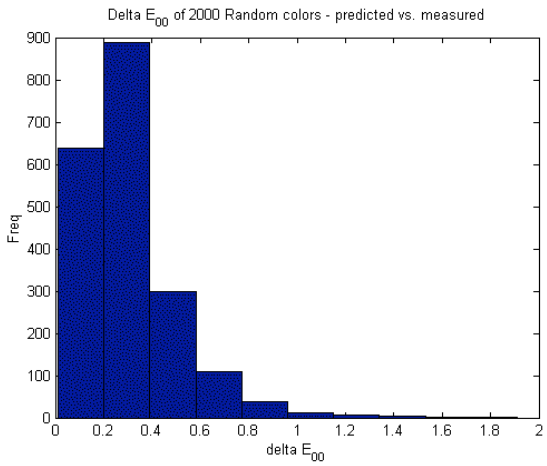


Figure A.8 - CIEDE2000 histogram of verification colors

Figure A.8 (76) shows the histogram of the ΔE_{00} for the 2000 colors and it is seen that the majority of differences are below 0.5, which is well below a JND. The next figure plots the L*a*b* and a*b* of colors above ΔE_{00} 0.5. It is seen that there are no discernable trends in the data; the colors appear uniform throughout CIELAB space.

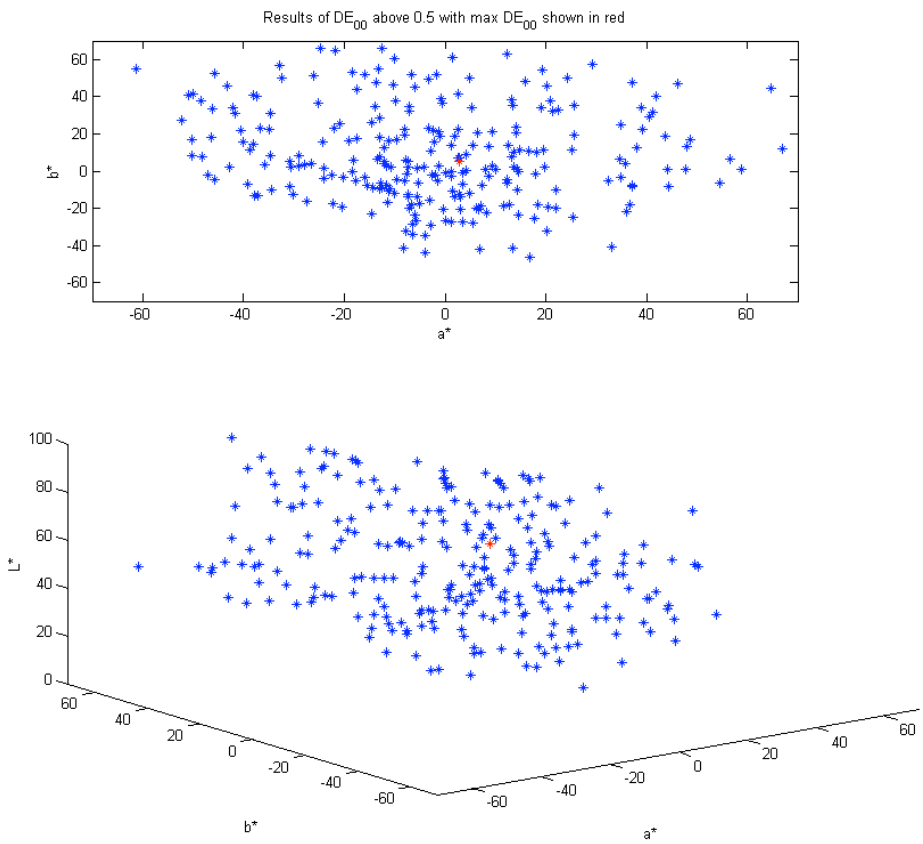


Figure A.9 – CIELAB plot of colors with ΔE_{00} above 0.5

Conclusions

The above data show that the characterization model utilized performed well. Although there was some issues in the channel and spatial independency test, the overall error was relatively low. Therefore, based on the results this display was well characterized and used for the experiments in Project 1. The next report is on the 37" LCD using the same characterization process and TRC-matrix model.

Appendix B

Comparison between PR650 measurements and the LMT

Measurements were taken of maximum red, green and blue primaries along with white and black using the PhotoResearch 650 SpectraColorimeter. A range of angles from 0º to 40º in 5º increments was taken of the primaries along with a sparse sampling of the primary and neutral scale ramps. The process of measuring different angles of a color shown on the LCDTV was to aim the PR650 at the center of a color patch and then move the 650 to get each angle. The results for each color patch show that as the angle increased, the color, in terms of CIELAB, became darker and less saturated. This was the case for all colors except black, in which the opposite happened.

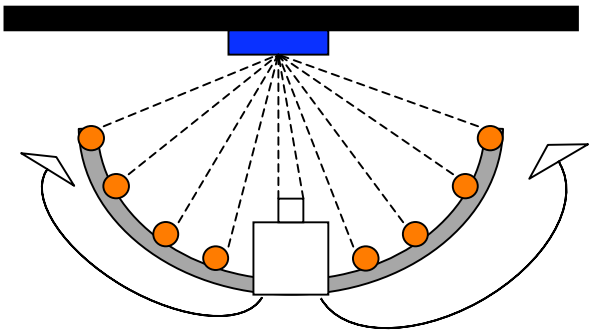


Figure B.1 - Schematic of measurement setup.

The diagram above shows a schematic of the measurement setup. The PR650 was moved along an arc so that it was always a constant distance from the LCDTV and each of the dots represent an angle where the furthest dot away from the 650 on the arc is 40º and the small blue square represents the color patch. Each dotted line represents the path from each angle on the arc to the same spot on the color patch.

In addition, the primary ramps were measured sparsely at the 0º mark using 15 nodes and then interpolated to attain all 256 values. CIELAB values were then calculated and compared to the ramps measured by the LMT. Note that the LMT values are in illuminance units but the values are scaled to luminance units and for the CIELAB calculations the same whitepoint is used for both instruments. The graphs in Figures B.2 (79) below are shown for the red primary but the other colors follow similar trends and are seen at the end of the report.

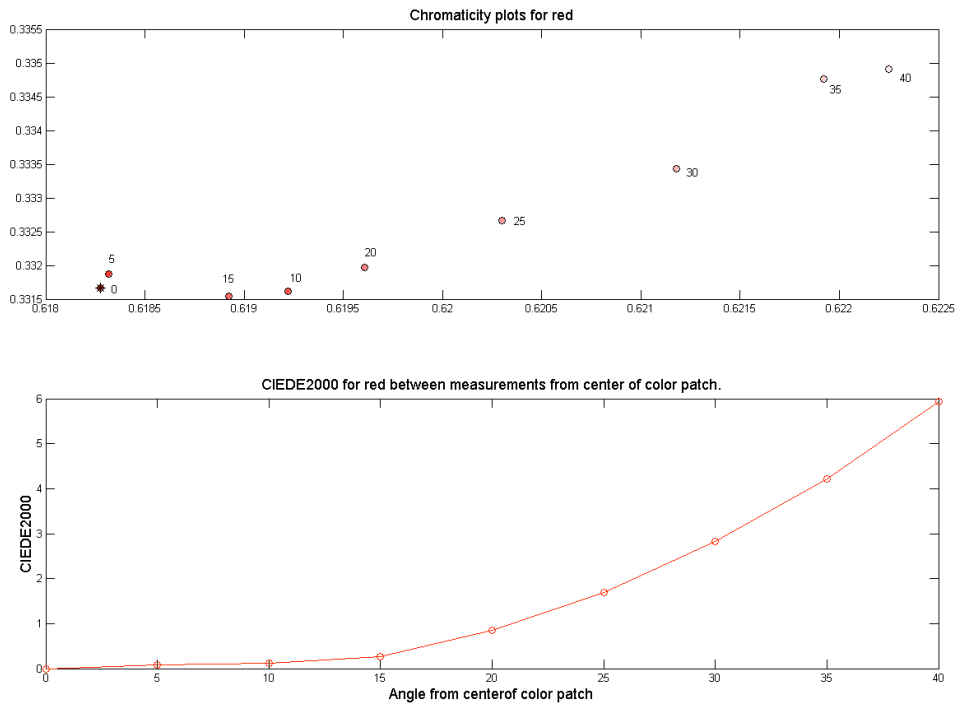


Figure B.2 - Chromaticity plot and ΔE00 for different angles for red primary

The top plot in Figure B.2 (79) shows the chromaticity coordinates for the red patch as the angle increases and the bottom plot is the ΔE_{00} between the 0º and all other angles, (where 0º is the PR650 centered on the patch). Note that each data point is the average of two measurements. Figure B.3 (80) below shows the differences in CIELAB between 0º

and the other angles with the 0º shown as a red triangle and the lines connecting to the other angles. The text in each graph is the angle that the marker represents. The first plot is a* vs. b*, (a* on x-axis and b* on y-axis), the second plot shows L* vs. C^*_{ab} and the third is a 3D plot of L*, a*, b* with L* on the z-axis, a* on x-axis and b* on the y-axis. These three plots show that as the angle increases, red becomes less saturated, (as seen by the decrease in C^*_{ab} and the shift in a* towards 0 – or the neutral point), and darker as seen by the decrease in L* in plots 2 and 3.

Figure B.4 (81) is an overall plot of the chromaticity coordinates for each primary and white as a function of the angle. It is seen that there is not much of a shift as the angles change. (The difference between Figure B.4 (81) and the plot in Figure B.2 (79) is the Y-axis scale.)

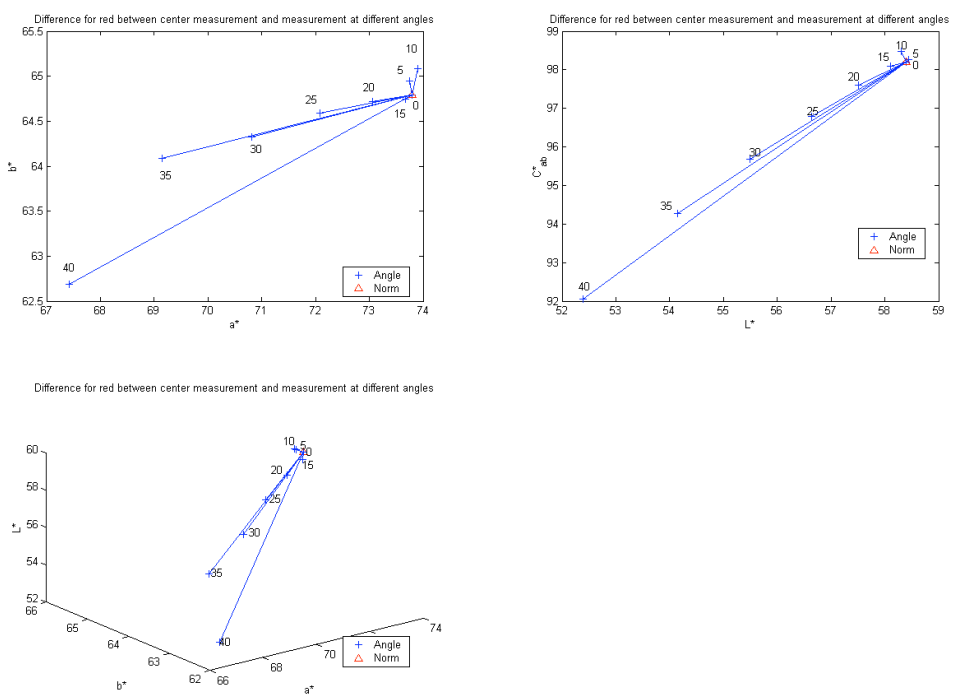


Figure B.3 - CIELAB error plots between 0º and all other angles

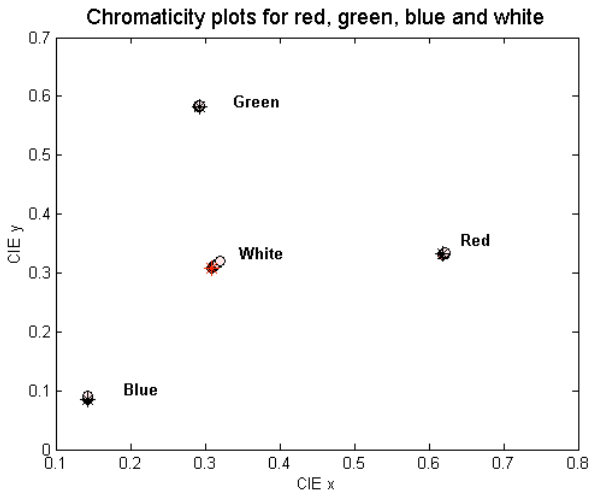


Figure B.4 - Chromaticity plot for red, green, blue and white for each angle

Next are the results for the primary and neutral ramp. The nodes were chosen by first taking the derivative of each primary ramp and then selecting 15 positions that were local minima and maxima. These were believed to be the breakpoints that exist for each ramp. The result is more sampling in the dark region, (lower digital counts) and less sampling in the higher digital count region. Figure B.5 (82) below is an example for the green ramp.

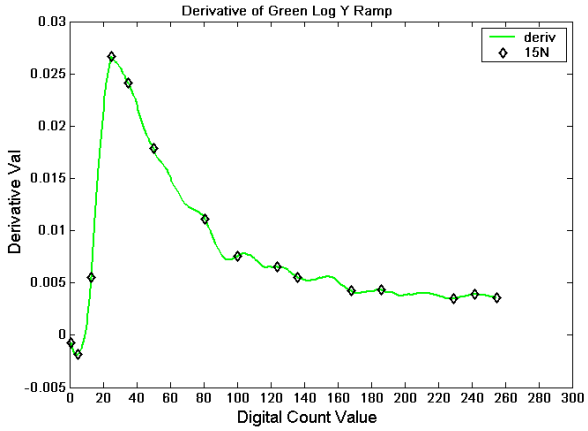


Figure B.5 - Derivative of green ramp with the 15 nodes measured

The next figure, B.6 (83), shows the difference between the PR650 and the LMT for the neutral scale ramp. Note that the trend is the same for the neutral scale as with the

primary ramps. The LMT is the dotted line and the PR650 is the solid line. A similar plot is shown for the red primary ramp in Figure B.7 (84).

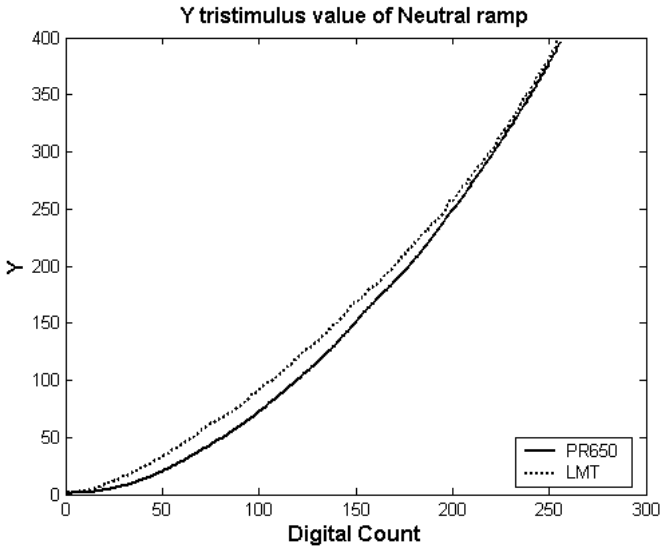


Figure B.6 - Y value of the Neutral Scale ramp from the LMT and PR650

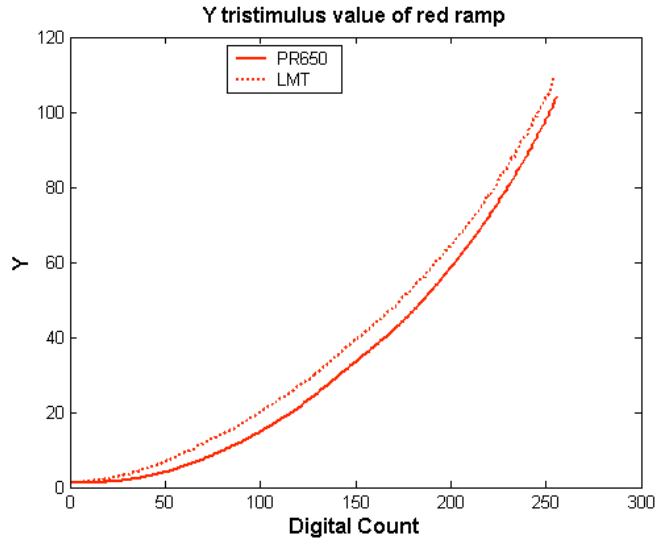


Figure B.7 - Y value for the red ramp from the LMT and PR650

Lastly, the gamma value of the neutral scale ramp, (Y tristimulus value), was calculated for the LMT as 1.6 and for the PR650 as 1.8. Figure B.8 (85) below is a plot of the graphs and the fitted gamma value for reference.

