

Human-centered display design : balancing technology & perception

Citation for published version (APA):

Murdoch, M. J. (2013). *Human-centered display design : balancing technology & perception*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Industrial Engineering and Innovation Sciences]. Technische Universiteit Eindhoven. <https://doi.org/10.6100/IR760469>

DOI:

[10.6100/IR760469](https://doi.org/10.6100/IR760469)

Document status and date:

Published: 01/01/2013

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

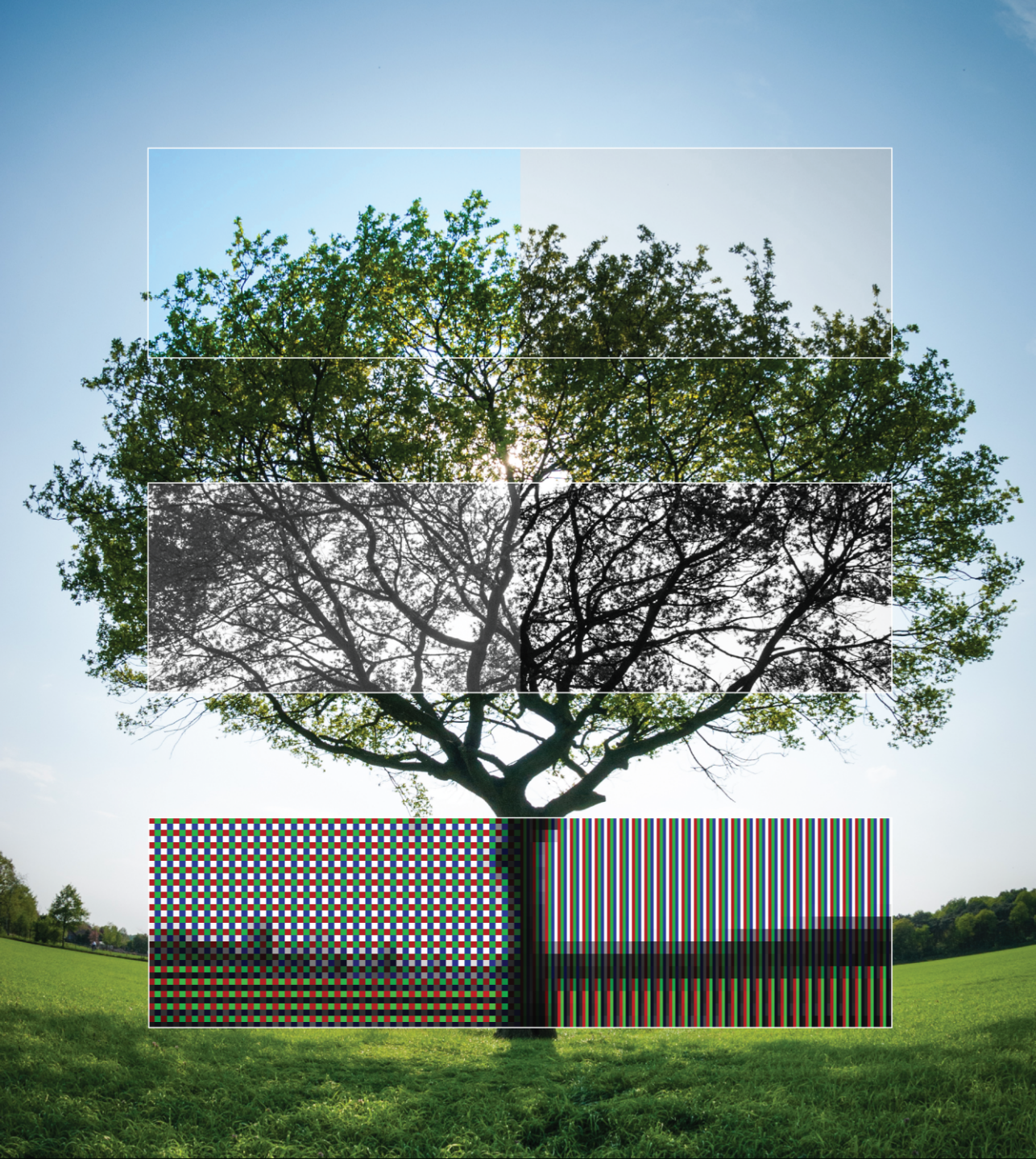
www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.



Human-Centered Display Design

Balancing Technology & Perception

Michael J. Murdoch

Human-Centered Display Design

Balancing Technology & Perception

Michael John Murdoch

Cover photo and design by the author. Overlays on the image illustrate in an exaggerated fashion some of the display image processing concepts discussed in the thesis, notably variations in colorfulness (top, related to Chapters 5 and 6), changes in black-level shown in monochrome with the sun coming through as a glare source (middle, referring to Chapter 2), and RGBW quad and stripe pixel patterns (bottom, as discussed in Chapter 3). The original photograph was made in Waalre, The Netherlands, in 2011.

© 2013 Michael John Murdoch

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the copyright holder.

A catalogue record is available from the Eindhoven University of Technology Library
ISBN: 978-90-386-3484-5

Keywords: visual perception, image quality, display, LCD, OLED, RGBW, color gamut, high dynamic range, HDR

This version was prepared for digital distribution via PDF

Human-Centered Display Design

Balancing Technology & Perception

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Eindhoven,
op gezag van de rector magnificus prof.dr.ir. C.J. van Duijn,
voor een commissie aangewezen door het College voor Promoties,
in het openbaar te verdedigen
op dinsdag 19 november 2013 om 16:00 uur

door

Michael John Murdoch

geboren te Illinois, Verenigde Staten van Amerika

Dit proefschrift is goedgekeurd door de promotiecommissie:

voorzitter: prof.dr. A.G.L. Romme
1^e promotor: prof.dr. I. Heynderickx
2^e promotor: prof.dr. W.A. IJsselsteijn
leden: prof. M.D. Fairchild PhD (Rochester Institute of Technology)
 prof.dr. H. de Ridder (TUD)
 prof.dr.ir. G. de Haan
 prof.dr.ir. J.B.O.S. Martens

Human-Centered Display Design

Balancing Technology & Perception

DISSERTATION

to receive the degree of Doctor
at the Eindhoven University of Technology,
under the authority of the Rector Magnificus prof.dr.ir. C.J. van Duijn,
before a committee designated by the Doctorate Board,
to publicly defend
on Tuesday November 19 2013 at 16:00

by

Michael John Murdoch

born in Illinois, United States of America



*for Dina
& her baby brother Leo*

Contents

1	Introduction: Technology & Perception	1
1.1	Visual Perception	2
1.2	Display Technology	10
1.3	Relating Perception to Technology	17
1.4	Outline of this Thesis	24
1.5	Bibliography	25
2	Perceived Contrast in High Dynamic Range Displays	29
2.1	Introduction	29
2.2	Models	31
2.3	Laboratory Setup	33
2.4	Experiment	36
2.5	Experimental Results	40
2.6	Modeling Results	44
2.7	Conclusions	47
2.8	Bibliography	47
3	Utilizing Metamerism in RGBW OLED Displays	49
3.1	Introduction	49
3.2	Additive Model of RGBW	50
3.3	White Equivalence	51
3.4	White Replacement	52
3.5	When W Is Not Quite White	54
3.6	When W Is Not White at All (RGBX)	55
3.7	Algorithm Summary	55
3.8	Subpixel Arrangements	56
3.9	Visual Impressions	57
3.10	Discussion	58
3.11	Bibliography	60
4	Balancing Color and Power in RGBW OLED Displays	63
4.1	Introduction	63

4.2	Image-Processing Options	65
4.3	Image Quality Effects	70
4.4	Discussion	75
4.5	Conclusions	76
4.6	Acknowledgments	78
4.7	Bibliography	78
5	Preferred Color Gamut for Reproducing Natural Image Content	79
5.1	Introduction	79
5.2	Experiment 1	81
5.3	Experiment 2	86
5.4	Conclusions	93
5.5	Bibliography	95
6	Preferred Color Gamut Boundaries for Wide-Gamut and Multi-Primary Displays	97
6.1	Introduction	97
6.2	Methodology	99
6.3	Results	106
6.4	Discussion	110
6.5	Conclusions	115
6.6	Acknowledgements	116
6.7	Bibliography	116
7	Discussion	119
7.1	High Dynamic Range LCDs	119
7.2	RGBW OLED Displays	121
7.3	Wide Gamut Displays	122
7.4	Comments, Connections, & Extrapolations	124
7.5	Concluding Remarks	130
7.6	Bibliography	131
	Summary	133
	Acknowledgements	137
	List of Publications	139
	Curriculum Vitae	143

Introduction: Technology & Perception

We are increasingly surrounded by displays: televisions, movie screens, computer monitors, cameras, smart phones, tablets, GPS devices, electronic billboards, digital signage, dashboards, control panels, and more. Whether for clear, efficient information communication or for relaxing, immersive entertainment, displays require visual quality in addition to clarity and robustness. While the development of a multitude of display types has taken decades, the current pace of improvement in all aspects, including image quality, size, price, operating cost, speed, and interactivity, is truly remarkable. Human-centered improvements to displays are the most important of these, focusing on the experience of the human end user, and thus human-centered display design is the topic of this thesis.

Research in the area of display design includes two major themes: visual perception and display technology. The former is primarily interested in understanding and describing the human visual system, and secondarily in applying models of this understanding to display design. The latter, display technology, is driven forward with time by innovations in materials and electronics, sometimes seemingly without regard to the needs of any human viewer. The aim of this thesis is to show that these two areas of focus cannot be independent, and that a mutual understanding of both themes and the relationship between them enables display design that is more oriented toward the needs of the viewer, and therefore better.

Display design already necessarily leans firmly on the basics of visual perception. A keystone of color reproduction is metamerism, the phenomenon by which two color stimuli of vastly different spectral content can appear identical to the human eye, which has been used in every color imaging system in history. The lucky correlation between the power-law intensity response of CRT displays and the human perception of brightness persists in standard image encodings even as the CRT has become obsolete. Persistence of vision, in which a pulsating stimulus (for example a television flickering along at 60 Hertz) is perceived as temporally constant, is another visual phenomenon that has been consistently utilized by “moving pictures.”

However, there are plenty of examples of technological developments (and marketing statements promoting them) that do not credibly enhance visual perception. For example, the push from standard-definition to high-definition television (HDTV) provided

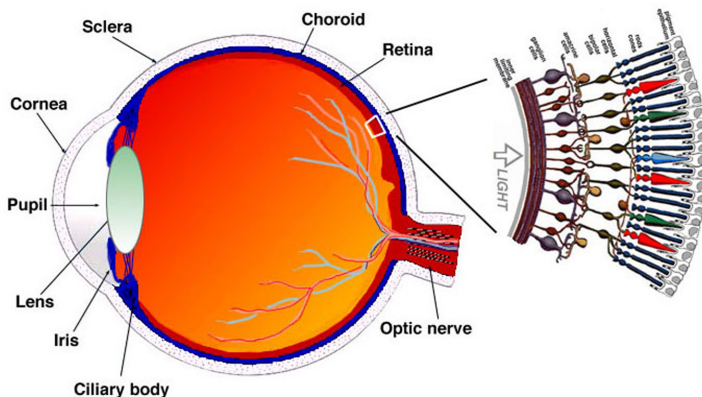


Figure 1.1: Cross-section of the human eye. At left, major components of the human eye are labeled. At right, a close up of the retina shows the layers of neurons through which light passes before reaching the light sensitive rod and cone cells. Image used with permission from webvision¹.

a welcome improvement in sharpness thanks to the added resolution, but further growth in pixel count to UHDTV and beyond may never be seen as an improvement by anyone in a normal-sized living room. Physical contrast in LCDs has indeed slowly improved with time, but advertised contrast ratios and especially viewing angle ranges have inflated somehow even further to the point that they have become meaningless.

In the coming chapters, a series of examples of display technology advancements approached with a user-centered understanding of human perception are presented. This introductory chapter aims to provide the reader with essential background knowledge on the topics of visual perception (Section 1.1), current and developing display technology (Section 1.2), and the tools and techniques of assessing and modeling visual image quality (Section 1.3). Relevant references are provided to guide the reader to further detail on the many underlying topics.

1.1 Visual Perception

The human visual system (HVS) is able to interpret many characteristics of the light which reaches it from the world around us. The most basic is intensity, in physical terms power or luminance, which we see as brightness. Spatially-varying intensity yields contrast and allows us to see lines, shapes, and forms. Spectrally-varying intensity gives us discernible colors. These concepts and more are detailed in Wandell's excellent text [1]; here, only very brief highlights are included.

Figure 1.1 shows a cross-section of the human eye, the instrument through which we see the world. In this diagram, light enters from the left, passing through the cornea, pupil, lens, and the fluid-filled eyeball itself (vitreous humor). Finally light reaches the retina, which is densely populated with photosensitive cells that send signals on to the visual cortex for further interpretation. Rod cells, highly sensitive to a wide band of the color spectrum, are the most numerous, making up 10^8 of the retina's cells. The three types of cone cells are much less sensitive than rods, but their different spectral sensitivity enables color discrimination when enough light is available. There are about 5×10^6 cones in the retina, of which about 6% are S cones, sensitive to short-wavelength light, and the remainder are split between M (middle-wavelength) and L (long-wavelength) cones in a ratio varying from 1:16 to greater than 1:1 in different people [2]. Cones are concentrated near the center of the visual field in the region called the fovea, where we see with greatest detail, while rods are more numerous in the surrounding retina. Note that a final type of photosensitivity is provided by melanopsin receptors in a small fraction of the retina's ganglion cells. This contributes to the body's regulation of circadian rhythm, but because such sensitivity is non-visual, it lies outside the scope of this thesis.

Adaptation is an extremely important and complicated phenomenon which occurs both in the retinal cells and in the visual cortex. Temporal luminance adaption to different levels of illumination is what enables "night vision," and is why going quickly from a dark to a bright environment can be temporarily blinding. Because of luminance adaptation, there is no fixed translation between luminance (physical intensity) and brightness (perceived intensity): brightness always depends on the intensity of surroundings and the recent temporal intensity history.

Chromatic adaptation is due to changes in the relative sensitivity in the cones and the cortical interpretation of them, and is the reason why a spectrally-white object appears white both in dim indoor light (warm, yellowish in color) and in outdoor shade (cool, bluish in color). Perceived color also is heavily dependent on the colors of the surroundings and the retina's recent temporal history. The subject of color appearance and modeling thereof is addressed in detail by Fairchild [3].

1.1.1 Luminance, Contrast, & Glare

Luminance is the perceptually-weighted physical intensity of light coming from an object to a point in space, for instance a human's eye. Luminance is quantified in physical units of candela per square meter, and thus is easily measured; however, the perception of luminance is of course much trickier to pin down due to temporal and spatial effects and adaptation within the HVS. Visual perception actually depends more on luminance differences, or contrast, as explained in Peter Barten's book [4].

For controlled, periodic luminance variations, specifically sine waves spatially varying in luminance, the sensitivity of the HVS may be modeled as a function of average

¹ <http://www.webvision.med.utah.edu>

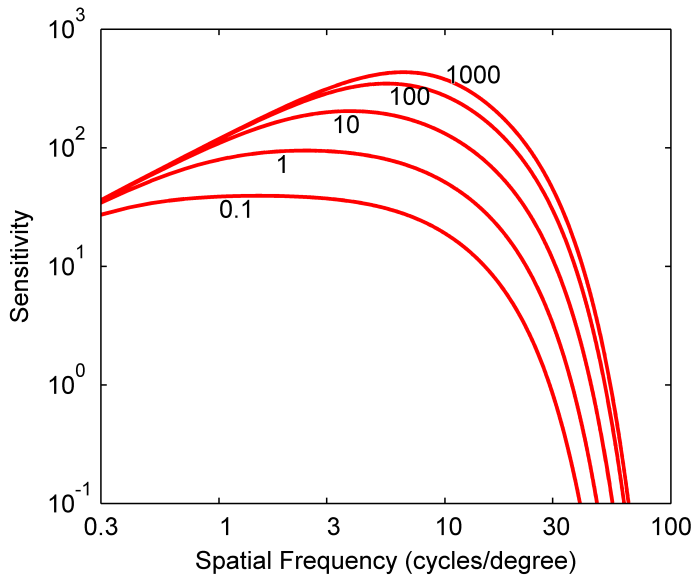


Figure 1.2: Contrast Sensitivity Function. Human sensitivity versus spatial frequency for different levels of luminance (labeled, in cd/m^2), according to Barten's empirical fit of multiple experimental data sets. As the luminance of the stimulus increases, sensitivity is higher and peaks at higher frequencies. In all cases, the maximum perceivable frequency is around 50 cycles per degree.

luminance, amplitude of the sine wave modulation, and spatial frequency (in degrees of visual angle). Further, this model can be explained via signal-to-noise concepts that describe the behavior of the HVS [5]. The generalized description of contrast sensitivity, which is defined as the inverse of threshold modulation amplitude, is illustrated in the contrast sensitivity function (CSF) in Figure 1.2; these curves are an empirical fit to multiple contrast sensitivity data sets as explained by Barten in 2004 [6]. The dependence on spatial (angular) frequency is very strong, as shown, with a peak between 3 and 10 cycles per degree, and diminishing substantially around 50 cycles per degree. This model of visibility is used and extended in Chapter 2 of this thesis.

An important aspect of human vision, especially at the limits of visibility, is glare. Because the human eye is an imperfect optical system which scatters and absorbs a good fraction of the light entering it, the image projected by the eye's lens on the retina is blurred to some extent. This intra-ocular effect is known as glare, or veiling glare, and can be thought of conceptually as the eye's point-spread function. Glare and its visual effects were characterized in detail by Holladay [7] and Moon & Spencer [8] in the

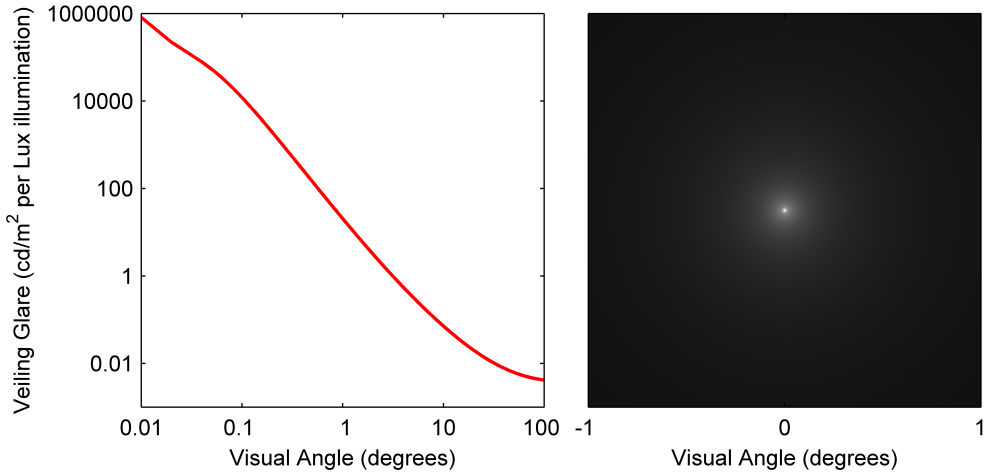


Figure 1.3: CIE total glare equation. Veiling glare as specified by the CIE total glare equation shown in two ways. At left, a log-log plot of veiling luminance versus visual angle, in cd/m^2 per lux illumination on the eye; at right, an image illustrating the circularly-symmetric glare as a point-spread function that is conceptually the visual spread of a pinpoint of bright light onto the retina.

first part of the twentieth century, and the basic shape of this effect is a steeply-peaked, circularly-symmetric falloff with a long tail, roughly proportional to $1/\theta^2$. A standardized version of a glare equation was published by the CIE in 1999 [9], which is shown in Figure 1.3 along with an image interpreting the equation as a point-spread function. Such a point-spread function is the somewhat-blurred spot the eye sees when looking at a very small point of light, the size and blurriness of which is proportional to the amount of light irradiating the eye. Glare is caused by bright light sources, but also by everything in the visual field – the brighter the object, the more glare it contributes – and the visual result is always reduced perceived luminance contrast, lower color saturation, and reduced visibility of details.

1.1.2 Color, Metamerism, & Additive Color Displays

The perception of color is best explained with a clear example; for the sake of this thesis, an obvious example is a color display, such as a television. All additive color displays take advantage of metamerism, the visual phenomenon by which two color stimuli with different spectral power content may be perceived as the same color. A color display generally has three color primaries, for example red, green, and blue, which can be modulated in intensity to create images. Thanks to metamerism, the HVS sees not the



Figure 1.4: Primary and composite images. Example showing the intensity images of (left-to-right) the red, green, and blue color channels along with the resulting full-color image (far right). The relative intensities of these primaries may be seen in different portions of the image. For example, the white stockings are high intensity in all three color channels. The pink dress consists of high-intensity red, low-intensity green, and mid-intensity blue, while the yellow-green grass consists of mid-intensity red and green and low-intensity blue.

constituent colors of red, green, and blue, but rather their visual sum: for example, a combination of nearly-equal amounts of red and green is seen as yellow, despite the absence of an actual yellow primary. In the same way, a wide range of colors is made available from a few (three, minimum) basic components.

The prevalence of consumer digital cameras and accessible computer graphics programs such as Adobe Photoshop means that many people understand that a given color may be a composition of red, green, and blue. In photos, in painting programs, and in text markup such as in HTML, specifying colors as RGB triads is quite common. This is due entirely to the additive color system that is a modern display. A display is additive because it starts with a black background and adds different amounts of intensity from its primary colors of red, green, and blue — note this is the opposite of paint or ink on paper, which is a subtractive color system starting with a white page and reducing the reflectivity selectively with colorants. In an additive RGB color system, combinations of the primaries always result in brighter, less pure colors; for example, red and green sum to yellow. A visual example of this is shown in Figure 1.4, where a full-color image is decomposed into its component primary intensity images, and the relative intensities of the primaries with respect to the final colors can be seen. The concept in this image is that each intensity image, shown in black-and-white, is the spatial-varying factor to be multiplied by its respective primary color.

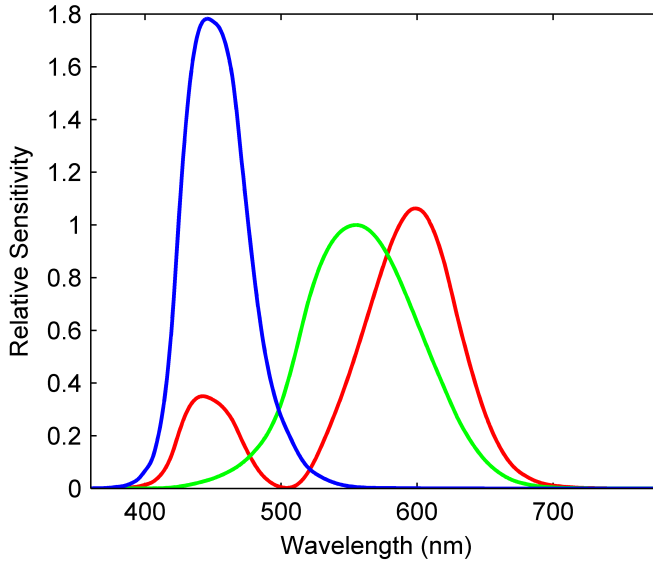


Figure 1.5: CIE 1931 standard colorimetric observer. The CIE 2-degree color matching functions \bar{x} , \bar{y} , and \bar{z} , shown as red, green, and blue, respectively. Note that the color matching functions do not correspond directly to these colors.

Colorimetry

At a more detailed level, it is important to realize that the color primaries are not abstractly red, green, and blue. In order for a display to perform predictably (read: accurately), they necessarily represent very specific examples of these colors, which must be unambiguously described and objectively measurable. Colorimetry, standardized over the years by the Commission Internationale de l'Éclairage² (CIE), provides a framework for communicating about color. A basic way of looking at color is through a “standard observer,” which has a theoretical color sensitivity based on aggregated measurements of real people. The CIE 1931 standard observer [10] is defined as a set of three wavelength-by-wavelength sensitivity curves \bar{x} , \bar{y} , and \bar{z} (plotted in Figure 1.5), which are used as weighting functions to numerically integrate a color stimulus (truly, a spectral power distribution). The result is a triad known as XYZ tristimulus values, which is a colorimetric description of the stimulus.

Two spectral power distributions which integrate to the same XYZ triad are by definition metamers. A graphical example of metamerism that is common in display

²In English, International Commission on Illumination: <http://www.cie.co.at>

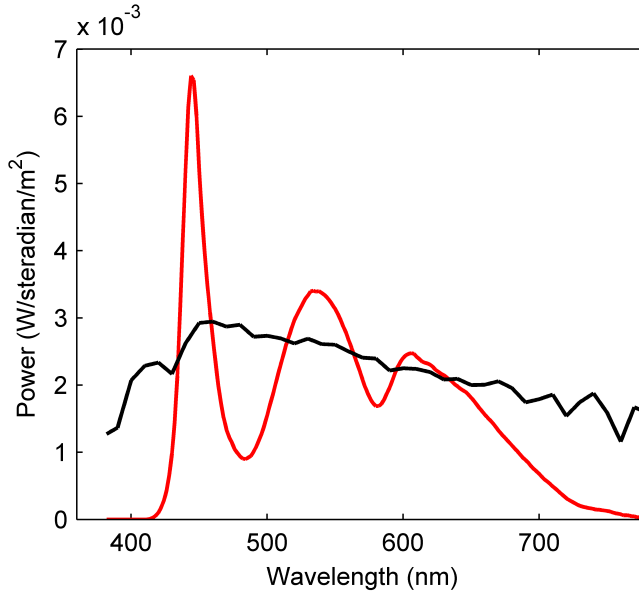


Figure 1.6: Metamerism example. A graphical example of metameric spectra: in black, the spectral power distribution of a standard D65 illuminant, and in red, the spectral power distribution of the sum of the red, green, and blue display primaries at appropriate intensity levels so that the result is the same XYZ tristimulus values as those of D65.

applications is shown in Figure 1.6. In this plot, the spectral power distribution of a standard D65 illuminant is shown along with the spectral power distribution of a display showing a color with the same XYZ tristimulus values as those of D65. Each curve is numerically integrated with the \bar{x} , \bar{y} , and \bar{z} sensitivity functions, resulting in identical XYZ tristimulus values – meaning, they are identical visual stimuli. Note that the display's spectrum is the sum of its red, green, and blue primaries; because the computation is linearly additive, it is equivalent to sum the power spectra before computing XYZ or to compute XYZ from each spectrum separately before summing the XYZs.

XYZ tristimulus values generally have photometric units, meaning they refer to the absolute visual intensity of the spectral power distribution they describe. The Y value alone represents luminance, which takes units such as candela per square meter, and correlates well (though non-linearly) with the HVS' perception of brightness. XYZ values may be scaled relative to a reference white, for example the luminance of maximum white in a display, in which case they are known as XYZ tristimulus factors (or percent tristimulus factors, if necessary). XYZ may be transformed to other quantities

for convenience, such as chromaticity spaces, which describe the chromatic composition of the color stimulus while ignoring the absolute intensity, or luminance. Note that this dimensional reduction provides convenience but also is potentially misleading! The prototypical chromaticity space is the CIE 1931 xy space, which is still used despite being superseded by the more perceptually accurate CIE 1976 uniform chromaticity scales, $u'v'$, the best choice for emissive displays and lighting. Chromaticity coordinates are typically shown in a “top-view” diagram, an example of which appears in Figure 1.7; chromaticity gamut boundaries shown on such diagrams are informative, but because they ignore the luminance dimension they should always be interpreted with caution.

Another commonly-used colorimetric description is the CIE 1976 $L^*a^*b^*$ color space, often abbreviated as CIELAB [11]. CIELAB was designed to mimic human visual perception, accounting for the nonlinear response to luminance in the achromatic L^* channel (Lightness), and putting colors in opponent channels a^* (reddish-greenish) and b^* (yellowish-bluish) which relate to the neural opponent process which is part of the HVS’ interpretation of colors. $L^*a^*b^*$ are computed from XYZ via a set of nonlinear formulae, always with respect to a “white reference” which should be the adaptation white for the viewer looking at the stimulus (in display applications, this is typically the display’s maximum white point). While $L^*a^*b^*$ provides Euclidean coordinates in the three-dimensional space, it is also common to use cylindrical coordinates, L^*C^*h : Lightness the same as in $L^*a^*b^*$, chroma C^* as a radial distance from the neutral L^* axis, and hue h as a polar angle. The L^*C^*h representation correlates well with how people think about color, with L^* related to brightness, C^* related to colorfulness or color purity, and h related to hue, often thought of as color names, and for this reason L^*C^*h is used frequently in this thesis. Note that chroma C^* specifically refers to colorfulness relative to a similarly-illuminated white, a slightly more specific definition than the general percept of colorfulness. Details of colorimetric quantities and formulae for transformations between these different color spaces are given by Hunt [12].

Display Color Standards

Accurate descriptions of color stimuli are essential in the specification of displays and the communication of colors to be displayed. Many contemporary displays utilize RGB color primaries that approximate the standard RGB set defined by the High-Definition Television (HDTV) specification ITU Rec. 709 [13], also used in the internet imaging standard sRGB [14]. The standard provides a formulaic computation from a RGB triad to colorimetric XYZ tristimulus values by defining an idealized display model, which of course is invertable to enable the preparation of images which will be displayed as intended. However, in the real world, most displays do not perform exactly like this model, which would objectively result in color errors because the display does not perform as intended. The magnitude of such errors depends on a lot of things including the display technology, color processing, and manufacturing variability, but in most cases they can be suppressed by display models customized to fit the expected behavior of a type of display

or, for more accuracy, the measured behavior of a specific display. The field of color management is concerned with modeling input- and output-device behavior, addressing device-dependent color spaces, and standardizing translations between color spaces in a controlled way. A practical guide is offered by Giorgianni & Madden [15].

A display's color capabilities are often described in terms of color gamut, meaning the range of colors made available by the display. Color gamut is a three-dimensional volume, thanks to the trichromatic nature of the HVS, and it is often interesting to look at color gamut in a color space which utilizes a luminance dimension and two chromaticity dimensions. Viewed from the top, thus effectively ignoring the luminance axis, the chromaticity gamut appears as the area enclosed by the display primaries in a chromaticity diagram. An example of this is shown in Figure 1.7, which is a CIE 1976 uniform chromaticity diagram with axes of u' (reddish-greenish) and v' (yellowish-bluish). The colored horseshoe gives an approximate orientation to the location of colors on the diagram and indicates the physical limit achievable by visible light. The triangle shows the sRGB red, green, and blue primaries and the chromaticity gamut they enclose. White colors lie in the middle, and the standard white point D65 is labeled. On a chromaticity diagram, points add linearly, meaning that the light resulting from a mix of the sRGB red and green primaries must lie along the line connecting them, located proportionally between them according to their relative intensity.

Often a display is asked to display a color which lies outside its gamut, most often because the gamut does not enable reproduction of the entire range of a standard encoding like sRGB. In this situation, a decision must be made to choose which in-gamut color to display in place of the requested out-of-gamut color. Options range from hard clipping, where out-of-gamut colors are simply clamped to the edge of the real gamut, to elegant mapping strategies which preserve color differences and smoothly compress colors as they near the gamut edges. In general, gamut-mapping, the intentional distortion of colors to account for differences in gamut, is used between different types of media (i.e., displays and printers) and between encoding standards and real devices, as described by Morović [16].

1.2 Display Technology

Displays, for entertainment and information, come in many forms. At the present moment, flat panel liquid-crystal displays (LCDs) are firmly dominant in the desktop- and television-sized display markets, having recently displaced cathode-ray tube (CRT) displays and fought off the advances of plasma displays and rear-projection displays. Further, LCDs have actually enabled whole classes of inexpensive and thin mobile devices. Organic light-emitting diode (OLED) displays have been in development for more than a decade and have made some inroads in mobile devices but so far have not been successful in desktop or television sizes. Of course, any attempt to say what is typical or contemporary in the fast-moving display industry becomes out-of-date nearly immediately, but for the

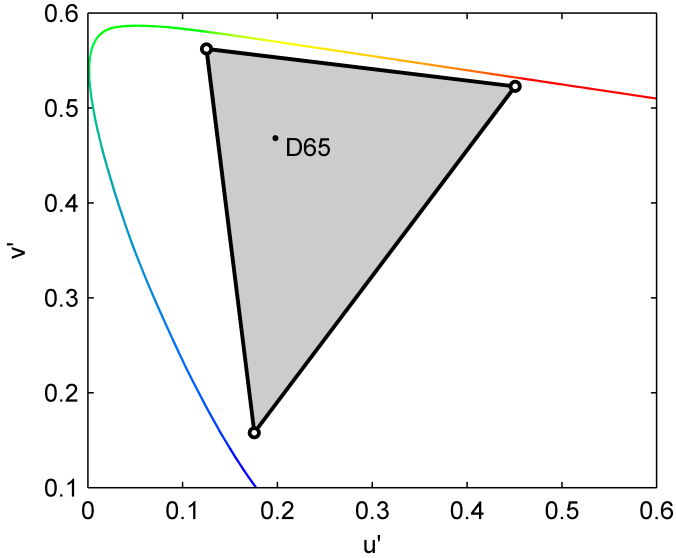


Figure 1.7: RGB system gamut. A CIE 1976 $u'v'$ chromaticity diagram showing the chromaticity gamut triangle of an additive RGB display system with the Rec. 709 primaries and display white point D65. In a chromaticity diagram, the sum of two physical light sources lies along the line joining them, proportional to the relative power of each. Thus, colors within the triangle are enabled by the display, while those outside are not. Note that a chromaticity diagram is only a top-view, ignoring the luminance dimension, and the chromaticity gamut boundary is only attainable at a specific locus of luminance values. The gamut volume shrinks to a white point at the bright end of the display's range.

purposes of this thesis a baseline display is considered to be a television-sized IPS LCD with a fixed fluorescent backlight and sRGB color gamut of vintage *circa* 2007. Such a display is described in the context of the ever-changing market by de Vaan [17], who also mentions several of the advanced display technologies studied in the coming chapters.

1.2.1 Liquid Crystal Displays

The LCD gets its name from its core element: a liquid crystal (LC) cell that can affect the polarization of light. Sandwiched between complementary polarizing filters and responsive to an electric field, the LC can attenuate light over an intensity range on the order of 1,000:1, thus providing an adjustable light valve to modulate the backlight behind it. Many different LCD implementations exist, varying in the placement of electrodes,

direction of the electric field, orientation of the front and back polarizers, and orientation of the LC material as constrained by rubbed polymer alignment layers. LC cells can be shaped like individual icons, as in a calculator or a digital watch, or they may be arrayed in large numbers in order to make a general-purpose display. Details of the components and design of modern LCDs can be found in textbooks by Lee et al. [18] and Chen [19].

The technology of the LC itself as well as the surrounding addressing and control hardware have been improving steadily since the first LCDs emerged in digital watches in the 1960s. A fascinating history of the development of LCD technology spanning several decades is provided by Kawamoto [20]. The classic implementation, or mode, of LCD, twisted-nematic (TN), uses electrodes sandwiching the LC material to turn the molecules perpendicular to the substrate. Depending on the orientation of the polarizers, the LC cell may be transparent in either the on or off state. TN can only attenuate light strongly in a perpendicular direction, resulting in light leakage off-axis and with it the characteristic contrast loss when viewed at different angles.

Other LC modes were designed to alleviate this angular dependence, including vertical-alignment (VA) and in-plane-switching (IPS). In VA mode, the LC molecules stand perpendicular to the substrate when voltage is off, producing a very dark black, and rotate to parallel to the substrate when voltage is on, allowing light to pass. IPS LC material is rotated entirely in parallel with the substrate by electrode pairs that are both on the same side of the LC cavity. Without voltage, the molecules align according to their rubbed polymer substrate, and when voltage is applied in the plane, molecules rotate further depending on distance from the electrodes and create a twist through the height of the LC cavity that affects polarization. The result is a dark black that is very insensitive to viewing angle. An IPS example using crossed polarizers is illustrated in Figure 1.8, showing one pixel open and one closed.

Driving the pixels of a display is a complicated feat. Small pixel arrays can be driven via passive matrix (PM) addressing, which involves row-wise electrode lines on one side and column-wise electrode lines on the other. These lines are activated one row at a time in a fast-scan (typically horizontal) direction and a slow-scan direction (vertical) at a pace that defines the displayed frame rate. Each pixel receives electricity when its combination of row and column lines are activated, so the pixel may be controlled only $1/R$ of the time, where R is the number of rows of pixels.

A modern television-sized LCD typically takes the format of high-definition, or HDTV, with 1920 columns of 1080 pixels each, usually with three (red, green, and blue) subpixels per pixel. Thus, there are about six million addressable pixels, each RGB trio of which is roughly half a millimeter square on a 42 inch diagonal display. Both because of the sheer number of pixels and because of many image quality benefits, large modern displays are driven with active-matrix (AM) addressing in which every subpixel has its own dedicated switching circuit consisting of a set of transistors and a charge-holding capacitor in a thin-film-transistor (TFT) backplane made of amorphous or polycrystalline silicon. AM displays have row and column lines for both power and control, and because of the TFT

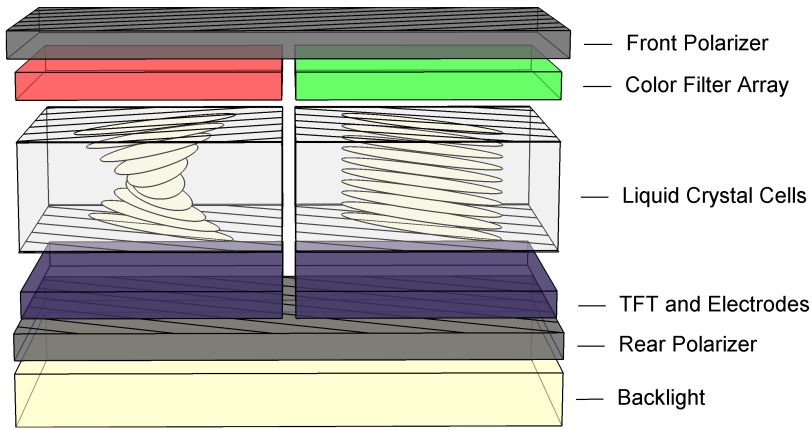


Figure 1.8: IPS LCD diagram. Illustration of the open (left) and closed (right) states of in-plane switching (IPS) liquid crystal cells. The backlight sends light upward from the bottom of the diagram, through the rear polarizer and the TFT control circuitry. The twisting (left) and aligned (right) LC cells are illustrated with ovals representing molecules. Both cells send light through the color filter layer, but only the light in the left cell, whose polarization state has been twisted by the LC, is allowed through the front polarizer.

circuits, each pixel can maintain its state between addressing cycles.

The mention of RGB subpixels is important. With rare exceptions, displays distribute their color primaries spatially, so that any physical point on the display is uniquely red, green, or blue. Because the subpixels are so small, they are integrated by the HVS and the spatial separation between them is not perceived. That being said, they are usually just on the edge of visibility, so that the physical subpixel layout can be important to image and especially text sharpness. In fact, the spatial separation provides an opportunity for enhanced resolution via subpixel interpolation algorithms, as discussed in depth in the thesis of Klompenhouwer [21]. Most current LCDs use a rectangular grid of square pixels divided into three equal portions corresponding to RGB, and subpixel addressing for text is provided standard in computer operating systems.

Even with all of the advancements to date, the liquid crystal display system remains imperfect. As a light valve-based system, it is inherently inefficient because the backlight

is brighter than the resulting image, and the backlight is always on, regardless of the requirements of the image content. The contrast range of the LC limits the displayable image contrast and tends to worsen as it is viewed off-axis because of its angular sensitivity. And, the color range provided by LCDs is limited either by the color primaries chosen or by the fixed boundary of standard image encodings such as Rec. 709. Of course, such challenges have kept research labs throughout the display industry busy brewing interesting new solutions.

1.2.2 Advancements in Display Technology

With the limitations of LCD well known, and the unstoppable progress of technology, many new technological ideas are beginning to find their way into the domain of displays. Of particular interest to the present work are segmented LCD backlights, which are addressable at a resolution lower than the LC panel, but higher than a single, fixed backlight; organic light emitting diode (OLED) displays, which obviate both the backlight and the LC light valve by generating light at each pixel directly; and wide color gamut, enabled either by wide-gamut RGB displays or multiprimary displays. There are many other display advancements that will not be discussed in this thesis, such as high frame rates, ultrahigh resolution, sequential-color, angular-dependence reduction, etc., all of which could be approached with the methods described herein.

Segmented Backlights

The very idea of a fixed backlight, which is simply “on” regardless of the image being shown on the display, is recognizably inefficient, yet with older backlight technology, which could not switch quickly, it was the only possibility. In the mid-2000s, fast-switching backlight technology, both fluorescent and LED-based, enabled dimmable and scanning backlights and started a trend toward backlight segmentation and addressability with advantages for both image contrast and motion portrayal [22].

A major step forward was made by Seetzen et al. [23, 24], in the form of a segmented backlight, first as a proof-of-concept involving a projector as a backlight, and then as a more practical implementation using an array of LEDs to turn the backlight into a relatively low-resolution display behind the high-resolution LC panel. With this system, low-frequency patterns of light and dark are made entirely by the segmented backlight, and the high-frequency details are filled in with the LC. The result is extremely high contrast capability on a large scale, but much more limited (equal to the LC’s inherent contrast range) on a small scale. Thus, choosing the number, shape, and luminance profile of the backlight segments is not trivial; in fact, it is essential that such choices be addressed with a perceptual understanding as promoted by this thesis.

OLED Displays

The first OLED cell, a single “pixel” of light-emitting material, was reported by Tang and Van Slyke of Eastman Kodak Company in 1987 [25]. It took another decade until this concept could be made into an actual RGB OLED display, both with passive-matrix [26] and active-matrix [27] addressing. Compared to LCD, OLED illustrates the conceptual extreme of segmented backlights: a backlight segment for every subpixel, in which case the backlight segment becomes the pixel, obviating the liquid crystal entirely. With this step comes some interesting advantages and lower system complexity, albeit with some new complications.

One complication is the difficulty of accurately patterning and aligning separate red-, green-, and blue-emitting OLED materials onto the substrate during manufacturing. An alternative arrangement, manufactured with an unpatterned white-emitting OLED material and using a color filter array similar to that used in LCDs to create colors, was possible. But, this W-RGB system was rather inefficient because most of the light being generated was thrown away by the RGB filters. A major efficiency gain was found with the use of pixels of unfiltered white in addition to RGB, making a four-primary W-RGBW system, as explained by Arnold et al. [28]. Because they generate light at each pixel only when needed, the advantages of using RGBW is much greater with OLEDs than with light-filtering LCD or projection displays. The efficient and easy to manufacture W-RGBW OLED architecture persists today as the first television-sized OLED displays have begun to emerge. Utilizing the fourth primary properly is a perfect example, explained in this thesis, of how a technological advancement provides an opportunity that may be used well or poorly depending on how perception knowledge is applied.

Wide Color Gamut

Regardless of the display technology, the choice of color primaries determines the color gamut, or the range of colors that a display is capable of reproducing. Choosing purer colors enlarges the chromaticity gamut, and adding additional primaries beyond the requisite three can open up new areas of color space in the gamut. The effects on the chromaticity gamut of these two approaches are shown in a chromaticity diagram in Figure 1.9. Using very saturated (pure) RGB primaries has less impact on the overall display design than adding additional color channels. However, because the chromaticity gamut remains triangular, the purity must become extreme in order to gain appreciable gamut area, which also generally means lower luminance for the same power input. Wide-gamut RGB displays may be made with RGB LED backlights [29] or with more selective color filters combined with either white LEDs or fluorescent backlights.

Realizing that a triangular chromaticity gamut can be made more round and spacious by adding more vertices, a multiprimary display may be made. Conveniently, extremely saturated RGB primaries are not necessary when they are augmented by additional primaries. However, using more primaries complicates the display design by requiring

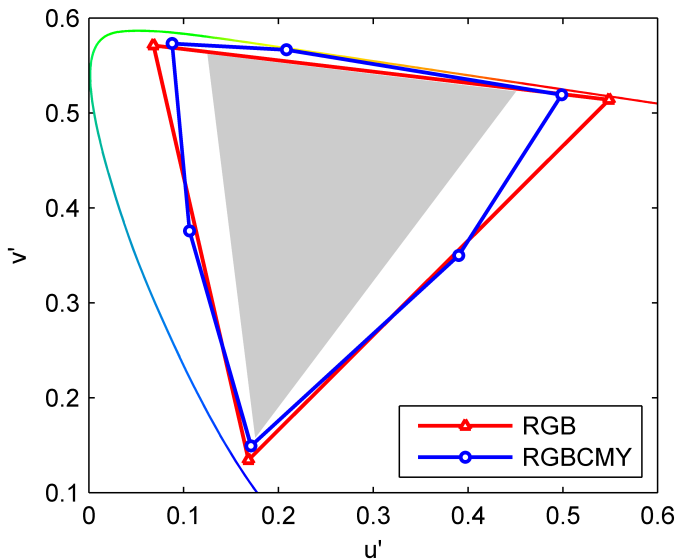


Figure 1.9: Wide gamut systems. A CIE 1976 $u'v'$ chromaticity diagram showing the chromaticity gamuts of two hypothetical wide-gamut displays providing more coverage than the Rec. 709 gamut shaded in gray. The red triangle shows a wide-gamut RGB display gamut that has expanded area because of extreme saturation of each primary. The blue polygon shows the resulting gamut of a 6-primary display system with more moderately saturated RGB primaries but which expands its gamut through the use of additional cyan, magenta, and yellow (CMY) primaries.

physical subpixels for these new colors [30], either reducing resolution in the process or increasing pixel density and cost, and it necessitates additional image processing to address these new color channels. As discussed by Roth et al. [31], either approach to wide gamut generally comes at the cost of lower efficiency, and because the resulting display has primaries different from the color primaries expected by typical image encodings, careful image processing is needed to properly use it. Thus, wide-gamut displays provide another excellent topic for the approach discussed in this thesis.

1.3 Relating Perception to Technology

In the development of any technology that will be directly used by humans, it is important to understand how it will be used, what will be important to the user, and how the limits of the senses – vision, hearing, and touch, primarily – affect what is visible, audible, and tangible in the technological implementation.

How bright must a brakelight be to clearly signal that a car is stopping? How quiet should a CD player mechanism be to not distract from the music? How tall must the little bumps on the “F” and “J” keys be to facilitate a typist’s hand orientation without looking?

Display technology is intended primarily for the visual sense, and thus the perception topics being discussed are all visual. Further, requiring more than simple detection tasks like those required for brakelights, displays are designed to convey images for information and entertainment. Thus, the *image quality* of a display (or actually, of the entire imaging chain that ends with a display) commands primary interest. What is image quality exactly? A range of interpretations can be found. At the technical, signal-processing end of this range, some definitions refer to the “fidelity” of an image with respect to an original or the real world; such definitions seem well-suited to describe degradations of image content. At the other end of this range, very human-centered definitions refer to general “excellence” or mention the context or observer expectations. In this category we find a favorite definition of image quality from the PhD thesis of Judith Dijk [32], whose concise honesty necessarily focuses on the human and acknowledges that most human responses are tempered by internal expectations:

Image quality is the extent to which the image corresponds to the internal expectation of the human observing the image.

Indeed the range of choices for defining image quality implies some structural relationships. For example, de Ridder and Endrikhovski [33] propose that the relative importance of aspects such as fidelity, usefulness, and naturalness is affected by the context in which an image is viewed. It seems logical that the most human-centered definition of image quality includes the other aspects in its cognitive sum, including fidelity, context, expectation, and purpose. In the next sections, a structured way of thinking about how the technological choices involved in display design can affect the perceived image, and thereby the impression of overall image quality, is presented.

The next step is to quantify image quality, which may only be done via the subjective opinions of real human observers, each of whom may have their own biases and ideals. To many physical scientists, the idea of quantifying a subjective response such as human opinion seems impossible, but in actuality overall quality or any of the underlying per-

ceptual attributes may be measured easily enough with well-controlled experiments and proper methodologies.

1.3.1 Quantifying Perceptual Attributes

Image quality is generally understood to be a mental integration of the quality of multiple contributors. In his text, Peter Engeldrum [34] provides a helpful framework, describing perceptual attributes as perceived impressions of the characteristics of an image which contribute to overall image quality. Examples include sharpness, colorfulness, and brightness, which explains why he playfully calls perceptual attributes “the nesses.” By definition, perceptual attributes are perceived by an observer, but indeed they may be related to measurable quantities. For example, the sharpness of an image may be predicted based on the physical characteristics of the pixels or ink dots that comprise it. It is important to remember that people are seldom unanimous in their opinions: practically speaking, there is a distribution of responses over a population of observers, and even over multiple observations by the same observer. For these reasons, perceptual attributes are necessarily quantified by asking a sampling of people their opinions in a controlled way and aggregating their responses. Engeldrum’s text outlines many of the relevant concepts of psychometric methods and analyses.

Psychometrics

An example that is relevant to Chapter 2 is the visibility of differences in the black level of images. As is often the case for fine perceptual measurements, images are presented in randomized pairs, one image the reference and the other varying in the attribute of interest – in this case, the luminance of the blackest portion of the image. The observer’s task is to choose the image with the darker black, and because it is “forced-choice,” if the difference is not seen he or she must choose one randomly. If the difference is very large, then observers will reliably see it and indicate so; likewise, if the difference is zero or close to it, then observers will reliably not see it and be forced to guess. In the interesting region between, the frequency with which observers see the difference increases with the size of the difference. The monotonic relationship between the intensity of a stimulus and the probability of detection is a psychometric function, an example of which is shown in Figure 1.10. As shown in the figure, a sigmoid shape, in this case the normal cumulative distribution function, is transformed to the probability range [0.5, 1.0], which is typical for difference-detection tasks such as paired-comparison experiments. The visibility threshold, which defines a just-noticeable difference (JND) is found at the stimulus level where $p = 0.75$. If instead a basic detection task is used, in which the probability of detection when the stimulus is very low actually approaches zero, then the function range is [0.0, 1.0] and the threshold is at $p = 0.50$.

The methods used to generate such a relationship vary depending on the attribute under study and the efficiency and accuracy required. In many cases, rather than the entire psychometric curve, what is needed is simply an estimate of a threshold. This can

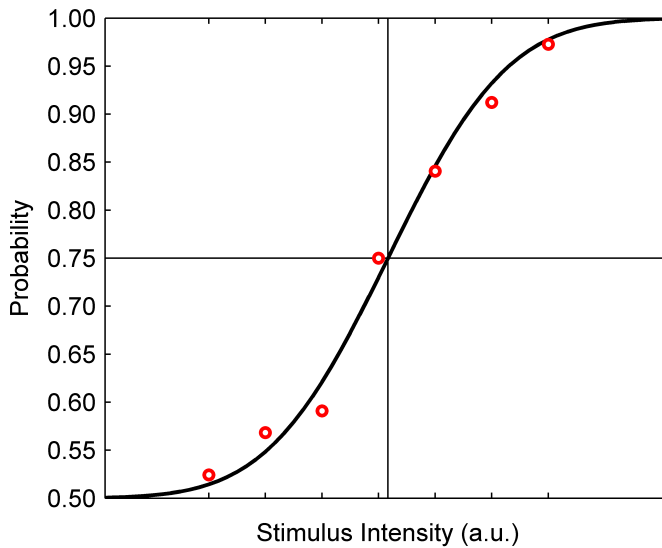


Figure 1.10: Psychometric function for difference detection. This example curve shows the probability of detection versus the intensity or amount of a stimulus for a difference-detection task such as is common in imaging. The red dots show, for various levels of the stimulus, measured frequencies of detection averaged over multiple observers. The black curve is a normal cumulative distribution function (erf) fitted to these data. At low levels of the stimulus (far left), the probability of detection is 0.50, because observers reliably don't see it but guess correctly half the time. At high levels of the stimulus, observers see it consistently ($p = 1.00$). The stimulus intensity corresponding to $p = 0.75$ is taken as the visibility threshold – the point at which half the observers see it, or, similarly, the point at which a given observer has a 50-50 chance of seeing it.

be made with various levels of effort and accuracy. For example, a preference threshold can be estimated quickly and roughly with a tuning task in which each observer looks at single-stimulus images in sequence and adjusts an image parameter freely until he or she is satisfied – this is especially useful when there are many images to be considered and efficiency is important. For a more precise measurement of a threshold, a series of forced-choice paired comparisons driven by staircase logic, in which the magnitude of the stimulus or difference is decreased when the observer correctly indicates seeing it and increased when he or she doesn't, converges to a threshold with fewer measurements than would be necessary to estimate the whole psychometric function.

In some cases, rather than a difference or detection threshold, scale values on an interval scale are desired to quantify differences between stimuli. Observers can rate stimuli directly on a discrete rating scale, providing a quick response that is fairly accurate when averaged over many observations and which can be treated with care as interval data. Or, with more effort and also more precision, paired comparisons can be made among all combinations of a set of images. Over many observations, a frequency matrix summing how many times each image is chosen over each other image can be translated to an interval scale corresponding to relative quality. Such psychometric methods are used in a variety of experiments in the course of this thesis, and further details are explained in each chapter as necessary.

Objective Metrics

A measured perceptual response may be described or modeled by an objective metric, which is a quantity that may be computed directly from either image data or system parameters. Objective metrics may describe individual perceptual attributes or approximate the overall image quality of a system or image. They are invaluable for system design because they provide a venue for the reuse of psychometric data, giving a perceptual correlate directly rather than requiring constant cycles of experimentation as the system evolves. A good example (though one that is not used further in this thesis) of an objective metric which relates physical display characteristics to a single perceptual attribute is motion picture response time (MPRT), which describes motion blur in a display. MPRT can be either derived from the physically-measured blurred edge of a moving object on the screen [35] or computed from display characteristics [36], and it has been shown to correlate well with perceived motion blur [37]. Thus, MPRT has proven useful in characterizing the motion blur characteristics of different displays – replacing the older metric of liquid crystal response time (LC-RT) which describes only the speed of the LC itself and doesn't correlate very well with perceived blur – and accurately showing the value in motion-blur reduction of pulsed and scanning backlight techniques.

Unfortunately there are many examples of poor objective metrics, meaning that they do not correlate well with a perceptual attribute. One such metric, used to describe the size of a display's chromaticity gamut, is %NTSC. This metric is simply the ratio of the area of the display's chromaticity gamut in CIE 1931 xy chromaticity space to

the area of the standard NTSC gamut, expressed as a percentage. Indeed the metric increases monotonically with display gamut area, but it has a number of flaws: the xy chromaticity space is not perceptually uniform, which means the relative value of different colors is misrepresented; it does not take into account gamut overlap, which means that only size matters, not how much of the NTSC gamut is covered; and the NTSC baseline itself was obsolete nearly immediately after it was enshrined into American law in 1953 by the National Television System Committee (NTSC) [38]. Ben Chorin et al. [39] provide criticism and an alternative; yet, %NTSC is commonly listed in TV and monitor specifications and even used to market wide-gamut displays with typical bigger-is-better oversimplification for the consumer market. Chromaticity gamut size and shape preferences are studied later in this thesis, and %NTSC will not be used to quantify or describe the gamut results.

At one level higher, objective metrics are sometimes used to quantify overall image quality rather than single perceptual attributes, but these quality metrics must be used with care as they typically correlate with perceived quality only over a small range of system parameters or only within a given class of algorithms. An example is peak signal-to-noise ratio (PSNR) [40] which can be used to quantify the visual artifacts resulting from image compression algorithms. PSNR correlates well with perceived quality over a range of compression amounts, but is known to fail when comparing across compression methods.

Objective quality metrics like PSNR don't generalize well because they don't include all relevant imaging system parameters and don't sufficiently model the characteristics of the human visual system. Accounting for the cumulative effects of different perceptual attributes requires an image quality model.

1.3.2 Image Quality Models

Throughout the history of the development of imaging systems, there has been a desire to predict image quality from knowledge of system components. However, image quality models have consistently proved difficult to build. There is a general understanding that overall image quality is an integrated sum of many perceptual attributes, each of which may be affected by the components of the imaging system, yet there remains no general, robust formula for image quality as a function of its parts. What we do have are approaches to image quality modeling that may be applied to specific systems or situations. Note that a related branch of image quality research studies computational measures for image quality that apply trained machine-learning algorithms to images to predict the human-perceived quality, but this empirical fitting approach is less interesting to the present thesis than system quality models.

One conceptually excellent approach provided by Engeldrum is the Image Quality Circle [34], shown in a diagram in Figure 1.11. Engeldrum realizes that the ideal goal is to know how technology variables, meaning characteristics of the materials or electronics

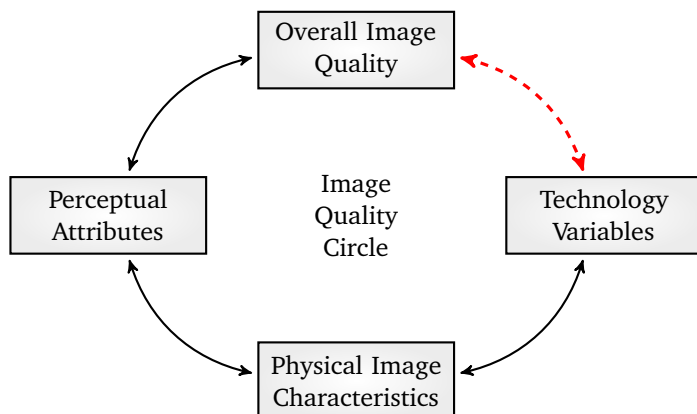


Figure 1.11: Image Quality Circle. The Image Quality Circle as described by Engeldrum. The relationship between the right and upper blocks is often desired, as this enables the choice of technology based on overall quality, but it is impossible to model generally. The long way around, from right to lower to left to upper, may be generalized, and thus is worth the effort to understand.

of the system, affect overall image quality, but that studying this relationship directly is not generalizable because the details of exactly what about the perceived image is affected by the technology are not uncovered. Without such details, every technology update requires new human observations of image quality, resulting in an endless cycle as technology changes.

Instead, the Image Quality Circle approach is to take the “long way” around the circle, by first understanding how technology choices affect physical (measurable) image characteristics, then understanding how physical image characteristics are seen by observers in terms of perceptual attributes or “nesses,” for example sharpness and colorfulness, and finally understanding how the “nesses” sum to overall image quality. A valuable aspect of this structure is that the latter relationships (between perceptual attributes and overall quality, and between physical image parameters and perceptual attributes) can be technology-independent. For example, if it is understood how perceived image contrast affects overall image quality, it does not matter whether the contrast is affected by a new liquid-crystal design, a new reflective coating on the display, or something else. This full-circle approach was utilized successfully for characterizing display artifacts and communicating specifications in a meaningful way by Teunissen [41].

A very rigorous image quality framework is provided by Keelan [42]. In his full-system approach, the contributions to overall quality of individual attributes and artifacts (i.e., sharpness, blockiness, etc.) are measured separately, anchored to a known, physical

scale. Then, a multivariate system image quality model is constructed which can compute image quality as a “sum” of each of these perceptual attributes. Crucial to this approach is the binding of every experimental result to a consistent anchor of known quality such as a set of ruler images of measured overall image quality. Keelan makes heavy use of a sharpness ruler containing images whose overall quality is affected by sharpness – typically, starting with a “perfect” image and degrading each sample in the ruler with increasingly aggressive blur filters. Other artifacts can then be compared by human observers against the sharpness ruler, with the observer choosing the “lesser of two evils,” for example weighing jpeg artifacts versus blurriness.

In order for this approach to succeed, it is important that all of the contributing attributes and artifacts are known, or else the overall model will be incomplete and quality possibly over-estimated. Importantly, it begins from the perspective that there is an ideal imaging system that may be approximated in practice at a lower quality level, which is obviously helpful in assessing the impact of things like production variability, errors, and cost-down system changes – but, it seems more difficult to apply this perspective to ever-improving imaging systems such as displays. Keelan’s rigorous, time-consuming approach may be well worth the effort for the long-term development of an imaging system whose failure modes are well-known. But, for many practical questions, it is unnecessarily burdensome, and because it cannot easily include trade-offs unrelated to image quality, such as cost or power consumption, even a detailed system model may never be sufficiently complete.

1.3.3 A Practical Approach to Display Design

For the assessment of the added value of a new technology, a practical, straightforward understanding of how system quality or preference is affected by the main variable(s) of that technology is sufficient, and this may be found efficiently and communicated effectively. Technology improvements come in a long series over time, not all at once, so they may be studied individually as they come. And, often many aspects of the whole imaging system are fixed and standardized, so it is beneficial to focus only on the part that is being changed. Because the implementation of a new technology is typically parameterized in some way, the design decisions involve not simply *whether*, but *how* and *to what extent* to use it for greatest human-centered advantage.

It can be helpful to look back to the Image Quality Circle in Figure 1.11 and realize that in many cases only a portion of this full circle is needed. The approach taken throughout this thesis may be described generally as follows:

- Determine the most important perceptual attribute affected by the technology variable under study.
- Measure or model the physical image characteristics that are modulated by the technology variable.
- Measure or model observers’ perceptual response to a relevant range of visual

stimuli modulating the physical image characteristics corresponding to the technology variable.

- Model the measured perceptual response as a function of the technology variable, and communicate the modeled relationship to enable an informed technology choice.

The result of such an approach is always a direct answer to the question about whether or how a new technology should be best implemented. In almost every case, the answer is a relationship, for example a description of the image quality degradation associated with a key parameter of a new power-saving scheme. In these cases, a simple answer is not enough, but the whole relationship must be communicated clearly to enable decisions later in the product-development chain weighing such a visual quality cost against other benefits not associated with visual quality. The approach is efficient, focusing on identifying the most important perceptual effects without unneeded structural burdens, and effective, providing a useful understanding of the relationships in order for technology choices to be made.

1.4 Outline of this Thesis

This thesis includes a detailed look at the perceptual implications of the application of three advanced display technologies. In each of these examples, the primary perceptual attribute is identified and a relationship between it and the image characteristics that affect it is created through psychophysics and subsequent modeling.

High dynamic range LCDs: Segmented-backlight LCD technology can provide arbitrarily high contrast over large spatial areas, but always with some restricted amount of contrast over small spatial areas. Chapter 2 describes the quantification and modeling of the human visual system's ability to see details in dark regions while impaired by glare from neighboring bright regions, which enables the design of segmented backlight tailored to provide as much contrast as is visible, but no more.

RGBW OLED Displays: OLED displays utilizing white-emitting OLED materials can be made substantially more efficient with the use of an unfiltered W primary in addition to the typical R, G, and B primaries. Chapter 3 outlines a general algorithm for the color processing required for a fourth primary of arbitrary color with no visual impact on the display's color reproduction, relying on the additivity of light and the visual principle of metamerism. This algorithm was used successfully in several generations of OLED display prototypes and is embodied in some form in current OLED displays for mobile devices and televisions. Even greater power efficiency can be gained by reducing the maximum luminance of the RGB primaries, and a study that relates the resulting loss of image quality to the reduction in primary luminance

is described in Chapter 4. This relationship may be used to select a tradeoff between power efficiency improvement and image quality reduction.

Wide Gamut Displays: Display technology has advanced sufficiently to enable the creation of arbitrarily-large color gamut. Such capability could be immediately advantageous as part of an entirely new imaging system; however, because a world of existing image and video infrastructure conforms to typical industry standard color encodings such as Rec. 709, taking advantage of wide gamut requires boosting the color of the input images. Chapter 5 explains a preference-based approach in which the gamut boundaries are inferred from the desired or tolerated amount of color boost in a perceptual study using a wide-gamut RGB LCD. This approach is extended in Chapter 6 in an additional study, through the use of hue-preserving boost algorithms and a multiprimary display, resulting in a relationship between gamut size and preference over the observer population. This relationship enables the quantification of the relative value of regions in the chromaticity gamut, and thus the added value of any proposed wide-gamut display.

The three topics are presented in this order, which is not chronological, with a small nod to one of my early mentors in imaging system modeling and design. He always insisted that we get the neutral tonescale straightened out first before working out the details of color. It seems natural then to start with contrast, whose luminance-only details were studied in neutral black-and-white, move on to accurate color and slightly distorted color for RGBW OLED, and finally on to enhanced color for wide-gamut displays. Each chapter includes its own list of relevant references.

Chapter 7 wraps up the thesis with a review of the major points presented and some conclusions regarding the effectiveness of the proposed framework in display design. A retrospective critique and a forward-looking proposal are provided.

1.5 Bibliography

- [1] Brian A. Wandell. *Foundations of Vision*. Sinauer Associates, Sunderland, Massachusetts, 1995.
- [2] H. Hofer, J. Carroll, J. Neitz, M. Neitz, and D. R. Williams. Organization of the Human Trichromatic Cone Mosaic. *Journal of Neuroscience*, 25(42):9669–9679, 2005.
- [3] Mark D. Fairchild. *Color Appearance Models*. Wiley-SID Series in Display Technology. John Wiley & Sons, Ltd., West Sussex, 2nd edition, 2005.
- [4] Peter G. J. Barten. *Contrast Sensitivity of the Human Eye and its Effects on Image Quality*. SPIE, Bellingham, 1999.

- [5] Peter G. J. Barten. Physical Model for the Contrast Sensitivity of the Human Eye. In *SPIE*, volume 1666, pages 57–72. SPIE, 1992.
- [6] P.G.J. Barten. Formula for the Contrast Sensitivity of the Human Eye. In Yoichi Miyaki, editor, *SPIE-IS&T Electronic Imaging*, volume 5294. SPIE and IS&T, 2004.
- [7] L. L. Holladay. The Fundamentals of Glare and Visibility. *Journal of the Optical Society of America*, 12(4):271–319, 1926.
- [8] Parry Moon and Domina Eberle Spencer. The Visual Effect of Non-Uniform Surrounds. *Journal of the Optical Society of America*, 35(3):233–248, 1945.
- [9] CIE. Research note: Disability glare. *Vision and Colour: Physical Measurement of Light and Radiation*, 135(1), 1999.
- [10] ISO/CIE. ISO 11664-1:2007(E)/CIE S 014-1/E:2006: Colorimetry — Part 1: CIE Standard Colorimetric Observers, 2007.
- [11] ISO/CIE. ISO 11664-4:2008(E)/CIE S 014-4/E:2007: Colorimetry — Part 4: CIE 1976 L*a*b* Colour Space, 2008.
- [12] R. W. G. Hunt. *The Reproduction of Colour*. Fountain Press, Kingston-upon-Thames, 5th edition, 1995.
- [13] ITU-R. BT.709 : Parameter Values for the HDTV Standards for Production and International Programme Exchange, April 2002.
- [14] Michael Stokes, Matthew Anderson, Srinivasan Chandrasekar, and Ricardo Motta. W3C: A Standard Default Color Space for the Internet - sRGB, November 1996.
- [15] Edward J. Giorgianni and Thomas E. Madden. *Digital Color Management: Encoding Solutions*. Addison-Wesley, Boston, 1998.
- [16] Ján Morovič. *Color Gamut Mapping*. Wiley-SID Series in Display Technology. John Wiley & Sons, Ltd., West Sussex, 2008.
- [17] Adrianus J.S.M. de Vaan. Competing Display Technologies for the Best Image Performance. *Journal of the Society for Information Display*, 15(9):657–666, 2007.
- [18] Jiun-Haw Lee, David N. Liu, and Shin-Tson Wu. *Introduction to Flat Panel Displays*. Wiley-SID Series in Display Technology. John Wiley & Sons, Ltd., West Sussex, 2008.
- [19] Robert H. Chen. *Liquid Crystal Displays: Fundamental Physics and Technology*. Wiley-SID Series in Display Technology. John Wiley & Sons, Inc., Hoboken, 2011.
- [20] Hirohisa Kawamoto. The History of Liquid-Crystal Displays. In *Proceedings of the IEEE*, volume 90, pages 460–500. IEEE, 2002.