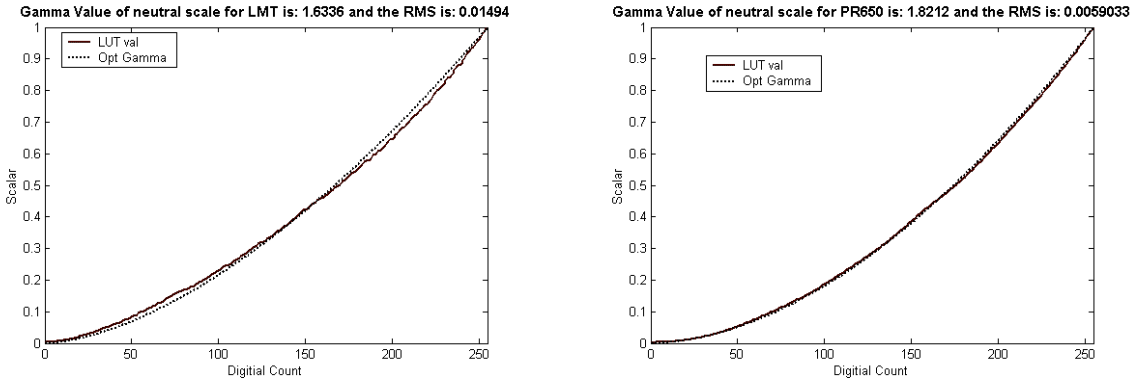


Figure B.8 - Curves fitted with gamma values for neutral scale ramps

Additional Plots

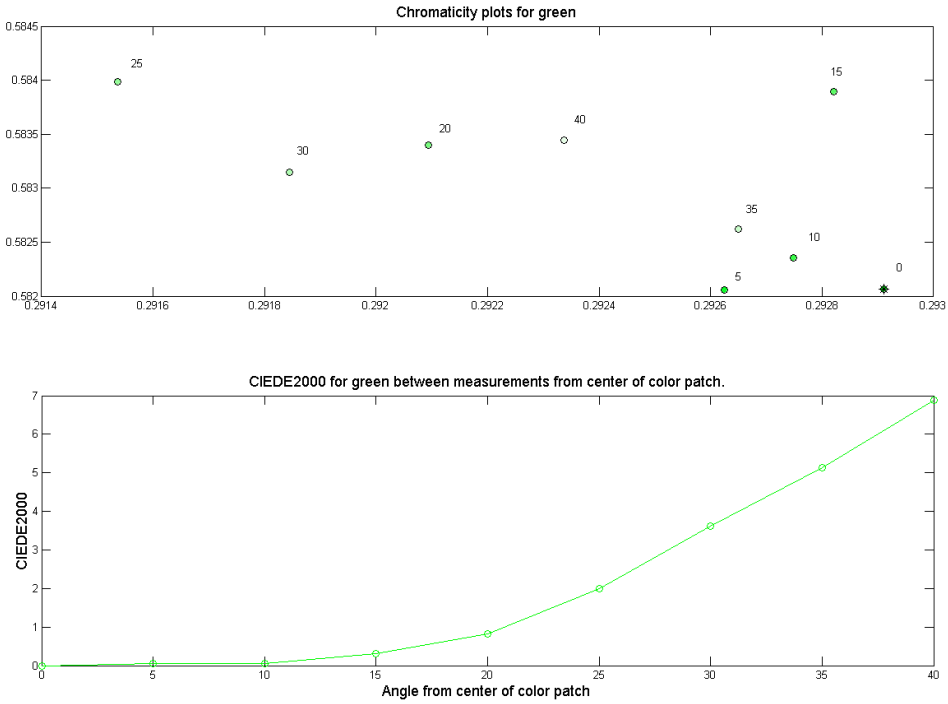


Figure B.9 - Chromaticity plot and ΔE00 for different angles for green primary

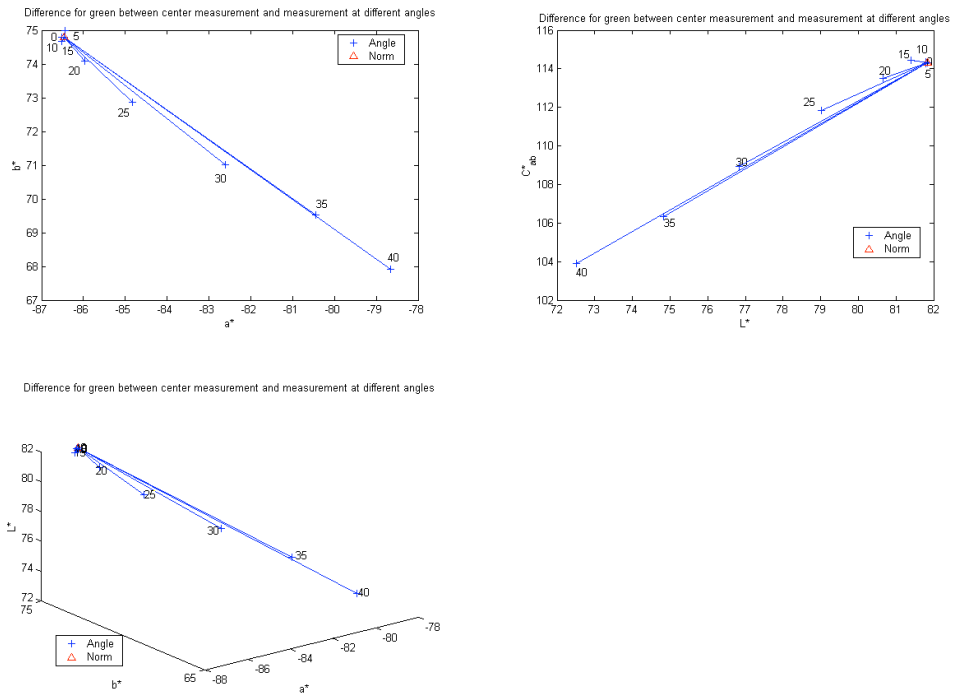


Figure B.10 - CIELAB error plots between 0º and all other angles for green primary

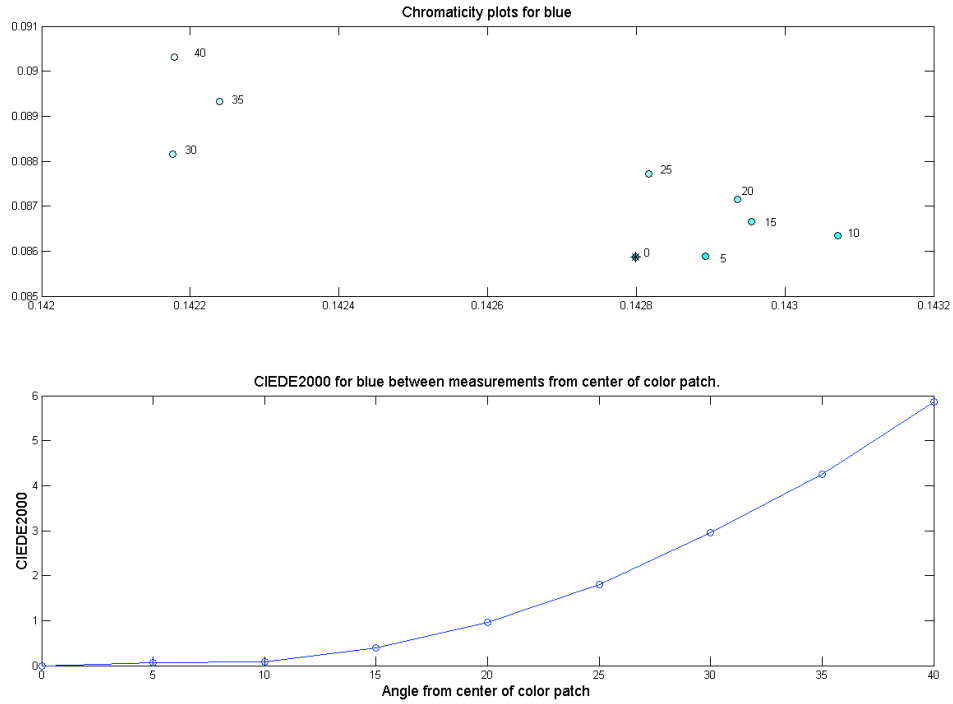


Figure B.11 - Chromaticity plot and ΔE00 for different angles for blue primary

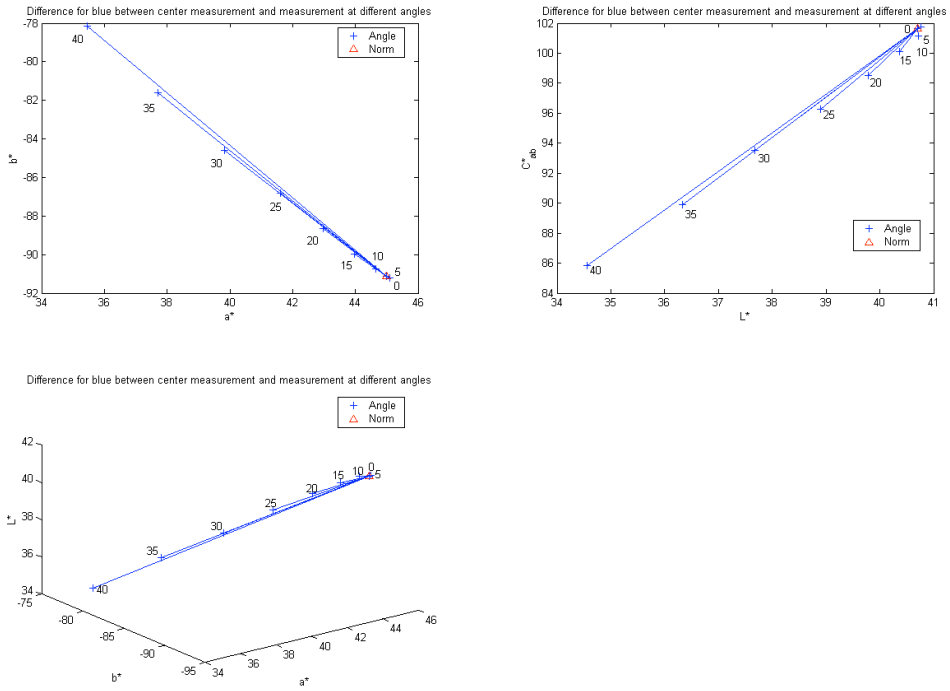


Figure B.12 - CIELAB error plots between 0º and all other angles for blue primary

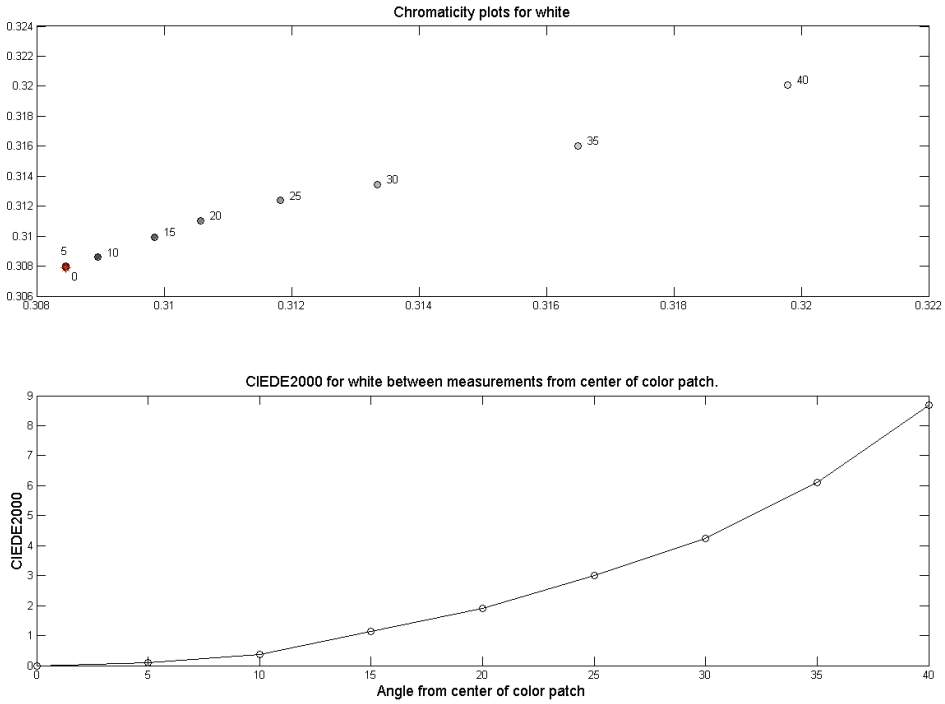


Figure B.13 - Chromaticity plot and ΔE00 for different angles for neutral scale

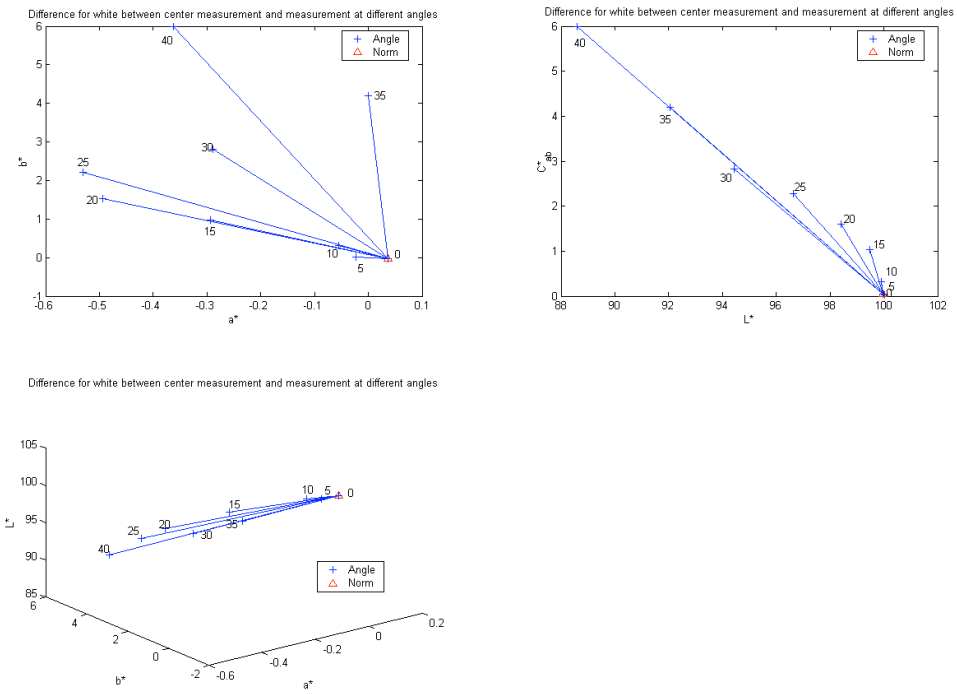


Figure B.14 - CIELAB error plots between 0º and all other angles for the neutral scale

Appendix C

Measurement Precision of 37-inch and 30-inch LCDTVs using the LMT C1210 Colorimeter

The initial process of determining precision of the two displays and LMT was to average measurements from multiple channel and spatial independent tests. For the 37" display there were 3 trials of each test run over a 5 day period and for the 30" display there were 5 trials per test over a 6 day period. The channel independency test measured a white, red, green, blue and black patch over a gray background. This was done originally to test whether the white patch equaled the [red+green+blue] patches. The spatial independency test measured a white patch over the following colored backgrounds: white, gray, red, green, blue and black. This test was originally performed to determine if colors at different locations around the screen affected the area of interest, (for the purposes here that would be a white circle at the center of the screen). Because these tests were run on different days, due to other reasons, they provide a good database for calculating the precision of our measurement system. An additional test was to measure both a white and black patch 10 times each. This would provide a sort of "best case" scenario, because the LMT is only measuring over a short time period and the same color is being shown consecutively. The results are given in units of ΔE_{00} between the individual measurements and the average of each color. For example, the red color measurements are averaged in the channel independency test and ΔE_{00} calculated between this average and all the red color measurement values in the channel independency test. A second metric is the overall MCDM, (mean color difference from

the mean), and this can be used to get a single number for each test, (channel and spatial), or for an overall value for each display.

Note that each measurement taken by the LMT is actually an average of three measurements, so for example the 10 consecutive white measurements were actually 30 measurements. Also, unless otherwise noted the background of the patches was medium gray. All this data will allow a good idea of the best achievable precision of measurement that can be expected. It is hypothesized that the lowest error will be for the white and black measurements taken consecutively. These are similar to "without replacement" values that are used in traditional precision and accuracy tests of measurement devices. The worst error should be in the white measurement on different backgrounds. This is due to the variability over time that our setup will experience, again these are similar to the "with replacement" in traditional precision and accuracy tests.

Initially, a precision measurement on the 30" LCDTV showed large variation of the measured colors when compared to the 37", as seen in Tables C.1 and C.2.