-
- [21] Michiel Klompenhouwer. *Flat Panel Display Signal Processing: Analysis and Algorithms for Improved Static and Dynamic Resolution*. PhD thesis, Technische Universiteit Eindhoven, 2006.
 - [22] Pierre De Greef, Hendriek Groot Hulze, Jeroen Stessen, Hans Van Mourik, and Seyno Sluyterman. Adaptive Scanning, 1-D Dimming, and Boosting Backlight for LCD-TV Systems. *Journal of the Society for Information Display*, 14(12):1103–1110, 2006.
 - [23] H. Seetzen, L. Whitehead, and G. Ward. A High Dynamic Range Display System Using Low and High Resolution Modulators. In *SID Symposium*, volume 34, pages 1450–1454. SID, 2003.
 - [24] Helge Seetzen, Wolfgang Heidrich, Wolfgang Stuerzlinger, Greg Ward, Lorne Whitehead, Matthew Trentacoste, Abhijeet Ghosh, and Andrejs Vorozcovs. High Dynamic Range Display Systems. In *ACM SIGGRAPH*, pages 760–768. ACM, 2003.
 - [25] C. W. Tang and S. A. VanSlyke. Organic Electroluminescent Diodes. *Applied Physics Letters*, 15(12):913–915, 1987.
 - [26] Y. Kijima, N. Asai, N. Kishii, and S. I. Tamura. RGB Luminescence from Passive-Matrix Organic LED's. *IEEE Trans. on Electron Devices*, 44(8):1222–1228, 1997.
 - [27] M. Stewart, R. S. Howell, L. Pires, M. K. Hatalis, W. Howard, and O. Prache. Polysilicon VGA Active Matrix OLED Displays-Technology and Performance. In *Electron Devices Meeting, 1998. IEDM '98. Technical Digest., International*, pages 871–874, 1998.
 - [28] A. D. Arnold, P. E. Castro, T. K. Hatwar, M. V. Hettel, P. J. Kane, J. E. Ludwicki, M. E. Miller, M. J. Murdoch, J. P. Spindler, S. A. Van Slyke, K. Mameno, R. Nishikawa, T. Omura, and S. Matsumoto. Full-Color AMOLED with RGBW Pixel Pattern. *Journal of the Society for Information Display*, 13(6):525–535, 2005.
 - [29] Gerard Harbers and Christoph Hoelen. LP-2: High Performance LCD Backlighting using High Intensity Red, Green and Blue Light Emitting Diodes. In *SID Symposium*, pages 702–705. SID, 2001.
 - [30] Young-Chol Yang, Keunkyu Song, SooGuy Rho, Nam-Seok Rho, SungJin Hong, Kang Beo Deul, Munpyo Hong, Kyuha Chung, WonHee Choe, Sungdeok Lee, Chang Yong Kim, Sung-Hee Lee, and Hyong-Rae Kim. Development of Six Primary-Color LCD. In *SID Symposium*, pages 1210–1213, 2005.
 - [31] Shmuel Roth, Nir Weiss, Moshe Ben Chorin, Ilan Ben David, and Ching Hung Chen. Multi-primary LCD for TV Applications. In *SID Symposium*, pages 34–37, 2007.
 - [32] Judith Dijk. *In Search of an Objective Measure for the Perceptual Quality of Printed Images*. PhD thesis, Technische Universiteit Delft, 2004.

- [33] Huib de Ridder and Serguei Endrikhovski. 33.1: Invited Paper: Image Quality is FUN: Reflections on Fidelity, Usefulness and Naturalness. *SID Symposium Digest of Technical Papers*, 33(1):986–989, 2002.
- [34] Peter G. Engeldrum. *Psychometric Scaling: A Toolkit for Imaging Systems Development*. Imcotek Press, Winchester, 2000.
- [35] Youichi Igarashi, T. Yamamoto, Y. Tanaka, J. Someya, Y. Nakakura, M. Yamakawa, S. Hasegawa, Y. Nishida, and Taiichiro Kurita. 31.2: Proposal of the Perceptive Parameter Motion Picture Response Time (MPRT). *SID Symposium Digest of Technical Papers*, 34(1):1039–1041, 2003.
- [36] Michiel A. Klompenhouwer. 51.1: Temporal Impulse Response and Bandwidth of Displays in Relation to Motion Blur. *SID Symposium Digest of Technical Papers*, 36(1):1578–1581, 2005.
- [37] J. Someya and Y. Igarashi. A Review of MPRT Measurement Method for Evaluating Motion Blur of LCDs. In *11th International Display Workshops*, pages 1571–1574. SID, 2004.
- [38] 47 C.F.R. § 73.682 (20) (iv). TV transmission standards, 10-1-07 1953.
- [39] Moshe Ben Chorin, Dan Eliav, Shmuel Roth, Assaf Pagi, and Ilan Ben-David. 21.3: New Metric for Display Color Gamut Evaluation. *SID Symposium Digest of Technical Papers*, 38(1):1053–1056, 2007.
- [40] ITU-T. J.247: Objective Perceptual Multimedia Video Quality Measurement in the Presence of a Full Reference, 2008.
- [41] Kees Teunissen. *Flat Panel Display Characterization: A Perceptual Approach*. PhD thesis, Technische Universiteit Delft, 2009.
- [42] Brian W. Keelan. *Handbook of Image Quality: Characterization and Prediction*. Optical Engineering. Marcel Dekker, New York, 2002.

Perceived Contrast in High Dynamic Range Displays

Abstract

A perceptual experiment was conducted to measure the visibility of black-level differences in the proximity of a bright glare source. In a controlled viewing environment, visual difference thresholds were adaptively measured using dark, shadow-detail images shown on a high dynamic range liquid crystal display while an external LED lamp was used to induce intra-ocular glare over a small range of eccentricities. This high-contrast situation is relevant to HDR displays which may have bright regions in displayed images as well as to viewing environments which include lamps or other light sources. The resulting difference thresholds are modeled with a combination of the CIE total glare equation, the DICOM contrast visibility model, and a new estimate of adaptation luminance.¹

2.1 Introduction

Technological advances continue to allow brighter and higher contrast displays. One example of such technology is the class of high dynamic range (HDR) segmented-backlight LCDs as introduced by Seetzen et al. [1]. The general concept of these displays is a two-part system, consisting of a low-resolution, segmented backlight (i.e., a grid of addressable LEDs) behind a high-resolution liquid crystal display. With a segment's LED(s) turned off, the luminance in that region can be essentially zero, while with the LED(s) on and the LCD in the transmissive state, the luminance can be very high (tens of thousands of cd/m^2), thus enabling an arbitrarily high contrast ratio. Such high contrast may be achievable at the same moment at a large scale, such as across the width of the display. However, because the backlight segments are larger than the display pixels, neighboring pixels can never achieve such a high contrast. Within a small region of the display, it

¹This chapter has been published in the Journal of the Optical Society of America A, Volume 29, Number 4, 559-566 (2012): "Veiling Glare and Perceived Black in High Dynamic Range Displays," Michael J. Murdoch and Ingrid E. J. Heynderickx.

behaves much like a fixed-backlight LCD, in which the contrast is limited by the dynamic range of the LCD itself. Thus, local luminance errors, known as halos, may result. Halos may be either lighter or darker than the intended luminance, but in general are areas of reduced luminance contrast.

Fortunately, early in the development of HDR displays it was recognized that the limits of the human visual system meant that such localized contrast range reduction actually did not matter to the perceived image when viewed from a normal viewing distance. The main reason for this is glare within the eye, caused by intra-ocular reflections and scattering, which lowers contrast on the retina in the proximity of bright light sources. The International Commission on Illumination (CIE) has published a general equation for the effect of intra-ocular glare, which has a general shape of $1/\theta^2$, where θ is the angle between the glare source and the eye's fixation point in degrees [2].

Several authors have used descriptions of glare to create design rules for HDR displays and to predict visibility of luminance errors. Seetzen et al. [1, 3] cited the glare-induced impairment of detail perception near high-contrast boundaries as the limitation of the human visual system that could be exploited in the construction of such displays, but they did not quantify the reduction in visibility of image detail. McCann and Rizzi [4] used the CIE glare equation to compute retinal contrast expected from high dynamic range transparency stimuli. Uniform black and white patches resulted in retinal images which were very nonuniform and relatively lower in contrast, yet qualitatively observers still described them as uniform in appearance. Thus, they recognized that even with well-modeled retinal contrast, the prediction of appearance is still elusive.

Several studies have focused on the optimal number and luminance profile of backlight segments. Swinkels et al. [5] measured user preference simulating different numbers of backlight segments using a double-LCD, and found that preference scores increased up to about 2500 segments for a 30" display viewed at 2.4m. Langendijk and Hammer [6] studied the relationship between physical black level and the spatial frequency of image content for segmented-backlight LCDs with different numbers of addressable segments. They applied the CIE glare equation to estimate the effective rise in black level caused by glare from the image itself, and they concluded that an LCD with at least 2048 backlight segments was comparable to an ideal HDR display (i.e. without the spatial limitations of backlight segments). Langendijk's modeling-based conclusions and Swinkels' experimental findings are similar, and they are corroborated by the successful examples of segmented-backlight LCD prototypes.

Contrast range is determined by both the black and white luminances, and contrast reduction often presents itself as a "lifted" black level. Because of the spatial characteristics of segmented-backlight LCDs, the display black level can vary across the display and with image content. Mantiuk et al. [7] studied the visibility and contextual interpretation of black under varied ambient illumination. They measured threshold black level differences and found that if the visual surround was increased in luminance or visual size, these threshold differences increased sharply. However, they did not provide a general model.

The goal of the present paper is to address a general answer to the question: how black is black enough, at what distance from a bright region of the image? Two components are critical to this understanding: the behavior of physical glare within the eye and the visibility of subtle differences in the image which are affected by the glare. Glare has been studied and modeled extensively. Visibility has been modeled in low-contrast situations and in limited experimental conditions. Yet so far, a convincing combination of both has not been made. This paper first explains these two models, then details an experiment designed to test visibility over a variety of black-level and glare conditions. Finally, a model incorporating glare, visibility thresholds, and a new, necessary adaptation component is provided and verified.

2.2 Models

2.2.1 Glare Model

Because of the imperfect optics of the human eye, not all incident light is properly projected onto the retina. Light may be lost in absorption, blurred by the imperfect lens, or scattered and/or reflected by the eye's components. Glare refers to the optical scattering and reflection of light within the eye that spreads incident light onto the “wrong” part of the retina, the main effects of which are the lowering of contrast and the hindrance of the visibility of shadow details. The term veiling glare describes the perceived effect, which is much like viewing the world through a thin white veil. Veiling glare has been quantified experimentally as equivalent veiling luminance (EVL). For example, if a viewer observes a dark alleyway at night, the presence of a streetlamp will cause veiling glare, which hinders the perception of shadows. The equivalent veiling luminance is literally what it says – the amount of luminance added uniformly to the physical scene (i.e., the thin white veil) that would result in an equivalent contrast reduction. Veiling glare always lowers contrast, defined most generally as a ratio between light and dark, because it is added uniformly to all light and dark regions of the scene.

CIE 135/1 [2] defines several glare equations which are functions of the visual angle between the glare source and the eye's viewpoint, as well as of the age and eye coloring of the observer. The CIE Total Glare Equation, which combines empirical measurements of human optical performance from very small (arc minutes) to very large (100°) visual angles, is roughly similar to $1/\theta^2$ over a broad range of θ , with the age and pigment parameters having relatively minor effects at angles above 2 degrees. The equation provides EVL in cd/m^2 per unit of illumination on the eye in lux, meaning it can be thought of as providing the angle (and age and pigment) dependent scale factor to convert illuminance incident on the eye from a glare source into equivalent veiling luminance.

For a streetlamp example, the lamp may be easily treated as a point source at a single visual angle from the fixation point, and the resulting EVL can be added to the luminance of the scene as measured from the location of the eye. The EVL of multiple

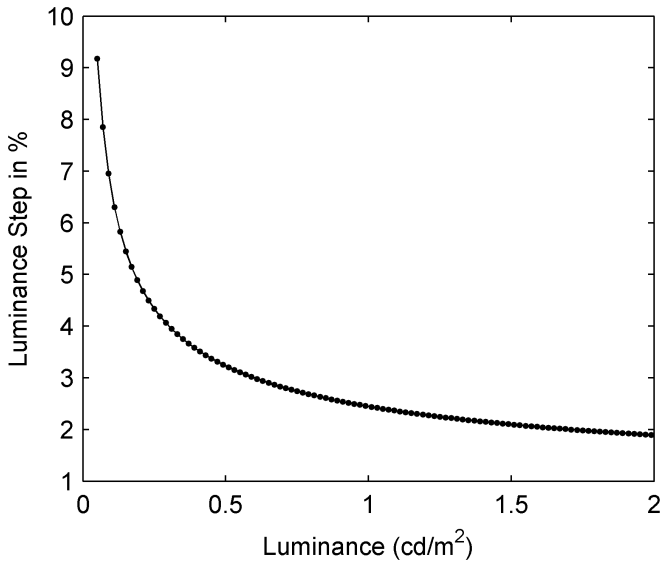


Figure 2.1: DICOM curve. DICOM perceptually-uniform luminance differences, in percent luminance, as a function of average luminance. Beyond the right edge of this plot, the steps reach a minimum of about 0.8% at 200 cd/m², and then rise slowly with higher luminance.

light sources can be added linearly, and extending this the EVL contribution of every point in the scene may be considered by integrating over the entire field of view. For display applications, the EVL resulting from the illuminance reaching the eye from every point on the display may be integrated, and if desired, the EVL due to illuminance from the surrounding room can be added.

2.2.2 Contrast Visibility Model

The human visual system's ability to discern luminance differences has been quantified by the notion of contrast sensitivity, i.e. by finding the detection thresholds of luminance differences at different spatial frequencies. Barten [8] provided a well-known signal-to-noise based model of contrast sensitivity that provides a good fit to a variety of earlier experimental data sets. The model gives the modulation threshold as a function of spatial frequency, luminance, and image (or object) size. As such, it is well-suited to synthetic patterns such as sinusoidal gratings, but it remains difficult to apply to natural images and scenes.

Barten's model was adapted to a very specific visual task in the creation of the DICOM standard display function [9], designed to ensure that the steps between quantized gray levels in medical imaging systems are distributed in a perceptually uniform way. DICOM used a contrast-detection task involving a simple sinusoidal grating of 4 cycles per degree of visual angle, cropped to 2x2 degrees and surrounded by a uniform luminance equal to the mean of the sinusoid. The Barten model was used to predict the visibility threshold in terms of amplitude of the sinusoid at different average luminances ranging from 0.05 to 4000 cd/m². Over a wide luminance range, contrast sensitivity peaks near 4 cycles/degree, so DICOM conceptually describes a conservative estimate of the contrast threshold, i.e. the size of the smallest detectable luminance difference, as a function of average luminance. In the low-luminance range especially interesting to the present work, the threshold luminance differences are between 2 and 10% of the average luminance, as shown in Figure 2.1. Even though it was designed for a very specific imaging target, DICOM provides a model of visibility that may be applied with caution to other situations.

2.3 Laboratory Setup

The laboratory setup consisted of three components: a high-contrast display, a glare source, and a viewing box. The display was used to create visual stimuli for the paired comparison staircase task. Near the display, in the same visual field, a LED luminaire was used to create a bright point of light to act as a glare source, inducing veiling glare in the eyes of the observer. The glare source was decoupled from the display in order to allow extremely high luminance differences between the glare and the display black, as well as to eliminate any possibility of flare within the display itself. The arrangement simulates either a situation in which the glare is part of the image on an HDR display or a situation with an external glare-inducing light source. To control light reflections, eliminate ambient light, and create a controlled viewing situation, a viewing box was built around the display and glare source with black baffles to trap stray light and a chinrest to fix the observer's position.

2.3.1 Double-LCD Monitor

The experiment utilized a FIMI-Philips 18" SXGA (1280×1024) monochrome medical imaging monitor, with two liquid-crystal (LC) panels in series in front of an adjustable fluorescent backlight. The display's luminance output was carefully characterized, and its spatial and temporal inconsistencies were measured and controlled. Spatial luminance errors did not exceed 2%, and temporal fluctuations were controlled to a standard deviation of 0.15%.

The main feature of the double-LCD is its extremely high contrast. The two LC panels in series provide over 5 log units of luminance range, with a contrast ratio of approximately 40,000:1. The display's backlight was set to provide a relatively low white point of 14.67 cd/m², which resulted in an extremely dark black point of 0.0003

cd/m². Note that this black level was reliably measured above the minimum luminance of the PR-680L spectroradiometer, 0.0002 cd/m². Based on extensive characterization measurements, a monotonic path in the 2D code value space (two 8-bit monochrome LC panels) was chosen which provided a distribution of luminance levels roughly uniform in $\Delta\text{luminance}/\text{luminance}$ and gave 511 discrete values, effectively 9 bits. Desired luminance values were mapped through this path to select drive values for each image.

At a viewing distance of 75 cm, the display had 46.7 screen pixels per degree of visual angle. The center region of 20x20 degrees (about 950 pixels square) was characterized for spatial uniformity correction and used for the experiment. An image showing a 1-degree grid on a white background was used for aligning the display measurements. A view of the experimental set-up, including the visible portion of the display screen, the grid image, image stimuli positions, and the external glare source, is shown in Figure 2.2.

2.3.2 Glare Source

A Philips Lexel LED DLM1100 downlight module, capable of 1000 lumen, was used as a glare source near the display. Using a separate glare source ensured an extremely large luminance contrast between the display and the glare source, allowing the emulation of the highest-brightness, highest-dynamic range displays without the complications of segmented backlights and potential uncertainties in light distribution. The LED module had a flat diffuser surface which provided a uniform circular spot of about 6 cm in diameter. For the experiment, this diffuser was masked with opaque black paper to provide a small circular spot with area of one square degree of visual angle. The device's light output was characterized by measuring the luminance of the diffuser surface with the PR-680L spectroradiometer. Luminance levels of 1,000 and 10,000 cd/m² with a color temperature of 4400 K, equal to the white point of the display, were used in the experiment.

2.3.3 Viewing Box

Because of the very low luminance levels used in the experiment, the experimental setup was very sensitive to ambient light and reflections. Additionally, because of the desire to model the complete visual field, an uncomplicated, preferably zero-luminance surrounding was desired. For these reasons, a viewing setup was created with a black box, constructed of matte black foam board, surrounding the viewer's peripheral vision and incorporating light traps to control stray light and reflections, thus controlling the entire visual field. The surfaces visible to the viewer were all angled so that they were not illuminated by the display or glare source. The result was a completely black visual field with the exception of the 20-degree square display and the glare source immediately below it. The viewing position was constrained at 75cm with a chinrest. A top-view of the interior of the box itself is shown in Figure 2.3.

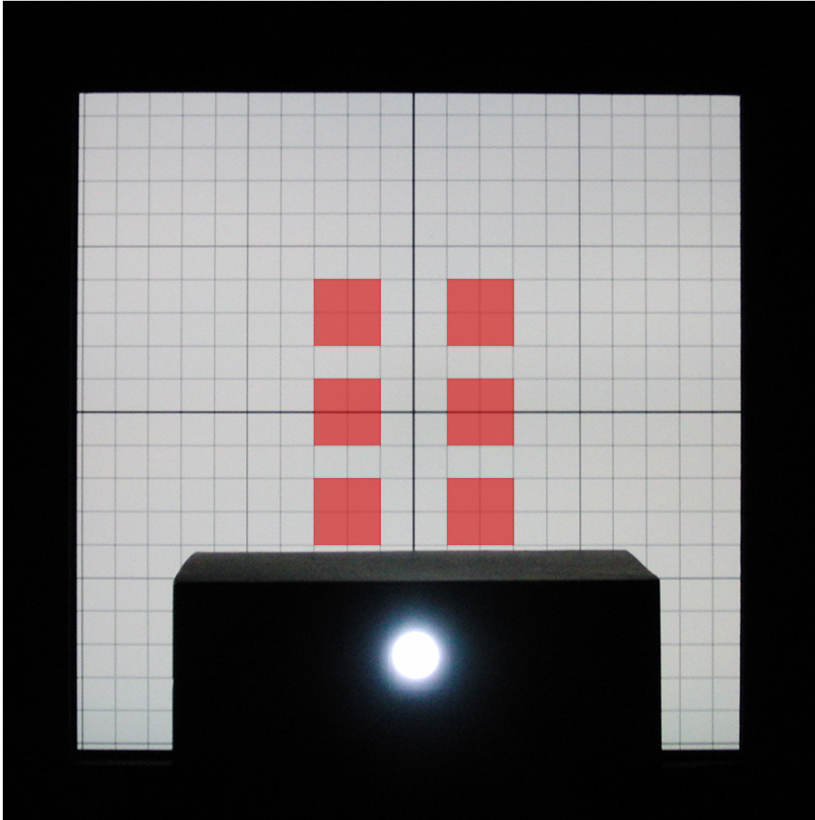


Figure 2.2: Observer's view. Photograph of the double-LCD taken from the viewpoint of the observer; showing the 1-degree grid image on the center 20x20 degree region of the screen and the lit LED glare lamp at the bottom. The red shaded squares show the possible positions of the image pairs at distances of 4, 7, and 10 degrees of visual angle from the glare source. The apparent “bloom” around the glare source which is a result of flare in the camera looks remarkably similar to the intra-ocular glare seen by the observers in the experiment.

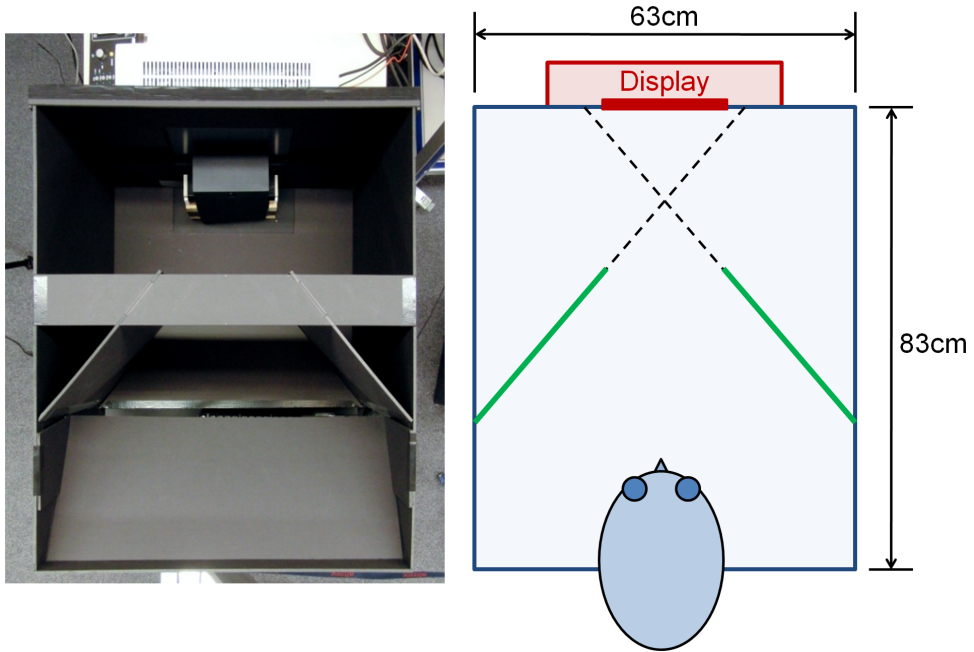


Figure 2.3: Viewing Box. Top view of the viewing box with its top removed in a photograph (left) and in a diagram (right), showing the baffles (bold green) which were angled so that the surfaces facing the observer were not illuminated by the display. Similar baffles were used above and below the observer's head, as can be seen in the photograph. Also visible in the photograph, directly below the display, is the box housing the LED glare source. The wide horizontal strip visible is a structural member holding the baffles in place.

2.4 Experiment

An adaptive staircase methodology with paired comparisons was used to measure thresholds for black level difference detection for several combinations of four experimental factors: glare source luminance, glare source angle in the visual field, image luminance and image content.

2.4.1 Adaptive Threshold Testing

The experiment was conducted using a two-alternative forced choice (2AFC, a.k.a. paired comparison) methodology behind which a staircase rule was used to adaptively present

stimuli to each observer. This means that in successive image pair presentations, one image, randomly the left or the right, was always a reference image with maximally dark black level. The other image (the sample) had a lifted black level, and the observer's task was to choose the image with the darker black. At the beginning of each staircase the sample image had a very obviously lifted black level, nearly assuring a correct response. With each correct choice of the reference image, the sample black level was lowered closer to the reference image. With each incorrect choice, meaning the observer could not distinguish the two black levels presented, the sample black level was raised. In general, this type of staircase methodology converges to a threshold estimate. In this experiment, a weighted up-down staircase [10] was used, meaning that the up and down step sizes were unequal; setting their ratio to 1:3 forced the convergence to the X_{75} point, the sample black level at which the observer was correct 75% of the time. In a paired-comparison experimental set-up the X_{75} point can be assumed to be the level where an observer has a 50% chance of actually seeing the black level difference – which is the definition of a 50% just noticeable difference (JND).

Each staircase in the experiment was allowed to proceed to 7 inversions, or changes in direction due to the sequence of incorrect and correct responses. The step size was halved after the first and fourth inversion to allow quick, rough convergence at the beginning and precision at the end. The first three inversion points were discarded, and the remaining four were averaged to result in the threshold.

2.4.2 Design

The experiment used four fixed factors in a partial factorial design. The factors were the luminance of the glare lamp, the visual angle between the glare lamp and the image pair, the average luminance of the images being presented, and the six images. Because for obvious reasons the position of the glare source was fixed, the visual angle between the image content and glare source was varied by vertically displacing the image pairs on the display. Figure 2.4 illustrates the experimental conditions in a 3-dimensional space. Conditions 2 and 3 can be thought of as baseline cases wherein the glare-free black level threshold was determined for each image luminance. Conditions 4a, 4b, and 4c consisted of visual angle variations at the high-glare, lower image luminance combination, and likewise 5a, 5b, and 5c vary visual angle at higher image luminance. Conditions 6 and 7 included the lower glare luminance level at a single visual angle. Finally, Condition 1 (not shown in the figure, with factor levels the same as 5b) was used as a training staircase to familiarize observers with the test setup and paired comparison methodology, and was not analyzed.

With the exception of Condition 1, which was always presented first in the experiment, the presentation order of the remaining conditions was balanced over observers to avoid any systematic influence of learning effects or observer fatigue. Within each condition, staircases for all six images were conducted in an intermingled, randomized fashion until they all reached completion, which made it unlikely that the observers could

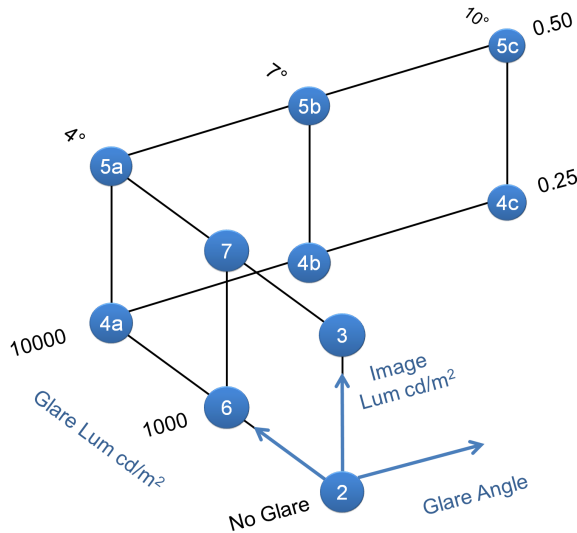


Figure 2.4: Experimental Conditions. Diagram of the experimental conditions on axes of glare angle, glare luminance, and image luminance. Six images were presented with each condition. Not shown, Condition 1 was a short training staircase with factor levels matching 5b; it was excluded from the analysis.

understand or manipulate the staircases' progress. Further, all images and visual angles comprising Conditions 4a, 4b, and 4c were intermingled together, and likewise with those comprising Condition 5. No time constraints were made on the observations, but observers were given a few minutes to adapt to the low image luminance at the beginning of the experiment, and they were instructed not to look directly at the glare source to avoid retinal after-images from reducing their visual sensitivity.

2.4.3 Image Stimuli

The six source images used in the experiment were chosen to include both diagnostic patterns and pictorial content. Because these were small images, 2x2 degrees of visual angle, the pictorial content was chosen to be cropped image details, rather than entire images. The intent was to represent a situation where the image detail was in a shadow region of a larger image which potentially included high-luminance regions that induce glare. As shown in Figure 2.5, there were two cosine patterns, *cosine* at 4 cycles per degree and *cosine2* at 1 cycle per degree. The 4 cyc/deg, 2x2 degree cosine pattern is the same as is prescribed by the DICOM standard [9]. The image *curls* is a close-up of curly dark hair, containing high-frequency, high-contrast detail, and *palm* is a tight crop of a



Figure 2.5: Images used in the experiment. Upper, left to right: cosine, cosine2, curls. Lower, left to right: eye, nose, palm. Each image was presented at a size of two degrees of visual angle square, which can be approximated by viewing this page from 60cm.



Figure 2.6: Image processing steps. To create black level variations, first the images were normalized to the same mean luminance. Next, the normalized image was scaled to move the mean luminance to the average luminance required for the experiment, and finally, the black level was raised by scaling the luminance down and shifting it up, compressing the range while preserving the maximum luminance value.

human palm and fingers, with low frequency detail and large bright regions. The image eye is a crop of a black dog's eye, with a low-key rendering and bright highlights on the eye and fur. Nose is a relatively low contrast view of a light-haired dog's face.

The images were all normalized to have the same mean luminance. Starting with linear luminance values in the interval $[0, 1]$, they were divided by their mean luminance values, respectively. This resulted in, for example, eye having a much brighter maximum luminance than palm, but ensured that all images had the same integrated luminance across their 2×2 degree size, encouraging a uniform state of adaptation and keeping total eye illuminance fixed as images were displayed in sequence during the experiment. For each image, at each prescribed average luminance level, a series of 49 images varying in black level luminance was pre-computed: these comprised all the possible stimuli for a

staircase. The steps in this processing path are shown in Figure 2.6. The 49 black levels L_i were logarithmically spaced, ranging from 0.0001 to 1. The black level adjustment was done with an affine transformation of the linear luminance, compressing the luminance range and shifting it higher.

After these steps were performed in linear luminance, the images were transformed using the inverse model of the double-LCD display to 8-bit device code values, including the front- and rear-LC split and the spatial uniformity correction. Because of the uniformity correction, different output images were created for each spatial location on the display: both for the left and right versions used in each paired comparison (which also mirrored the content left-to-right) and for the different vertical displacements used to vary the visual angle from the glare lamp. The background rectangle of 0.05 cd/m^2 , also corrected for uniformity, was included with the saved image stimuli. During the experiment, the image pairs were composited by a Java program that handled the staircases and image display. The 8×10 visual degree background rectangle was centered on the display, with the image pair separated by a 2-degree gutter.

2.4.4 Observers

In total, 23 people participated in the experiment, and the data analysis included 22 of them. One observer's data showed a surprisingly high mean threshold in black level difference. Because he explained after the experiment that he found the glare visually uncomfortable, and indeed it appeared that his results didn't depend on glare angle like the rest of the population did, his data were discarded. Of the remaining 22 participants, there were 20 observers who completed the whole experiment, and two additional observers who provided partial, balanced results. There were 6 females and 17 males ranging in age from 22 to 59, with a median of 32 years. Characteristics of their eyes were recorded because of the potential effect on intra-ocular glare. All reported normal [corrected] visual acuity: five wore glasses, five wore contact lenses, and 12 had uncorrected (naked eye) vision. Thirteen had eye color categorized as light (blue, green, or grey), and nine were categorized as dark (brown and dark brown). Thirteen of the group hailed from The Netherlands or Belgium. Other nations of Europe, Asia, and North America were represented by individuals.

2.5 Experimental Results

The result of each staircase (for each observer for each condition for each image) was a single number, i.e. the computed threshold in units of step levels, which were logarithmically spaced in luminance. Each threshold was converted to an actual log luminance, and all further computations were done in this space. At the end of the analysis, the mean values were converted to linear luminance. Mean results per condition are shown in Table 2.1.

Table 2.1: Table of experimental results. Summary of black level thresholds measured in the experiment for each condition. For each condition, the factor levels are shown with mean thresholds in both logarithmic and linear luminance (cd/m^2).

Cond	Experimental Factor			Mean Thresh. Log Lum	Mean Thresh. Lum
	Glare Lum	Glare Angle	Image Average Lum		
2	0	–	0.25	-2.04	0.0092
3	0	–	0.50	-1.88	0.013
4a	10000	4	0.25	-1.40	0.039
4b	10000	7	0.25	-1.67	0.022
4c	10000	10	0.25	-1.84	0.014
5a	10000	4	0.50	-1.26	0.055
5b	10000	7	0.50	-1.53	0.029
5c	10000	10	0.50	-1.67	0.022
6	1000	4	0.25	-1.88	0.013
7	1000	4	0.50	-1.67	0.022

To evaluate the significance of these results, a univariate ANOVA was calculated with SPSS. In the ANOVA the threshold log luminance was the dependent variable, and the glare luminance, glare angle, image luminance, and image content were the fixed factors. Additionally, observer was included as a random factor. The model included all main effects and two-way interactions. All factors, except observer ($p = 0.09$), were significant.

Comparing all the experimental factors, the strongest effects found were for glare luminance ($p < 0.001$, $\eta_p^2 = 0.90$) and glare angle ($p < 0.001$, $\eta_p^2 = 0.89$), followed by image luminance ($p < 0.001$, $\eta_p^2 = 0.41$) and finally, much weaker, image content ($p < 0.001$, $\eta_p^2 = 0.19$). Several interaction effects were found to be significant (in order of effect size): image \times observer, image \times glare angle, and image \times glare luminance; however, the effect sizes for these were all smaller than the weak main effect of image itself.

Figure 2.7 shows the measured black luminance difference threshold as a function of visual angle for all ten experimental conditions, averaged over image content and observer. The conditions with a high glare luminance are shown in blue squares and red circles, for the high and low image luminance, respectively, and the conditions with a low glare luminance are shown in magenta triangles and green diamonds. The no-glare conditions are shown at the far right of the plot. The pronounced increase in luminance

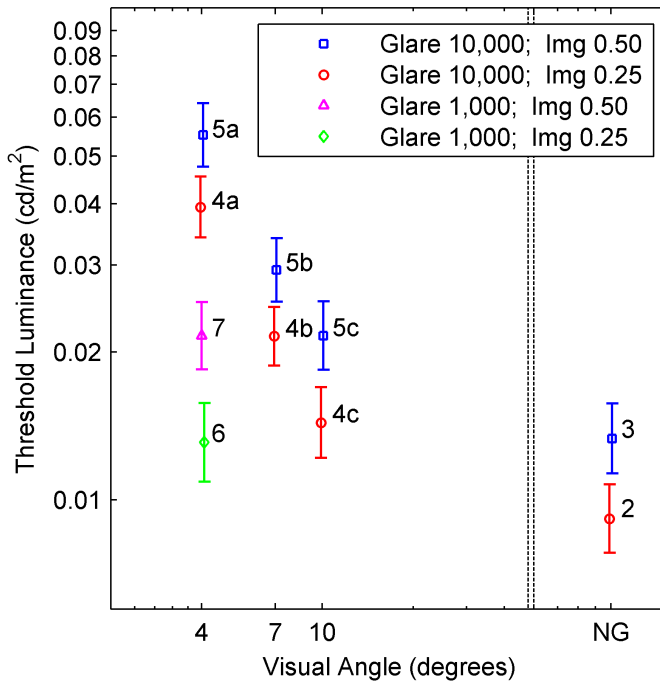


Figure 2.7: Experimental results. Black luminance thresholds with 95% confidence intervals for all ten experimental conditions (labeled). High-glare (10,000 cd/m^2) conditions for two average image luminance levels (0.50 cd/m^2 in blue squares; 0.25 cd/m^2 in red circles) and low-glare (1,000 cd/m^2) conditions for two average image luminance levels (0.50 cd/m^2 in magenta triangles; 0.25 cd/m^2 in green diamonds) are shown as a function of visual angle from the glare source. The no-glare condition is labeled on the x-axis as “NG.”

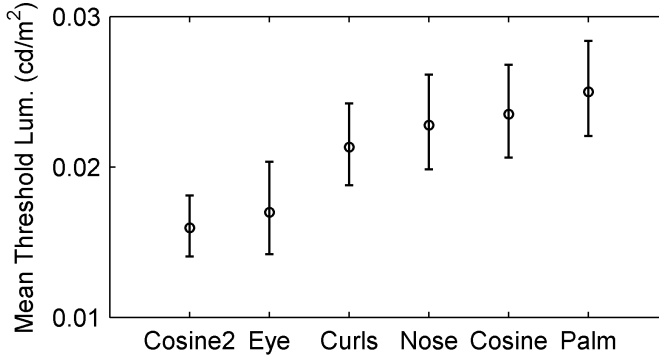


Figure 2.8: Experimental results per image. Mean black luminance thresholds with 95% confidence intervals for the six images used in the experiment. The pair of images cosine2 and eye is significantly lower than the group of the four remaining images; however, eye and curls are not significantly different from one another.

difference at small angles, especially for the conditions with a high glare luminance, shows the clear impairing effect of glare proximity on black level discrimination, and is similar to the $1/\theta^2$ behavior of the CIE glare equation. At the low glare luminance at 4 degrees from the glare source, a smaller apparent rise in the black threshold is seen.

Effect of Image Content

Image content also had an effect according to the ANOVA, and the thresholds for each image are compared in Figure 2.8. As the figure suggests, a Tukey post-hoc analysis confirmed that *cosine2* and *eye* formed one group with a black level difference threshold significantly lower than the group of the four remaining images. However, *eye* and *curls* were not found to be significantly different from each other. There is no obvious trend with spatial frequency or pictorial vs. synthetic image content. Rather, the experimenters observed that the low-threshold images both had relatively large black areas, which were useful in discerning black level differences slightly more critically than was possible with the other images.

Observer Variation

Looking only at observers who saw all experimental conditions, another ANOVA was computed, this time with condition as a fixed factor and observer as a random factor. Both were found to be significant, and the estimated marginal means for each observer were obtained. The marginal means were used as the dependent variable in a subsequent

ANOVA with binary independent variables representing the observers' characteristics of glasses, contact lenses, male, young (≤ 32 years), and dark eyes. This analysis found contact lenses the only significant factor, with a medium effect size ($p < 0.01$, $\eta_p^2 = 0.42$). Just not significant was glasses ($p = 0.067$, $\eta_p^2 = 0.24$). There was no significant effect of being male ($p = 0.24$) or young ($p = 0.71$) or having dark eyes ($p = 0.45$). It seems rational that the extra optical surfaces and scattering of contact lenses provided this significant upward change in measured thresholds, and that eyeglasses provided a similar trend.

2.6 Modeling Results

An excellent fit of the experimental data was made using a combination of literature models and an empirical description of adaptation luminance. The DICOM model, based on Barten's CSF, predicts the visibility threshold of a luminance difference at a given average luminance, for low-contrast images in an average surround condition. The present experiment, however, used high-contrast images and a non-uniform surround, which means that the DICOM model is somewhat misused. In fact, applying DICOM directly to the no glare conditions for the average image luminances predicted thresholds about 20% higher than were measured. It is a step further to apply DICOM to the glare conditions, requiring an assumption – that the average luminance can be computed by adding equivalent veiling luminance to the average image luminance. Thus, here enters the glare model. EVL from the glare lamp is the biggest component, but glare from within the image and the background cannot be ignored. For each experimental condition, these glare components can be added in order to compute an effective average luminance, for which DICOM can be used to predict the corresponding visibility threshold. Doing this, the results were good, with $R^2 = 0.90$ for the mean data ($R^2 = 0.24$ over all observations), but showed a systematic under-prediction of the thresholds at smaller visual angles and over-prediction for the no-glare conditions. The model was further improved by accounting for adaptation luminance.

It was noted that the average image luminance, even when adding equivalent veiling luminance, was not the best input into the DICOM relationship because of the non-homogeneous field of view. To improve the model, a simple conceptual model of luminance adaptation was made, and its parameters were fit to the experimental data. The luminance of everything in the field of view was weighted by a pair of circularly-symmetric Gaussian sensitivities and integrated. One Gaussian was chosen to be very narrow, corresponding conceptually to foveal sensitivity, and the other much wider, to take into account the surroundings, including the equivalent veiling luminance and the glare source itself, which was in the field of view. The resulting adaptation luminance was used with DICOM to predict luminance difference thresholds, as shown in the following equation. Working from the inside out, first the the EVL of the image (including its surround; truly, the entire visual field) is computed using the CIE glare equation, and added to the image; then the adaptation luminance (Y_{adapt}) is computed using the

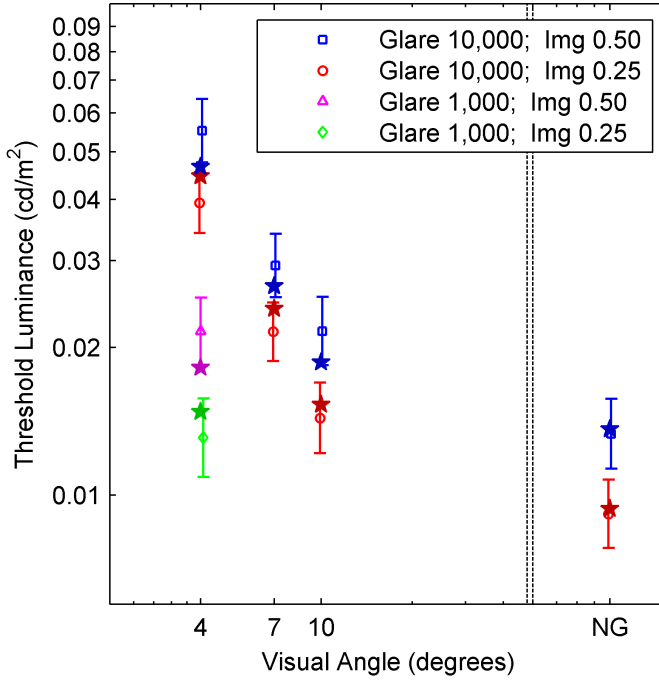


Figure 2.9: Experimental results and improved model. Experimental mean luminance thresholds with 95% confidence intervals for all experimental conditions shown as a function of visual angle from the glare source (the no-glare condition is labeled on the x-axis as “NG”). The predictions of the improved model taking into account adaptation luminance are shown as stars near the corresponding experimental values. The model fit to the mean values resulted in a $R^2 = 0.95$.

empirical model from this experiment; and finally the DICOM model is used to compute the corresponding difference threshold.

$$threshold = \text{DICOM} \left(Y_{adap} (image + \text{EVL} (image)) \right) \quad (2.1)$$

The standard deviation of the two Gaussians and a coefficient for their linear combination were optimized to fit the data. The resulting improved model fits the mean data very well ($R^2 = 0.95$; $R^2 = 0.25$ over all observations). More specifically, the improved model improves the fit of both the lower threshold in the no glare conditions and

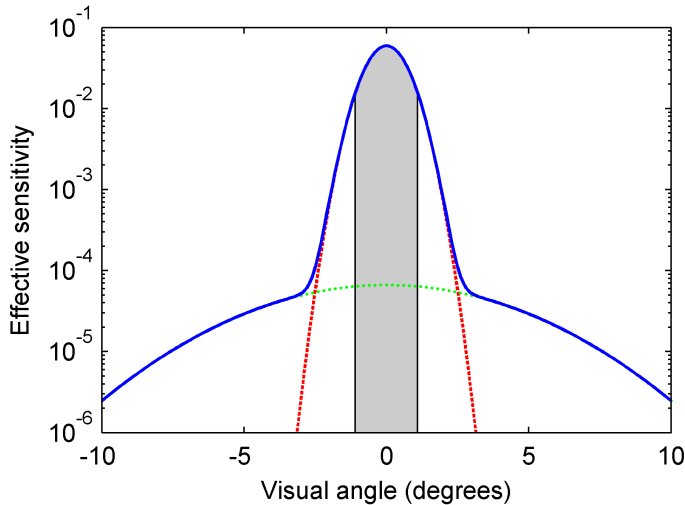


Figure 2.10: Spatial adaptation sensitivity. Spatial sensitivity curve used in the computation of adaptation luminance as a function of visual angle. The dashed red curve (std. dev. 0.67) is weighted 99.35% in linear combination with the dotted green curve (std. dev. 3.9) weighted 0.65% to create the solid blue curve. Looking at the area beneath the curve, the central 92%, based on Moon and Spencer's suggestion, is shaded gray, corresponding to a 2.2 degree foveal region.

of the steep behavior at smaller visual angles to the glare source. However, it also predicts a convergence of the different image luminance conditions at small angles with the glare source which was not observed in the experiment. The model converges simply because the modeled EVL contribution of the glare lamp dominates that of image luminance with close glare proximity, so apparently something minor is still missing from the model. The experimental results and fitted points (stars) are shown in Figure 2.9. The empirically-derived spatial sensitivity curve is shown in Figure 2.10. It is a linear combination of Gaussians, one with standard deviation of 0.67, weighted 0.9935, and another with standard deviation of 3.9, weighted 0.0065. This distribution of sensitivity was reached via a model fit, but it corresponds well to an approximation suggested by Moon and Spencer [11]. They suggested that the luminance reaching the fovea contributes 92%, and the luminance of the surrounding area integrated by a $1/\theta^2$ function similar to the glare equation contributes the remaining 8%. Comparing the present model to the Moon and Spencer relationship, there is essentially a difference in foveal size approximation, for which literature values from 1 to 2 degrees exist. The central 92% of the present empirical fit corresponds to a foveal size of 2.2 degrees, as shown by the shaded region in Figure 10.

The sensitivity curve as a combination of Gaussians then appears a reasonable description of foveal/background adaptation.

2.7 Conclusions

In this study, a tightly-controlled paired comparison experiment was conducted to measure the impairment of the visibility of dark details in a displayed image due to veiling glare, and this impairment was successfully modeled with components corresponding to intra-ocular glare, luminance adaptation, and contrast sensitivity. As measured, the black level difference threshold increases (meaning observer sensitivity decreases) with any of the following, in order of importance: glare source distance (visual angle) decrease, glare source luminance increase, or average image luminance increase. Relative to an extremely low reference black level of 0.0003 cd/m^2 , a threshold black level was found to be as small as 0.0092 cd/m^2 with no glare source present and as large as 0.055 cd/m^2 with a small glare source of $10,000 \text{ cd/m}^2$ at 4 degrees of visual angle from the black level comparison.

The black level difference threshold was successfully modeled with a combination of (1) the CIE glare model, which describes the intra-ocular scattering within the human eye as a function of visual angle to the glare source, (2) an adaptation luminance computed as a spatially-weighted function of the luminance seen in the visual field, and (3) the DICOM model, which predicts the visibility of luminance differences as a function of luminance. The resulting model can be used to predict whether a luminance difference within an image in a given environment with glare caused by the image itself and/or other light sources will be visible, which can be beneficial to designing displays as well as viewing situations.

2.8 Bibliography

- [1] H. Seetzen, L. Whitehead, and G. Ward. A High Dynamic Range Display System Using Low and High Resolution Modulators. In *SID Symposium*, volume 34, pages 1450–1454. SID, 2003.
- [2] CIE. Research note: Disability glare. *Vision and Colour: Physical Measurement of Light and Radiation*, 135(1), 1999.
- [3] Helge Seetzen, Wolfgang Heidrich, Wolfgang Stuerzlinger, Greg Ward, Lorne Whitehead, Matthew Trentacoste, Abhijeet Ghosh, and Andrejs Vorozcovs. High Dynamic Range Display Systems. In *ACM SIGGRAPH*, pages 760–768. ACM, 2003.
- [4] John J. McCann and Alessandro Rizzi. Retinal HDR images: Intraocular glare and object size. *Journal of the Society for Information Display*, 17(11):913–920, 2009.

- [5] S. Swinkels, R. Muijs, E. Langendijk, and F. Vossen. Effect of Backlight Segmentation on Perceived Image Quality for HDR Displays. In *IDW*, volume 2, 2006.
- [6] Erno H.A. Langendijk and Martin Hammer. 14.4: Contrast Requirements for OLEDs and LCDs Based on Human Eye Glare. In *SID Symposium*. Society for Information Display, 2010.
- [7] Rafał Mantiuk, Scott Daly, and Louis Kerofsky. The Luminance of Pure Black: Exploring the Effect of Surround in the Context of Electronic Displays. In *Proc. of Human Vision and Electronic Imaging XXI, IS&T/SPIE's Symposium on Electronic Imaging*, pages 7527–31. IS&T/SPIE, 2010.
- [8] Peter G. J. Barten. Physical Model for the Contrast Sensitivity of the Human Eye. In *SPIE*, volume 1666, pages 57–72. SPIE, 1992.
- [9] DICOM. Digital Imaging and Communications in Medicine (DICOM), Part 14: Grayscale Standard Display Function. *National Electrical Manufacturers Association*, PS 3.14, 2008.
- [10] C. Kaernbach. Simple Adaptive Testing with the Weighted Up-down Method. *Perception & Psychophysics*, 29(W):227–229, 1991.
- [11] Parry Moon and Domina Eberle Spencer. The Visual Effect of Non-Uniform Surrounds. *Journal of the Optical Society of America*, 35(3):233–248, 1945.

Utilizing Metamerism in RGBW OLED Displays

Abstract

Displays that employ RGBW primaries have demonstrated greater power efficiency than similar displays with only RGB primaries. Unfortunately, RGBW systems with spatial light modulators, such as LCD flat panels and DMD projectors, have typically traded color accuracy for improvements in power efficiency. This paper presents a color-processing algorithm for emissive RGBW OLED displays that preserves colorimetric accuracy while still reaping the efficiency benefits of RGBW. RGBW extensions of additive RGB color models are discussed, along with a methodology for deterministically choosing RGBW solutions. A flexible image-processing path is illustrated that may be optimized for power efficiency, uniformity, and color gamut.¹

3.1 Introduction

Additive displays have long depended on RGB primaries to synthesize color mixtures. Recently, displays of a variety of technologies using RGBW primaries have emerged, promising improved efficiency through higher luminance and/or lower power consumption. Color reproduction in these displays often suffers as a result of the desire to boost efficiency, but this tradeoff is not always necessary, depending on the display technology and color processing algorithm applied. “Perfect” color reproduction is always a matter of perspective. Here, it is asserted that in a display system, perfection means accurate reproduction of the color specification encoded in the signal it receives.

Many display types, such as LCD and digital micromirror device (DMD), rely on spatial light modulators to attenuate a backlight, projector bulb, or other always-on, full-field light source. In these displays, efficiency is determined by how much of the light generated by the always-on light source is transmitted to the viewer, leading to

¹This chapter is based on research presented at the International Congress of Imaging Science, 448-451 (2006), “Perfecting the Color Reproduction of RGBW OLED,” Michael J. Murdoch, Michael E. Miller, and Paul J. Kane, and disclosed in United States Patent 6,897,876 (2005), “Method for Transforming Three Color Input Signals to Four or More Output Signals for a Color Display,” Michael J. Murdoch, Michael E. Miller, and Ronald S. Cok.

drive schemes that maximize the use of all four RGBW primaries to synthesize the display white point. Algorithms for driving displays with spatial light modulators typically add an amount of luminance from the W primary that is correlated with the amount of input RGB, thus augmenting the light output. An example is provided by Sampsel for DMD displays [1]. Such algorithms result in color reproduction error, displaying at least some colors less saturated and/or lower in luminance as compared to the color reproduction of an otherwise similar RGB display.

More sophisticated algorithms for light modulator displays mitigate the color reproduction error by modifying the corresponding RGB intensities where possible; however, this approach cannot both repair the effect for all colors and maintain the efficiency improvement. Examples are given by Kunzman & Pettit, who describe a DMD RGBW implementation which preserves color accuracy for some colors [2]; and by Lee et al., who give an algorithm for a TFT LCD RGBW display in which white is added to colors in different amounts to make the color error less objectionable [3].

Emissive displays, such as OLED displays, utilize an array of light-emitting subpixels, meaning that the efficiency of the display is dependent on the efficiencies of the subpixels in use. A filtered white RGBW (W-RGBW) OLED uses an independently controlled white emitter at each subpixel site with color filters for each of the RGB subpixels and no filter for the W subpixel. Because they are unfiltered, a W-RGBW OLED display's W subpixels are much more efficient than its RGB subpixels, so efficient drive schemes utilize the W primary as much as possible and the RGB primaries as little as possible. It has been shown that a W-RGBW OLED panel requires half the power, on average, of an otherwise similar W-RGB OLED panel, without color error [4]. The present paper outlines an algorithm for accomplishing this combination of power savings and color accuracy.

3.2 Additive Model of RGBW

The light output of many color display systems can be modeled using a combination of a set of nonlinear characteristic curves and a linear primary matrix. The familiar primary, or phosphor, matrix is used to describe the linear addition of color in the display by computing the XYZ tristimulus values that a given linear RGB intensity input triad will produce. This is simply a linear combination, as in Eq. 3.1.

$$\mathbf{P}_{3 \times 3} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3.1)$$

The columns of the 3×3 primary matrix $\mathbf{P}_{3 \times 3}$ are typically filled with the XYZ tristimulus values of each primary, scaled such that input linear RGB intensities (1, 1, 1) result in the tristimulus values of the desired display white point. This implicitly defines the maximum luminance, or unit intensity, for each primary.

The primary matrix relationship provides an essential feature through inversion, allowing the prediction of the necessary RGB triad to provide a desired XYZ tristimulus output. For a display with stable primary chromaticities and without any crosstalk or loading effects, the model works very well, and any XYZ tristimulus specification within the RGB gamut is reproduced accurately. XYZ specifications outside the RGB gamut result in RGB intensity values outside the interval $[0, 1]$, which still are useful for modeling but are not physically realizable in the display.

A four primary system can be modeled similarly: the output is the linear combination of four primaries' contributions instead of three. Likewise, the 3×3 primary matrix may easily be extended to a 3×4 matrix, as in Eq. 3.2. The 3×4 primary matrix $\mathbf{P}_{3 \times 4}$ is formed from the 3×3 RGB matrix appended with a fourth column holding the tristimulus values of the W primary, such that an input linear RGBW intensity quad results in an XYZ triad. In this arrangement, two questions are immediately apparent: how to invert the non-square matrix $\mathbf{P}_{3 \times 4}$, and how to normalize its fourth column.

$$\mathbf{P}_{3 \times 4} \begin{bmatrix} R \\ G \\ B \\ W \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3.2)$$

Normalization for a 3×3 primary matrix is well understood: using Eq. 3.1, unit intensity in all three RGB primaries results in the XYZ tristimulus values of the display white point. Proper normalization for the additional W column in $\mathbf{P}_{3 \times 4}$ is less obvious and depends on how the display will be used. This discussion will progress focusing on what is best for W-RGBW OLED displays, which might not be what is best for displays using spatial light modulators. Some reasons for this distinction will be discussed later.

Unfortunately, the 3×4 RGBW primary matrix is not invertible, with the practical implication that given desired XYZ tristimulus values, there is not a unique RGBW solution; rather, there are many that will give equivalent results. A goal of this paper is to outline a method for choosing intelligently and deterministically from the possible solutions.

3.3 White Equivalence

Important to the use of a W primary in an additive display is the concept of white equivalence. Metamerism is the phenomenon whereby two spectrally dissimilar stimuli integrate to the same XYZ tristimulus values, implying that a viewer with normal color vision would see them as the same color, assuming similar viewing conditions. A W-equivalent RGB intensity triad is a combination of RGB intensities that produces a metamer of some amount of W primary intensity. The normalized W-equivalent RGB intensity triad, \mathbf{W}_{RGB} , is scaled such that the maximum of the RGB intensities is unity. This works for any

color; as long as the chromaticities of the W primary are within the RGB gamut, all three W_{RGB} intensities are positive.

This normalization can be used to define the unit intensity of the W primary, and thus the scaling of the fourth column of the 3×4 RGBW primary matrix. This ensures that color resulting from the peak intensity of the W primary can be equivalently, i.e., metamERICALLY, reproduced using only the RGB primaries. Note that a convenient case arises when the chromaticities of the W primary are the same as those of the desired display white point; in this case, the W_{RGB} values are (1, 1, 1).

3.4 White Replacement

Because in a W-RGBW OLED display, the W subpixels are much more efficient than are any of the RGB subpixels, an effective concept is that of white replacement. W intensity is equivalent to a combination of R, G, and B intensities, thus the W subpixel can be used in place of a combination of R, G, and B subpixels. Conceptually, this means removing the neutral luminance from an RGB triad of subpixels and transferring it to the W subpixel for an equivalent result. A visual example with one image decomposed into its R, G, B, and W components is shown in Figure 3.1.

A bounding example of this is to compute for each image pixel the $\min(R, G, B)$, which may be thought of generally as neutral luminance, subtract it from each of the R, G, and B values, and assign it to W. This is termed 100% white replacement, as all possible neutral luminance has been transferred from the RGB to the W subpixel. Similarly, some fraction, termed the white mixing ratio (WMR), of the neutral luminance may be transferred. Equations 3.3 and 3.4 show the transfer from RGB to W, resulting in R' , G' , B' , and W. Use of varying WMR values offers a range of solutions while maintaining a metameric match to the original color. They range from $\text{WMR} = 0$, corresponding to a strictly RGB solution that does not utilize W, to $\text{WMR} = 100\%$, corresponding to the transfer of as much neutral luminance as possible to the W subpixel. When the W subpixel is more efficient than the RGB subpixels, a WMR of 100% achieves the highest possible display efficiency.

$$W = \text{WMR} \cdot \min \left(\begin{bmatrix} R \\ G \\ B \end{bmatrix} \right) \quad (3.3)$$

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix} - W \quad (3.4)$$

Some similarity might be seen between this algorithm and CMYK printing with undercolor removal (UCR) or gray component replacement (GCR), in which dark colors

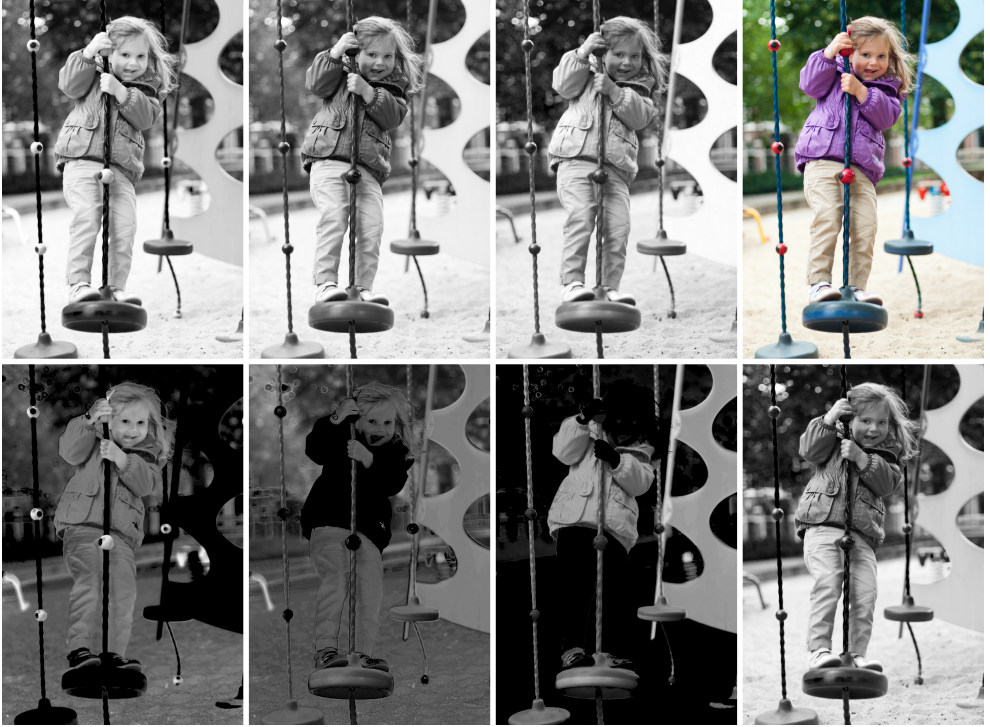


Figure 3.1: Comparison of RGB and RGBW. In the upper row are normal intensity images of (left-to-right) the red, green, and blue color channels along with the resulting full-color image (far right). In the lower row are intensity images of (left-to-right) the red, green, blue, and white color channels using as much white as possible ($WMR = 1$). The resulting full-color image is identical to the top-right image.

are formed using black ink to replace large amounts of CMY ink. In printing, this is done for cost, to avoid physical problems with ink quantity, and/or to prevent color errors that arise when forming neutrals from CMY combinations. Significantly, white replacement in an additive RGBW display can be justified entirely by efficiency and need not change the color reproduction at all.

White replacement using WMR between 0 and 100%, assuming that the chromaticities of the W primary are the same as those of the display white point (the combination of unit intensities of the RGB primaries), results in equivalent color reproduction. Of course, if the W-equivalent RGB values are not equal, the W primary is not the same color as the white synthesized from the original three primaries, and the color subtracted from RGB will not be equivalent to the color produced by W. In this case, further compensation must be made.

3.5 When W Is Not Quite White

In a display, “white” is generally defined by the display white point, which might be set manually to meet a specification in a colorimetric space such as xyY, or which might be the physical result of a light source with light modulators and/or color filters. In a display that utilizes a broadband source and color filters, it is generally most efficient to choose a white point equivalent to the broadband color. This is true of both filtered-white OLED displays and backlit or projection spatial light modulation displays. However, regardless of the display technology, the inherent broadband color might not be the same as the desired display white point. When this is true, the basic white replacement algorithm will introduce color error; for example, a W primary that is yellowish compared to the display white point will introduce a yellow bias to the extent that it is used to replace RGB intensity that is equivalent to the white point.

The solution is to account for the color of the W primary when transferring luminance from the RGB subpixels to the W subpixel, using the concept of W equivalence introduced earlier. The three W-equivalent RGB intensity, W_{RGB} , values are used to scale the RGB input before the minimum is computed, as in Eq. 3.5.

$$\begin{bmatrix} W_R & 0 & 0 \\ 0 & W_G & 0 \\ 0 & 0 & W_B \end{bmatrix}^{-1} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} R_n \\ G_n \\ B_n \end{bmatrix} \quad (3.5)$$

The result of this scaling is to transform the RGB intensity values, which are by definition normalized such that an equal-RGB triad produces a color having the chromaticity coordinates of the display white point, to W-normalized RGB, or $R_n G_n B_n$, in which an equal- $R_n G_n B_n$ triad produces a color having the same chromaticity coordinates as the W primary. In the $R_n G_n B_n$ space, the minimum is computed and the WMR fraction of the minimum is subtracted, resulting in $R_n' G_n' B_n'$ values as in Eqs. 3.3 and 3.4. Subsequently,

a renormalization is performed to return to the white point-normalized RGB space, as in Eq. 3.6. Earlier this process was conceptually described as a transfer of “neutral” luminance, an intentionally ambiguous descriptor; truly, it is a transfer of luminance of the color of the W primary, and this is made explicit through the normalization process.

$$\begin{bmatrix} W_R & 0 & 0 \\ 0 & W_G & 0 \\ 0 & 0 & W_B \end{bmatrix} \begin{bmatrix} R'_n \\ G'_n \\ B'_n \end{bmatrix} = \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (3.6)$$

Again, note the convenient case in which the W primary shares the chromaticities of the display white point. In this case, W_{rgb} is (1, 1, 1) and both of the above transformations become identity matrices. Using a W primary close to the display white point, the transforms are likely to be close to identity, providing a small but important correction.

3.6 When W Is Not White at All (RGBX)

The normalization and W-equivalency concepts above are also applicable to systems with a fourth primary that is not near white, in general termed RGBX, where the X can be cyan, yellow, or another color. As long as the X is still within the RGB gamut, the method works without modification. If the X is outside the RGB gamut, one required modification is a change to the definition of W-equivalent RGB intensity. It is useful to define instead the X-equivalent RGB intensity, X_{RGB} , whose values are now scaled such that $\max(|X_{RGB}|)$ is unity. Taking the absolute value is necessary because mathematically reproducing the XYZ tristimulus values of an out-of-gamut X primary requires a negative amount of intensity from at least one of the RGB primaries. A second modification comes in the computation of the $\min(R_n, G_n, B_n)$ value. The negative value or values in X_{RGB} should be used in the normalization step, making some of the normalized $R_n G_n B_n$ intensity values negative. However, these negative values must be excluded when computing the minimum $R_n G_n B_n$ value. Thus, the minimum of the non-negative $R_n G_n B_n$ values should be computed.

Another simple extension can be made to handle more than four primaries. The replacement algorithm can be applied multiple times in series, minimizing power draw by successively transferring luminance to more efficient primaries. Each replacement step transforms three intensities to four; therefore, in subsequent steps when more than three are present, the largest three values should be used. The result is a multi-step transfer of luminance from the least to the most efficient primaries.

3.7 Algorithm Summary

The general white replacement algorithm is shown in the form of a flow chart in Figure 3.2 from linear RGB intensity to linear RGBW intensity. The starting point is linear RGB intensity values in the device-dependent RGB primary space of the display itself.

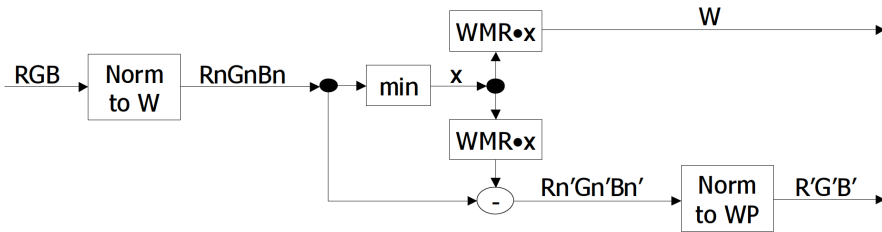


Figure 3.2: RGBW algorithm. Flow diagram for RGB to RGBW via white replacement.

Importantly, an sRGB-encoded signal must be linearized with the proper gamma decoding transform, then rotated from the ITU-R Rec. 709 RGB primaries [5] to the display RGB primaries before proceeding with this algorithm. The first step provides the normalization of RGB from the display white point to the display's W pixel, giving (R_n, G_n, B_n) , then the min function computes x from the normalized RGB values, specifically taking the minimum of the non-negative values, thus delaying the clipping of out-of-gamut colors. The $\min x$ times the WMR becomes the W intensity value as well as the value subtracted from the normalized RGB values. Thus, the WMR parameter controls the amount of luminance transferred from the RGB subpixels to the W subpixel. Finally, the (R_n', G_n', B_n') are normalized back to the display white point, resulting in the output values of $R'G'B'$ and W. After the algorithm is complete, typical steps must be taken to account for the actual display characteristics, for example mapping the linear $R'G'B'W$ intensity values according to the display's nonlinearity.