Table C.1 - Measure of below colors on different days (Channel Independency test)

	37"	**30"**
White	0.09	2.29
Red	0.19	1.80
Green	0.12	1.69
Blue	0.12	1.42
Black	0.34	0.37
MCDM	**0.17**	**1.52**

Table C.2 - Measure of White on below colored backgrounds on different days (Spatial Independency Test)

	37"	**30"**
Black	0.38	2.10
Red	0.40	2.14
White	0.40	2.18
Green	0.38	2.05
Gray	0.39	2.11
Blue	0.38	2.15
MCDM	**0.39**	**2.12**

Table C.3 - Comparison of 30 and 37 displays of consecutive white and black measurements in ΔE_{00} units

White	37”	30”	Black	37”	30”
Min	0.00	0.01		0.00	0.00
Max	0.03	0.07		0.06	0.06
Average	0.02	0.04		0.02	0.02
Std Dev	0.01	0.02		0.02	0.02

The results in Tables C.1 and C.2 were not satisfactory for psychophysical experiments on the 30” display and therefore the reasons for such variation were investigated. There could be several reasons for the results: engineering differences between the two displays, measurement error, or environmental differences to name a few. It is not believed that ambient, (such as lights coming on during measurements), conditions led to the errors. The measurement process was the same for both displays, with the only differences are locations of the displays and that the 30” is controlled by a Mac and the 37” is controlled by a PC. The results of consecutive measurements were similar for both displays as seen in Table C.3. These results indicate that the LMT/Display combination for both displays is precise over a short range and for similar colors shown consecutively. In order to rule out the LMT as a potential source of error, it was measured against a source other than the LCD displays. This was accomplished using a light bulb that was accurately controlled by a voltage regulator and verified independent of the LMT by a photodiode. Measurements were made for a 24 hour period, once each minute. The results show that the LMT had an average ΔE_{00} of 0.05 from the average. Table C.4 shows the statistics from the measurements.

Table C.4 - Statistics for light bulb measurement to check precision of LMT

	ΔE_{00}
Min	0.00
Max	0.12
Average	0.05
Std Dev	0.02

These results indicate that the LMT is indeed quite accurate and is on the same order of magnitude as the consecutive white/black patch measurements shown on Table C.3. The differences in the measurement for the 30" display were not random. There appeared to be trends that it followed during the 6 day period. The figure below is a plot of the Y tristimulus value for each color in the channel independency test. The trend was the same for the X and Z values and therefore not shown here.

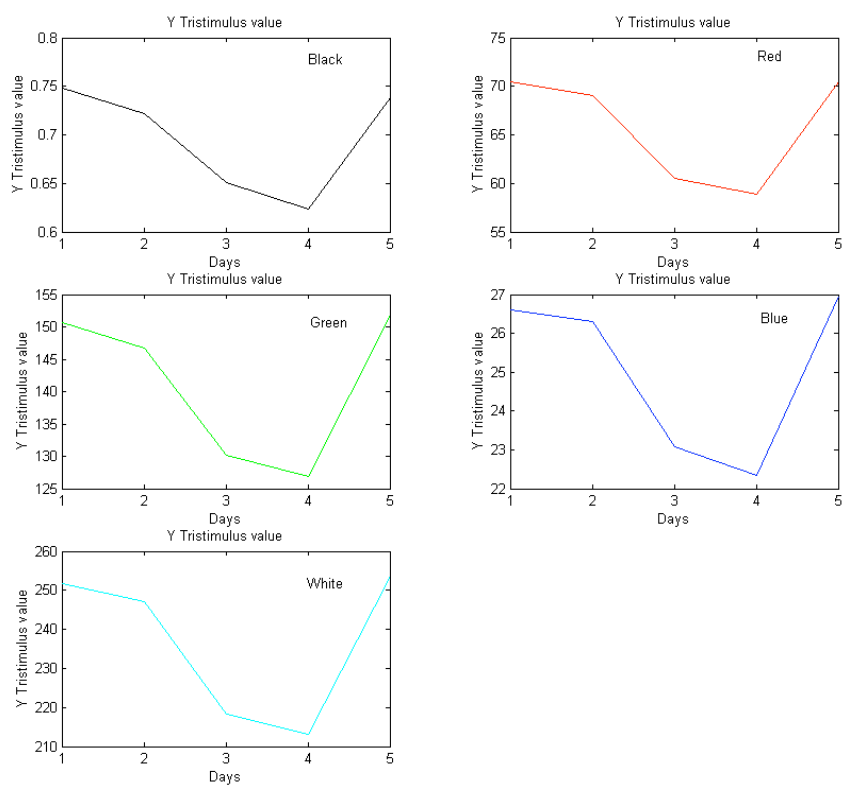


Figure C.1 – Y value plotted as a function of days for the colors measured in the Channel Independency Tests. There were 5 tests measured over a period of 6 days

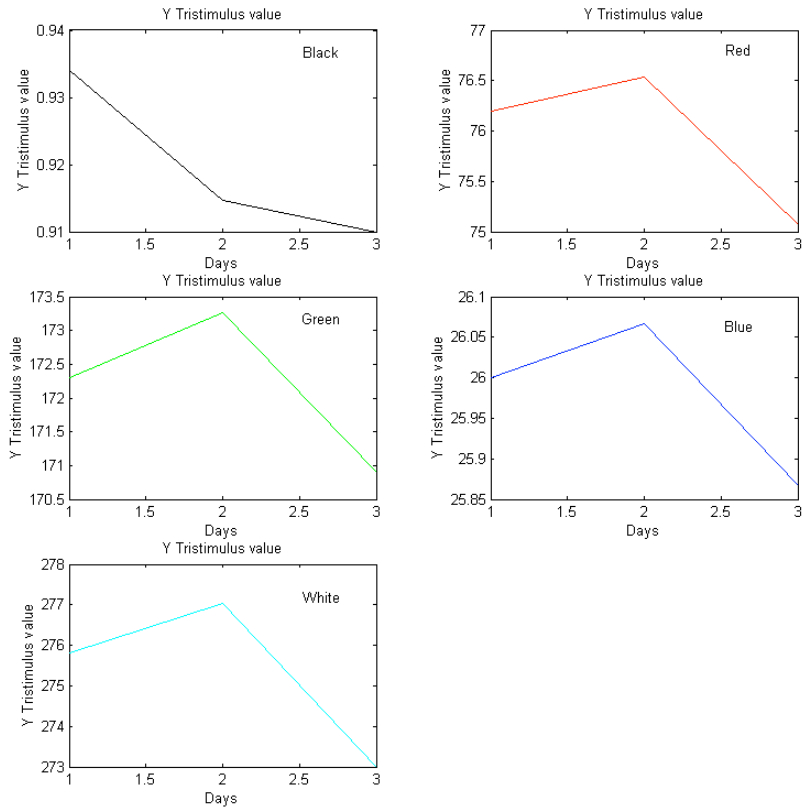


Figure C.2 – Y value for same colors as Fig. 1 but on 37" display. There were 3 tests over a 5 day period.

The two figures show that there are trends in the data for the two displays. However, it is possible these trends are from a lack of proper warm-up. As an example, in Figure C.1 data points 3 and 4 could be an indication of measurements taken before the LCD was stabilized. Also note the difference in scales between Figure C.1 and C.2; Figure C.2 shows variation over a much smaller region than those in Figure C.1. Since there were only a few points to evaluate in the graphs of Figures C.1 and C.2 it was decided to make more measurements of the 30" and look for trends in the data. First the warm-up time of the LCD was evaluated since this could be one possible source of error. It was found that the LCD takes approximately 1 hour to stabilize, as seen in Figure C.3 below.

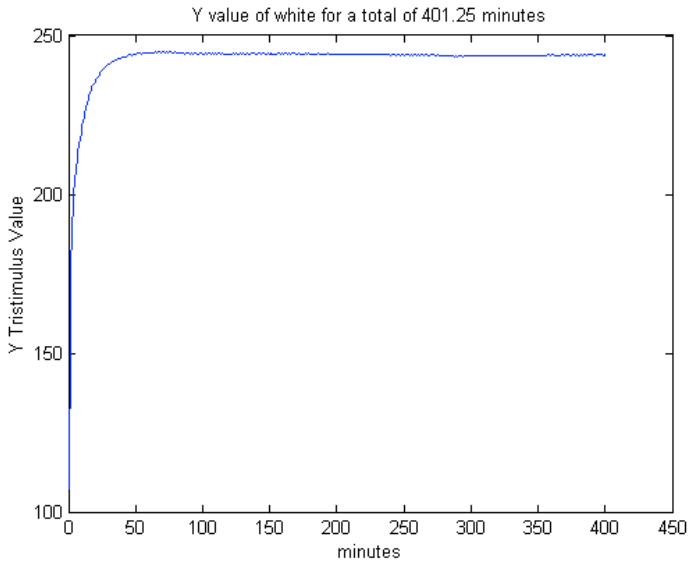


Figure C.3 – Y tristimulus value of white patch measured every 30 seconds over 6 hour period

Next, measurements of white, red, green, blue and black were made every hour over a weekend and then again over two days during the week. This was to ensure there were no environmental issues affecting the measurements. It is possible to have fluctuations of voltage that are present either over the weekend or during the week.

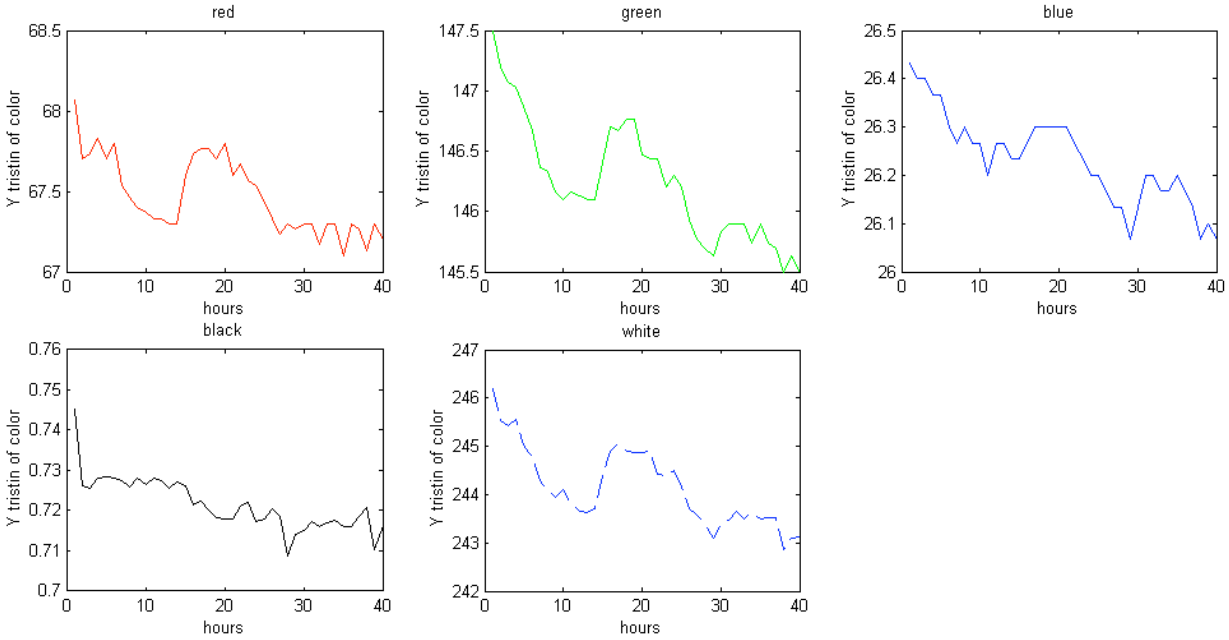


Figure C.4 – Y tristimulus values of each color over 40 hours during a typical weekend

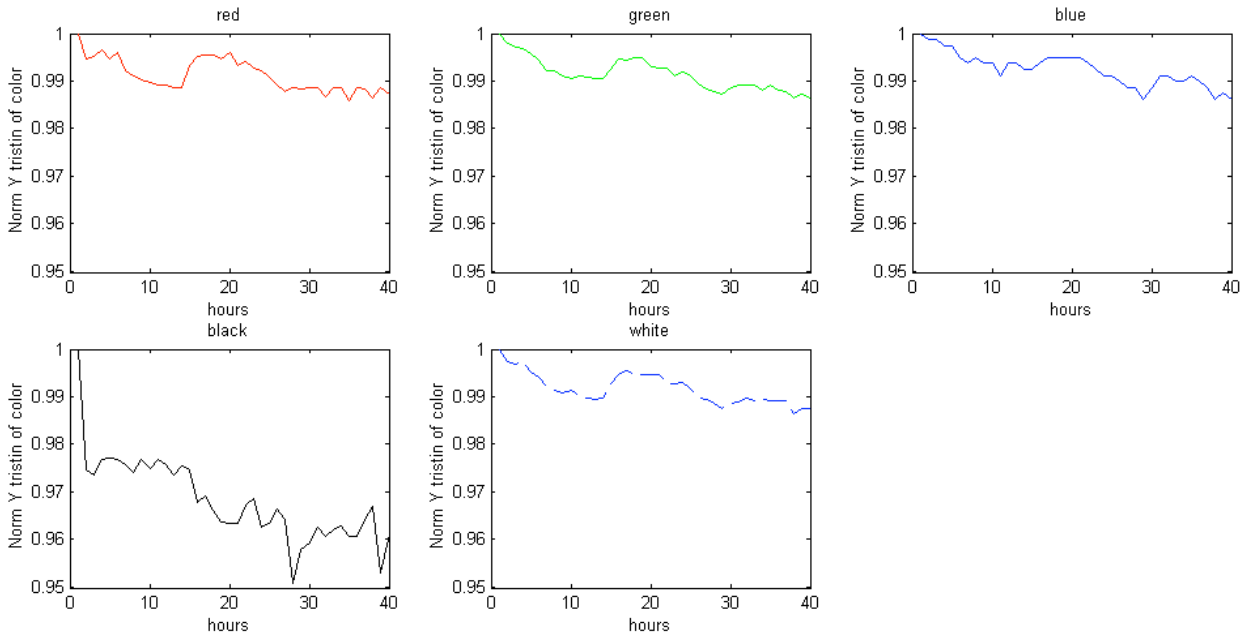


Figure C.5 – Normalized Y tristimulus value of measured colors over a weekend

Figure C.5 shows the normalized Y tristimulus value in order to demonstrate the relative differences between the colors and it can be seen that the colors do not vary much. Although the black patch showed the largest variation in Figure C.5, this does not reflect a large color difference as seen in Figures C.8 and C.9. The next two figures are similar to Figures C.4 and C.5 except they represent the measurements taken during the week, over a 28 hour period instead of over 40 hours.

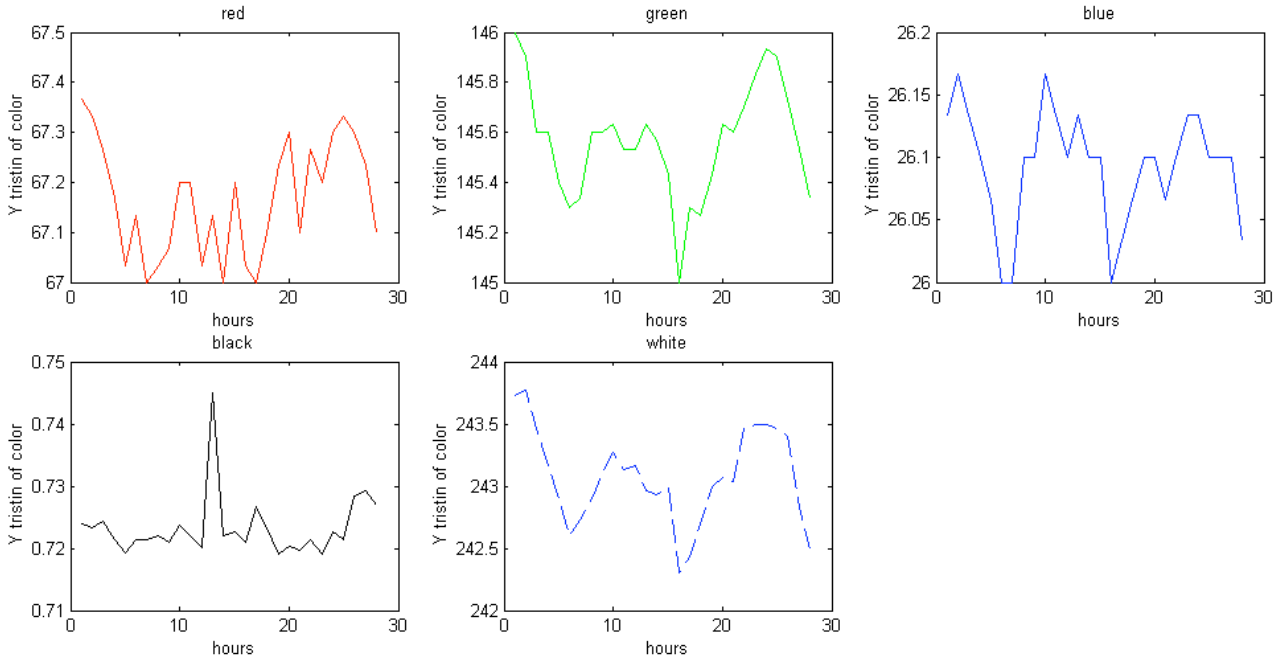


Figure C.6 – Y tristimulus values of colors measured every hour over two weekdays