3.8 Subpixel Arrangements

In a typical RGB display, each of the addressable, logical pixels is made up of three spatially separate subpixels, one of each color. These are most often arranged in a stripe mosaic wherein the three narrow subpixels comprise a square pixel, but may also be arranged in a delta (or delta-nabla) pattern of triangular pixels. At higher pixel resolution, of course, the difference disappears, but at lower resolution the stripe pattern is slightly better for text, and is much more common for monitors and televisions, while the delta pattern can be beneficial for natural images and thus may be seen in camera displays. A perceptual measure for the effects of structure of the pattern and related blur on pictorial content is presented by Nijenhuis & Blommaert [6], and further detailed by Martens [7]. Klompenhouwer [8] discusses the enhancement of text and image sharpness with proper use of subpixel interpolation, which is useful with any pixel pattern. The pattern variants are shown in Figure 3.3.

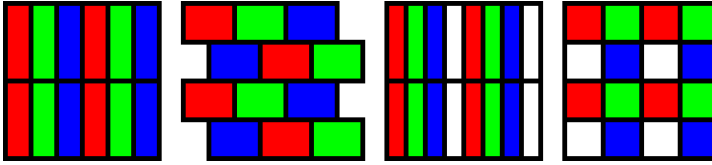


Figure 3.3: Subpixel arrangements. Each of the above panels shows a 2-by-2 grid of four logical pixels. Leftmost is the RGB stripe arrangement with three tall, narrow subpixels comprising a square pixel. Second is the RGB delta pattern, in which each pixel is a triad of nearly-square subpixels, R & G on top with B below, or vice-versa. The third panel shows the RGBW stripe, with the subpixels correspondingly narrower to fit a fourth subpixel in the still-square pixel. Rightmost is the RGBW quad arrangement, with each subpixel occupying one quadrant of the pixel.

Adding a fourth primary to a display system also requires adding a fourth subpixel. Two possible RGBW subpixel mosaics are shown alongside the RGB patterns in Figure 3.3. The RGBW stripe pattern is similar to RGB stripe, except the subpixels are squeezed even narrower to fit a fourth subpixel into the square pixel. The RGBW quad pattern intuitively comprises four square subpixels per pixel. Because the green and white subpixels contribute relatively more luminance to the pixel, and thus more visual weight, they are placed as far as possible from each other – in the stripe pattern this means interspersing them with red and blue, while in the quad pattern this means diagonally opposed. In terms of sharpness and the value of subpixel interpolation, the RGBW stripe behaves similarly to the RGB stripe. RGBW quad is more similar to the RGB delta, but without the diagonal offset, the visible pattern is more gridlike. Comparing RGBW stripe and quad, the stripe’s additional horizontal subpixel resolution is more beneficial in text while the quad’s equal horizontal and vertical subpixel resolution only shows benefits in images. In all RGBW cases, proper subpixel interpolation may be limited by high WMR values because in such cases the W subpixel carries so much of the pixel’s luminance. Such limitations would also be expected when using additional constraints such as those employed by Klompenhouwer.

3.9 Visual Impressions

The RGBW algorithm described relies on colorimetry and metamerism, which robustly model human color vision, and provides what is in principle a perfect match between RGB and RGBW images. For this reason, no perceptual study was necessary. Regardless, the success of the algorithm was confirmed visually both with simulated images and display prototypes. Simulated images were digitally written onto large format Ektachrome transparency film with all of the subpixels and black inactive areas drawn at actual size. Because of the black areas and the array of primary colors, these transparencies were

viewed at a normal display brightness by placing them on a very high brightness backlight. The high resolution and high dynamic range of the film proved very effective for this task.

A range of images were made with all of the subpixel arrangements seen above, and comparisons between the RGB originals and RGBW versions with different WMR values and white pixel colors were made. The visual match in terms of color reproduction was excellent, and as expected image sharpness and detail were affected to some extent by both the WMR parameter and the choice between stripe and quad. As soon as OLED RGB and RGBW prototypes were available, the colorimetric accuracy of the white-replacement RGBW algorithm was confirmed [4].

Illustrating the result of the RGBW algorithm on paper is difficult because the images are intended to match exactly: a colorimetric simulation will be trivially perfect, while an attempt to reproduce the RGBW subpixels will suffer according to the imperfect color reproduction of the printer. However, it is still instructive to look closely at a detail level not normally visible to the eye: Figure 3.4 shows what a magnified portion of an example image looks like with various pixel patterns and WMRs. Image A shows large, full-color square pixels separated by black lines for clarity. The typical RGB stripe (B) is visually similar to the RGBW stripe with a low WMR of 0.33 (D), while the RGBW stripe adopts a slightly different character as the W subpixels are relatively brighter with a maximized WMR of 1 (F). The RGBW quad images (C & E) look quite different from the stripes and also show a difference in the dominance of the W channel with different WMR values. In general an intermediate value for WMR appears the smoothest because luminance is distributed among all subpixels. Maximizing WMR (and maximizing the power efficiency benefit of RGBW) means that at least one of the subpixels in each pixel remains completely black, resulting in more contrast between subpixels and, at least at low spatial resolution, a more overt pixel pattern. All of these nuances become less visible as the pixel dimensions of the display get smaller or the viewing distance gets longer.

3.10 Discussion

A flexible algorithm for maximizing the efficiency of emissive W-RGBW OLED displays while preserving colorimetric accuracy has been presented. It is useful regardless of the color of the fourth primary and in cases with more than four primaries, working on the principle of white replacement. Using this principle, luminance is transferred from less efficient subpixels to more efficient subpixels without changing the display's color reproduction.

It is perhaps debatable whether “perfect” color reproduction for a display system is defined by colorimetric accuracy. If it is assumed that the incoming color signal is indeed device independent and rendered as desired by elements upstream in the image-processing chain, then it is easy to argue that reproducing the color faithfully as encoded is correct. A display designer may decide to “enhance” an incoming image through color modification, among other things, and the usefulness of such improvement is outside the scope of this

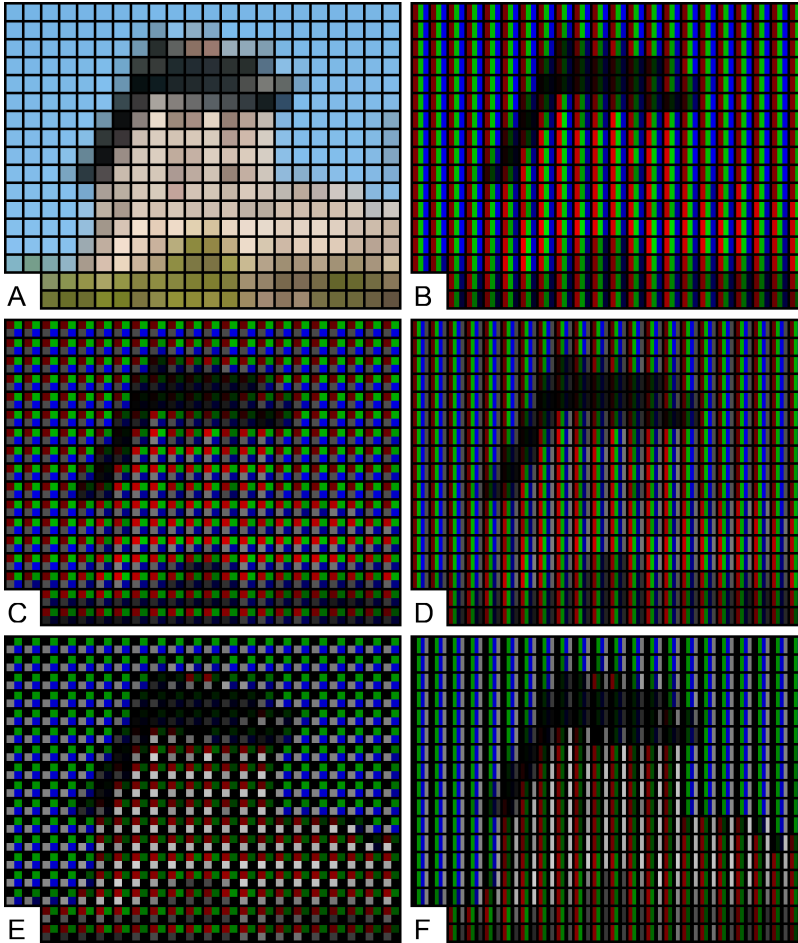


Figure 3.4: RGB & RGBW pixel patterns. Image A shows a full-color original image magnified so the pixels are visible. B shows the same image with an RGB stripe pattern that is typical of LCD displays. C shows an RGBW quad pattern with WMR of 0.33, while D shows an RGBW stripe pattern with WMR of 0.33. E shows RGBW quad with WMR of 1.0 and F shows RGBW stripe with WMR of 1.0. Images B-F appear very dark on paper because the pixels are separated spatially and only reflect a small amount of the light hitting the page, relative to white. Images like these on Ektachrome transparencies viewed on a strong backlight appear as bright as a normal display.

discussion. It is assumed, however, that lowered saturation and/or luminance of colors through the use of a high-efficiency primary is not an improvement, and that accurate color is a preferred result.

An important distinction can be made between RGBW algorithms that use white to augment luminance and the one discussed here that uses white to efficiently replace neutral luminance. Algorithms in the former category trade away color accuracy, in at least some colors, for higher efficiency. The white replacement algorithm preserves color reproduction and does not modify the white point luminance, yet it provides greater efficiency by favoring the subpixel with the highest efficiency. Interestingly, specific display technologies clearly steer the choice of which style of algorithm is to be employed. Emissive displays such as W-RGBW OLEDs draw power proportionally to their light output. This makes them well suited to take advantage of white replacement because the efficiency of the light output of the individual subpixels is important. Utilizing a higher-efficiency W subpixel in lieu of lower-efficiency RGB subpixels results in lower power consumption. Also, OLEDs can be driven to very high luminance levels, so there is little need to use W to augment luminance.

Displays such as backlit LCDs and DMD projectors employ spatial light modulators and an always-on light source, thus drawing a constant amount of power regardless of the modulated light output. Because of this, RGBW is most often used to augment luminance, accepting some amount of color degradation. For these devices, white replacement as proposed in this paper does not provide much benefit; in fact, it can provide a net loss. The use of additional primaries typically reduces the relative spatial aperture ratio or temporal fraction provided for each color, and restricting the use of the RGB results in lost light and power.

The efficiency benefit realized in a W-RGBW OLED display system depends heavily on how often the W subpixel is utilized in place of RGB. This means that the nature of the content displayed has a large effect. In pictorial applications, neutral and near-neutral colors are extremely frequent, providing a large benefit: in fact, about twice the efficiency on average using a set of typical consumer digital camera images [4]. Other applications and content might provide different levels of W primary utilization, and likewise different efficiency benefits. Designing an RGBW display system requires co-optimizing a large set of parameters, including the WMR parameter offered by the present algorithm, the physical pixel layout, including aperture ratio, and the chromaticities of the primaries themselves. Details on the effects of these variables on image quality, display lifetime, and power consumption are provided in [9].

3.11 Bibliography

- [1] Jeffrey B. Sampsell. US Patent 5,233,385 White Light Enhanced Color Field Sequential Projection, Aug 3, 1993 .

- [2] A. Kunzman and G. Pettitt. 10.3 White Enhancement for Color Sequential DLP. *SID Symposium Digest of Technical Papers*, 29(1):121–124, 1998.
- [3] Baek-Woon Lee, Cheolwoo Park, Sangil Kim, Taehwan Kim, Youngchol Yang, Joonhak Oh, Jeongye Choi, Munpyo Hong, Dongsik Sakong, Kyuha Chung, Seongdeok Lee, and Changyong Kim. 40.5L: Late-News Paper: TFT-LCD with RGBW Color System. *SID Symposium Digest of Technical Papers*, 34(1):1212–1215, 2003.
- [4] A. D. Arnold, P. E. Castro, T. K. Hatwar, M. V. Hettel, P. J. Kane, J. E. Ludwicki, M. E. Miller, M. J. Murdoch, J. P. Spindler, S. A. Van Slyke, K. Mameno, R. Nishikawa, T. Omura, and S. Matsumoto. Full-Color AMOLED with RGBW Pixel Pattern. *Journal of the Society for Information Display*, 13(6):525–535, 2005.
- [5] ITU-R. BT.709 : Parameter Values for the HDTV Standards for Production and International Programme Exchange, April 2002.
- [6] M. R. M. Nijenhuis and F. J. J. Blommaert. Perceptual Error Measure for Sampled and Interpolated Images. *Journal of Imaging Science and Technology*, 41(3):249–258, 1997.
- [7] Jean-Bernard Martens. *Image Technology Design: A Perceptual Approach*. Kluwer Academic Publishers, Norwell, MA, 2003.
- [8] Michiel Klompenhouwer. *Flat Panel Display Signal Processing: Analysis and Algorithms for Improved Static and Dynamic Resolution*. PhD thesis, Technische Universiteit Eindhoven, 2006.
- [9] J. P. Spindler, T. K. Hatwar, M. E. Miller, A. D. Arnold, M. J. Murdoch, P. J. Kane, J. E. Ludwicki, P. J. Alessi, and S. A. Van Slyke. System considerations for RGBW OLED displays. *Journal of the Society for Information Display*, 14(1):37–48, 2006.

Balancing Color and Power in RGBW OLED Displays

Abstract

Organic light emitting diode (OLED) displays employing white light emitting OLEDs can be made much more efficient by employing a RGBW color filter array rather than the typical RGB. Self-luminous OLEDs have a unique advantage over light-modulator displays such as LCDs because the W channel can be used to efficiently replace the light of the combination of RGB while preserving colorimetric accuracy. A generalized algorithm for the required color processing is described. Because RGBW OLEDs can demand high peak currents to present images with bright, highly saturated colors, alternative image-processing methods that reduce the peak current and power of these displays are described. The image-quality impact of these algorithms are explored to develop a final image processing algorithm.¹

4.1 Introduction

Organic Light-Emitting Diode (OLED) displays can potentially be constructed at lower cost and lower power consumption than competitive display technologies, such as LCD. These advantages arise from the construction of solid-state displays that emit light directly, rather than by modulating light from a high-intensity light source. One format for manufacturing OLED displays in a particularly cost-effective manner involves the use of an unpatterned white light-emitting layer in combination with color filters. Such a format does not require patterning of emissive organic materials, allowing the display to be constructed with high yield on large substrates. Forming such a display with RGB color filter arrays can result in displays with unacceptably high power consumption, but fortunately, it is possible to

¹This chapter is based on research published in the SID Symposium Digest of Technical Papers, Volume 39, Issue 1, 791-794 (2008), "Distinguished Paper: RGB to RGBW Conversion for OLED Displays," Michael J. Murdoch and Michael E. Miller, and in the Journal of the Society for Information Display, Volume 17, Number 3, 195-202 (2009), "RGB-to-RGBW conversion with current limiting for OLED displays," Michael E. Miller and Michael J. Murdoch.

reduce the power consumption of the display by a factor of two or greater by including unfiltered, white subpixels [1, 2] to create a display having an RGBW subpixel array.

Taking full advantage of the highly efficient, unfiltered white subpixels involves using them as much as possible in place of the less efficient RGB subpixels without distorting the color information within the image. We previously presented such an algorithm [3], which subtracts a portion of the neutral luminance from an RGB pixel signal and provides this luminance with the W subpixel. For neutral and near-neutral colors, the white subpixels provide most of the luminance while the RGB subpixels are employed only at low intensities, leading to moderate power consumption and current levels. For bright, highly saturated primary colors, more participation from the RGB subpixels is required, so the efficiency improvement is minimized and power and current increase. Further, when producing bright secondary colors, more than one of the RGB subpixels within a pixel may be employed, requiring higher peak current values. High peak current values are generally not desirable as they can translate to the need for higher voltage power supplies and power supply lines, larger drive TFTs, and short-lived RGB subpixels. Therefore, it is desirable to identify methods to reduce the peak current of the RGB subpixels.

The uses of RGBW subpixel arrays have been discussed for LCDs and include algorithms for adding more white luminance to form an image than the sum of the RGB primaries [4, 5, 6]. In general, this literature demonstrates that it is possible to render acceptable images by applying RGB primaries that sum to a luminance that is lower than the peak luminance of the W subpixel. However, these algorithms often distort the tone scale of the neutral content in an image without affecting the tone scale of the saturated colors. This manipulation both reduces the saturation of many colors within the image and provides a higher luminance image on the RGBW display than would be created on a comparable RGB display. Under these conditions, it has been demonstrated that most users prefer the increased luminance RGBW rendition to an original RGB rendition [4]. However, these results do not provide much insight into the effect of the various manipulations of color when the overall goal is to provide an RGBW display with a peak luminance equal to that of the reference RGB display but at substantially lower power consumption.

The current paper explores image-processing methods for rendering images onto RGBW OLED displays while limiting the peak current to the RGB subpixels. Such a manipulation provides images having equal peak luminance values but with different renderings of saturated or near-saturated colors. To assess the usefulness of these algorithms, their effect on perceived image quality and the peak power consumption of RGBW OLED displays are discussed.

4.2 Image-Processing Options

To reduce the peak values of the RGB subpixels, two basic options are considered. In both cases, the peak luminance levels of the RGB primaries are limited. In the first case, this luminance loss is simply accepted, providing overall lower power consumption. In the second, a portion of the lost luminance is replaced with luminance from the W primary, resulting in lower saturation and consuming slightly more power than required when limiting the RGB primaries.

4.2.1 RGB Limiting

In an RGBW OLED display system, the maximum intensities of the R, G, and B channels are selected such that when all three are lit together, they produce the display white point, for example, a luminance of 200 cd/m² with CIE 1931 chromaticity coordinates of (0.3127, 0.3290), or D65. In an RGBW system, the RGB subpixels will typically not simultaneously be driven to their maximum currents, as the W subpixel will be used in their stead to produce some or all of the luminance when displaying neutral colors. Despite this fact, accurate color reproduction of all colors in the additive RGB gamut requires that the RGB channels be able to produce the display white point.

Explaining this further requires the use of a 3×3 primary matrix ($\mathbf{P}_{3 \times 3}$). Assuming the Rec. 709 (sRGB) primaries and white point (D65), and working in percent luminance factor, this matrix is:

$$\mathbf{P}_{3 \times 3} = \begin{bmatrix} 41.2 & 35.8 & 18.0 \\ 21.3 & 71.5 & 7.2 \\ 1.9 & 11.9 & 95.1 \end{bmatrix} \quad (4.1)$$

A linear, additive color model says that the color stimulus, in terms of CIE 1931 XYZ tristimulus values, produced by a display with a given $\mathbf{P}_{3 \times 3}$ and a triad of RGB intensity values, is determined by the following equation, which helps explain the meaning of the nine elements of the matrix:

$$\mathbf{P}_{3 \times 3} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (4.2)$$

The first column contains the XYZ tristimulus values of the full-intensity R channel, the second the G channel, and the third the B channel. The middle row of the $\mathbf{P}_{3 \times 3}$ provides percent luminance factors for the R, G, and B channels. For example, the R channel at full intensity has a luminance Y of 21.3% of peak display luminance, i.e., the luminance of the white point of the display.

An RGBW system adds a fourth column to the primary matrix, the XYZ tristimulus values of the W channel. In a simple case, the W primary is the same color as the display white point, meaning it alone can be used to produce white. This simple case is convenient and maximally efficient but difficult to achieve in practice. However, assuming that the W primary is the same color as the display white point, the W channel intensity is normalized such that the RGB channels can reproduce its color. In this case, it means the W column of the matrix $\mathbf{P}_{3 \times 4}$ is exactly the XYZ tristimulus values of the D65 white point:

$$\mathbf{P}_{3 \times 4} = \begin{bmatrix} 41.2 & 35.8 & 18.0 & 95.0 \\ 21.3 & 71.5 & 7.2 & 100 \\ 1.9 & 11.9 & 95.1 & 108.9 \end{bmatrix} \quad (4.3)$$

meaning that there is not a unique solution for converting an input RGB to an RGBW signal. Additional constraints such as maximizing the value of the W signal can be applied to choose among the possible solutions. In RGBW OLED displays, color accuracy can be maintained while maximizing the use of the W signal by calculating the W intensity as the minimum of the R, G, and B intensity values and calculating the transformed R, G, and B intensity values by subtracting the W intensity value from the input R, G, and B intensity values. This manipulation reduces the magnitude of the R, G, and B intensity values required to form most colors, as any color not on the gamut boundary is formed by a combination of one or more of the RGB primaries and the white primary. Unfortunately, bright colors near the color gamut boundary will have nearly the same peak RGB intensity values within both the RGB and RGBW systems.

In the above discussion, the RGB luminance values sum to 100% of the display white point luminance. However, it can be useful if these values are limited to some smaller amount, for example 75%. In such a case, there is no way to produce a full-luminance red of 21.3% of the white point luminance, as the red maximum is now 75% of 21.3, or has a 16.0% luminance factor. Given this limit, it is interesting to consider what can be done in the range from 0 to 75% of the red to eliminate the need for higher red intensity values. Various signal manipulations can be performed to achieve this goal, including clipping, scaling, nonlinear compression, desaturation, and adaptive limiting. Some of these options are obviously better than others, as they are more likely to retain important information within the original images. Within this paper, we will limit our discussion to nonlinear compression and desaturation of the input image signal. Further, while the limit that is applied can be the same or different for each of the color channels that are manipulated, in this paper it is assumed that a common value will be used for all three channels.

The algorithm for limiting the peak luminance of the color channels utilizes nonlinear compression and is applied to the RGB intensity values of the RGBW signal. This nonlinear compression is applied by calculating a scale factor using the equations:

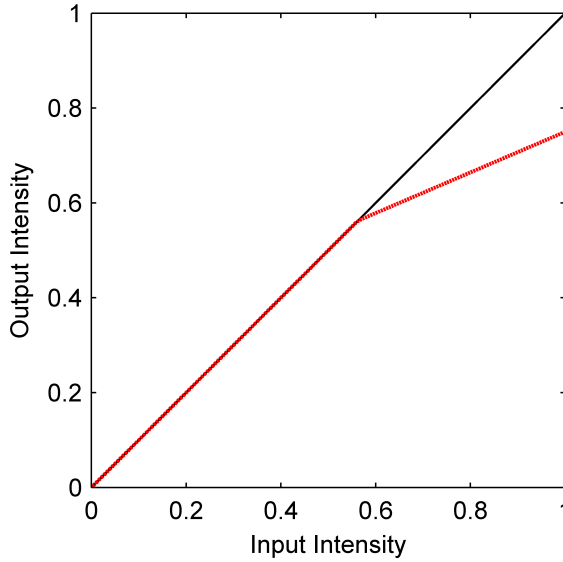


Figure 4.1: RGB intensity limit. Example image intensity limit transforms unlimited (black) and limited (red). The limited case shows a limit of 0.75 and a threshold of 0.75 times the limit.

$$s_R = \begin{cases} 1, & \text{if } R \leq R_{thresh} \\ \frac{m(R - 1) + R_{lim}}{R}, & \text{if } R_{thresh} < R \leq 1, \end{cases} \quad (4.4)$$

$$\text{where } m = \frac{R_{lim} - R_{thresh}}{1 - R_{thresh}}$$

where s_R is the scale factor for the red color channel, R is the red channel intensity value (i.e., code value transformed to remove any nonlinear encoding), R_{lim} is the maximum allowed percent luminance, and R_{thresh} is a threshold below which input values are unchanged and above which the input code values are reduced.

Within this algorithm, R_{thresh} is specified as a proportion of the limit value R_{lim} to ensure that the threshold value is always less than or equal to the limit value. This manipulation provides a tone scale for saturated colors as depicted by the red line in Figure 4.1. This line can be compared to the black line indicating the original 1:1 tone



Figure 4.2: RGB limitation. Illustration of nonlinear compression with RGB limit of 1, 0.67, 0.33, and 0 with a threshold of 0.75 of the RGB limit.

scale. Note that within the current implementation, these tone scales overlap for input intensity values less than 0.5. However, for input image intensity values greater than 0.5, the output intensity values are reduced as compared to the original, unlimited intensity values. Through slight modifications of the previous equations, it is also possible to reduce the slope of the tone-scale values below the threshold; however, this manipulation was not explored within the present experiments. It is also important to recognize that colors near neutral will be formed by a combination of the saturated color component rendered with the limited tone scale and the neutral content that will be provided by the W subpixel, which will be rendered with the original, unlimited tone scale. Therefore, this algorithm will modify input saturated colors having high intensity values and will partially modify less-saturated colors having high-intensity values, with the degree of modification being dependent upon the saturation of the color.

The scale factors are computed for the R, G, and B channels individually. However, if each of these scale factors is applied to the individual color channels, colors within the image will undergo hue rotations, which are generally undesirable. Therefore, the minimum of these three scale factors can be computed and applied to all three of the input R, G, and B intensities within each pixel to acquire output intensities with the same hue as the input intensities.

A pictorial example of the output is shown in Figure 4.2. The yellow sweater in the image has a high amount of red and a slightly lower amount of green luminance. The non-linear algorithm compresses R by a scale value between 1 and the RGB limit, depending on the input intensity. The calculated scale value for the G channel is smaller than for the R channel because the G intensity is lower than the R intensity. To preserve the yellow hue, the G and R intensity values are scaled by the same amount (the minimum of calculated scale values), such that the R:G ratio is maintained. Hue preservation is demonstrated in Figure 4.2. This algorithm maintains hue and saturation, but because the R, G, and B channels are scaled, the luminance of all colors having both a high intensity



Figure 4.3: Desaturation. Illustration of desaturation of RGB with ν parameter values of 1, 0.67, 0.33, and 0.

and high saturation is reduced.

4.2.2 Desaturation

The previous algorithm modified the RGB channels without adjusting the W channel. As a result, when tighter RGB limits are applied, the relative luminance of portions of the image is reduced; often providing images with an unnatural appearance as saturated colors become dim while the peak white of the display is unchanged. A more natural manipulation can include desaturating all colors within the image, as shown in Figure 4.3.

Desaturation is a broad term, but here it refers to a full or partial suppression of the color information without disturbance of the luminance information in an image. This manipulation is applied before the RGB-to-RGBW transformation by multiplying the RGB intensity values by a 3×3 desaturation matrix (\mathbf{D}), and the degree of color suppression can be parameterized as shown below, with the parameter ν indicating the proportion of saturation retained:

$$\mathbf{D} = \nu \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \frac{(1 - \nu)}{100} \begin{bmatrix} 21.3 & 71.5 & 7.2 \\ 21.3 & 71.5 & 7.2 \\ 21.3 & 71.5 & 7.2 \end{bmatrix} \quad (4.5)$$

The matrix \mathbf{D} is a linear combination of two 3×3 matrices: an identity matrix and a luminance matrix. The values in the luminance matrix are replicates of the second row of the sRGB primary matrix $\mathbf{P}_{3 \times 3}$, which are the Y percent luminance factors of the R, G, and B channels. Thus, an RGB signal multiplied by the luminance matrix will result in three identical luminance values, each equal to the sum of the luminance contribution of the three channels. The luminance matrix, because it is expressed in percent luminance factor, is divided by 100. With high values of ν , the identity matrix dominates, meaning \mathbf{D} has

no effect on the RGB image. With low values, the luminance matrix dominates, leading to the luminance image as shown on the far right in Figure 4.3. Intermediate values blend the two endpoints smoothly. Notice that if v is assigned a value, such as 0.75, and the image is transformed into the RGBW color space as described previously, the resulting maximum R, G, and B intensity values will be equal to the assigned value of 75% and will reduce the peak current of the R, G, and B channels of the OLED to 75% of the values required to produce the display white point using these three color channels.

4.3 Image Quality Effects

To evaluate the impact of these algorithms on perceived image quality, a psychophysical evaluation was performed.

4.3.1 Experimental Setup

A total of 23 observers having 20/40 or better visual acuity and normal color vision, as assessed using Ishihara's Tests for Colour Deficiency [7], rated the image quality of seven scenes rendered with 23 combinations of the two previously discussed algorithms.

The seven scenes were each selected to include an area of bright, highly saturated color, which was affected by each of the algorithms. The input images were rendered into sRGB color space before being manipulated. The image-processing chain for these sRGB images included the following steps.

1. Application of an inverse gamma curve to convert the gamma-encoded sRGB values to relative intensity values.
2. Application of the desaturation algorithm to preserve between 40 and 100% of the original saturation.
3. Conversion of RGB intensities to RGBW intensities following the algorithm in Chapter 3.
4. Application of the RGB limiting algorithm to limit the peak intensities of the RGB channels in the RGBW image to a value between 20 and 100% of the original RGB intensities.
5. Conversion of RGBW intensities to gamma-encoded values with primaries matching the target CRT.

During the experiment, the images were displayed on a CRT calibrated to have a code value to luminance response that matched the sRGB specification with a peak luminance of 80 cd/m². The experiment was completed in a dark room to avoid any

effects of screen reflections and external glare sources, as these generally reduce the visibility of differences between experimental conditions.

While a factorial of desaturation and RGB-limiting might seem appropriate, in fact RGB limiting has no effect on an image if its limit value is lower than the desaturation parameter value because desaturation is applied first. Therefore, the RGB limit values were restricted to be equal to or lower than the desaturation value for all experimental conditions.

During the experiment, the participants performed two different tasks. In the first task, they viewed each image (7 scenes by 23 manipulations) under single stimulus conditions and rated it on a five-point category scale of image quality, with descriptors of “like extremely,” “like moderately,” “neither like nor dislike,” “dislike moderately,” and “dislike extremely.” During an initial analysis, this set of data was analyzed by applying Thurstone’s law of categorical judgment [8] to create an interval scale of image quality based on these descriptors. However, further detail on quality differences was desired, so a second task was included in the experiment.

In the second task, forced-choice paired comparisons were provided to the observers with the instruction to select the preferred image. The pairs comprised three groups of three images (thus, nine pairs), each group made up of neighboring conditions in high-, medium-, and low-quality regions of the parameter space. The pairs data were used to determine scale factors between just-noticeable differences (JNDs) in image quality and the desaturation and RGB limit parameters in each quality region, and the three scale factors were taken as slopes used to construct a transform between the interval image quality scale values and JNDs of image quality.

4.3.2 Results

The results of this evaluation are shown in Figure 4.4. Dots within this figure represent the 23 sample points within the design. Contour lines represent lines of equal image-quality degradation in JNDs of image quality. This figure indicates that the average image in this experiment underwent about seven JNDs in image-quality degradation as the saturation is reduced from 100% to about 40% and over six JNDs in image-quality degradation as the peak RGB luminance is reduced from 100% to 20%.

This figure should not be interpreted as indicating there are only 6–7 levels of detectable difference between images over the range of levels investigated in this study, but instead that there are only 6–7 distinct levels over which 50% of the participants would reliably rate one image as being lower in image quality than the other. In fact, some participants preferred certain images to be rendered with a desaturation value less than 100%.

While the desaturation algorithm occasionally resulted in images with a more natural appearance, the RGB-limiting algorithm generally had less impact on overall

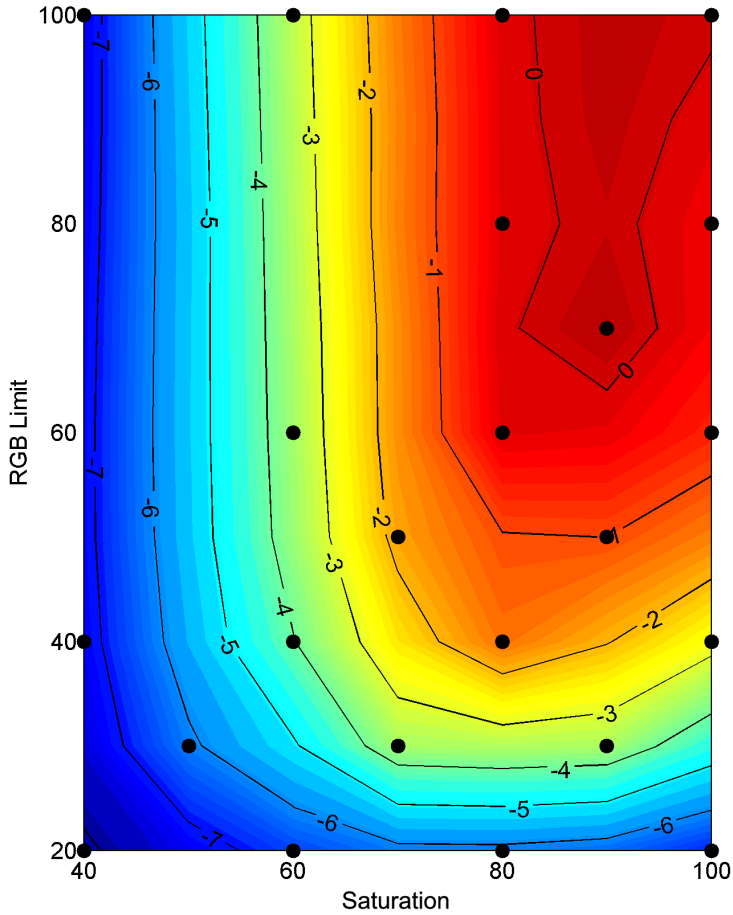


Figure 4.4: Experimental results. Contour plot of image quality on parametric axes of saturation and RGB limit, represented as JNDs of image-quality loss relative to the original (located at the top right). Dots within this figure represent the 23 sample points tested in the experiment. Contour lines represent lines of equal image-quality degradation in JNDs of image quality.

image quality. This is likely due to the fact that the RGB-limiting algorithm affected only small regions within each image while the desaturation algorithm affected all colors in each image other than the neutral scale.

A small interaction was also present between the two algorithms. Image quality near that of the original image occurred for RGB-limiting values between 100% and 67% for a desaturation value of 90%. However, image quality was degraded if the peak RGB intensity value was limited to 80% of the intended RGB intensity value without applying desaturation. This result shows that the combination of slight desaturation with the limiting algorithm allows some of the luminance of the saturated colors to be transferred to the white subpixel and avoids the more extreme loss in brightness and a perceived loss of detail that occurs when the RGB intensity value is reduced without desaturation. This insight was explored further as discussed in Section 4.4.

Each of the algorithms affect the number of color differences produced by the final system and change the appearance of some colors within the initial images. Therefore, a basic colorimetric difference metric can be used to model the impact of these manipulations on image quality. To assess such a metric, the images were converted into CIELAB color space, and ΔE^* values were calculated between corresponding pixels of each processed image and the original. Root-mean-squared (RMS) ΔE^* values were computed over all pixels in each image. The averages of these values were then computed across the seven scenes. The image-quality results are plotted as a function of these averaged ΔE^* values in Figure 4.5, with each colored group of points having a constant RGB limit value and increasing amount of desaturation as they trend to the right.

As shown in Figure 4.5, the RMS ΔE^* values averaged across scenes correlate relatively well with changes in image quality when the images have been highly desaturated. However, for low levels of desaturation, at the left end of each group, this metric overpredicts the change in image quality; i.e., even though the RMS values increase as low amounts of desaturation are introduced into the images, the participants did not indicate that image quality was reduced by this manipulation. In fact, when significant RGB limiting is applied to the images, image quality was perceived to increase for intermediate amounts of desaturation. Although these results would at first appear to be counterintuitive, reasons for this behavior may be hypothesized. First, the category scaling experiment was conducted using single stimulus conditions. Therefore, it is possible that the participants adapted to small changes in saturation of the images as they were displayed and did not perceive the change in saturation. Alternatively, they may have perceived these changes but may not have seen these changes as having a significant effect on image quality. These explanations could account for the initial plateau in these curves as a function of desaturation.

To understand the slight increase in image quality with increasing desaturation for small desaturation values, it is important to understand that the RGB-limiting algorithm decreased the slope of a portion of the tone-scale curve, decreasing contrast within regions of the image. This decrease in contrast reduced the visibility of detail in portions of some

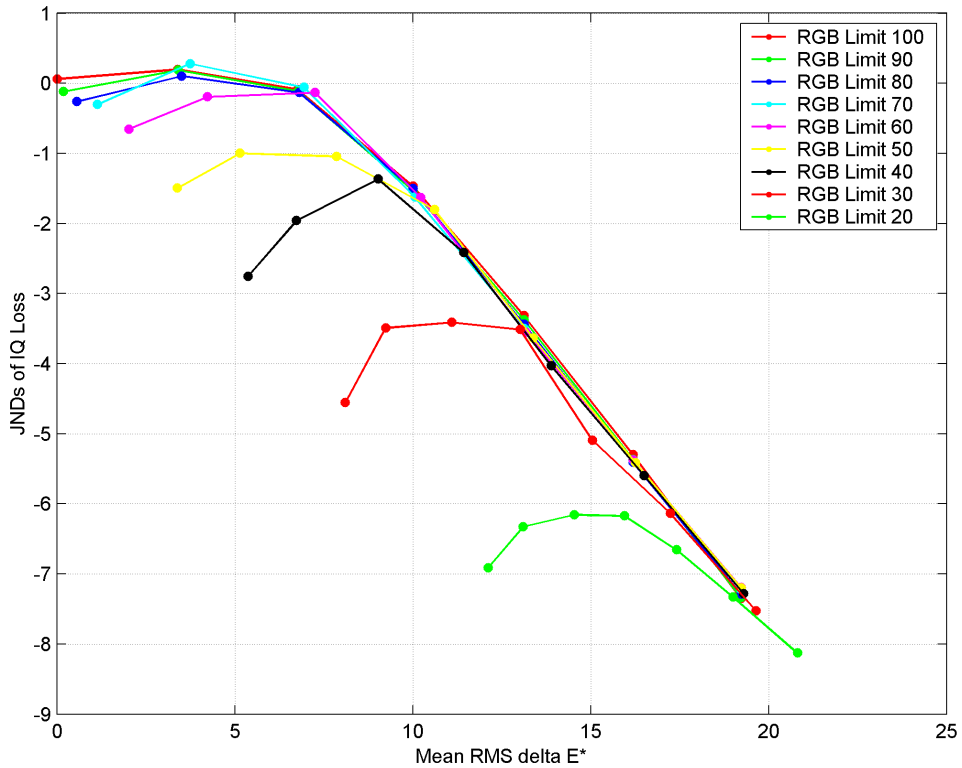


Figure 4.5: Results against color error. JNDs of image quality loss shown versus RMS ΔE^* averaged over images. Each colored line shows a series of desaturation values at a fixed RGB limit value, with lower desaturation (fully saturated images) at the left end of each series and higher desaturation at the right.

scenes. However, the reduction in contrast was effectively reduced when the image was desaturated as a greater proportion of the image luminance was transferred to the white channel. Therefore, this increase in image quality with a reduction in saturation demonstrates a bias towards sacrificing color fidelity over losing spatial detail. This hypothesis is supported by the fact that the participants preferred larger amounts of desaturation with increases in RGB limiting.

4.4 Discussion

As was shown in the previous section, both RGB limiting and desaturation can effectively reduce power consumption and increase the lifetime of an OLED display. Further, the combination of these algorithms improves image quality under some conditions as compared to images that are produced with the same peak RGB intensities through the application of only one of the approaches. Therefore, it is important to explore the possibility of a unified image-processing algorithm to provide enhanced image quality for a given level of power consumption or lifetime improvement. Such an algorithm might allow the RGB primaries to be limited while replacing a portion of their luminance with light from the white subpixel, thus effectively selectively desaturating the image only where the RGB limit has greatest effect.

With this in mind, an improvement to the algorithm presented here is suggested. Returning to Equation 4.4, a single scale factor s was taken as the minimum of those computed for each of the R, G, and B image channels. The difference $(1 - s)$ then represents the proportion by which the luminance is reduced by RGB limiting. The decrease in luminance that occurs can be computed by multiplying $(1 - s)$ by a 1×3 matrix containing the primary luminance factors (L_R , L_G , L_B) for the display [i.e., (0.213 0.715 0.072) in an sRGB display example] and the linear intensity values for the respective color channels and summing the result. The intensity of the W subpixel can then be increased to compensate for this luminance loss. However, it is not necessarily desirable to replace all of the lost luminance with luminance from the W subpixel as this manipulation can result in severe desaturation of colors that undergo a large amount of limiting. Therefore, another free parameter η is added, indicating the proportion of the luminance loss to be replaced by the W subpixel. The resulting equation for the intensity to be added to the white subpixel would then be:

$$W_{add} = \eta(1 - s) \begin{bmatrix} L_R & L_G & L_B \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4.6)$$

W_{add} then represents the proportion of the lost luminance intensity to be added to the luminance intensity of the W subpixel. This increase in W subpixel luminance will result in desaturation of only the pixels for which RGB limiting is applied and will

not result in desaturation of other pixels in the image as occurred within the earlier image-quality experiment that was discussed.

Figure 4.6 shows a two-dimensional grid of images that were created by applying this algorithm. Levels of RGB limiting (including 1.0, 0.67, and 0.33 at a constant threshold value of 0.75 of the limit value) are shown across the horizontal dimension of the grid and various levels of luminance replacement (0.0, 0.5, and 1.0) are shown along the vertical dimension of the grid. As shown in this figure, spatial detail is lost particularly in the yellow sweater as the RGB limit value is reduced when the luminance replacement value is zero (i.e., the upper-right image). Increasing the white replacement value increases the visibility of the spatial detail within the yellow sweater while simultaneously reducing the color saturation within the sweater. Note, however, that the saturation of other colors within the scene is virtually unaffected as they lie below the threshold of the RGB-limiting algorithm. Therefore, this algorithm effectively reduces the peak current to the RGB subpixels while providing the ability to trade loss of detail for loss of saturation in only the portions of the image having bright, highly saturated colors. Therefore, this algorithm supplies the ability to provide higher quality images with very small losses of detail. Images having an RGB limit of 0.7 and a W replacement proportion in the neighborhood of 0.3 clearly are improved in quality over the images having an RGB limit of 0.7 and a saturation limit of 0.9. Such an algorithm provides the ability to reduce the peak current of RGBW OLEDs to 70% of what would otherwise be required with a very small loss of image quality.

4.5 Conclusions

Algorithms have been discussed for limiting peak current to the RGB subpixels in an RGBW-OLED display. The effects of these algorithms on perceived image quality have been demonstrated. Based upon these results, it appears possible to limit the peak luminance, and therefore the peak current, to the RGB channels of an RGBW OLED to 70% of their peak with only a very small impact on the overall perceived image quality of the system. However, this manipulation requires a modification of the tone scale for saturated colors as well as desaturation of high-intensity, highly saturated colors. This manipulation requires that the W subpixel be applied to replace a portion of the luminance of the saturated colors that would otherwise be produced by the RGB subpixels.

A metric based upon changes in CIELAB values was shown to correlate with the image-quality results for only high levels of desaturation. Additional metrics will be required to fully explain the image-quality results provided within this paper.

This reduction of the peak current in RGBW OLED displays to 70% of their typical aim values allows the peak power of the display to be reduced. This reduction results in many efficiencies, including reduction of the area devoted to power TFTs and power lines. Further, the peak current densities are reduced for the RGB subpixels, reducing the overall power consumption of the display.

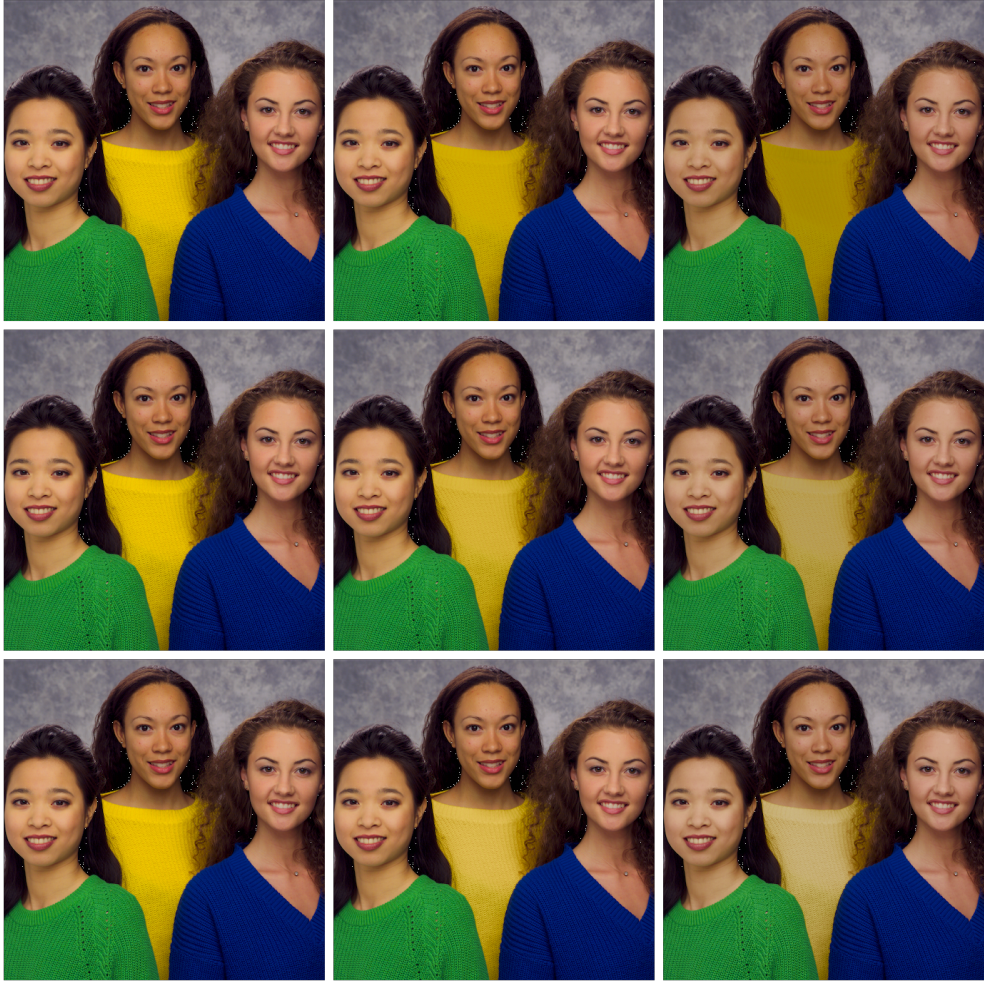


Figure 4.6: RGB limit with white-light replacement. A pictorial example of the effect of RGB limiting with white-light replacement. Horizontally, these images have RGB limit values of 1.0, 0.67, and 0.33. Vertically, the luminance replacement portion η values are 0.0, 0.5, and 1.0.

4.6 Acknowledgments

We thank our colleagues at Eastman Kodak Company for their guidance and assistance. We especially acknowledge the assistance of Paula Alessi, Paul Kane, and Cynthia Pellow.

4.7 Bibliography

- [1] A. D. Arnold, P. E. Castro, T. K. Hatwar, M. V. Hettel, P. J. Kane, J. E. Ludwicki, M. E. Miller, M. J. Murdoch, J. P. Spindler, S. A. Van Slyke, K. Mameno, R. Nishikawa, T. Omura, and S. Matsumoto. Full-Color AMOLED with RGBW Pixel Pattern. *Journal of the Society for Information Display*, 13(6):525–535, 2005.
- [2] J. P. Spindler, T. K. Hatwar, M. E. Miller, A. D. Arnold, M. J. Murdoch, P. J. Kane, J. E. Ludwicki, P. J. Alessi, and S. A. Van Slyke. System considerations for RGBW OLED displays. *Journal of the Society for Information Display*, 14(1):37–48, 2006.
- [3] M. J. Murdoch, M. E. Miller, and P. J. Kane. Perfecting the Color Reproduction of RGBW OLED. In *Proceedings International Congress of Imaging Science*, pages 448–51, 2006.
- [4] L. Wang, Y. Tu, L. Chen, K. Teunissen, and I. Heynderickx. Trade-off Between Luminance and Color in RGBW Displays for Mobile-Phone Usage. In *SID Symposium*, volume 38, pages 1142–1145, 2007.
- [5] C. Brown-Elliott and T. Credelle. High-Pixel-Density Mobile Displays: Challenges and solutions. In *SID Symposium*, volume 37, pages 1984–1986, 2006.
- [6] Baek-woon Lee, Kyongtae Park, Alexander Arkhipov, Sungtae Shin, and Kyuha Chung. L-2: Late-News Paper: The RGBW Advantage for AMOLED. *SID Symposium Digest of Technical Papers*, 38(1):1386–1389, 2007.
- [7] Shinobu Ishihara. *Ishihara's Tests for Colour Deficiency*. Kanehara & Co, Ltd., Tokyo, 24 plates edition, 1999.
- [8] W. S. Torgeson. *Theory and Methods of Scaling*. Wiley, New York, 1958.

Preferred Color Gamut for Reproducing Natural Image Content

Abstract

With the development of wide-gamut display technology, the need is clear for understanding the required size and shape of color gamut from the viewers' perspective. To that end, experiments were conducted to explore color gamut requirements based on viewers' preferred level of chroma enhancement of standard-gamut images. Chroma preferences were measured for multiple hues using single-hue images, and a corresponding hue-dependent preferred chroma enhancement was successfully applied to natural, multi-hue images. The multi-hue images showed overall success, though viewers indicated that reds could be decreased even further in colorfulness, and yellows could be increased, which may argue in favor of multi-primary displays. Viewer preferences do vary within the population, primarily in overall chroma level, and the differences can be largely accounted for with a single parameter for chroma level adjustment that includes the preferred hue dependence. Image content dependencies were also found, but they remain too complex to model. The hue-dependent chroma preference results can be applied to display design and color enhancement algorithms.¹

5.1 Introduction

Recent technology development has enabled the creation of displays with extremely wide color gamut. This progress has been driven partly by manufacturers' desire for technological leadership in the marketplace and partly by consumers' appreciation of enhanced color performance. Specifications for display color rendering such as Rec. 709, used in the sRGB standard for web graphics and in the HDTV standard for television, define a reference display with a standard color gamut [1]. Wide-gamut displays have

¹This chapter has been published in the *Journal of the Society for Information Display*, Volume 18, Number 12, 1111-1118 (2010): "Preferred and Maximally-Acceptable Color Gamut for Reproducing Natural Image Content," Michael J. Murdoch, Dragan Sekulovski, Robert de Volder, and Ingrid E. J. Heynderickx.

the ability to render more saturated colors than the reference display, a capability that is generally advantageous to the viewer, according to several perceptual insights. It is known that not all colors occurring in nature can actually be reproduced on standard displays [2]. However, even a display capable of rendering all natural colors may not be satisfactory. Displays must physically exaggerate colors in order to create the same impression of colorfulness that would be seen in reality, due to the relatively lower luminance of the reproduction [3]. Further, it has been shown that viewers prefer colors to be slightly more saturated than what they realize is natural [4, 5]. Thus, creating realistic and hyper-real color reproduction is impossible on standard displays and requires a wider color gamut. Recent work has addressed the gamut size requirements that may be beneficial to the viewer.

A display's color gamut, which is the range of colors that the display is capable of producing, is the first of three important aspects in the discussion of wide gamut color. Second is the source image, which, if stored using a typical color encoding standard such as Rec. 709, implicitly includes the limit of its color palette as defined by its encoding primaries. Third is the set of choices that make up the image rendering by the display system, or the gamut mapping that determines how the starting image is modified to take advantage of the wide color gamut. For the near future, it is reasonable to assume that the input image would comply with an existing standard-gamut representation such as Rec. 709, meaning that research has focused on the gamut mapping details and gamut size requirements.

Starting from standard-gamut images, wide-gamut color must be created by extending the existing information. The most straightforward way of doing so is by just using the same drive signals (SDS), as if naively accepting the color gamut difference and resulting color expansion. However, this approach is difficult to generalize because of its dependence on the physical display details, and examples in the marketplace produce unevenly-expanded colors that in some cases are pushed "too far," in the authors' opinion. Other, more intelligent approaches are possible, some of which were considered by Muijs et al. [6]. This study showed preference for an adaptive mapping that extended the chroma of a given color non-linearly in a direction dependent on the lightness of the color, but it also found differences based on the extent to which the mapping algorithms used the wider color gamut. That is, the mapping algorithms' success was confounded with the gamut limits involved. Approaching the problem differently, looking first at gamut limit requirements, Laird et al. performed an experiment in which the preferred chroma was measured for images with varied hue and lightness [7]. This study, which included two images at 6 hues and 3 lightness levels, found chroma preference levels that were indeed hue-dependent, and that overall were beyond Rec. 709, but still lower than the limits provided by the wide-gamut display used in the experiment. Hue dependence was also found in an experiment by Hisatake et al. [8], which used natural images with different dominant colors. They additionally concluded that because preferred chroma was higher in the direction of the primaries than in the direction of the secondaries, a wide-gamut three-primary display has more added value than a multi-primary display. A

later experiment by Sakurai et al. [9] found that a similarly moderate increase in chroma was preferred for a larger set of more complex natural images, but did not investigate hue dependence. Several of these studies [6, 7, 9] identified a strong dependence on image content but included relatively small numbers of images and did not attempt to model the effect.

To build on these previous hue- and content-dependent results, two experiments were performed that investigate color gamut preference. Experiment 1, which used natural images with very narrow hue ranges, measured the preferred and maximally-acceptable levels of chroma for selected hues. Experiment 2 measured chroma preference using a large set of natural images with typical, complicated hue combinations, which were modified using the results of Experiment 1, and subjectively tested the acceptability of their inter-hue relationships.

5.2 Experiment 1

5.2.1 Experimental set-up

The goal of this experiment was to measure the preferred level of chroma and the maximum accepted level of chroma at different hues, using a variety of images, as will be explained below. These two measurements were performed in subsequent tuning tasks by 42 observers. Half were male and half female, and they ranged in age from 21 to 61 years with a mean age of 31 years. All of the observers were employed by Philips in various disciplines; 24 were Dutch, and the rest were primarily from Europe or Asia. All had normal color vision as tested with the Ishihara color blindness test.

Stimuli were displayed on a 40" RGB LED-backlit, wide-gamut display with a 9000K whitepoint and a maximum luminance of 341 cd/m². The CIE 1931 xy chromaticities of the primaries were (0.699, 0.291) for red, (0.188, 0.700) for green and (0.146, 0.063) for blue. Compared to the Rec. 709 standard primaries, the wide-gamut red was significantly more pure and more purple, the green was significantly more pure and more cyan, and the blue was very similar. Thus, the wide chromaticity gamut was unevenly enlarged relative to Rec. 709, more so in cyan-green and magenta-red hues, and less so in blue and yellow hues. Image modifications were made as described below, resulting in target colorimetric values per pixel, from which display drive values were computed. The display's output was characterized with a Photo Research PR680 spectrophotometer, and the measurements were used to create an additive colorimetric forward model: the CIE XYZ contributions of the constant display "black" minimum and each of the color channels as measured individually at each drive value. For each target colorimetric value in the images, optimal drive values were computed iteratively, using the forward model. This method proved quite accurate, measured to have an average color error of $0.3 \Delta E_{ab}^*$ over a set of 10,000 random color patches, which confirmed that the display was both additive and channel-independent. The experiments were performed in a viewing laboratory, with the white wall behind the display illuminated at 22 lux without illuminating the display

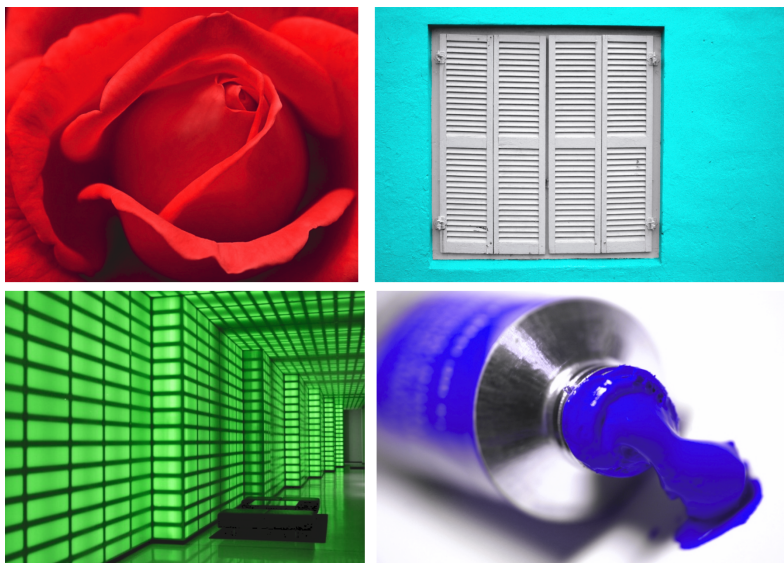


Figure 5.1: Single-hue images. Single-hue images used in Experiment 1: rose, window, room, and paint, shown rendered at exemplary hues.

surface. Observers were seated behind a table at a viewing distance of 3 m from the display. The images shown varied in size slightly, but all were about 600 pixels tall, and therefore did not fill the display's 1360x768 pixels. The remaining screen border was filled with solid gray at the 9000K color temperature and a CIELAB lightness value of 50.

Four images with very narrow hue ranges are shown Figure 5.1. Each of these images was modified by changing its CIELAB hue to each of seven different hues of interest: those of the three wide-gamut display primaries R, G, and B (CIELAB $h = 37$, 149, and 308, respectively); the display secondaries Y, C, and M (CIELAB $h = 104$, 191, and 335); and the Rec. 709 R primary (CIELAB $h = 41$). Both red hues were used because, despite their similar hue angles, they appeared very different to the experimenters: Rec. 709 R appearing as the familiar display red, and the wide-gamut display red distinctly purplish and less pleasant. The hue rotation procedure, which included remapping the originally-sRGB pixels' CIELAB lightness and chroma to match the max-chroma cusp of each destination hue, was explained in more detail in a previous paper [10]. The resulting 28 image-hue combinations, along with eight additional single-hue or narrow-hue-range images, plus five duplicates, made a total of 41 images that were each presented to the observers in the experiment.

The chroma variations used in the tuning task were processed in two steps. First, in a SDS rendering, the RGB images were input to the wide-gamut display forward model, using its primaries and gamma nonlinearity, resulting in CIE 1931 XYZ values, and converted to CIELAB LCh using the display's D90 white point for a reference white. Then, to create the chroma variations, the CIELAB chroma values were multiplied by factors ranging from 0 to 4 in steps of 0.05, implying that the variation with a multiplier of 1 corresponded to the same drive signal applied to the wide-gamut display. In all 81 chroma variations, the corresponding lightness and hue values were kept constant. The modified CIELAB LCh images were transformed back to RGB using the colorimetric model of the wide-gamut display and stored. Because all of the original images included pixels that reached maximum chroma, in all chroma variations with a multiplier larger than 1, some LCh values were pushed beyond the wide-gamut display gamut and were hard-clipped to the maximum chroma at the corresponding L and h.

The experiment employed a tuning methodology in two parts, each part using all 41 image-hue combinations. First, observers were asked to select the level of colorfulness that they preferred for each image. Starting at a random chroma level, they were able to adjust the chroma of each image in discrete steps, upward and downward using the arrow keys, until they were satisfied. Second, starting again at a random chroma level, they were asked to select the highest level of colorfulness that was still acceptable for each image. In both parts, images were presented in random order, and no time constraints were given. The tuning task progressed quite quickly, typically 10-20 seconds per tuning, and the complete experiment took about 20 minutes.

5.2.2 Results

The main results are the preferred and maximum accepted levels of chroma at each of the seven hues. Figure 5.2 shows a boxplot of both, per hue, pooled over images. The chroma level is shown on the y-axis as a percentage of the wide-gamut display chromaticity gamut boundary at each hue. For reference, the relative size of the Rec. 709 standard chromaticity gamut at each hue is shown by the small black marks beside the boxes. It is clear that for most hues, at least half of the preferred chroma levels are beyond Rec. 709 chromaticity gamut boundary, which already shows the added value of a wide-gamut display. There is also an apparent hue dependency, and especially interesting is the remarkable difference in preferred chroma between the two red hues used, despite their proximity in CIELAB hue angle. This difference could be due to a memory color effect, because the purplish wide-gamut red is noticeably different from the standard Rec. 709 display red, which people may be accustomed to seeing. The median maximally-accepted chroma reaches the maximum of the wide-gamut display's capability in several hues (blue, yellow, cyan, and magenta), but in the other measured hues it does not, which shows that this wide-gamut display provides unacceptably-wide color in some regions of color space.

It is possible that the preference results for B and Y are limited by the display. At both hues, the display does not offer much additional chroma beyond the Rec. 709 chromaticity

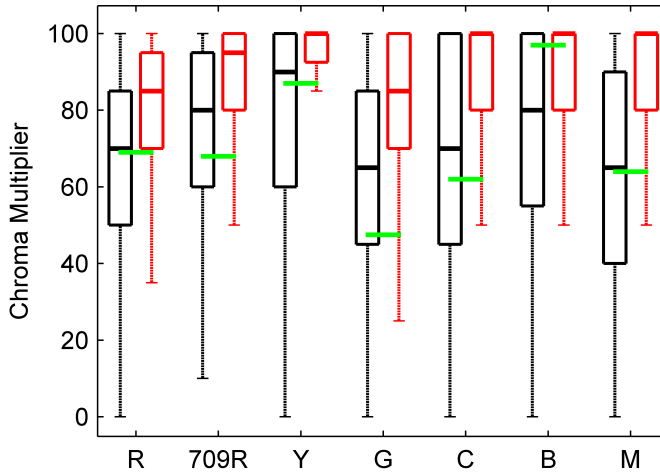


Figure 5.2: Experiment 1 results. Boxplot of the preferred level of chroma (black) and the maximum accepted level of chroma (red) at each of the seven hues, on a scale that represents the percent of the wide-gamut display’s available chroma at that particular hue. Each box indicates the the 25th, 50th (median), and 75th percentiles of observations, and the long whiskers reach to the min and max values observed. The results may be compared to the green tick marks overlaying the boxes, which indicate the per-hue size of the Rec. 709 chromaticity gamut on the same scale. The x-axis label “R” refers to the wide-gamut red ($h=37$), whereas the axis label “709R” refers to the Rec. 709 red ($h=41$).

gamut boundary, which means that the ability to evaluate expansion at these hues is limited. Further, the preference results at both hues appear clipped at the maximum available chroma. Looking again at the overall hue dependence, the median preferred chroma levels, with respect to the Rec. 709 boundary, range from a compression of roughly 20% in blue to an extension of 30% in green. This hue dependence was further confirmed with an ANOVA performed with preferred chroma level as the dependent variable, image content (meaning the four named original images that were rotated in hue) and hue as independent variables, observer as a random factor, and including a 2-way interaction between image content and hue. The results showed that the preferred chroma depended on image content ($F=38.1$, $df=3$, $p<0.001$, $\eta_p^2 = .482$) and on hue ($F=14.9$, $df=6$, $p<0.001$, $\eta_p^2 = 0.267$), and was different for the different observers ($F=4.97$, $df=41$, $p<0.001$, $\eta_p^2 = 0.608$).

Considering the content dependency of the preferred chroma level, post-hoc testing showed that this effect was driven primarily by a different preferred chroma level for each of the images window and paint. The image window had the lowest preferred chroma level, inside the Rec. 709 gamut for all hues. This low preferred chroma was perhaps due to the image subject for which the expected real-life chroma level was low. A rough stucco wall with white shutters might evoke the pastel-toned architecture of the Mediterranean or Caribbean, and therefore might not “look right” with a high-chroma rendering. Likewise, the image with the highest preferred chroma level, paint, showed an object with a high expected chroma: a pure, glossy artist’s pigment. The results for the image paint were beyond the Rec. 709 chromaticity gamut for almost all hues.

A significant effect of observer on preferred chroma level was seen in the ANOVA, meaning that different people preferred different amounts of chroma. This was further evaluated by means of a principal component analysis, which showed that 65% of the variance was explained by a first component representing an overall chroma increase. In other words, if an observer preferred a rather high chroma level for one of the stimuli, his chroma preference was also high for most of the other stimuli; however, the increase was not equal for all hues, a fact that was later applied in Experiment 2. The second and third principal components were significantly weaker, explaining 9.3% and 8.6% of the variance, respectively, but they roughly represented distortions of the hue circle of the first component. The second component showed red opposing yellow and cyan, and the third component showed yellow opposing blue. Thus, a single parameter for overall image chroma is sufficient to satisfy the vast majority of the differences in preference in the observer population, while the subsequent, weaker components describe the per-hue customization each observer would prefer, if given the additional degrees of freedom. Based on the ANOVA, the effect of content on the preferred chroma level was larger than the effect of hue. However, understanding the effect of content remains a major challenge. An attempt was made to model preferred chroma level as a linear function of objective image statistics, such as average and maximum chroma, size and number of high-chroma regions, and local contrast in lightness and chroma channels. A combination of these objective measures seemed promising based on the four images used in Experiment 1, but the model did not perform satisfactorily when applied to additional images in Experiment 2. A more complicated model, involving object and context recognition and cognitive modeling, can be envisioned but is far from fruition. Setting content aside, effectively averaging out its effect, attention was focused on the effect of hue, which can be straightforwardly implemented in a mapping algorithm. Such a mapping, essentially a hue-dependent chroma multiplier, was created and studied in Experiment 2.

5.3 Experiment 2

5.3.1 Experimental set-up

The second experiment tested to what extent the hue-dependent chroma preference, which was determined from single-hue images, was applicable to natural, multi-hue images. This was done in two parts: a tuning task, very similar to Experiment 1, and a painting task, in which observers were asked to paint, using the mouse cursor, regions that were too high or too low in chroma. The same wide-gamut display and viewing conditions from the previous experiment were used. Experiment 2 was performed by 43 observers, of which 27 were male and 16 female. Their ages ranged between 20 and 64 years with a mean age of 30 years. Twelve of the observers had also performed Experiment 1.

The tuning task utilized single-hue images rotated to a variety of different hues, including the same paint image used in Experiment 1 plus two new images. Additionally, the image set contained 51 natural, multi-hue images. The eleven images shown in Figure 5.3 are representative of the multi-hue set, and also comprise the subset of images used in the painting task. The multi-hue images were selected for variety of content, naturalness, and a predominance of high-chroma colors. The intentional bias toward high-chroma colors (and the avoidance of sensitive colors like skin tones) was meant to weight the experiment toward the gamut boundaries and to reduce the effect of mid-chroma colors. As mentioned earlier, the first priority in gamut expansion was to find the gamut size requirements; later experiments will improve rendering for mid-chroma, memory colors, and skintones.

The three single-hue images were rotated to different hues in the same way as in Experiment 1. The ten destination hues were the same seven given above, plus BC and RM tertiaries (CIELAB h of 240 and 10) and Rec. 709 G (136). These intermediate values were selected to more completely sample the hue circle and confirm the interpolation. The 30 single-hue images and 51 multi-hue images were all processed to a variety of chroma levels for presentation in the tuning task. Based on the results of Experiment 1, which provided preferred chroma levels at seven hues, a smoothly-interpolated hue-dependent chroma multiplier was created for each chroma level.

5.3.2 Hue-Dependent Chroma Multiplier

The hue-dependent chroma multiplier was implemented in the second step of image processing. With sRGB images as a starting point, the first step was, similar to Experiment 1, rendering the RGB images using the wide-gamut display primaries, resulting in boosted chroma for all pixels because of the high purity of the wide-gamut primaries compared to sRGB's Rec. 709 primaries. In contrast to the previous experiment, however, the standard sRGB nonlinearity was used, rather than the display's native non-linearity, to preserve tonal accuracy in the multi-hue images. As a second step, the images were varied in overall chroma level using the hue-dependent chroma multiplier, creating a set of chroma



Figure 5.3: Multi-hue images. Examples of the 51 multi-hue images used in Experiment 2. These 11 images were used in the painting task.

variations that were then used in the tuning task.

The principal components analysis performed in Experiment 1 yielded a first component that was conceptually an “overall” chroma level preference, to which all seven hues contributed a similar, but slightly different amount. The raw factor loadings in the analysis can be interpreted as the relative contribution of each hue to the overall preferred chroma level. The preferred level (mean preference) and the corresponding relative contribution per hue were used as offset and gain to create a chroma multiplier relationship for each hue. Figure 5.4 shows the seven linear relationships as a function of overall chroma level, a scale on which 1 is the mean preference from Experiment 1. The different slopes visible show that as the overall chroma level increases, the chroma multipliers at some hues increase slightly faster than at other hues.