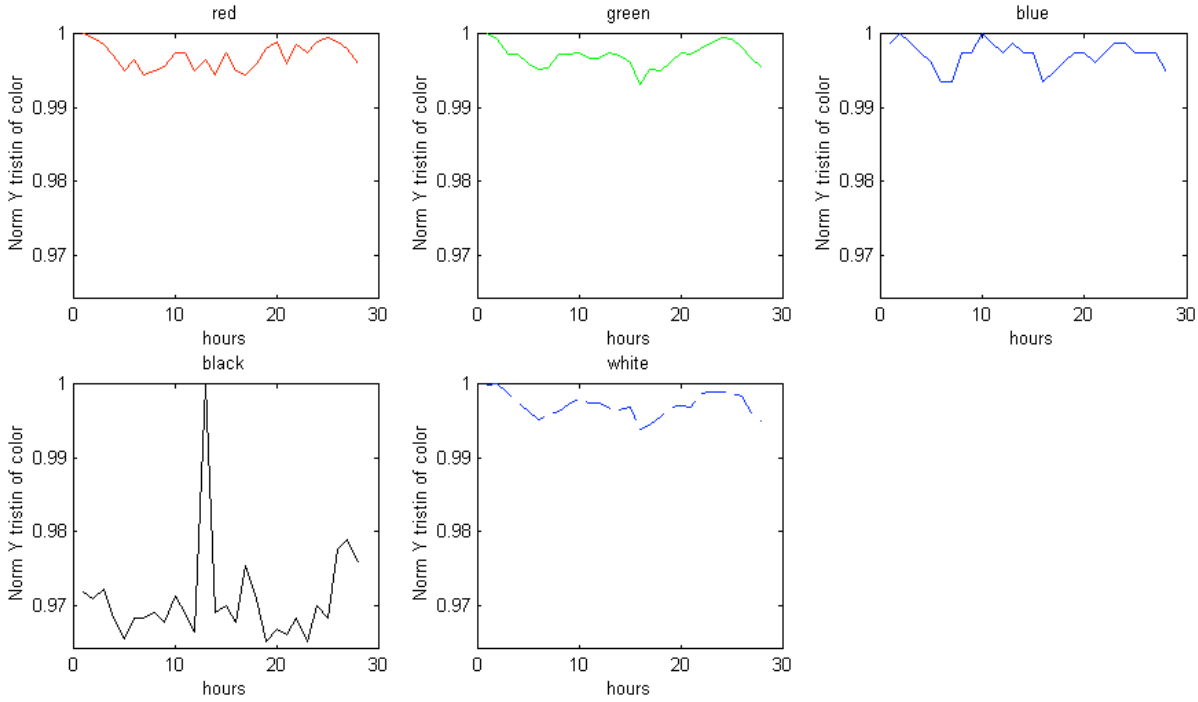


Figure C.7 – Normalized Y value of each measured color during two weekdays

Figures C.6 and C.7 show similar trends as Figures C.4 and C.5. There is a possible outlier in the black measurements but again, this does not translate to large color

differences. Figures C.4 through C.7 show that the measurements are similar during the time periods measured.

The color differences, in terms of MCDM, are shown for the weekend and weekday measurements. It seen that the color differences are quite small, with white having the largest ΔE_{00} for both sets of data.

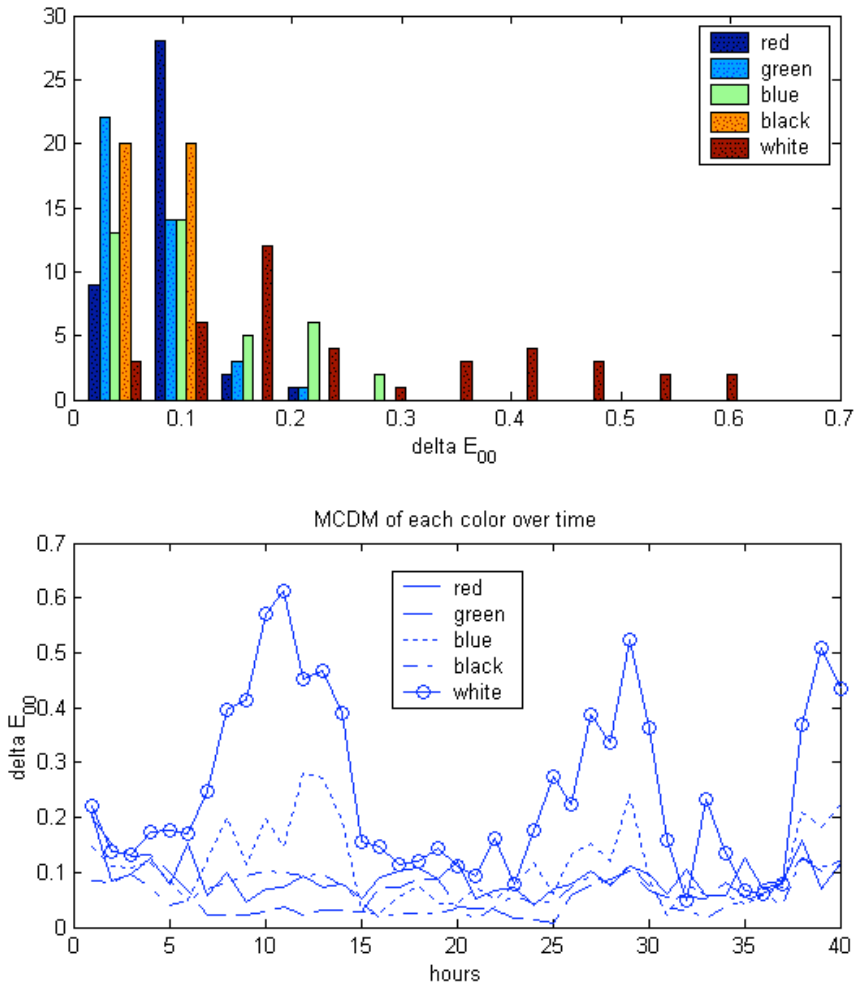


Figure C.8 – Histogram of ΔE_{00} and plot of ΔE_{00} over time for each color measured over the weekend

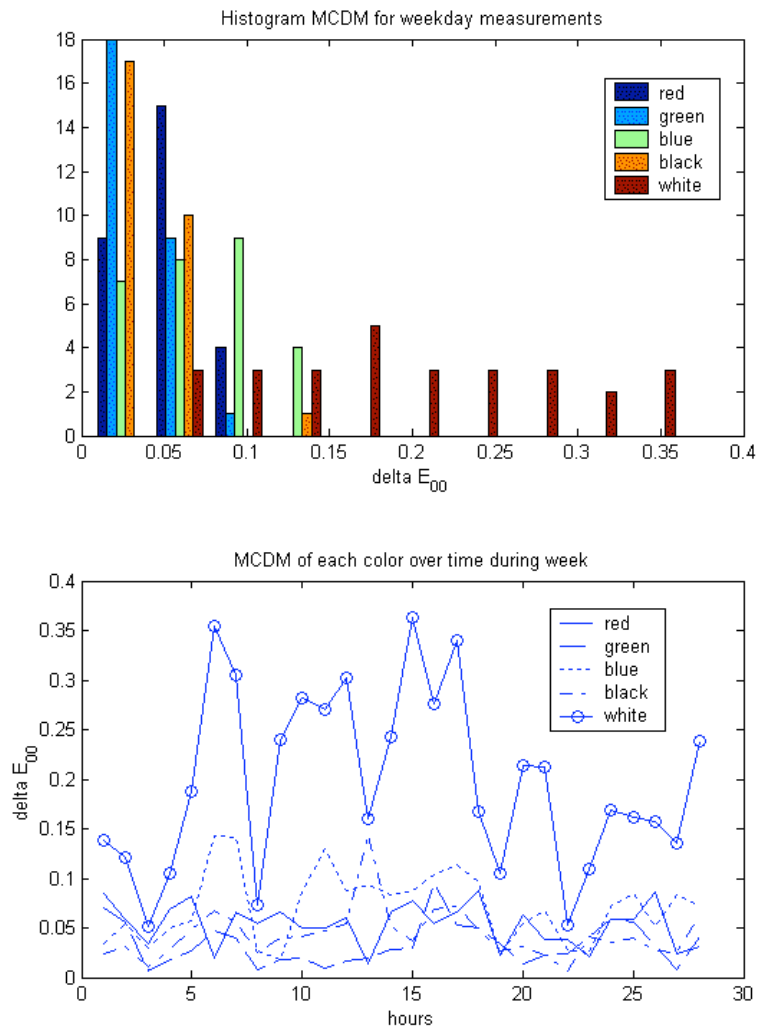


Figure C.9 – Color differences over two weekdays

The trends are very similar in the graphs shown in Figures C.8 through C.9; the major difference is the MCDM during the weekdays is smaller than the weekend. The major outcome shown in Figures C.4 through C.9 is that the 30" LCDTV does not have large color differences as was shown in Tables C.1 and C.2.

Next measurements of white, red, green, blue and black were made on three different days, after allowing the LCD to warm up at least 2 hours. There were multiple comparisons made: each day to all other days and each day to the average of the weekend measurements. (Note that the weekend measurements were used instead of the weekday

or the warm-up measurements, after a point in time, since the MCDM of the colors were larger and thus would provide a worst case scenario).

Table C.5 - Color difference between each day and the weekend

	Wed-Wknd	Wed-TH	TH-Wknd	Wed-Fri	Fri-Wknd	Fri-TH	**Avg**
Red	0.05	0.12	0.07	0.26	0.21	0.15	**0.14**
Green	0.05	0.09	0.04	0.16	0.12	0.07	**0.09**
Blue	0.08	0.05	0.07	0.13	0.07	0.14	**0.09**
Black	0.16	0.01	0.17	0.14	0.21	0.15	**0.14**
White	0.17	0.14	0.25	0.48	0.63	0.48	**0.36**
Avg	**0.10**	**0.08**	**0.12**	**0.24**	**0.25**	**0.20**	**0.16**

Table C.6 shows similar results for the 30" display when compared to the 37" as was originally thought.

Table C.6 - Comparison between 37" and 30" LCDTV

	37	30
White	0.09	0.36
Red	0.19	0.14
Green	0.12	0.09
Blue	0.12	0.09
Black	0.34	0.14
MCDM	**0.17**	**0.16**

Finally, the average white of the weekend measurement set was compared to the warm-up set after 20, 30, 45, 60 and 200 minutes. The longer the LCD is allowed to warm-up the lower the color differences become.

Table C.7 - Average weekend measurements of white to warm up measurements of white at specified times

Minutes	ΔE_{00}
20	1.95
30	1.27
45	0.69
60	0.58
200	0.03

Table C.7 provides intuitive results, meaning it is not a surprise that the LCD has to warm up before taking measurements due to the large effect that a light source has on the color primaries. However, Table C.7 does show that the results in Tables C.1 and C.2 were

likely caused by measurements taken prior to the 30” LCDTV stabilizing. The recommendation of this report would be to allow 2 hours for the 30” and 37” LCDTVs to stabilize prior to performing any experiments or taking measurements. Additionally, this report shows that the precision of the measurement system used throughout this and other reports allows, on average, a ΔE_{00} of 0.2 for both displays.

Appendix D

Characterizing LCD using TRC-matrix model

Summary

A characterization process, including a channel and spatial independency test was performed on the Sharp 37" LCDTV. Subsequently, a model was built using a TRC-matrix model similar to the one described previously for the CRT. The channel independency test showed that measured full white had a greater luminance than the sum of tristimulus values measured for the primaries and the ΔE_{00} between the measured and sum was 1.47. Interestingly, this is in contradiction to the observations for the CRT. The spatial independency test showed a difference between the white patch on a black background and all other backgrounds at an average ΔE_{00} of 0.06. Although the differences were small there was a noticeable effect for the white on the colored backgrounds to take on the particular backgrounds' opposite hue, (same as the effect on the CRT). As expected, the TRC-matrix model did not provide accurate results, in terms of ΔE_{00} between measured tristimulus values of 2000 random colors and the calculated tristimulus values. It is hypothesized that a 3D CLUT will be needed to account for complex pixel interactions. More details of this process are described in the next section.

LCD Characterization

The procedures described below follow those outlined previously for the CRT and therefore will not be detailed here; for any procedural questions refer to the previous section. In addition, this model was not expected to provide optimal results in terms of high accuracy in the prediction of displayed colors but a quantification of any errors the model brought out were desired and this model provided the motivation needed for a

different method of predicting colors accurately. At the end of this report details of subsequent data collection with the aim of building a 3D LUT are provided.

The characterization process included performing a channel and spatial independency test. These tests were run three times on different days. This allows for an average to be taken of the results and the ability to quantify the precision of the measurement process. That average of the three trials is shown below as the final result of the two tests. A metric of precision was established using these six trials and another test which measured white and black repeatedly, (described in the previous section, "Measurement Error of the System").

Table D.1 (19) and D.2 (20) shows the results of the channel independency test. It is clear that there is a dependency with the respect to the blue channel. An interesting outcome is that the measured white is greater than the sum of the primary channels. This is especially true for the blue channel. It is believed that there is additivity creating the greater luminance of the white, whether due to more twisting of the LCs or a greater potential being created when all pixels in a given area are completely turned on. This implies crosstalk between the three subpixels, where each R,G,B subpixel affect each other. One outcome of this leads to a white color becoming whiter.

Channel Independency Test

Table D.1 - Channel independency test, differences in tristimulus measurements

Channel	X	Y	Z

Red	139.73	75.93	9.88
Green	82.48	172.16	36.40
Blue	48.65	25.98	267.37
**********			
R+G+B	270.87	274.07	313.64
Meas White	272.62	275.28	321.88
Perc Diff	0.64%	0.44%	2.56%
	Average Percent Difference		1.22 %

Table D.2- Channel independency test, differences in CIELAB units

Channel	L	a	b	ΔE_{00}
Meas. White	100	0	0	0
R+G+B	99.83	-0.34	1.43	1.47

Spatial Independency Test

The spatial independency test is an indication of the ability of the display to supply enough power to each sub-pixel in order to maintain constant luminance of the white when other colors are being displayed. Similar to the CRT, there could be a limited power supply available to all pixels, so it is feasible that a similar color shift as was found on the CRT be found on the LCD as well.

Table D.3-Spatial independency test, differences based on white patch with black background

Differences between measurements of white on different backgrounds.

Background	X	Y	Z	L	a	b
Black	272.37	275.03	321.58	100.00	0.00	0.00
Red	272.13	274.83	321.42	99.97	-0.02	-0.02
White	271.88	274.51	320.99	99.93	0.02	-0.01
Green	272.32	274.91	321.50	99.98	0.05	-0.01
Gray	272.17	274.79	321.29	99.97	0.03	0.00
Blue	272.23	274.85	321.36	99.97	0.03	0.00

Table D.3 (21) shows the different CIELAB units of the measured white on different backgrounds. It is noted that there are differences outside the measurement error of 0.02 MCDM. It is clear from Tables D.1 (19) to D.3 (21) that there are influences from other background colors on the measured white, albeit very small influences.

Table D.4- CIELAB differences

Differences from white patch on black background		
L	a	b

Black	0.00	0.00	0.00
Red	0.04	-0.02	-0.02
White	0.07	0.02	-0.01
Green	0.03	0.05	-0.01
Gray	0.03	0.03	0.01
Blue	0.03	0.05	0.01
Average	0.04	0.02	0.00

Table D.5 - CIEDE2000 between white on black background and all others

Background	ΔE_{00}
Black	0.00
Red	0.05
White	0.07
Green	0.08
Gray	0.05
Blue	0.08
Average	**0.06**

The graphs seen below show the changes of the white patch with respect to the background color. Similar results were found for the different primaries in the LCD as in the CRT, where the white on the red shifted towards cyan, the green towards magenta but the blue did not shift towards yellow as expected.