Applying the chroma multipliers based on seven measured hues to natural, multi-hue images requires interpolating the results to intermediate hues, and that was done via cubic interpolation of the multipliers themselves as a function of CIELAB hue angle, constrained for continuity and smoothness around the hue circle. Figure 5.5 shows the smoothly interpolated result for a range of overall chroma levels. The impact of the different gain factors per hue can be seen in the different vertical spacing of the colored dots (at each

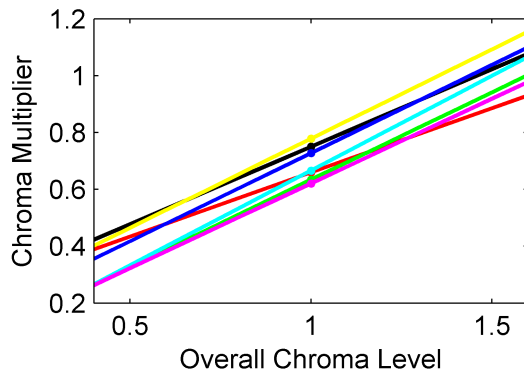


Figure 5.4: Chroma multiplier slopes. Each line shows the relative slope derived from the principle component analysis, colored according to the hue it represents, with the exception of the black line, which corresponds to Rec. 709 R.

of the seven hues) and the shape change of the curve at different overall chroma levels. These interpolated relationships were used as lookup-tables in the creation of the chroma variations: for a given overall chroma level, each image pixel’s hue determined the chroma multiplier to be applied to that pixel’s CIELAB chroma.

The processing chain was designed for color enhancement, meaning that color accuracy was not the goal, but it is worth pointing out that chroma variations were not the only changes introduced. Because the wide-gamut primaries were different in color than the original images’ Rec. 709 primaries, the first processing step introduced minor hue and lightness errors throughout the hue range. In the previous experiment, where only single-hue images were used, this was not an issue, but for multi-hue images, the inter-hue relationships became somewhat distorted. That being said, the distortion was introduced at the beginning of the processing path, and the experimental chroma variations held them constant, ensuring that hue and lightness errors did not correlate with the results. Regardless, to be confident that the distortions did not introduce distracting artifacts, the experimenters surveyed all of the experimental images at a “normal” level of chroma and agreed that the renderings were subjectively pleasant and sufficiently natural to be reasonable starting points for the chroma variations.

5.3.3 Painting Task

Generally applying the hue-specific results of Experiment 1 requires an assumption that the chroma preferences for isolated hues would be maintained when multiple hues are present in the same image. Rather than exhaustively testing hue combinations to

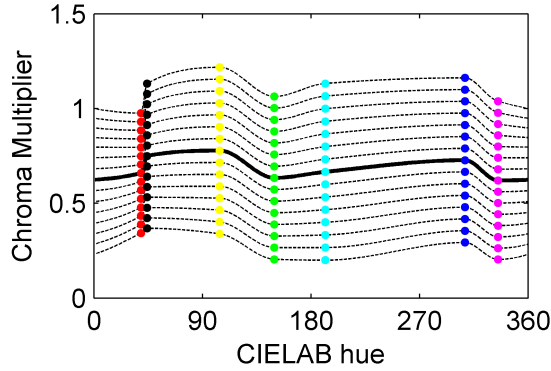


Figure 5.5: Resulting chroma multipliers. Interpolated chroma multipliers for different overall chroma levels as a function of CIELAB hue angle. The solid line represents the mean preference level, and the colored dots show the seven measured hues.

prove this assumption holds, a secondary task was devised with the multi-hue images in which the observers were asked to qualitatively indicate their satisfaction with inter-hue relationships. Each observer was shown an image at the chroma level they had personally selected moments earlier, and they were asked to paint, using a mouse cursor, areas they felt were too high in colorfulness, resulting in a binary mask image. Then the same task was repeated, but they were asked to paint areas that were too low in colorfulness, making a second mask image. In both cases, they were free to paint nothing, indicating satisfaction with the colorfulness of all areas of the image. The painting task was done with a subset of eleven of the multi-hue images, shown in Figure 5.3, and the resulting masks were saved for analysis.

5.3.4 Results

In the tuning task of this experiment, the observers were asked to select their preference among images varying in overall chroma level, which incorporated a different linear scaling function for every hue. The overall chroma level is a convenient parameter; however, its interpretation is complicated slightly by the gamut limitations of the display used in the experiment. The full amount of chroma enhancement required by the overall chroma levels was not always achievable, depending on the hue content of the images, meaning that the actual chroma of colors in the experiment was sometimes clipped. Generally, as the overall chroma level was increased, perceived colorfulness increased as well, to a point, then the colorfulness increase “slowed down” as more image pixels became clipped. To describe the chroma modulation while properly accounting for hue-

and image-dependent clipping, a corrected chroma multiplier (CCM) was computed. CCM was the effective chroma multiplier averaged over all non-neutral pixels (chroma > 12.5) of an image. CCM included the nonlinear effect of clipping, and as such, might be different for a given overall chroma level depending on an image's color distribution. It was defined such that it corresponded well visually to the colorfulness increase of the images. Values were normalized such that a CCM of 1 corresponded to the mean preferred chroma level, per hue, from the results of Experiment 1. Hence, the expected result in Experiment 2, ignoring differences due to image content, was a CCM of 1.

The results of the two experiments may be compared. Figure 5.6 shows a boxplot of preferred CCM for the ten different hues of the image paint, which was used in both experiments. The median CCM is virtually independent of hue for the seven hues used in Experiment 1, meaning that the previous result is confirmed and showing that the hue-dependent chroma multiplier performed as expected. Interestingly, the one hue that deviates is yellow, which was already suspected to be unreliable due to the limitations of the wide-gamut display. Also, the three additional hues added in this experiment behave similarly, so it is apparent that the smooth interpolation of the measured hues was reasonably successful. All the median CCM values are larger than the predicted value of 1, which is somewhat surprising, but perhaps explained by different image content and the known but unmodeled content dependence of preference. To make one point of comparison, the mean chroma preference for the image paint was computed for both experiments in terms of CCM. The mean CCM for this image in Experiment 1 was 1.16, while the mean in Experiment 2 was 1.26; this small increase is possibly due to observer differences.

Pooling over all images, Figure 5.7 shows a boxplot of preferred CCM per hue. The first ten boxes correspond to the ten hues used for rendering the three original single-hue images of Experiment 2, and the final box represents the CCM pooled over all 51 multi-hue images. The median CCM value is fairly independent of hue, demonstrating that the hue-dependent chroma multiplier concept is also applicable to new content, and even (on average) to images containing multiple hues. However, again the median CCM is consistently larger than 1, meaning that the overall preference was for a higher chroma value in Experiment 2 than in Experiment 1. The reasons for this may include different content and different observers. There is a strong dependence on content, and it is interesting to note that for complex, multi-hue content, the preferred CCM more closely matches that of the images with higher preferred chroma in Experiment 1, namely paint and room.

An ANOVA was performed on the single-hue stimuli only, with the preferred CCM as dependent variable, the image content and hue as independent variables including their 2-way interaction, and the observers as random factor. The analysis yielded a significant effect of content ($F=72.4$, $df=2$, $p<0.001$, $\eta_p^2=0.633$), of hue ($F=3.29$, $df=9$, $p<0.001$, $\eta_p^2=0.073$) and of observer ($F=4.62$, $df=42$, $p<0.001$, $\eta_p^2=0.697$) on preferred chroma level. The hue-dependent chroma multiplier was designed to make the preferred

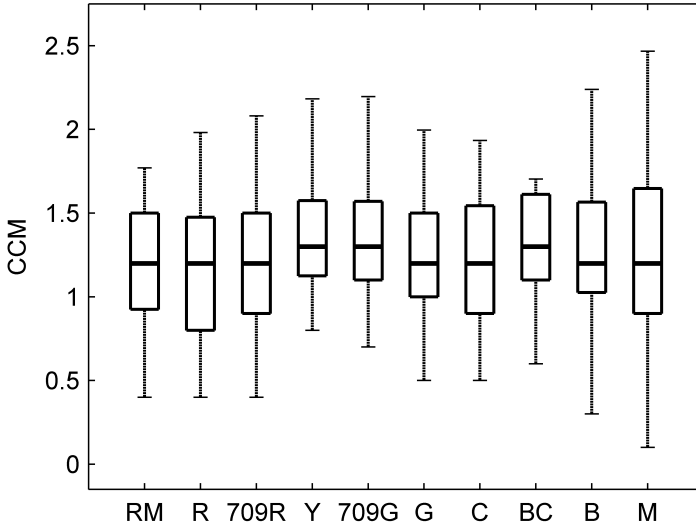


Figure 5.6: Partial reference results. Boxplot of preferred chroma (expressed in CCM) per hue for the image paint. Each box indicates the the 25th, 50th (median), and 75th percentiles of observations, and the long whiskers reach to the min and max values observed.

chroma independent of hue, but there remains a significant, though small, effect of hue on preferred chroma. Once again, suspicion falls on the Y hue, whose result stands out from the others. For this hue, the predicted chroma multiplier is too small, but this is likely due to the uncertainty caused by the chromaticity gamut capability of the wide-gamut display, which at the Y hue is not much wider than Rec. 709. In Experiment 1, there was a significant interaction effect between hue and observer, but it is not significant in Experiment 2. This improvement shows that the hue-dependent gain, on top of the offset, used to compute the hue-dependent chroma multipliers, successfully described inter-observer differences. The significant content dependence shown by the ANOVA again shows the importance of this factor, something that should not be forgotten even though it is not addressed by the current work.

The painting task provides some interesting results. For each of 11 images, each observer painted 2 masks, on one indicating which regions they considered too colorful, and on the other indicating which regions were not colorful enough. Most strikingly, 55% of all masks were blank, indicating satisfaction with the colorfulness of all hues. In cases where masks were painted, some inferences can be made about the strategy the observers

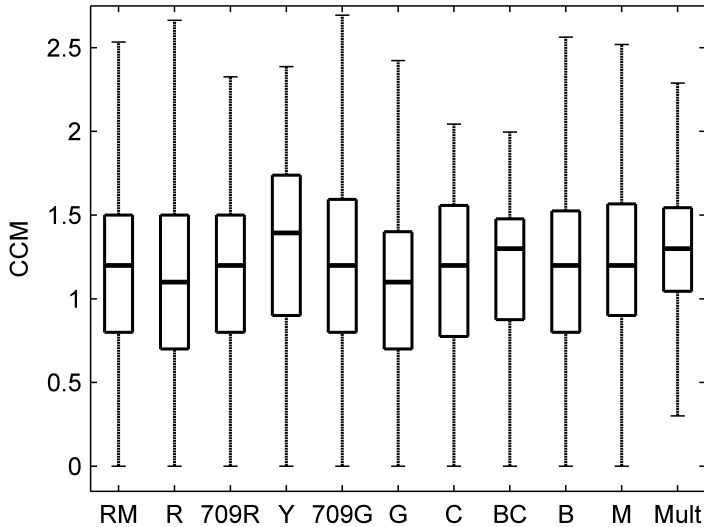


Figure 5.7: All preference results. Boxplot of preferred chroma (expressed in CCM) pooled over the three single-hue images per hue. Each box indicates the the 25th, 50th (median), and 75th percentiles of observations, and the long whiskers reach to the min and max values observed. The last box, labeled “mult,” represents the data pooled over all multi-hue images. The expected value based on the previous experiment was 1.0, while the present median result over all hues is 1.2.

used when selecting their preferred level of chroma in the tuning task. In general, if the one-dimensional tuning of overall chroma level did not provide satisfactory performance for all colors, the result from the observer’s perspective was a choice among an image with some colors over-saturated, some colors under-saturated, or some of both. The paintings indicate which strategy the observer chose; for example, if an observer painted regions indicating too much colorfulness and didn’t paint regions indicating too little, then he or she apparently chose an overall chroma level that avoided colors having too little colorfulness while tolerating too much colorfulness in other colors. The per-observer data shown in Figure 5.8 are the numbers of masks painted (out of 11 possible) to indicate image regions with too much color (left plot) and the numbers of masks painted (again out of 11) to indicate too little color (right plot). For many observers, the left plot shows more masks painted than the right, which means that these people preferentially tolerated an image with some over-saturation rather than one with some under-saturation. Overall, 44% of the too-colorful masks were painted, and 38% of the not-colorful masks were

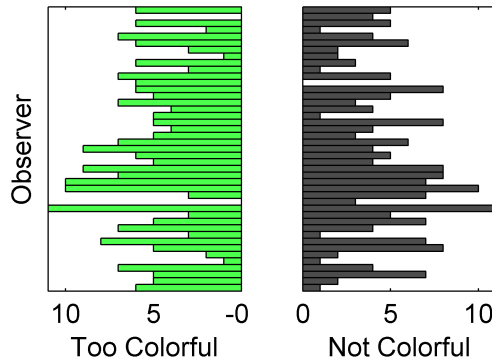


Figure 5.8: *Numbers of image masks painted by each observer. At left, the green bars show the number of images (of 11 possible) that were painted to indicate regions of too much colorfulness, and at right, the gray bars shows the number of images (again 11 possible) that were painted to indicate regions of too little colorfulness.*

painted, and almost all observers painted both kinds of masks, albeit not necessarily for the same image. In fact, observers very rarely painted both kinds of masks for the same image – this case occurred in only 8% of all image-observer combinations – which suggests that the observers were purposefully selecting their strategy on a per-image basis.

Some examples of the painted images are shown in Figure 5.9. Looking at the image content in the areas that were painted, the hue most commonly indicated too colorful was red, and the hues most commonly reported not colorful enough were yellow and lime-green. One interesting example is blue sky, which was painted in both directions: as too colorful when the original was a deep blue sky, and as not colorful enough when the original sky was hazy or pale blue. This result points to an optimum chroma for blue sky and illustrates the importance of memory colors and the contextual complexity of content dependence.

5.4 Conclusions

The two experiments discussed here provide useful results for the preferred wide chromaticity gamut boundaries for the rendering of standard gamut content. Observers generally preferred image renderings that were moderately higher in chroma than the standard rendering, while still remaining for most hues within the physical limitations of display technology available today. Observers were willing to accept chroma levels somewhat higher than the preferred levels, but for some hues, current wide-gamut dis-

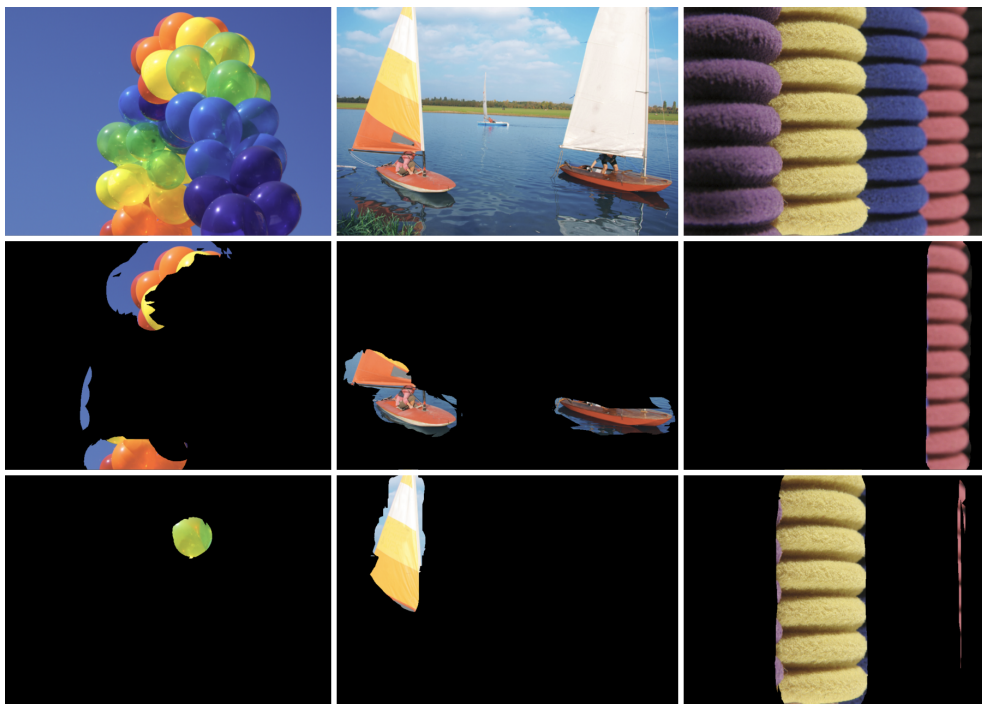


Figure 5.9: Painting task results. The top row shows the starting image for three examples, the second row shows the images masked with a thresholded sum of observers’ paintings that indicate regions too high in colorfulness, and the third row shows the images masked to indicate regions too low in colorfulness.

play technology provided unacceptably high chroma. The measured chromaticity gamut boundary preference depended on observer, hue, and image content. The content dependence was complex and not easily modeled, but the observer preference variation could be largely modeled with a single overall chroma parameter that incorporated first- and second-order hue dependence. This hue-dependent chroma multiplier was verified to be successful with real world images, in which observers appeared to systematically balance regions of too-high and too-low colorfulness, depending on the image content. Overall, they tended to prefer to err on the high side, more often selecting the higher-chroma image when the inter-hue balance of colorfulness was not perfect. This suggests that display design should be biased toward higher chroma, assuming the consumer is given an overall colorfulness adjustment control that takes into account the hue dependence described. The results include some uncertainty in yellow hues due to the chromaticity gamut limit of the display used for the experiments, so this should be confirmed in a

follow-up with a more capable display, such as a multi-primary display including a yellow channel. Additionally, results using multi-hue images suggested that the chroma preference level for red that was found using single-hue images was too high for general use, and should be verified.

5.5 Bibliography

- [1] ITU-R. BT.709 : Parameter Values for the HDTV Standards for Production and International Programme Exchange, April 2002.
- [2] Adrianus J.S.M. de Vaan. Competing Display Technologies for the Best Image Performance. *Journal of the Society for Information Display*, 15(9):657–666, 2007.
- [3] R. W. G. Hunt. *The Reproduction of Colour*. Fountain Press, Kingston-upon-Thames, 5th edition, 1995.
- [4] E.A. Federovskaya, H. de Ridder, and F.J.J. Blommaert. Chroma Variations and Perceived Quality of Color Images of Natural Scenes. In *Proc. SPIE*, volume 2411, pages 51–61, 1995.
- [5] Huib de Ridder. Naturalness and Image Quality: Saturation and Lightness Variation in Color Images of Natural Scenes. *Journal of Imaging Science & Technology*, 40(6): 487–493, 1996.
- [6] R. Muijs, J. Laird, J. Kuang, and S. Swinkels. Subjective Evaluation of Gamut Extension Methods for Wide-Gamut Displays. In *Proc. IDW*, pages 1429–1432, 2006.
- [7] J. Laird and I. Heynderickx. Perceptually Optimal Boundaries for Wide Gamut TVs. *Proceedings of the SPIE - The International Society for Optical Engineering*, pages 1–10, 2008.
- [8] Y. Hisatake, A. Ikeda, H. Ito, M. Obi, Y. Kawata, and A. Murayama. The Ergonomics Requirement for Reproducible Area of Color Chromaticity in Electronic Displays. In *Proc. IDW*, pages 2301–2304, 2007.
- [9] Masato Sakurai, Rodney L. Heckaman, Stacey E. Casella, Mark D. Fairchild, Takehiro Nakatsue, and Yoshihide Shimpuku. Effects of Display Properties on Perceived Color-Gamut Volume and Preference. *Journal of the Society for Information Display*, 16(12):1203–1211, 2008.
- [10] Dragan Sekulovski, Robert de Volder, and Ingrid Heynderickx. Preferred and Acceptable Color Gamut for Reproducing Natural Image Content. In *SID Symposium*, pages 1014–1017, 2009.

Preferred Color Gamut Boundaries for Wide-Gamut and Multi-Primary Displays

Abstract

Preferred chroma enhancement and its dependence on hue are studied in a two-part experiment using a wide-gamut multi-primary display. Earlier research showed a clear dependence on hue but was limited by the gamut of the display it employed; the present work builds on this while easing the gamut constraints. In the first part of the present experiment, a tuning task was used to refine the preference for chroma boost starting with standard-gamut (Rec. 709) images. The overall median preferred boost is roughly 20%, but it is not uniform over hues: the preferred boost for orange, yellow, green, and cyan colors is greater than that for blue, magenta, and red colors. Dependence on image content and observer is noted, though a content-independent chroma boost created by aggregating preference over many images performs well. An adjustment parameter for overall chroma which incorporates the hue dependence averaged over image content should be sufficient to handle the vast majority of inter-observer variance in preference. In the second part of the experiment, various chroma boost algorithms were evaluated through a paired comparison task. The prescribed hue-dependent chroma boost is preferred over all other variations, and all hue-preserving chroma boost variations are preferred over both colorimetrically accurate and naive same-drive-signal renderings. The results may be applied in display design to select gamut boundaries that maximize satisfaction over the observer population.¹

6.1 Introduction

Despite the recent entry of wide-gamut displays into the marketplace, basic questions of how wide the gamut must be and how best to address it remain insufficiently answered.

¹This chapter has been published in the journal *Color Research & Application*, doi: 10.1002/col.21780 (2012): "Preferred Color Gamut Boundaries for Wide-Gamut and Multi-Primary Displays," Michael J. Murdoch, Dragan Sekulovski, and Ingrid E. J. Heynderickx.

These questions are especially important as long as most input sources remain standard-gamut content (i.e. Rec. 709, the standard for HDTV broadcast and sRGB internet images [1]), because image enhancement is then required in order to use the wide gamut. While it is well known that people prefer images that are more colorful than what is recognized as natural or realistic [2, 3, 4], inconsistent and generally poor results are obtained by simply providing an arbitrarily wide gamut and addressing it naively.

The question of required gamut size has been addressed by several researchers. Rather than arguments based on surface colors or self-luminous colors, a practical approach is to study what range of colorfulness people prefer. Early looks at preference-based image chroma requirements were provided by Hisatake et al. [5] and by Laird and Heynderickx [6]; in both papers, a moderate boost in chroma was appreciated by viewers, depending on both hue and image content. The present authors previously reported experimentally-measured hue-dependent chroma preference for highly-chromatic images, with the intent of providing preference-based guidance on the size and shape of display gamut boundaries [7]. The results indicated a need for more color gamut in yellow hues which is addressed in the present paper.

The question of how best to address a wide-gamut display with incoming standard-gamut images has also been studied to some extent. Historically, gamut mapping research has focused on gamut compression, with special emphasis on squeezing the gamuts of additive color spaces (sRGB, Adobe RGB, etc.) into the gamuts of subtractive color systems, such as CMYK printers and printing presses. Gamut expansion accomplishes the opposite of this, though both are concerned with how far and in what manner to modify image colors. The characteristics of the destination wide gamut are of course of prime consideration, but just as important are the characteristics of the starting image, including its color space and its content and context. Preferences for the expansion method itself, meaning how the starting gamut is mapped into the available destination wide gamut, were reported by Laird et al. [8] They compared matrix-based enhancements, naive same-drive-signal (SDS) mappings, and hue-preserving linear and nonlinear chroma boosts, finding a preference for hue-preserving chroma boosts with a nonlinearity that enhanced high-chroma colors more than low-chroma colors. Additionally, the importance of treating some colors in special ways, such as protecting skintones from excessive boosting, is well recognized in gamut mapping. Chen et al. provide a good example in their method of skintone segmentation and hue mapping which avoids contouring at segment boundaries [9].

Though the questions of gamut boundaries and gamut expansion cannot be completely separated, this nonetheless has been the goal throughout the present authors' research. As reported previously, we intentionally utilized image content biased toward the outer edges of the gamut area, thereby allowing the use of linear chroma expansion, and we avoided skintones altogether [7]. In the present paper, we continue with both of these constraints. Figure 6.1 summarizes the results of our prior paper, showing three chromaticity gamuts: the standard gamut of the starting image encoding, Rec. 709; the

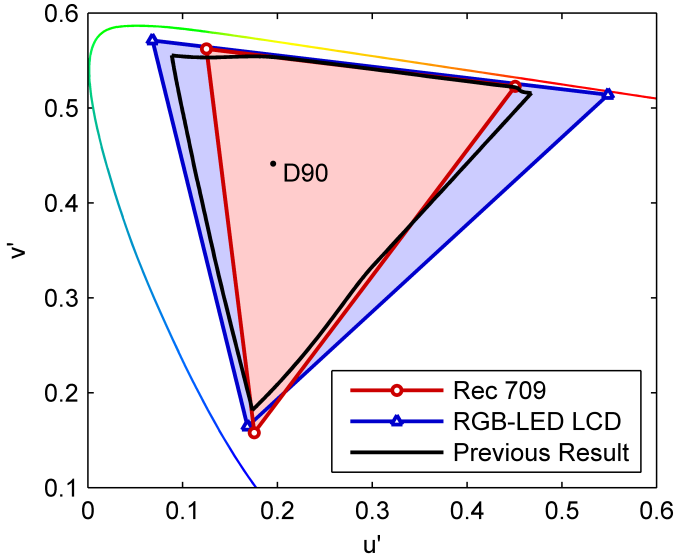


Figure 6.1: Previous result. The median result from the authors' previous work, shown as a black line on this top-view $u'v'$ chromaticity diagram, is a moderate expansion from the starting Rec. 709 chromaticity gamut, shown in red. The blue chromaticity gamut is that of the display used for the experiments, which does not offer significant expansion for blue or yellow hues.

median preferred gamut boundary, interpolated between sampled hues; and the available wide gamut of the RGB LED-backlit LCD that was used in the earlier experiments. It is apparent from the plot that the previous results were potentially limited by the capabilities of the wide-gamut display in yellow and blue hues. Additionally, viewers explicitly indicated they would prefer more yellow chroma than was available. To further test hue-dependent preferences while removing the limitations and potential bias of the wide-gamut RGB display, we specified a new 5-primary (red-green-blue-cyan-yellow: RGBCY) wide gamut display and crafted new viewing experiments.

6.2 Methodology

A two-part experiment was designed to take advantage of the color gamut capabilities of the multi-primary display. The first part of the experiment measured hue-dependent chromaticity gamut expansion requirements for natural images via a tuning task similar to the previous work, and the second part of the experiment incorporated paired comparisons

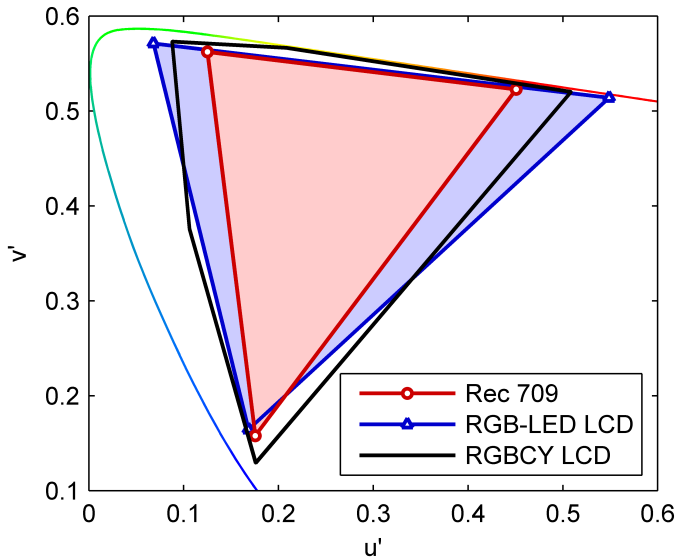


Figure 6.2: New multi-primary display. Chromaticity gamut comparison in a $u'v'$ chromaticity diagram of the starting image color space, Rec. 709 (red), the previous experiment's RGB display (blue), and the present experiment's RGBCY multi-primary display (black).

between the predicted preferred hue-dependent gamut expansion method and other methods.

6.2.1 Multi-Primary Display

The multi-primary display (MPD) used in the experiment was a prototype 5-primary (RGBCY) liquid-crystal display with a customized color filter array providing expanded chromaticity gamut, especially in the yellow and cyan hues. For comparison, Figure 6.2 shows the chromaticity gamuts of this RGBCY display, the Rec. 709 standard, and the RGB-LED LCD used in the previous experiment. The advantages in blue and yellow hues, compared to the previous RGB display, are valuable improvements for the present experiment, while the reduced capability in green and red is not seen as a disadvantage, according to the preferences measured previously.

The MPD was measured with full-screen color patches throughout its entire range using a Photo Research PR-680 spectroradiometer. The maximum luminance of the display with all 5 channels fully open was 457 cd/m^2 ; however, as discussed in the next

paragraphs the maximum white luminance used in the experiment was half of this value. The display was not calibrated in the sense that its behavior or LUTs were modified to make it behave like a standard monitor, but instead these characterization data were used as a forward model that accurately described its behavior. This behavior was used in the generation of a colorimetrically-accurate 3D-LUT that both inverted the forward model and performed the 3-to-5 channel multi-primary mapping.

Color rendering on a multi-primary display is nontrivial because each additional primary beyond RGB adds an additional degree of freedom that must be constrained in some way. Hinnen et al. explained methods of multi-primary mapping, possible constraints for the additional degrees of freedom, and the additional possibility of sub-pixel rendering [10]. For the purpose of this experiment, the main conceptual constraint was visual smoothness, i.e. the avoidance of sharp transitions in color gradients, so sub-pixel luminances were equalized as much as possible in the 3D-LUT implementation. Note that colors very close to or beyond the MPD gamut boundary were explicitly soft-clipped to avoid hard-edged artifacts.

In general, the use of more than three primaries in a filtered-light system such as an LCD results in lower luminances of the primaries relative to white, when compared to a typical 3-primary system, as discussed by Hinnen. This can produce unnaturally dark red and green colors, for example. Though it is known that a moderate amount of primary-dimming is acceptable [11], in order to avoid any distortion of the results of the present experiment, the choice was made to lower the display white luminance in software while leaving the pure primaries unconstrained. In the MPD, there were half as many green and blue subpixels as would be in a normal RGB display, so that by constraining white to half of its physical maximum luminance, the relative luminance ratios between the primaries and white were the same as they would be in a RGB display. Note that while Heckaman and Fairchild constrained white relative to pure colors in an RGB display with the goal of increasing colorfulness beyond the display's native performance [12], in our case the goal was to maintain colorfulness despite the use of more primaries. Another way to look at this is that the display efficiency was intentionally lowered by implementing a luminance limit in the video data. Thanks to this constraint, the 3D-LUT mapping from RGB to RGBCY was able to preserve luminance for all colors.

For the experiment, images on the MPD were shown with height of 600 pixels, corresponding to 15 degrees of visual angle from the viewing distance of 2.5m. The images were presented on a uniform gray background of L^* 50 at the display white point of D65, and the viewing lab utilized dim, indirect lighting amounting to 4 lux horizontal illumination on the desk surface.

6.2.2 Image Stimuli Preparation

There were two distinct sets of image scenes used in the experiment, one consisting of single-hue images rendered at a variety of CIELAB hue angles, and the other consisting of



Figure 6.3: Single-hue stimuli. The 6 single-hue scenes shown at some of the 12 hues used in the first experiment.



Figure 6.4: High-chroma stimuli. Representative scenes from the high-chroma image set.

natural images with very high-chroma content. Scenes from each set appear in Figure 6.3 and Figure 6.4. For clarity, the image content and source images will be referred to as scenes, and the rendered variations as images or image stimuli.

The scenes were selected from Creative-Commons licensed content available on web sites such as Flickr². Their processing pedigree and capture conditions are unknown, but this lack of control is typical for display applications and part of the reason for using a wide variety of scenes. The image processing chain makes the assumption that they are properly encoded as sRGB.

The six single-hue scenes as shown in Figure 6.3 were intentionally processed to have a very narrow range of hues. They began as regular sRGB images and were converted to CIELAB LCh so that the hue channel h could be constrained to the 12 target hue angles: CIELAB h values of 10, 39, 44, 65, 101, 125, 150, 192, 214, 270, 312, and 335. The target hues are the primary and secondary hues of relevant display systems (including, for example, the red primaries of both Rec. 709 and the RGB LED display from previous experiments) as well as some intermediate hues to further sample the hue circle. The hue-rotation procedure was the same as that explained in detail previously [7, 13]. To summarize, the h values were simply replaced with the target hue, and the L^* and C^* values modified to account for the difference in gamut cusp between the start and

²<http://www.flickr.com>

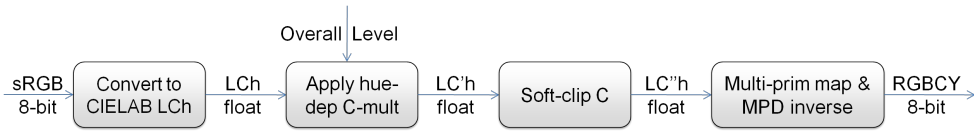


Figure 6.5: Stimuli processing. Flow diagram for the creation of chroma level variations for display on the MPD in the experiment. Input images were 8-bit per channel sRGB, which were converted to CIELAB LCh. The chroma channel C was modified using the chroma multiplier LUT for the desired overall chroma level, then soft-clipped to ensure it would be in-gamut on the MPD. The resulting LCh values were mapped to 8-bit RGBCY display drive values using the multi-primary mapping and MPD inverse 3D-LUT.

destination hues. These six starting scenes each at 12 hues made a total of 72 single-hue scenes.

The 52 high-chroma natural scenes used in the experiment were selected for their wide variety of content and high-chroma color palette. A preponderance of highly-saturated colors was chosen in order to bias the experiment toward the gamut boundaries, because the preferred expansion of colors near the borders of the starting gamut defines the requirements for a wide gamut display. In all scenes, humans and skintones were omitted, because no special treatment of skintones was implemented in the experiment. The worry was that observers would push the chroma expansion of a scene differently, probably less aggressively, with skintones involved.