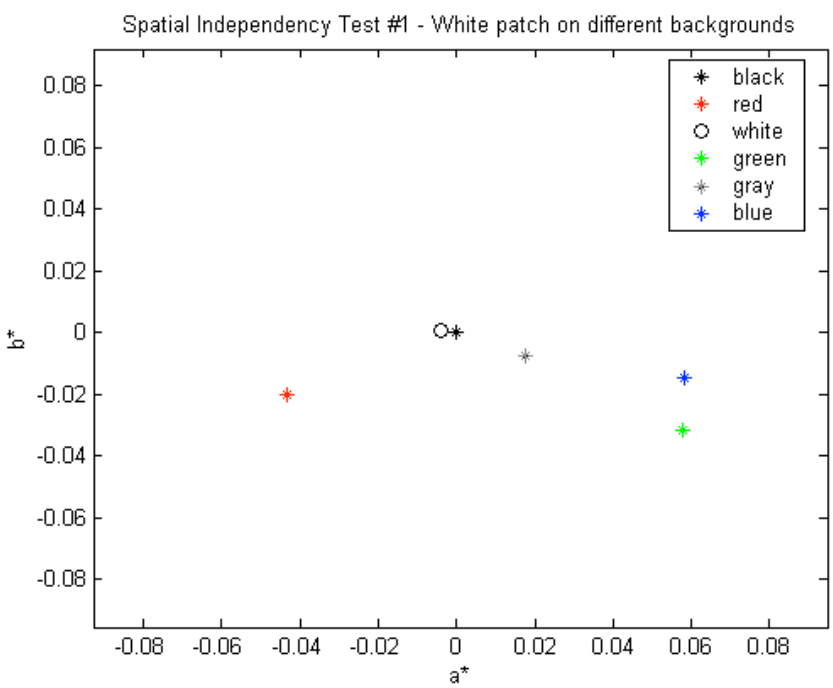


Figure D.1 - White patch on different backgrounds, marker colors are the background color

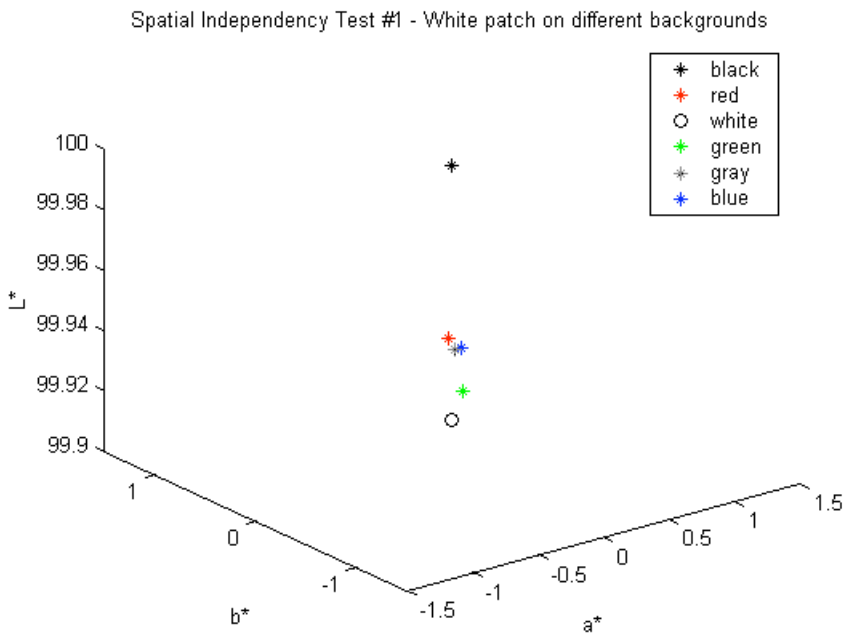


Figure D.2 - 3D view of L*a*b* of white on different backgrounds

It should be noted that these influences are quite small and it is very possible that the measurements are affected by noise in the system as much as any actual spatial effects from the display. (See previous section detailing the measurement precision of the system)

Building the TRC/Matrix Model

The next portion of the characterization involves the use of three 1D LUTs and a 3x3 matrix as was used with the CRT. It was not expected that this would provide as good a result as with the CRT. The graphs below show that the relationship between digital count and normalized Y, (called a scalar on the Y-axis in the graphs below), value is not a simple gamma function. Furthermore, Figure 4 shows that the gamma values were different for each channel. It is well known that LCD devices do not have an inherent gamma function as with CRTs but instead their function is more sigmoidal. (reference)

The transfer functions of LCD televisions are engineered to mimic those of CRTs, resulting in the curves seen in Figure D.3 (97) and D.4 (98).

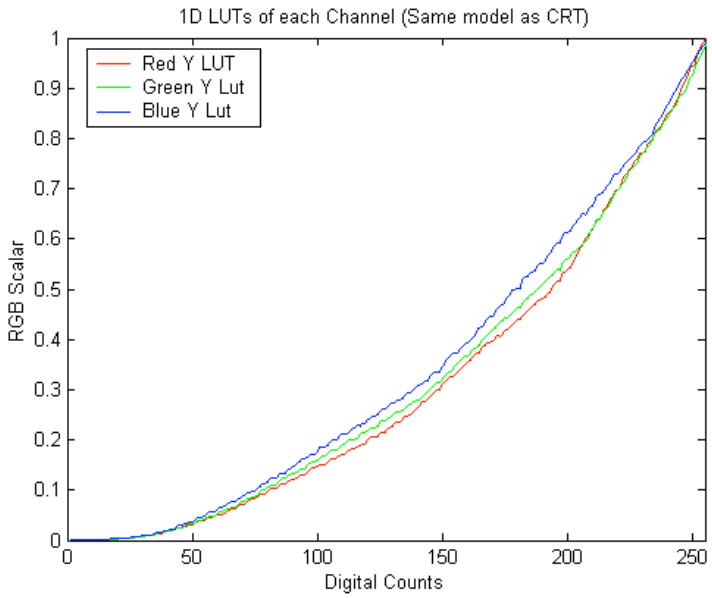


Figure D.3 - TRCs of 3 channels

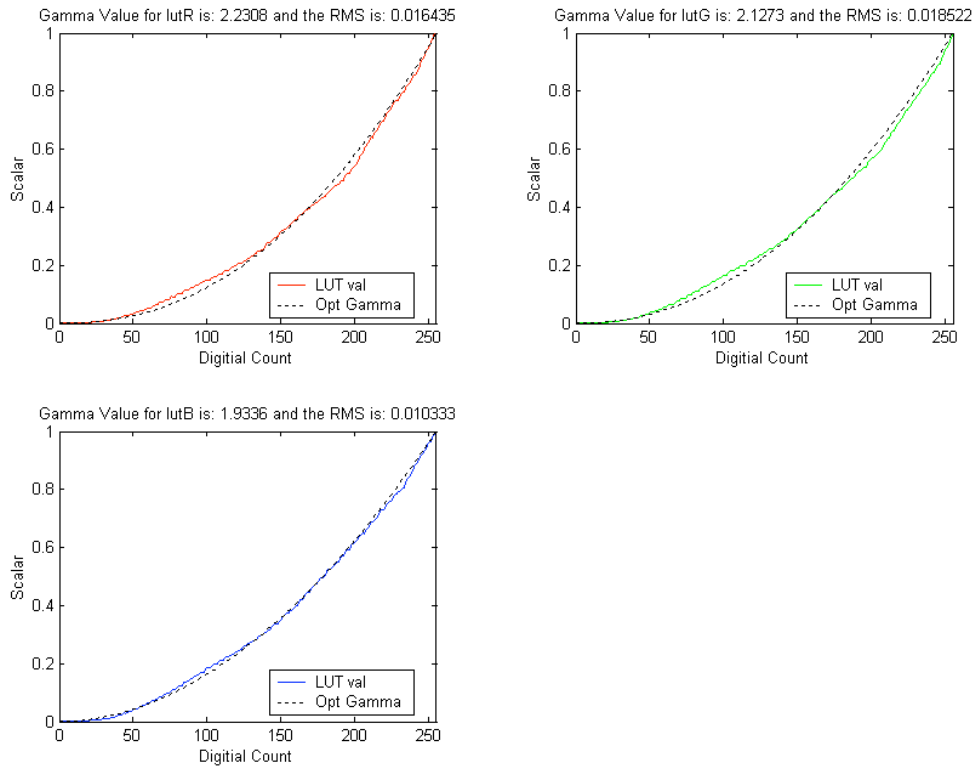


Figure D.4 - Gamma determination of 3 channels

Verification using TRC-matrix model

It was clear from the verification that the TRC-matrix model would not work as expected. Table D.6 (24) and the histogram in Figure D.5 (99) below shows that there was quite a large spread of errors. Even more troubling was the maximum ΔE_{00} value seen in Table D.6 (24).

Table D.6 - Verification results using CIEDE2000 between 2000 measured and calculated randomized colors

Verification of model using Delta E 2000		
MAX	MIN	AVG
9.47	0.05	2.64

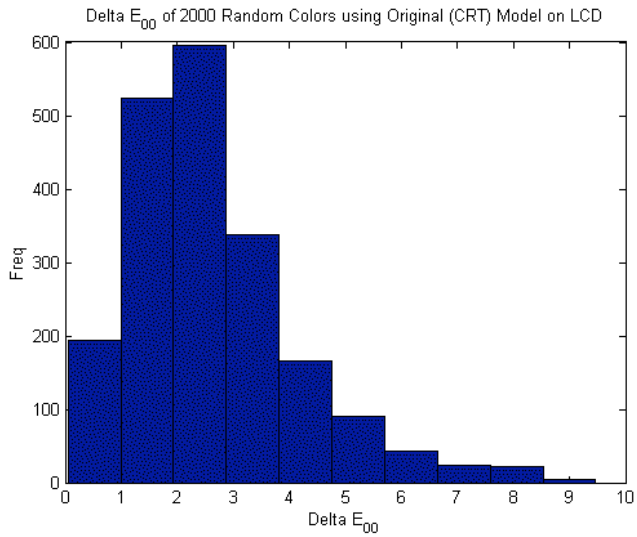


Figure D.5 - CIEDE2000 of verification results

The plot below shows that the largest errors were confined to a certain section of CIELAB space. The majority of these corresponding colors were relatively bright and in the red-yellow region. In contrast, there was no noticeable trends for colors with low ΔE_{00}, meaning that the colors with a low ΔE_{00} were spread uniformly across CIELAB space. At this point, it was decided to move on to a more complex characterization model instead of trying to optimize the current model.

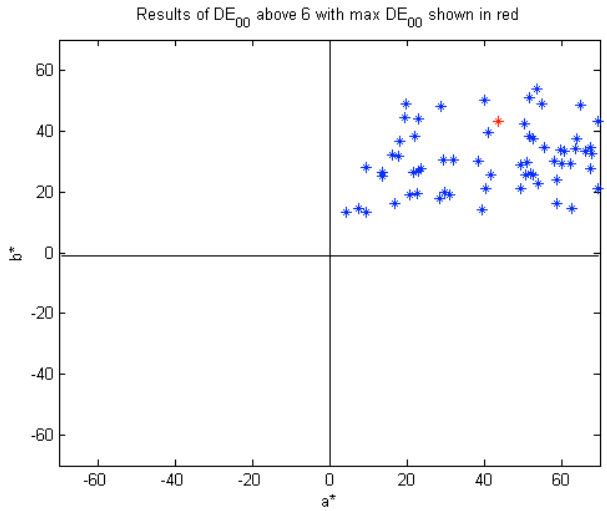


Figure D.6 - Collection of colors with highest CIEDE2000

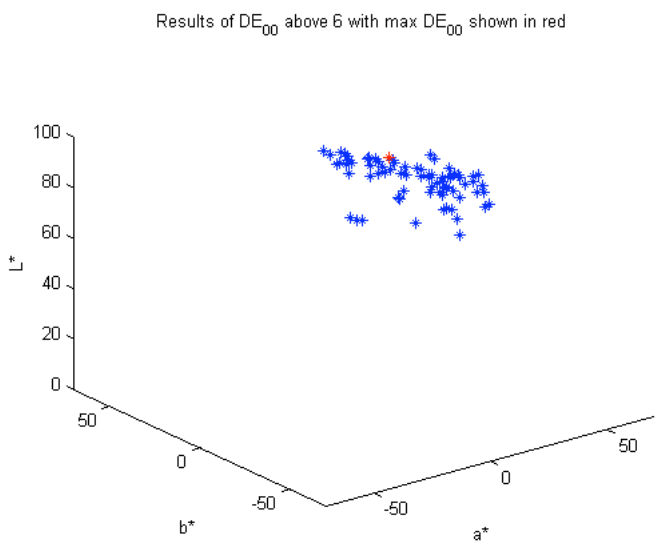


Figure D.7 - Colors with highest CIEDE2000 in 3 dimensions

Conclusion/Next Steps

As stated previously there is crosstalk between the three subpixels which leads to nonlinearity between each channel. This interaction cannot be modeled using 1D LUTs and/or a 3x3 matrix. There are a variety of ways to handle this interaction and it has been decided to build a 3D LUT that will allow direct look up between digital counts and tristimulus values (or CIELAB values). The process used to build the CLUT will be detailed next.

Appendix E

A New Method for LCD Characterization

Summary

A new approach is introduced for building accurate and compact color look-up tables (CLUTs) in order to describe the colorimetric performance of liquid crystal display (LCD) devices. In an experiment, a new technique was found to be useful in choosing original sample points to be used as nodes in the CLUT. The experiment demonstrated that an 8x8x8 CLUT produced through the new sampling method delivered comparable results to an 18x18x18 made from a different sampling technique. The technique will allow for future automation of choosing optimal samples for building LCD CLUTs.

Introduction

Due to a well-known crosstalk issue in LCD devices, simple color management models that treat channels as independent of each other are not able to accurately describe LCD colorimetric performance. Between the three sub-pixels comprising an LCD pixel, crosstalk results when voltage sent to one subpixel increases the voltage of the neighboring subpixel. The end result is that the liquid crystals (LCs) associated with one subpixel will pass more light when other subpixels are activated leading to channel dependency[57]. A three-dimensional CLUT is often used to characterize a three-channel device when an otherwise simple model is not available.

The first step in the process of building a three-dimensional CLUT is to choose samples from all available combinations of digital counts of the device. The digital counts are sent to the display and then measured by an instrument such as a colorimeter.

If the sampling protocol is properly designed, these measurements may be directly used as CLUT nodes. For the purposes of this experiment, a CLUT will take in RGB data and return CIELAB values.

As shown in Figure E.1 (102), three values are needed to locate values in a 3D CLUT. One is a value in the R direction, another in the G direction and a last one in the B direction. If the CLUT were fully populated (256 values per dimension for an 8-bit device) then the output from the CLUT would be a straight forward lookup with no interpolation necessary.

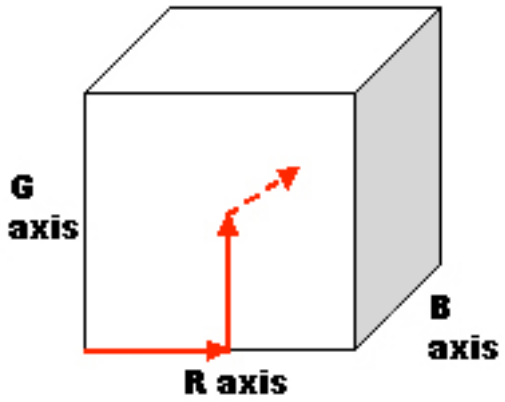


Figure E.1 - Finding 1 value inside a cube.

In most cases, though, the CLUT represents a sparse sampling of the device space. Thus, the lookup values do not typically point directly to output values, but instead fall within a subcube, as shown in Figure E.2. The eight vertices of the selected subcube are then used to interpolate the result. In order to obtain the lowest error, a good interpolation method is required. The number and placement of nodes becomes important for this reason. Typically linear interpolation is used and the more samples taken, the smaller the subcubes and the lower the chance that the assumptions of linearity are violated. If there is not a linear relationship between vertices and output values within a subcube then that

cube will have a high error. Such an error can be corrected by improving the placement or frequency of nodes.

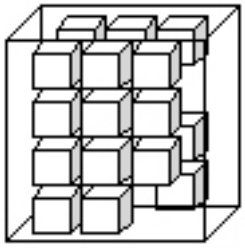


Figure E.2 - 1 3D LUT is actually made up of many small cubes.

CLUT Built Using Traditional Sampling

A traditional method for building a CLUT is to sample in equal increments of L*. That is consistent with the goal of minimizing color error in ΔE^*_{00}. A CLUT for a 37 inch LCD TV was built from this traditional method. In order to deliver low error from this approach, the CLUT size was chosen based on the largest number of samples that could be measured over a typical night, (10-12 hours). A 17x17x17 sampling fell within the time limit. To determine equal steps in L*, the CIELAB values of all 255 values for the three primary ramps were measured. Then equal L* increments were sampled along each ramp and the corresponding colorimetric measurement values recorded. The digital count values corresponding to those L* values for each primary ramp are used as the indices for CLUT nodes. Table E.1 shows the digital counts that corresponded to the even L* sampling.

Table E.1 - Digital count values corresponding to evenly sampled L* units.

	L* red	L* green	L* blue
1	0	0	0
2	36	30	43
3	47	39	55
4	58	50	66
5	70	61	77
6	82	72	88
7	96	84	102
8	115	99	118
9	133	116	134
10	146	135	150
11	161	152	163
12	181	168	176
13	197	188	193
14	210	208	209
15	225	223	226
16	241	240	241
17	255	255	255

For the 17^3 CLUT, the average ΔE_{00} was 0.58 from 2000 random colors. Table E.2 shows descriptive statistics for this CLUT.

Table E.2 - 17-cubed 3D LUT Results for the 37" of 2000 random colors.

Verification of model using ΔE_{00}			
MAX	MIN	AVG	STD
5.45	0.04	0.58	0.46

A maximum ΔE_{00} of 5.45 was surprising given the large amount of characterization samples that were taken. Figure E.3 illustrates how the error was manifested for samples associated with ΔE_{00} values above 3.0. The red triangle represents the measured value and the blue cross at the end of the attached line segment represents the interpolated value. The analysis revealed several groupings of high error colors. For example, there is one such grouping in the upper left corner of Figure E.3 representing saturated yellow-green colors.

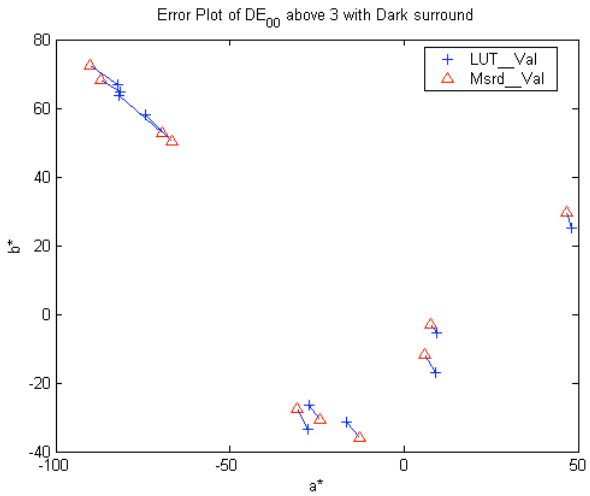


Figure E.3 – CIELAB values with ΔE_{00} above 3 for the 37". There were clouds of data that appeared systematic.

The colors in this group had similar ratios of RGB values, seen in Table E.3, in which the colors had low digital count values in the red and blue portion of the RGB color.

Table E.3 - RGB values of colors in upper left quadrant in Figure 3

R	G	B
15	159	11
11	250	10
15	169	12
17	240	25

Pushing the individual ramps through the CLUT and comparing them to their measured values created the errors shown in Figures E.4 through E.6. The digital count that corresponding to the highest ΔE_{00} in each figures falls within the first two nodes in each dimension. The use of the individual ramps for analysis of CLUT performance is highly accurate for pointing out the most important areas where error, in terms of non-linearity, is manifested within the CLUT.

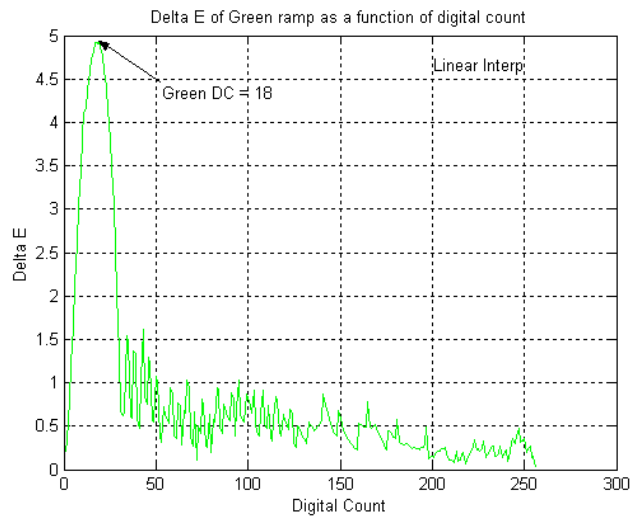


Figure E.4 – ΔE_{00} between measured and predicted Green ramp values of 37" display, with largest ΔE_{00} indicated by arrow.

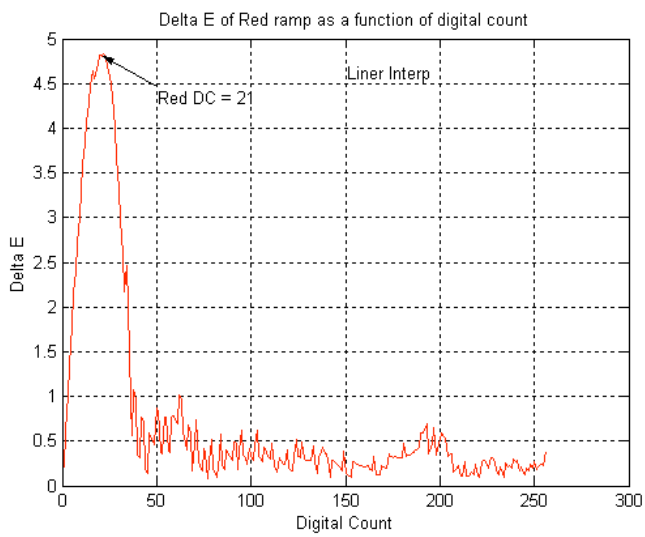


Figure E.5 – ΔE_{00} between measured and predicted Red ramp values of 37" display, with largest ΔE_{00} indicated by arrow.

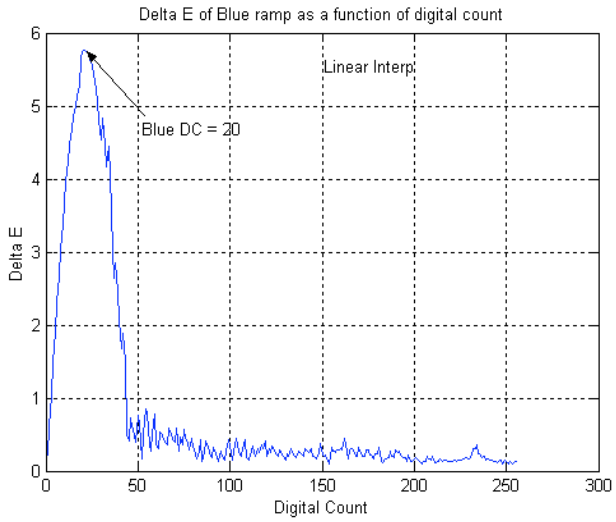


Figure E.6 – ΔE_{00} between measured and predicted Blue ramp values of 37" display, with largest ΔE_{00} indicated by arrow.

The measured primary ramps are plotted in log-luminance in Figures E.7-E.9. The values associated with the 18 nodes for each dimension are plotted on the curves. This shows where there is relative over- and under-sampling within each dimension.

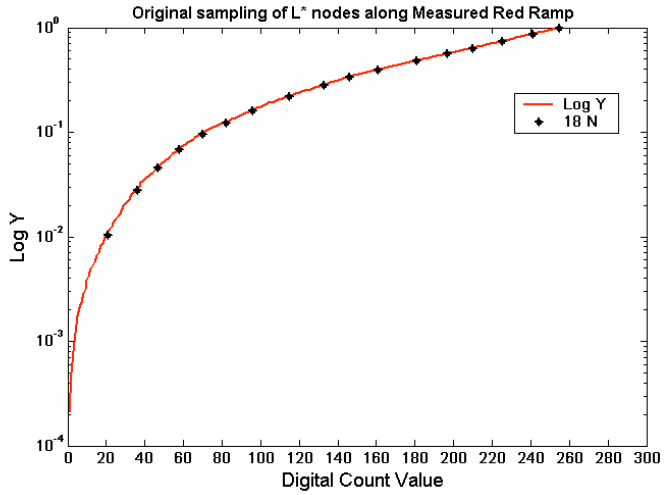


Figure E.7 – Log Y vs Digital count with original 18 nodes shown along Red dimension.

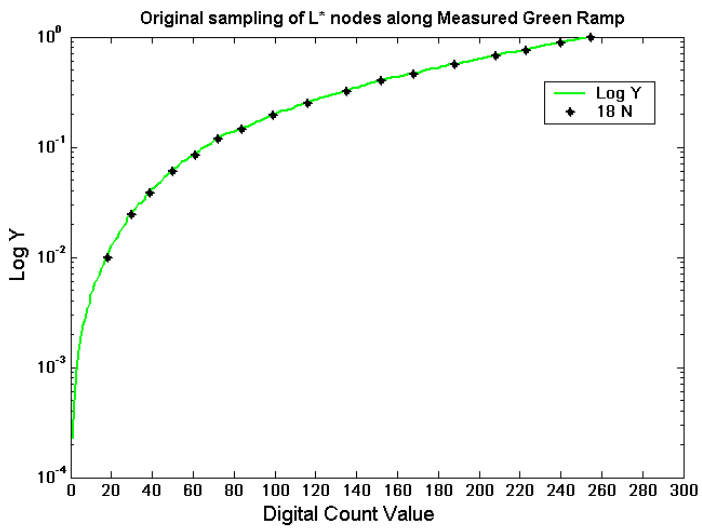


Figure E.8– Log Y vs Digital count with original 18 nodes shown along Green dimension.

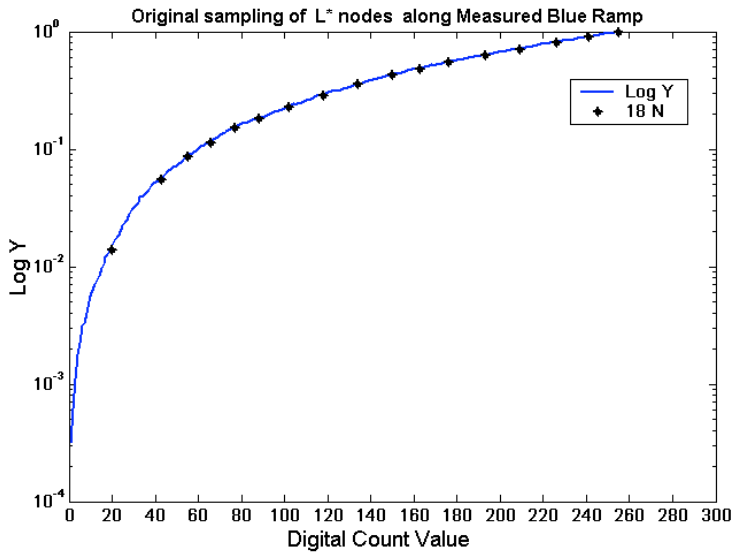


Figure E.9 - Log Y vs Digital count with original 18 nodes shown along the Blue dimension.
(Note that the graphs in Figs. E.7-E.9 do not show the sample point at digital count of 0)

Figures E.7-E.9 show over-sampling at higher digital count values, above 120, and lack of sampling in the region below 30. However, there is sufficient sampling in the bend of the ramp between approximately 30 and 120 digital counts. The final nodes selected are shown in Table E.4. These were chosen by visually judging that all the features of the ramps were fully captured by the sampling.

Table E.4 - Final 18 nodes used for 30" and 37" displays.

Node #	R	G	B
1	0	0	0
2	2	3	4
3	5	6	10
4	12	12	20
5	21	18	43
6	36	30	55
7	47	39	66
8	58	50	77
9	70	72	88
10	82	84	102
11	96	99	118
12	115	116	134
13	146	135	150
14	161	152	163
15	197	168	176
16	225	208	209
17	241	240	241
18	255	255	255

The adjustment of a few sample points considerably improved the prediction results of the CLUT for both the primary ramps and 2000 random colors. To check the accuracy of the node-adjusted 3D LUTS, the primary ramps were pushed through the final CLUT for the display. Results are illustrated in Figures E.10 – E.12.

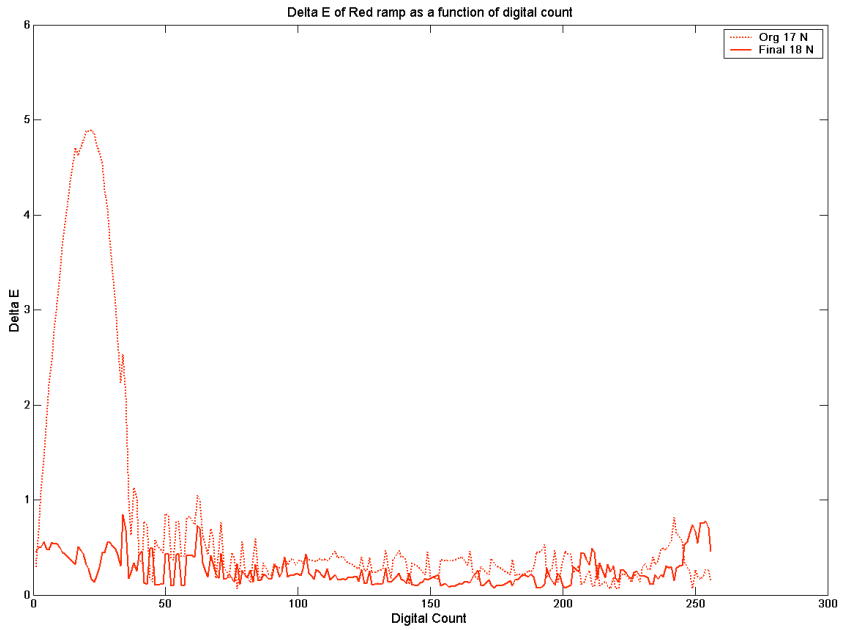


Figure E.10 – ΔE_{00} between measured and predicted Red ramp of 37" display for original 17^3 and final 18^3 3D LUTS.

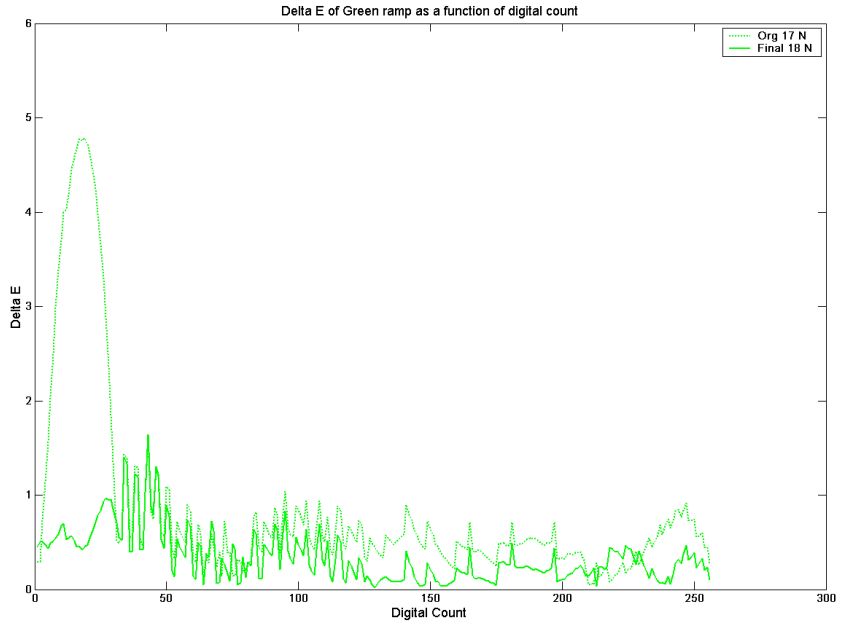


Figure E.11- ΔE_{00} between measured and predicted Green ramp of 37" display for original 17^3 and final 18^3 3D LUTS.

.

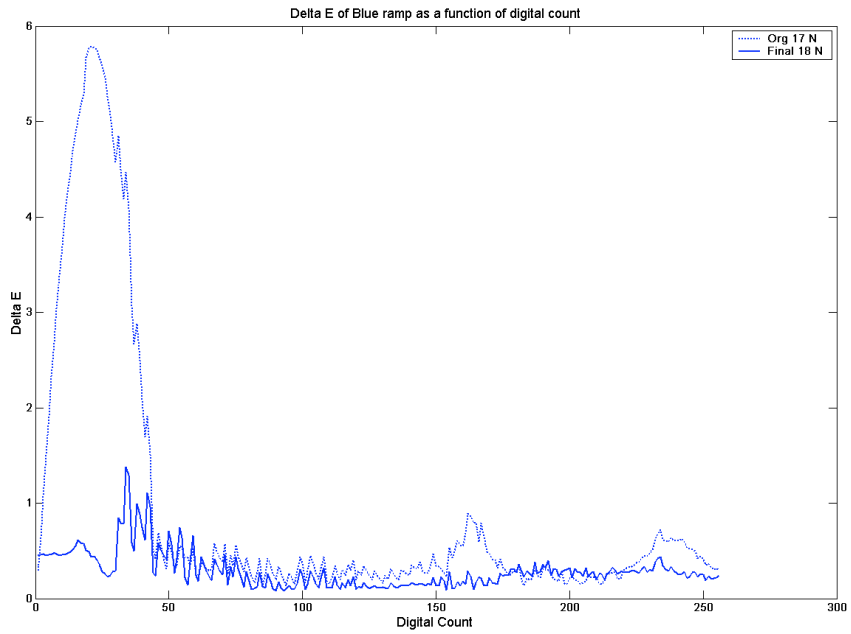


Figure E.12 - ΔE_{00} between measured and predicted Blue ramp of 37" display for orginal 17^3 and final 18^3 3D LUTS.

The histogram in Figure E.13 shows that most of the color predictions for 2000 random colors were under ΔE^*_{00} of 0.50. Table E.6 contains statistics for this approach. Maximum error was $2.06\Delta E^*_{00}$ and average error was $0.45\Delta E^*_{00}$.

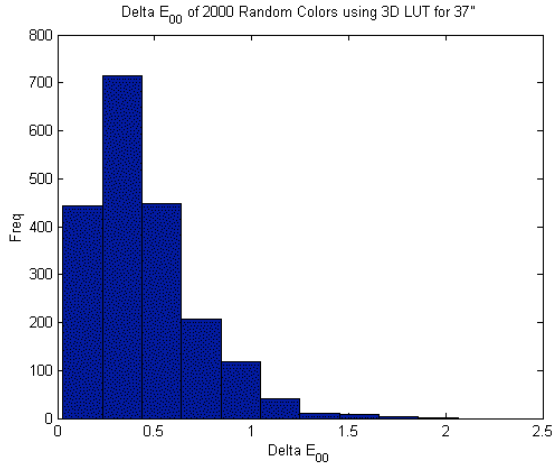


Figure E.13 - Histogram of 2000 random colors measured and predicted by the 37" display.

Experiment

The final method described above was applied to a 30” LCD TV with similar color rendering characteristics to the original 37” display. The same sample points were used as above. A CLUT was formed and then the primary ramps were sent through the CLUT and compared to the measurements. The results are seen in Tables E.5 and E.6. For comparison, Table E.6 includes final results from the 37” display previously characterized.

Table E.5 – ΔE_{00} results from 18 on primary ramps for 30”

	Red	Green	Blue	Avg
Max	0.75	0.72	0.78	0.75
Min	0.02	0.01	0.05	0.03
Avg	0.18	0.21	0.25	0.21
Std Dev	0.14	0.14	0.13	0.14

Table E.6 - Verification Results using 2000 random colors for the 37" and 30" display.

	30"	37"
Max	2.15	2.06
Min	0.02	0.03
Avg	0.40	0.45
Std Dev	0.21	0.28

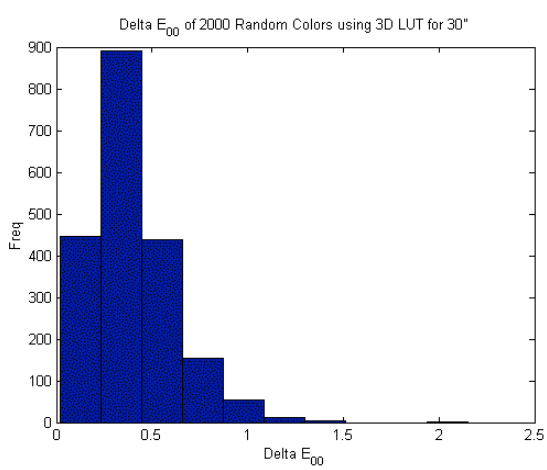


Figure E.14 – Histogram of 2000 random colors measured and predicted by the 30” display.

The histogram in Figure E.14 and the data shown in Table E.6 show that this CLUT provided very acceptable results since the majority of ΔE_{00} values fell below 0.5 for

predicting 2000 random colors. The above method provided good results in terms of accuracy in characterizing the display. However, if the goal is to automate the process and take the fewest number of sample measurements, then the above approach is too burdensome.

It was known that LCD primaries are engineered to mimic the shape of those in CRTs[55, 56]. Since the tone response of LCDs are naturally sigmoidal in shape[55, 56], the gamma function inherent in CRTs is approximated using a piecewise linear fit. This processes introduces discontinuities along each primary ramp at the breakpoint between adjacent linear segments. By taking the derivative of each measured value along a primary ramp it was expected that the breakpoints would show up as local minima or maxima, since derivatives show inflection points quite well.

After measuring primary ramps, the next step was to send the ramps through a low-pass differentiation filter that calculated instantaneous derivatives and then performed a smoothing operation. This process eliminated high frequency noise created through the derivative calculation but still retained important information gained from the derivatives. Afterward, the nodes were chosen as the local minima and maxima along the derivative ramp. Figures E.15-E.17 below show the result of calculating instantaneous derivatives for each ramp and the placement of node values.

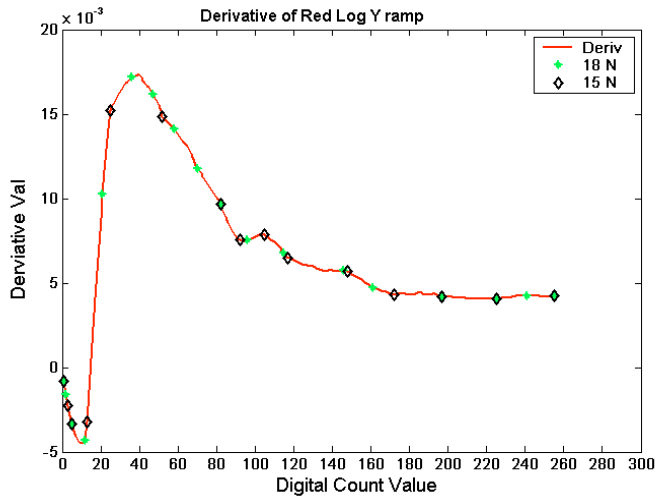


Figure E.15 – Red Derivative ramp showing placement of 15 and 18 node set.

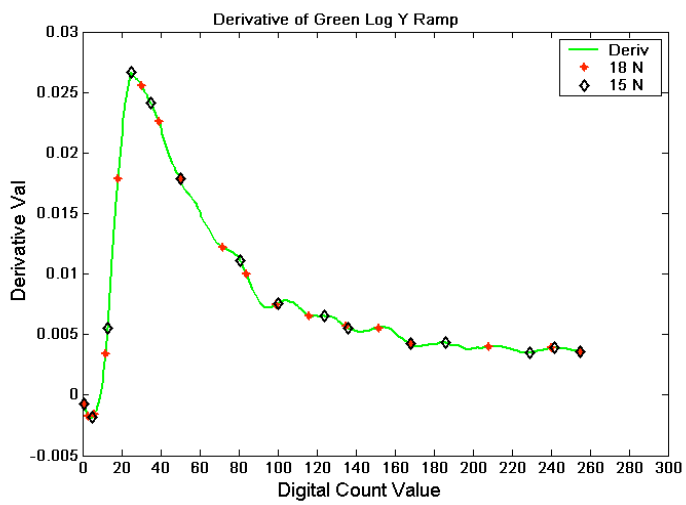


Figure E.16 - Green Derivative ramp showing placement of 15 and 18 node set.

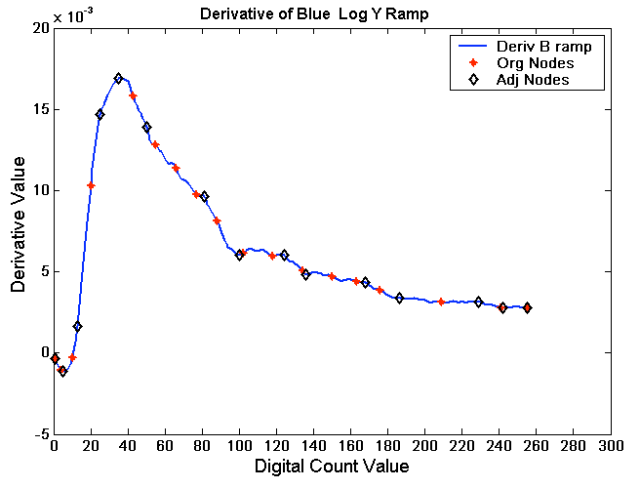


Figure E.17 - Blue Derivative ramp showing placement of 15 and 18 node set.

A 15x15x15 CLUT was built with these sampling points. As seen in Figures E.18-E.20, the results were quite acceptable.

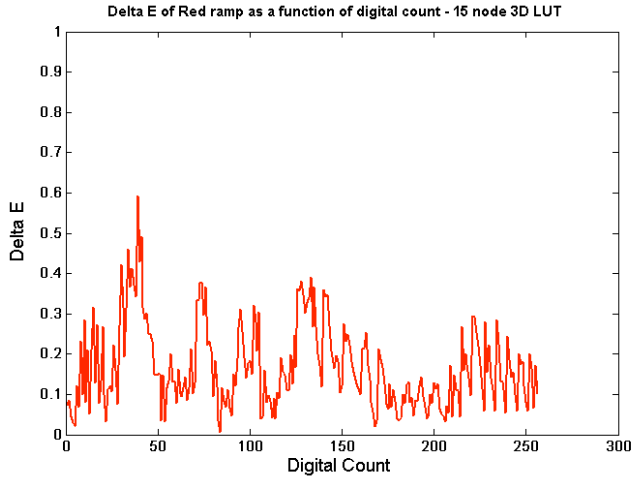


Figure E.18 - Prediction result of red ramp from 15 node 3D LUT

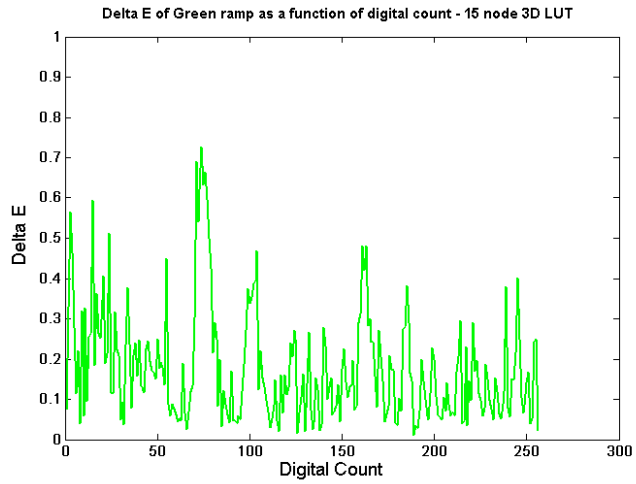


Figure E.19 - Prediction result of green ramp from 15 node 3D LUT

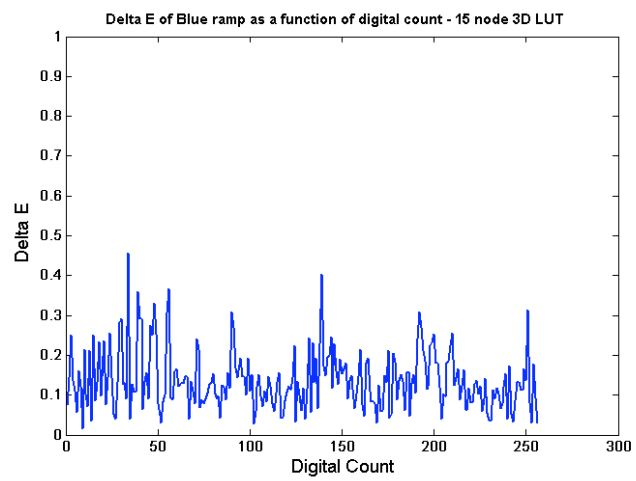


Figure E.20 - Prediction result of blue ramp from 15 node 3D LUT

The importance of the results shown in Tables E.7 and E.8, below, are that the CLUT formed from this latest set had fewer nodes but equal or better results than those from the previous, (Compare with Tables E.5 and E.6).

Table E.7 – ΔE_{00} results from 15 on primary ramps

	Red	Green	Blue	Avg
Max	0.59	0.73	0.46	0.59
Min	0.01	0.01	0.02	0.01
Avg	0.17	0.18	0.14	0.16
Std Dev	0.10	0.14	0.07	0.12

Table E.8 - Results from node-set 4 on 2000 random colors

	30"
Max	1.77
Min	0.02
Avg	0.36
Std Dev	0.20

The results above show that by reducing and carefully placing the measurement points along the ramp, an acceptable CLUT can be derived. Further reducing the number of sample points by approximately cutting them in half to 8 was attempted. This included the first and last points in each ramp (0 and 255), with 6 points selected in between.

Table E.9 – 1st 8 node set

Node #	R	G	B
1	0	0	0
2	12	4	12
3	25	12	25
4	40	25	40
5	94	95	100
6	103	142	137
7	173	170	183
8	255	255	255

The object was to get close to or even better results using 8 nodes in each dimension compared to the above 15 node set. The three derivative-ramps were carefully analyzed and most significant local minima and maxima were selected as seen in Figures E.21-E.23.

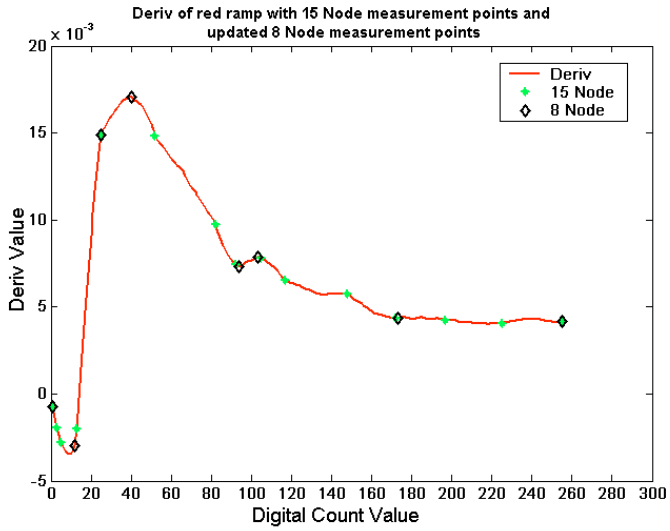


Figure E.21 - Derivative of red ramp showing placement of 15 and 1st 8 node-set points

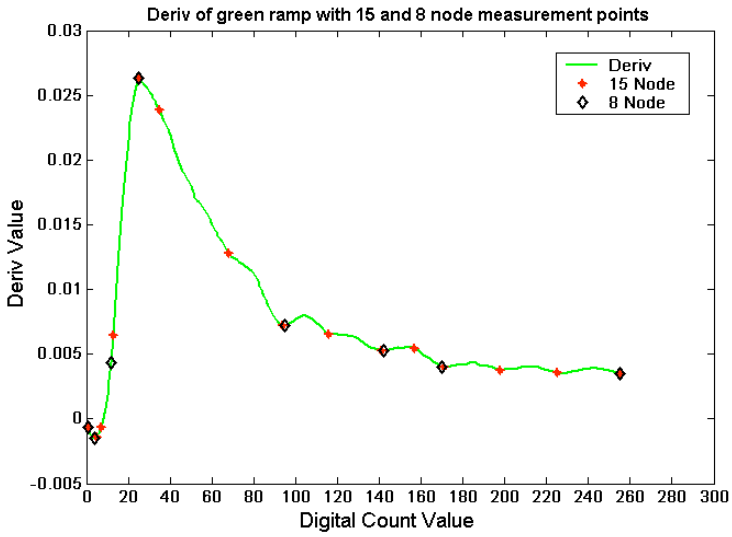


Figure E.22 - Derivative of green ramp showing placement of 15 and 1st 8 node-set points

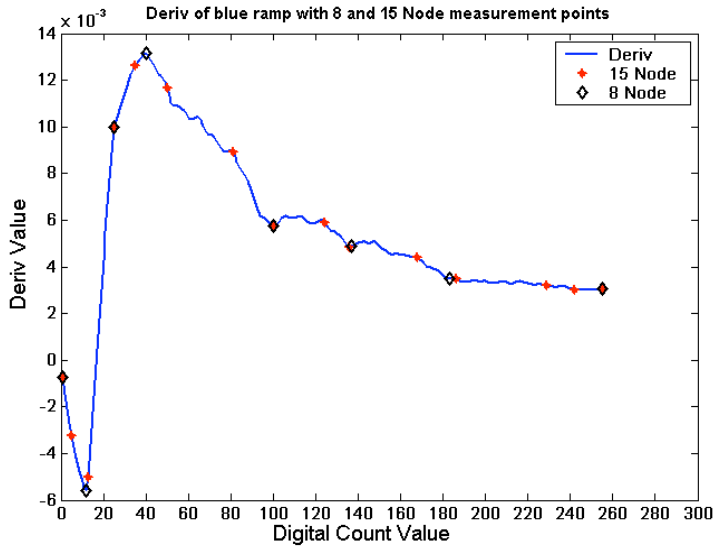


Figure E.23 - Derivative blue ramp showing placement of 15 and 1st 8 node-set points

After choosing the nodes and measuring them, the same analytical tool of sending the primary ramps through the LUTs was used to verify the performance of the CLUT. Results are shown in Tables E.10 and E.11.

Table E.10 – ΔE_{00} results from 1st 8 node set on primary ramps

	Red	Green	Blue
Max	0.62	0.96	0.75
Min	0.02	0.03	0.01
Avg	0.17	0.38	0.23
Std Dev	0.11	0.25	0.21

Table E.11 - Results from 1st set of nodes on 2000 random colors

	30"
Max	3.20
Min	0.01
Avg	0.45
Std Dev	0.28

It was found that there was a peak in ΔE_{00} at the mid-lower digital count region for each ramp as seen in Figures E.24-E.26 below. The node corresponding to the lowest ΔE_{00} was moved to the position on the ramp with highest ΔE_{00}. This was done in each dimension

and again created a new set of nodes shown in Table E.12, which was the final node-set for this experiment.

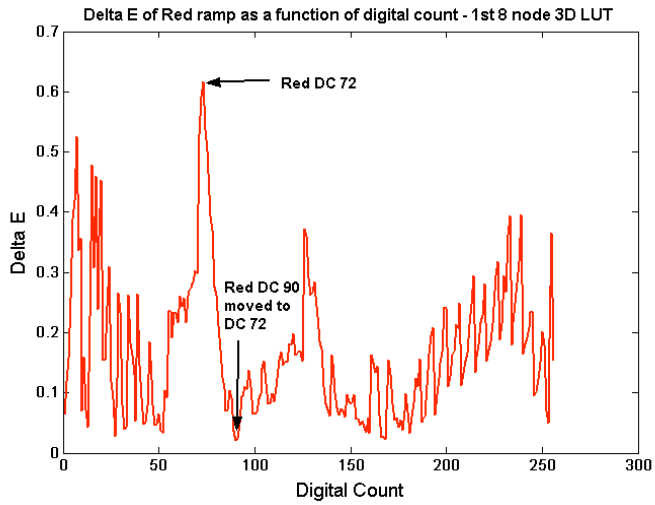


Figure E.24 - ΔE of red primary ramp using 1st set of 8 nodes - arrows point to the digital count values that were switched.

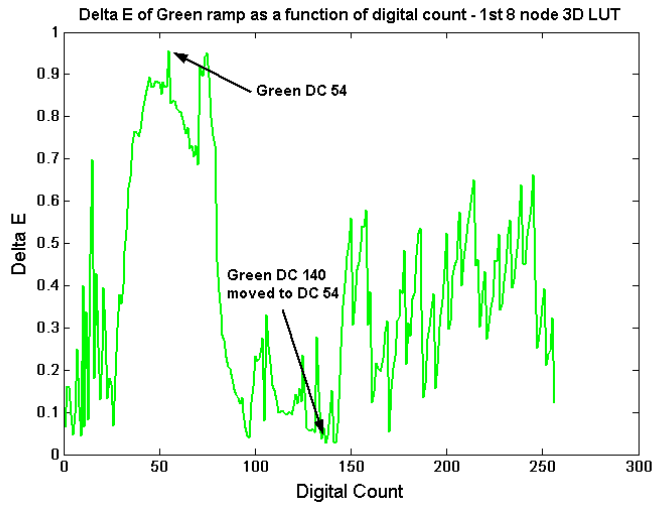


Figure E.25 - ΔE of green primary ramp using 1st set of 8 nodes - arrows point to the digital count values that were switched.

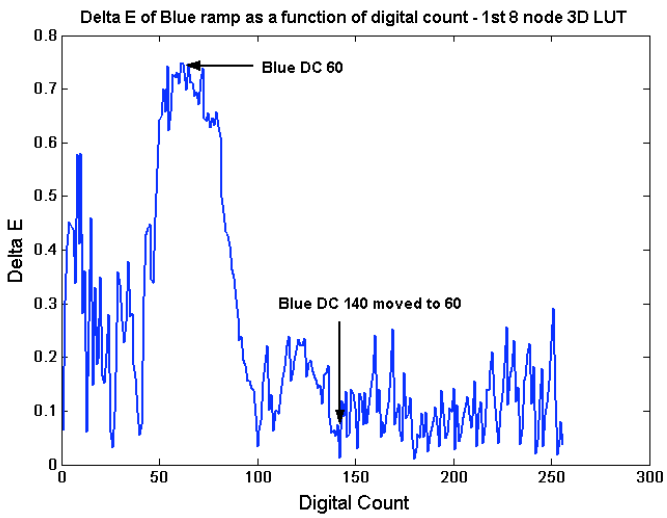


Figure E.26- ΔE of blue primary ramp using 1st set of 8 nodes - arrows point to the digital count values that were switched

Table E.12 – Final 8 node set

Node #	R	G	B
1	0	0	0
2	12	4	12
3	25	12	25
4	40	25	40
5	72	54	60
6	103	95	100
7	173	170	137
8	255	255	255

The ΔE_{00} results of each primary ramp using the final set of 8 nodes were similar to Figures E.18-E.20 and the noisiness of the data showed that the precision reached the measurement precision of the system. The final results, seen in Tables E.13 and E.14 show acceptable precision using only 8 nodes. Although the values in Table E.14 are slightly worse than those in Table E.8, the results are still well within acceptability range while using almost half the nodes.

Table E.13 – ΔE_{00} results from final set of nodes on primary ramps

	Red	Green	Blue
Max	0.56	0.74	0.82
Min	0.02	0.01	0.03
Avg	0.21	0.25	0.29
Std Dev	0.13	0.16	0.13

Table E.14 - Results from final set of nodes on 2000 random colors

	30"
Max	2.97
Min	0.03
Avg	0.53
Std Dev	0.26

Impact

This paper outlines an algorithm for future model development. It is believed that the method can be automated quite easily since calculating derivatives and finding maxima and minima are mathematical procedures. This method will allow an automatic, precise characterization of an LCD device with a minimum number of measurements. In addition, unlike past characterization procedures, this process is not strictly empirical as the primary ramps are analyzed first to find breakpoints and then the empirical method of shifting nodes based on the highest ΔE is accomplished to fine tune node placement.

Building the Inverse 3D LUTs

The inverse 3D LUTs take in a CIELAB value and return the corresponding RGB value. The inversion process is straight forward. Essentially, the procedure is to first run all 16 million RGB combinations through the forward LUTs, thus creating 16 million L*a*b* values, one for every RGB combination. The result of this step can be thought of as creating the gamut of a particular display. Next the CIELAB values are encoded; this process transforms the L* from 0-100 to 1-256 by multiplying each L* value by 2.55 and the a*b* values have 128 added to them, with the presumption that the original values in these dimension vary between -128 and 127. Essentially, this CIELAB encoding creates index coefficients that hold the RGB values. Now that each CIELAB value is in the same format as the RGB data the next step is to assign the RGB values to the CIELAB positions. This is the inverse LUT and to use this LUT one first encodes the LAB value,

as done before, and then sends this value through the inverse LUT. What gets returned is an index number that corresponds to an RGB value. This works quite well as seen in the table and figure below. Note that the results seen below are from the same 2000 colors displayed and measured on both displays.

Table E.15 - Statistics for ΔE_{00} between original RGB and RGB from Inverse 3D LUTs.

	30"	37"
Max	1.28	2.6
Min	0	0
Avg	0.28	0.30
Std Dev	0.19	0.19

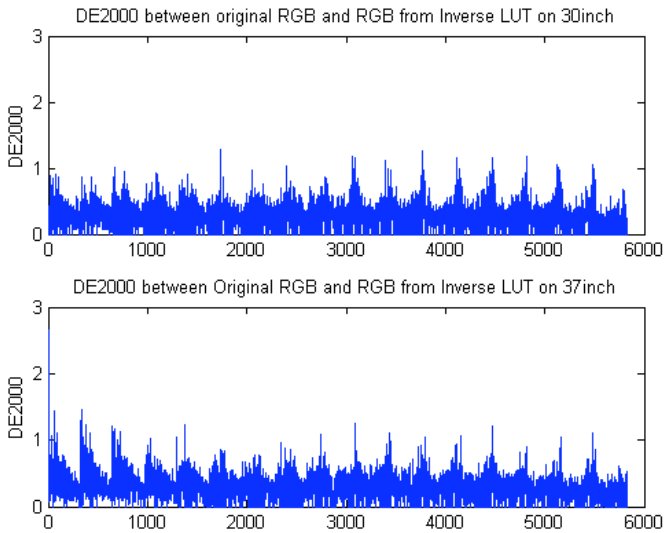


Figure E.27 – ΔE_{00} histograms of original RGB and RGBs from Inverse LUTs for both displays.

To obtain the results above, the same set of RGB values were displayed and tristimulus values measured for both displays. CIELAB values were calculated from the measured tristimulus values and then sent through the inverted 3D LUTs to get RGB values. The two sets of RGB values (for each display) were then sent through the forward 3d LUTs and the resulting CIELAB values compared using ΔE_{00}.

There were some problems using this simple method of pushing all RGB values through the forward LUTs. While the RGB values are discreet, at lower digital counts

one step in an RGB unit does not translate to one step in CIELAB. This leaves “holes” inside the gamut where multiple RGB values are mapped to a single CIELAB value. The process to fix these quantization errors inside the inverse LUT and account for values that might fall out of the gamut, (for this portion the inverse LUT will be referred to as a gamut), included 3 main stages: create a mask, replace holes one slice at a time and expand the gamut edge to fill the entire color space. The first step in stage one was to create a mask of the gamut where there were 1s everywhere except for where there were no values, i.e. holes. Note that this will also apply outside the gamut, (see Figure E.28 (23)). Next a copy of this gamut-mask was made and then all holes were filled using two image processing/morphological processes, (see Figure E.29 (24)).

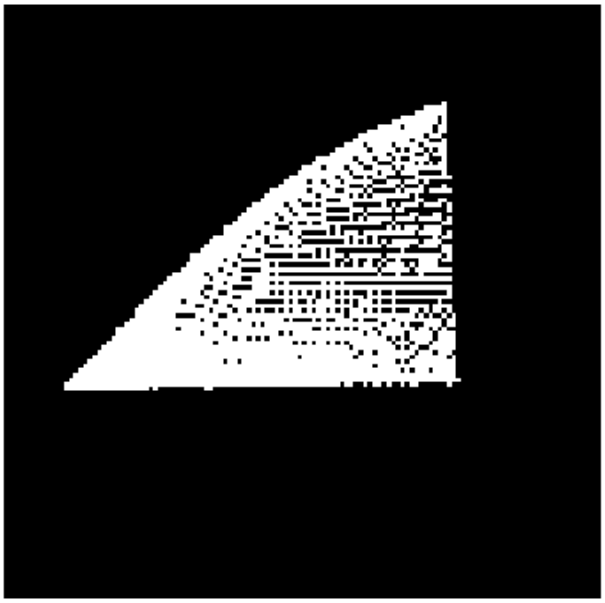


Figure E.28 – Gamut Mask at Slice 50 out 256 looking down L* axis at all a* and b* combinations. Original Mask showing all the holes at this particular slice.

Figure E.29 – Gamut Mask at Slice 50 (out of 256) looking down L* axis at all a* and b* combinations. Holes filled.

The next stage was to step downward, in the L* dimension, through the gamut one slice at a time beginning at slice number 255, (out of 256 slices). Each step looked for holes inside the slice by comparing the mask with holes, Fig. E.28 (23), to the mask without holes, Fig. E.29 (24); if found the holes were replaced with values at the hole location from the previous slice in the gamut. Figure E.30 (25) shows a slice out of the gamut which had many holes and Figure E.31 (26) shows those values filled. Next the values inserted into the current slice were smoothed using a smoothing filter with the 9 values around the pixel location. This smoothed value was then assigned to the actual gamut location where the hole had previously existed. This method only affected the values at the holes and did not affect any other values inside the gamut.

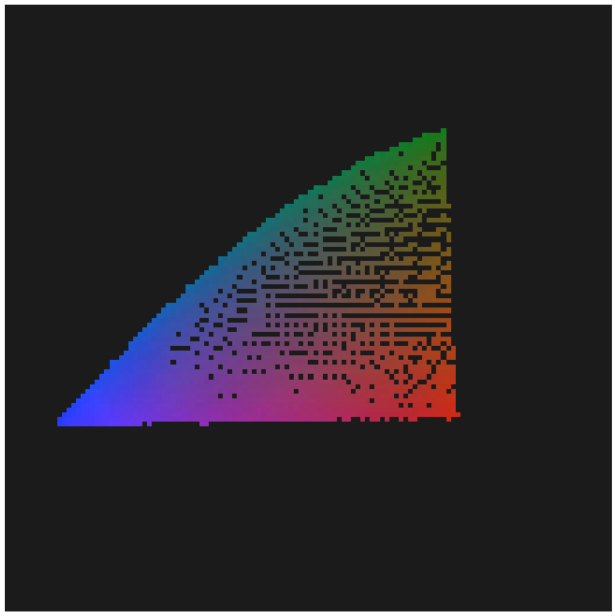


Figure E.30 – Original slice 50 from gamut.. Showing the holes at same location as gamut mask in fig 23.

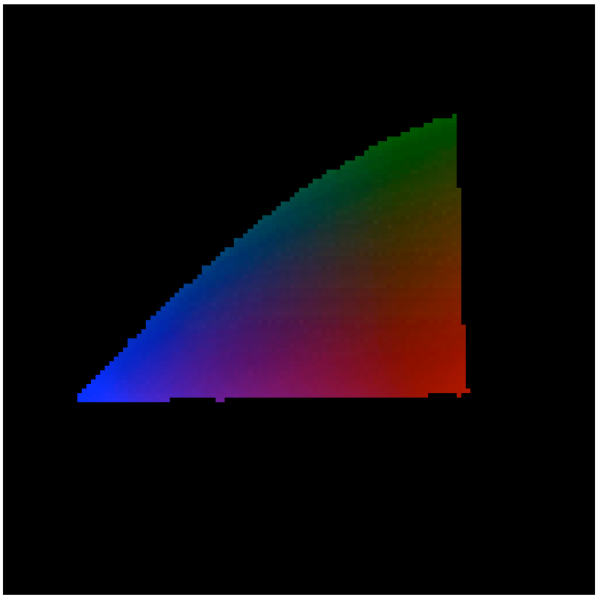


Figure E.31 – Gamut slice 50 with holes filled. Some values were much brighter than all other values in the slice.

In order to fix any edge effects due to the smoothing filter using values outside the gamut, the smoothed edge, shown in Fig. E.32 (27), was replaced with the original edge, shown in Fig. E.31 (26). The resulting slice that now exists in the gamut is shown in

Figure E.33 (28). Each of the above procedures was done for each of the 256 slices as was needed.

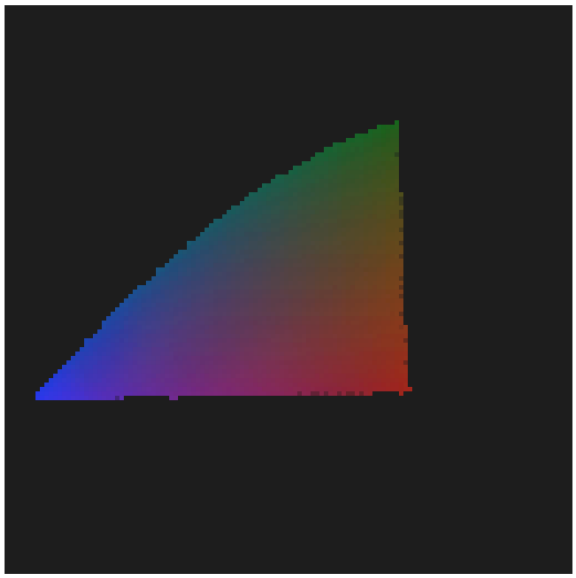


Figure E.32 – Gamut slice 50 with smoothed values at hole locations, notice smoothed edge as well.

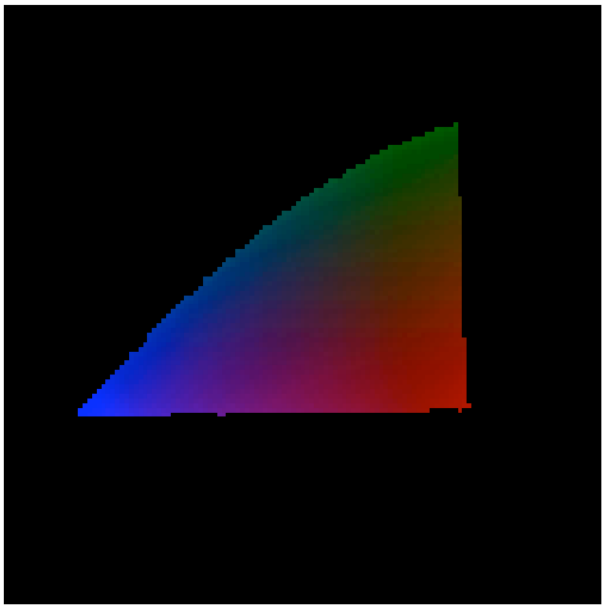


Figure E.33 – Gamut slice 50 with edge replaced.

These few steps fixed the quantization errors inside the gamut. The final stage was to handle any values that might fall outside the gamut. This ensures that the inverse LUT returns a value for any possible input. Instead of gamut-mapping a particular RGB value

that falls outside the gamut, the algorithm simply "grew" the gamut edge from the centroid value of L* = 50, a*=b*=0 to fill the entire 256^3 space. Now the inverse LUT has values at all locations in the 256^3 space, as seen in Figure E.34 (29), and is just a look-up-table.

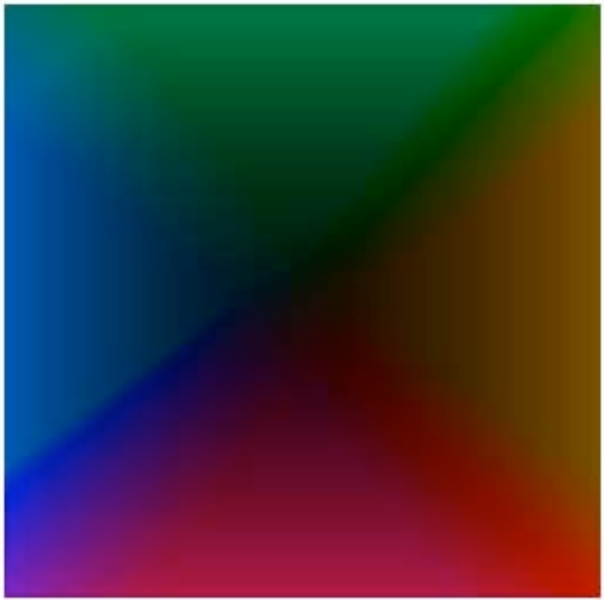


Figure E.34 – Gamut "mapped" at slice 50.

This inversion process provides accurate RGB values, (shown in Table E.15 and Figure E.27), quickly because it is essentially a large look-up-table. There was considerable time invested in the building and "touching up" of the inverted 3D LUTs but the results prove that the methods warranted the time.

If it were decided to continue with this research of building an efficient model the next step in this process would be to automate the determination of where to reasonably place measurement points along each ramp and measure the smallest possible points for the creation of the forward 3D LUTS. One possibility would be to examine the TRCs of each ramp in detail and look for discontinuities or "bends" along the curve. One method o automate this process would be to look at the derivatives of the TRCs. Afterwards, the

following step would be to research the smallest possible sub-sample of RGB values in which to build the inverse LUT. A clever interpolation scheme would be needed, along with a program to fix any quantization errors in the gamut and lastly a fast and efficient gamut mapping algorithm would be needed. It is certainly plausible that these steps could be automated and developed using a much smaller set of measurements than what was used in this report.