In preparation for the experiment, scenes were pre-processed to a set of image stimuli with a wide range of colorfulness for the observers to select among. The flow diagram illustrating the creation of the image stimuli is shown in Figure 6.5. The process used the hue dependence derived in previous experiments; however, the application of the chroma boost was improved. Previously, a hue error was knowingly introduced in the image processing because the RGB primaries of the RGB-LED backlit display were different in hue than the sRGB primaries and the goal was to maximize available color gamut. In the present experiment, because it was known that the maximum gamut was not needed, image hues were preserved accurately.

An overall level parameter was used as an index into a family of hue-dependent chroma multiplier LUTs corresponding to different levels of overall colorfulness. The levels ranged from 0 to 4 in steps of 0.1 units, on which scale 1.3 was the median preferred level of chroma enhancement based on previous experiments. At this overall level, the chroma multiplier per hue ranged from about 0.9 in some hues to about 1.3 in others. LUTs with overall levels from 0 to 3 are shown in Figure 6.6, and the median level of 1.3 is emphasized. Note that the hue-dependence (the deviation from a constant value) changes

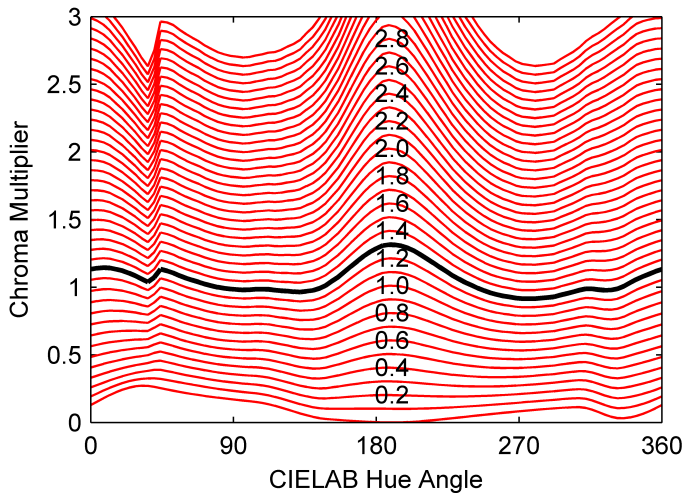


Figure 6.6: Hue-dependent chroma multiplier LUTs. These are shown for overall chroma levels ranging from 0 to 3 (labeled). The previous experiment median preferred level of 1.3 is emphasized by the heavy black line.

with overall colorfulness level, a feature based on a model of inter-observer preference variance obtained from previous experiments. In the implementation, for each pixel in an image and for a given overall chroma level, the LUT value corresponding to the pixel's hue was used as a multiplier for the pixel's CIELAB chroma C^* . Then, MPD digital drive values were computed as explained in the previous section.

For the second part of the experiment, five variations of the chroma enhancement were created, illustrated schematically in Figure 6.7. Three of the variations consisted of: the same hue-dependent chroma boost as in the tuning part of experiment, called preferred chroma (PC); flat chroma (FC), in which the chroma of all hues was boosted uniformly; and an inverse of the preferred chroma (IC) which boosted the chroma according to the opposite of the preferred hue dependence. Image stimuli were created for each of these three variations at all 41 levels of overall chroma level. The two additional variations were the original sRGB image (sRGB), correctly (colorimetrically) rendered for the MPD, and a version mimicking RGB same-drive-signal (SDS) rendering, which ignored the cyan and yellow primaries and treated the MPD as a wide-gamut RGB display, naively stretching the scene's colors to fill its gamut (arbitrarily scaling chroma and distorting hues according to the specifics of the wide-gamut RGB primaries). These two variations did not have different overall chroma levels.

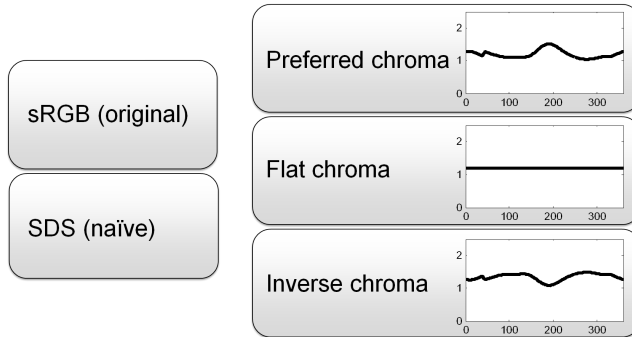


Figure 6.7: Chroma variations. These variations were prepared for the second part of the experiment. At left, the original image, colorimetrically displayed according to sRGB, and the SDS “naïve” gamut expansion to fill the RGB display’s gamut. At right, three variants of hue-preserving, hue-dependent chroma boost, showing mini-plots to illustrate the hue dependence: the preferred chroma, using the same hue dependence as in the first experiment; flat chroma, meaning hue-independent chroma scaling; and inverse chroma, a hue dependent mapping using the opposite relative scaling with respect to what was preferred.

6.2.3 Experimental Tasks

The first part comprised a tuning task using the method of adjustment, in which observers selected their preferred level of colorfulness among the set of pre-processed variations for each scene. Observers were asked to “tune” the level of overall colorfulness to their preference, and the levels presented corresponded to chroma-scaled versions of the predicted preferred chroma enhancement from prior experiments. Each scene was presented in randomized order by a Java program that also controlled the user input and data recording. For each scene, the initial image stimulus had a level of overall chroma randomly selected to be at or below the median preferred level from the previous experiments. Using the keyboard arrows, the observer could change the level of overall chroma in the scene: the left and right arrow keys moved one step, and the up and down arrow keys moved five steps (out of a total of 41 steps). The observer was free to adjust the overall chroma at will, higher, lower, and back-and-forth without any time restriction. When the observer reached a satisfactory overall level of chroma, he or she pressed “Enter” to register the preference and move on to the next scene. In practice, tuning a single scene took a typical observer only 20 to 30 seconds, and was considered a rather easy task.

In the second part of the experiment, which was conducted immediately after the first, paired comparisons allowed each observer to evaluate the preferred chroma

enhancement along with other methods of chroma enhancement. For this task, twelve scenes were selected from those used in the first part of the experiment. Because we found considerable observer variation in overall preferred chroma level in previous experiments, each observer's personal preferred chroma from the tuning experiment immediately preceding was used in the comparisons. The paired comparison task consisted of a series of 2-alternative forced choice (2AFC) presentations, in which a pair of image stimuli was shown side-by-side on the screen and the observer selected one of them using the left or right arrow keys. Observers were instructed to choose the image that they preferred overall, and if they could not tell a difference or did not have a preference, to select one randomly. No time constraints were employed. Five variants of chroma enhancement, yielding 10 pairwise comparisons for each of the 12 scenes, were presented to each observer in random order by a Java program that also recorded the results.

6.2.4 Observers

27 observers participated in the experiment, all of whom had normal color vision according to the Ishihara test and Landolt-C visual acuity of 1.0 or better. The group included 14 males and 13 females and had a median age of 29 in a range of 22 to 50 years. All of the observers were from the Philips Research environment, including students and regular employees, and they received no additional compensation for their participation.

6.3 Results

The results of the tuning task provide a picture of the dependence of chroma boost preference on hue, scene content, and observer. The parameter under the control of the observer during the tuning was the overall chroma level, which ranged from 0 to 4 in steps of 0.1 units. However, because the MPD had a physical gamut limit, some pixels in some images at high values of overall chroma level were clipped and thus not displayed at the correct level according to this parameter. In order to take clipping into account, an effective chroma multiplier (ECM) was computed, which for a given image is simply the ratio of output (including chroma boost and clipping) chroma to input chroma averaged over all pixels with non-zero chroma. The ECM was computed for each scene at each overall chroma level, and the basic relationship between ECM and chroma level resembled the clipping function, albeit scaled differently for each image stimulus according to the particular chroma distribution. The ECM metric was more representative of what the observers actually saw on the screen than the input parameter of overall chroma level, and therefore ECM was used in the statistical analysis of the results. Note that ECM was conceptually the same as the corrected chroma multiplier (CCM) that was reported in 2010 [7], but including all starting chroma values larger than zero rather than using the unnecessarily conservative constraint of chroma larger than 12.

6.3.1 Single-hue Scenes

One goal of the present experiment was to verify previous results with a more capable display, thus a detailed look was made at the hue dependence of the preferred chroma boost. The 6 single-hue scenes were each presented for tuning at 12 different hues. Using SPSS, an ANOVA was computed with preferred ECM as the dependent variable, scene and hue as fixed factors, and observer as a random factor. All factors and their two-way interactions were found significant ($p < 0.001$). Medium-sized effects were seen for observer ($\eta_p^2 = 0.589$), meaning that people prefer different levels of overall chroma; for scene ($\eta_p^2 = 0.554$), meaning that image content plays a large role in determining preferred chroma; and of course hue ($\eta_p^2 = 0.419$). The interaction between scene and observer was also medium in effect size ($\eta_p^2 = 0.507$), showing that different people respond to image content in different ways. The interaction of hue and observer was small in effect size ($\eta_p^2 = 0.218$), which means that when the observers' overall level preference is accounted for, the relative chroma boost per hue doesn't differ strongly. Finally, the interaction between hue and scene, while statistically significant, is extremely weak ($\eta_p^2 = 0.076$), meaning that the hue dependence does not strongly depend on image content.

The distributions of results according to hue are shown in detail in Figure 6.8, along with those for the high-chroma scenes. Though the red violin envelopes show that the distributions are not strictly normal, the observed mean and median values generally correspond closely for the single-hue scenes. The hue dependence indicated by the ANOVA is clearly seen, and indeed a Tukey post-hoc analysis showed several overlapping subgroups, with the groups of lowest 5 hues (10, 39, 270, 312, 225) and highest 5 hues (65, 101, 150, 192, 214) distinct.

The results of the present experiment may be compared with the previous experiment [7] as seen in Figure 6.9. In both experiments, biggest contributor to inter-observer differences is an overall level preference, thus it is not surprising that the mean ECM of the present experiment over the twelve measured hues, 1.012, was slightly lower than the mean of the previous result over its seven measured hues, 1.065. For this reason, both results are shown with their mean values subtracted. The present curve is a cubic interpolation of the results measured at twelve discrete hues (labeled, and at which 95% confidence intervals are shown), constrained to be smooth around the hue circle. The previous curve was interpolated between only 7 hues, but additionally incorporated the effect of chroma distortions introduced by the RGB LCD used previously: specifically, the curvature near 10 and 270 degrees that would not have appeared with simple interpolation between the 7 measured hues.

Comparing the previous and present results, there are some noticeable differences. Overall, the present result is slightly flatter (meaning smaller differences in preference with respect to hue). Further, the ECM for red and cyan colors is lower than previously observed, while for yellow and green it is higher. It appears that the predicted, but

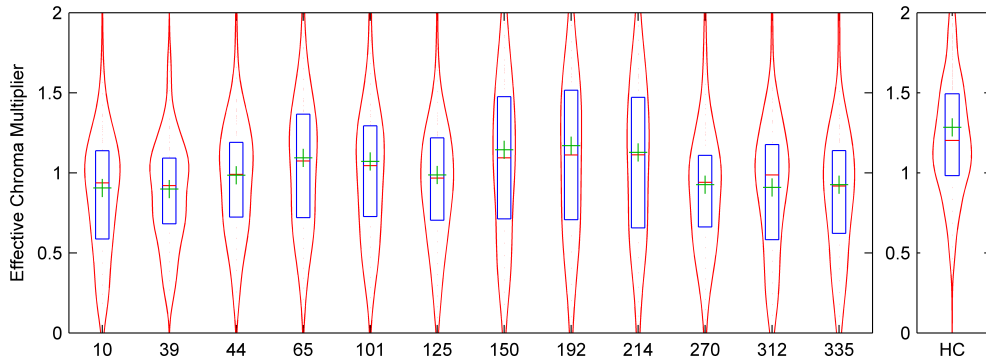


Figure 6.8: Per-hue results. Violin plots illustrating the distributions of preferred effective chroma multipliers observed in the experiment. The larger subplot shows distributions over observer (27) and scene (6) for each of the single hues indicated on the x-axis, while the smaller plot shows the distribution over observer (27) and all high-chroma scenes (52). The central blue boxes indicate 25th and 75th percentiles, with median as a red tick and mean as a green plus. The surrounding red violin envelopes represent kernel-smoothed estimates of the local probability density of the observations.

previously unmeasured bump at 10 degrees was not preferred, while the dip at 270 was preferred. Based on the confidence intervals, it may be concluded that the present result is indeed significantly higher in yellow and green hues, as was expected based on the limitations of the previous research. The results match very well between hue angles 214 and 335, but the deviations at 39, 44, and 192 are moderate. Interestingly, the distinct difference in chroma preference between the two red hues, 39 and 44, which was measured in previous experiments, is still clearly seen, even while the absolute levels differ. It appears that as expected based on the previous experiments, making more color gamut available in the yellow region was appreciated by the observers. Also as expected, the extra cyan color gamut enabled by the MPD was not used or appreciated by the observers.

As can be seen in the confidence intervals, the variation among observers and images is not constant with hue. That is, at some hues, such as yellow-green, blue, and purple, the range of preference is tighter, while at other hues, such as green and cyan, the range is wider.

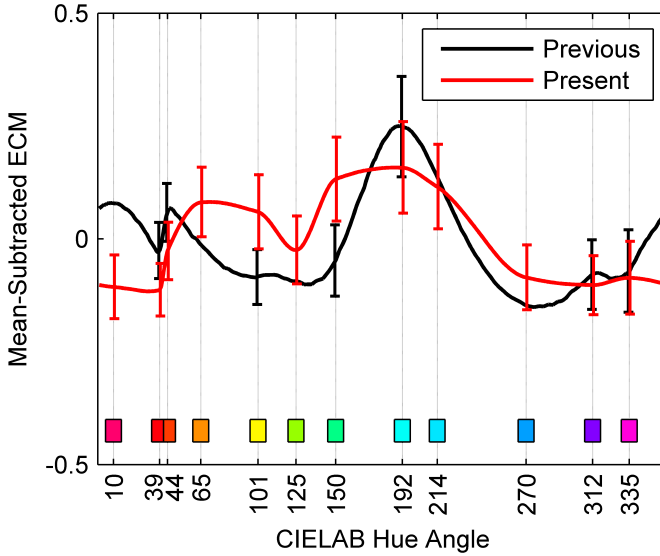


Figure 6.9: Previous and present results. Comparison of the hue dependence of ECM with that from the previous research. The present result (red) and previous result (black) are shown with their respective mean values (1.012 and 1.065) subtracted. The red curve was interpolated between the 12 marked hues, and the black curve was interpolated between 7 hues according to the behavior of the LCD used in previous experiment. Error bars at the measured hues show 95% confidence intervals.

6.3.2 Single-hue vs. Multi-hue Scenes

The single-hue scenes were important to the study of the effect of hue on preferred chroma boost, but they remain a bit contrived, and multi-hue, high-chroma scenes are much more representative of real television and internet image content. The two sets of scenes behave differently: the median ECM for multi-hue scenes is 20% higher than for single-hue scenes, and the mean is 28% higher. As computed using a 2-sample T-test, the two populations differ significantly ($p < 0.001$). Interestingly, where the mean and median ECM for single-hue scenes are nearly identical, 1.01 and 0.99, they differ more for the multi-hue, high-chroma scenes, 1.28 and 1.20, respectively, emphasizing the non-normality of the multi-hue ECM preferences.

6.3.3 Content Dependence in High-chroma Scenes

In previous experiments [7], modeling the content dependence of preferred chroma using objective measures of the scene content, such as average starting chroma, the size of highly-chromatic regions, and local contrast metrics, did not work satisfactorily. In the present experiment, the multi-hue scenes were manually categorized according to several criteria to see if high-level context provided insight into the variance in preferred ECM. 2-sample T-tests were computed to gauge the significance of the effect of each criterion on differences in mean preferred ECM. Note that compensation for significance errors due to multiple tests was not included, but this doesn't interfere with the purpose of discussion. The presence of luminous objects was significant ($p = 0.001$), with the mean ECM lower for luminous objects (mean = 1.17) than for non-luminous objects (mean = 1.30). A white background made a significant difference ($p = 0.005$), with preference for higher ECM when objects are presented against a white background (mean = 1.36 vs. 1.27). The presence of a rainbow, meaning a wide range of highly chromatic colors, led to slightly lower preferred ECM ($p = 0.019$, mean = 1.20 vs. 1.24). The presence of a blue sky was significant ($p = 0.039$), with people preferring a little less boost (mean = 1.23 vs. 1.30); however, naturalness of the image did not make a significant difference ($p = 0.77$). Also not significant was whether the subject was a close-up without much surrounding context ($p = 0.47$).

6.3.4 Paired Comparison Results

The paired comparison task was included to make a precise comparison among the different variations of hue dependence, and also to compare these to typical gamut expansion methods. The pairs data were analyzed using the method of Montag [14], which provides an estimate of significance based on the numbers of stimuli and observers. The results are shown in terms of quality scores, similar in concept to Z-scores, in Figure 6.10. The variation with preferred hue dependence (PC) was found significantly better than all other variations, and all three hue-preserving variations (PC, FC, IC) were found better than naive SDS; further, all four of these performed better than the original color reproduction (SRGB). According to Montag's method, all differences are significant with a $p = 0.05$ threshold except that between FC and IC.

6.4 Discussion

Looking at the results, the hue-dependent chroma boost based on previous experiments was largely verified with more observers and a more-capable multi-primary display. Importantly, the value of the hue-dependent boost over a hue-independent boost was proven by a significant preference in the paired comparison, and further, all of the hue-preserving chroma boost variations significantly outperformed a naive SDS rendering. We have now shown the value of hue-dependent chroma boost with both wide-gamut RGB and multi-primary displays. The MPD provided more yellow, cyan, and blue, and

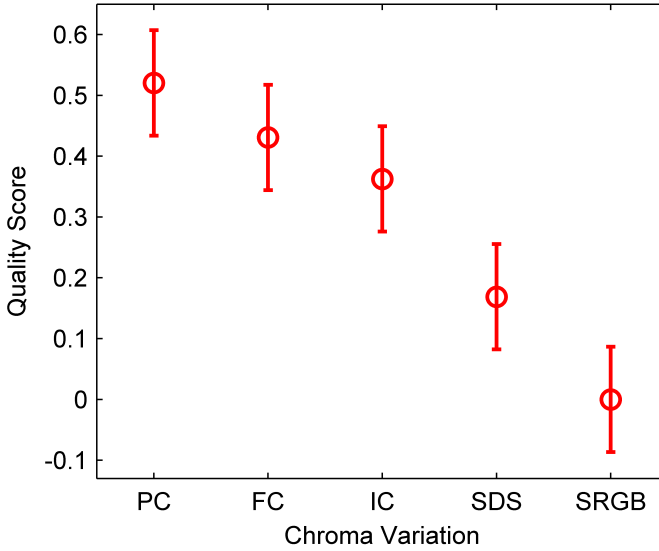


Figure 6.10: Chroma variation results. Quality scores for each of the hue dependence variations used in the paired comparison. PC refers to preferred chroma, FC to flat chroma, IC to inverse chroma, SDS to same-drive-signal, and SRGB to colorimetrically accurate color reproduction. Error bars are scaled such that if a mean value of one variation is outside the bars of another variation, the difference between those variations is significant. All differences are significant except for FC-IC.

less red and green than the RGB-LED LCD used in previous experiments, and observers appreciated the added yellow, and also preferred more green. Blue chroma preference was the same and red and cyan were lower in comparison to our previous experiment. Thus, it is not likely that chroma boost preference is simply a question of “headroom,” or distance from the display’s gamut boundary, and the implied correlation with clipping.

6.4.1 Preferred Hue-dependent Chroma Boost

The present results allow an update of the hue-dependent chroma boost provided in our previous paper [7]. As in that work, the preference behavior observed throughout the experiment was analyzed with a principal components analysis (PCA) over hue. The first principal component corresponds to the overall chroma level. The factor loadings indicate the rate of change at each hue with changes in the overall chroma level, and as was seen previously, these are also hue-dependent. Using the experimental mean as an offset and

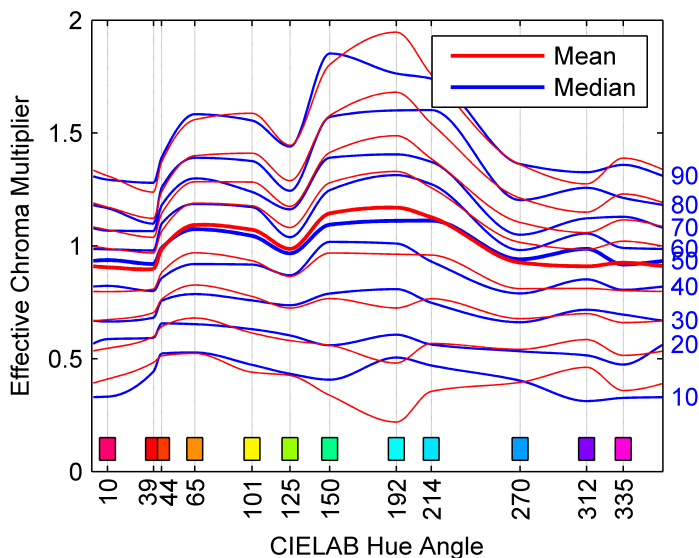


Figure 6.11: Hue-dependent chroma multiplier curves. These are shown in two ways: in blue, every tenth percentile, with the median (50th) indicated in bold. And, in red, the experimental mean (thick line) is used as an offset and PCA factor loadings as hue-dependent gains to create parameterized curves at overall levels which match the percentile-based curves.

the PCA factor loadings as slopes, a parametric family of hue-dependent chroma multiplier curves was created. These may be compared with their nonparametric analogues, as seen in Figure 6.11.

From these curves, a family of hue-dependent chroma multipliers can be created and used directly as LUTs to enhance the chroma of standard-gamut images. Using either mean-based or percentile-based curves is possible, the latter being more clearly related to proportions of the observer population, but having the handicap of depending on smaller and smaller subsets of the population as the extreme percentiles are approached. The mean-based curves depend on PCA, which linearly models the behavior of the whole population, and thus is both less likely to be affected by noise and more likely to miss complicated effects. With the present results, in which for example the 10th percentile includes only 3 observers, it seems prudent to utilize the parametric, mean-based results, though if a larger experimental data set were available, the percentile-based approach would warrant closer study. Both versions are shown in Figure 6.11 for comparison, the mean-based curves drawn to match the overall mean levels of the corresponding ten

percentiles. They match reasonably well, worse for the highest-variance hues such as cyan.

As was stated previously, preference for the chroma boost of the multi-hue, high-chroma scenes was 20-28% higher than for single-hue scenes. Because the former are much more similar to real-world images, the chroma multiplier should probably be implemented with a corresponding factor included. However, because the variation in preference over people is so large, we maintain that the user should still be given an overall colorfulness adjustment. In that case, the choice of gamut boundaries specifies only the maximum chroma, as any lower chroma is still attainable with an appropriate level choice.

6.4.2 Gamut Requirements

The hue-dependent chroma preference results may be directly applied to the design of a better display gamut for the display of enhanced standard-gamut (Rec. 709) image content. Scaling the standard chromaticity gamut boundary by the hue-dependent chroma multiplier curves, including the 20% higher median preference for high-chroma over single-hue scenes, provides expanded (and contracted) gamut boundaries, as shown in Figure 6.12. Calibrating the mean-based curves to percentile is intuitive here, because the resulting chromaticity gamut boundary provides, with properly adjustable overall boost, satisfaction for all observers whose preference is at or below that boundary. It is interesting to note that the Rec. 709 boundary lies at the 40th percentile along the B-R line and only the 30th percentile along the R-G and G-B lines: thus, the standard gamut satisfies less than half of the observer population. Some of the contours indeed exceed the gamut of the MPD and occasionally cross the spectral locus; this is an extrapolated result including the 20% boost and undoing the effect of clipping, and can be thought of as what observers would like to see.

Figure 6.12 can be used to assess the value of proposed display chromaticity gamuts, based on how much high-percentile area they enclose. For example, consider a choice between enhancing a green primary from Rec. 709 green in either the positive v' or the negative u' directions. Either enlarges the chromaticity gamut, but because the lines of equal-percentile are very closely spaced in the v' direction and less so in the u' direction, the v' -directed enhancement provides more value in terms of observer preference. This could be quantified by integrating the proposed chromaticity gamut area weighted by the cumulative percentiles, and thus provide objective comparisons.

6.4.3 Multi-primary Displays

The present results were obtained using a multi-primary display, but they are applicable to traditional RGB displays as well. If the goal is to maximize the value of the display chromaticity gamut with respect to the iso-percentile lines shown in Figure 6.12, this may be done with arbitrary arrangements of primaries. However, it is important to remember

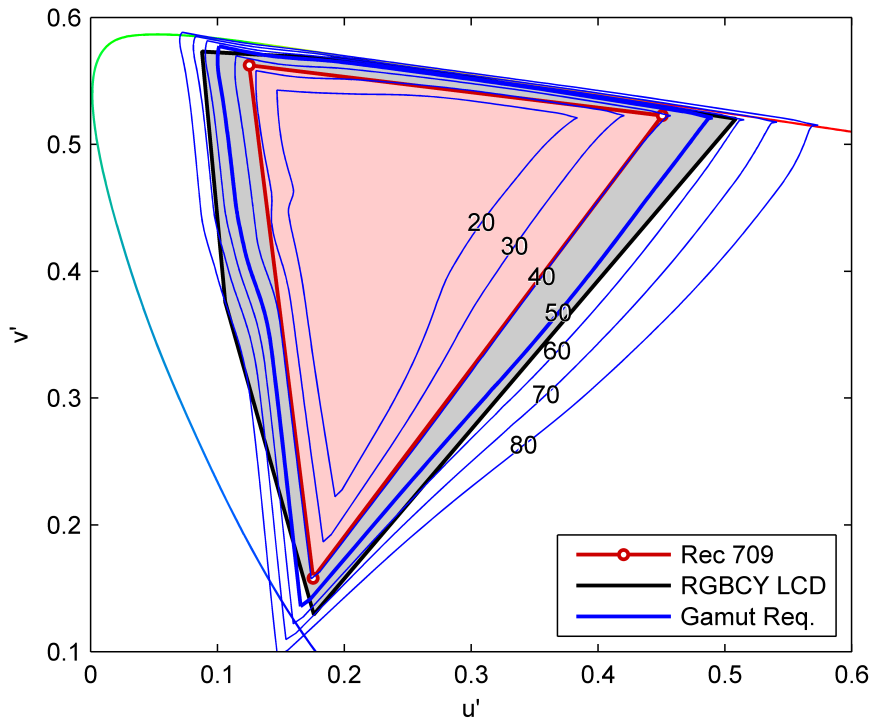


Figure 6.12: Gamut boundary requirements. Resulting from the present experiment, the blue curves show expanded (and contracted) versions of the standard Rec. 709 chromaticity gamut using the mean-based hue-dependent chroma multiplier curves corresponding to the percentiles labeled derived from the present experiment. The underlying red triangle indicates the standard Rec. 709 gamut.

that the present experiment with the MPD was done with a luminance limit on the displays' white point, which is different from the typical inherent behavior of LCD-based multi-primary displays. Thus, care must be taken to balance the display's primaries (or indeed to limit the white) so that they do not appear darker than anticipated, or else the results may not apply.

6.4.4 Looking Forward

The present experiment provides a strong result to guide the choice of chromaticity gamut area for displays. An implementation of hue-dependent chroma enhancement may be made as well, but it would be wise to include some improvements from elsewhere in the literature. For example, the simple linear scaling used in the present work may be improved with other variants – nonlinear, soft-clipping, skintone-preserving, lightness-adapting, and other creative algorithms – which may be used while still taking advantage of hue dependence. Another question alluded to earlier has to do with how best to use multi-primary displays, in which primary luminance is inherently lower than in RGB displays, for example weighing quality and efficiency.

An improvement could be made by accounting for the content dependence of chroma preference. The present experiment provides some clues: the preference for less chroma boost for luminous objects and rainbow colors could be either because of cognitive expectations or because of their higher starting chroma. Preference for higher chroma boost for images including a white background could be simply due to compensation for visual effects that might make an object on a white background appear less colorful, such as glare, where light from the bright background is scattered within the eye to contaminate the object's color, or simultaneous contrast, whereby a bright background makes an object appear darker and less colorful. The influence of this kind of contextual information is difficult to model, and for now it is effectively “averaged out” in the results, which is satisfactory. However, image understanding is becoming more powerful, and as automated machine learning-based image quality prediction comes of age, this problem may get easier.

6.5 Conclusions

Based on the present work it is clear that, starting with standard-gamut (Rec. 709) images, viewers prefer a chroma boost from the original color reproduction. For the median observer, this boost is roughly 20%. The overall boost preference is distorted according to hue: preferred boost for orange, yellow, green, and cyan colors is greater than that for blue, magenta, and red colors. This hue-dependence was determined based on single-hue images, but the present experiment additionally confirmed that it was significantly preferred in general for multi-hue, natural images.

Once the preferred hue-dependence of chroma boost is applied, the overall level of preferred chroma boost depends on both image content and observer variation. The

content dependence is not readily modeled; however, a content-independent chroma boost created by aggregating preference over many images performs well, on average. Individual observers' preference for chroma boost may be handled by a single overall colorfulness parameter, because the hue-dependence is relatively consistent over observers. Further, future display gamuts may be designed with respect to the variation in preference over observers, for example maximizing the proportion of the population whose preference is attainable within the gamut boundaries.

6.6 Acknowledgements

Kind thanks to Karel Hinnen for his help with the MPD image processing. Thanks as well to our observers who contributed their time and expressed their preferences.

6.7 Bibliography

- [1] ITU-R. BT.709 : Parameter Values for the HDTV Standards for Production and International Programme Exchange, April 2002.
- [2] E.A. Federovskaya, H. de Ridder, and F.J.J. Blommaert. Chroma Variations and Perceived Quality of Color Images of Natural Scenes. In *Proc. SPIE*, volume 2411, pages 51–61, 1995.
- [3] Huib de Ridder. Naturalness and Image Quality: Saturation and Lightness Variation in Color Images of Natural Scenes. *Journal of Imaging Science & Technology*, 40(6): 487–493, 1996.
- [4] Masato Sakurai, Rodney L. Heckaman, Stacey E. Casella, Mark D. Fairchild, Takehiro Nakatsue, and Yoshihide Shimpuku. Effects of Display Properties on Perceived Color-Gamut Volume and Preference. *Journal of the Society for Information Display*, 16(12):1203–1211, 2008.
- [5] Y. Hisatake, A. Ikeda, H. Ito, M. Obi, Y. Kawata, and A. Murayama. The Ergonomics Requirement for Reproducible Area of Color Chromaticity in Electronic Displays. In *Proc. IDW*, pages 2301–2304, 2007.
- [6] J. Laird and I. Heynderickx. Perceptually Optimal Boundaries for Wide Gamut TVs. *Proceedings of the SPIE - The International Society for Optical Engineering*, pages 1–10, 2008.
- [7] Michael J. Murdoch, Dragan Sekulovski, Robert de Volder, and Ingrid Heynderickx. Preferred and Maximally-Acceptable Color Gamut for Reproducing Natural Image Content. *Journal of the Society for Information Display*, 18(12):1111–1118, 2010.

- [8] Justin Laird, Muijs Remco, and Kuang Jiangtao. Development and Evaluation of Gamut Extension Algorithms. *Color Research & Application*, 34(6):443–451, 2009. 10.1002/col.20537.
- [9] Hung-Shing Chen, Te-Mei Wang, Shih-Han Chen, and Jin-Sin Liu. Skin-Color Correction Method Based on Hue Template Mapping for Wide Color Gamut Liquid Crystal Display Devices. *Color Research & Application*, 36(5):335–348, 2011.
- [10] Karel Hinnen, Michiel Klompenhouwer, Yingrong Xie, Ruben Rajagopalan, and Erno Langendijk. Multi-Primary Displays From a Systems Perspective. In *EuroDisplay*, 2009.
- [11] Michael E. Miller and Michael J. Murdoch. RGB-to-RGBW Conversion with Current Limiting for OLED Displays. *Journal of the Society for Information Display*, 17(3): 195–202, 2009.
- [12] Rodney L. Heckaman and Mark D. Fairchild. Expanding Display Color Gamut Beyond the Spectrum Locus. *Color Research & Application*, 31(6):475–482, 2006.
- [13] Dragan Sekulovski, Robert de Volder, and Ingrid Heynderickx. Preferred and Acceptable Color Gamut for Reproducing Natural Image Content. In *SID Symposium*, pages 1014–1017, 2009.
- [14] Ethan D. Montag. Empirical Formula for Creating Error Bars for the Method of Paired Comparison. *Journal of Electronic Imaging*, 15(1), 2006.

Discussion

This thesis has presented a practical approach to choosing technology based on visual perception, aiming to be both efficient and effective. Four steps were laid out in Section 1.3.3 using the language of the Image Quality Circle (see Figure 1.11) to optimize technology variables with respect to quality. These steps include determining the perceptual attribute (which contributes to overall image quality) that is most important to the technology, modeling the impact on the physical image that the technology makes, measuring and/or modeling observers' perceptual response to physical images corresponding to the technology, and communicating the overall relationship to guide technology selection. In this chapter, a retrospective look at the examples presented explains their fit into this framework clearly and emphasizes the relationships which were uncovered between technology and perception. Some comments toward future work are provided.

7.1 High Dynamic Range LCDs

A basic question regarding the development of high dynamic range liquid-crystal displays is *how much dynamic range is needed?* In general, higher dynamic range comes at greater expense, so a pragmatic answer to this question comes from an understanding of how much dynamic range is actually visible to the human observer. A detailed answer was provided in Chapter 2.

HDR LCDs are enabled by segmented-backlight technology, which offers not a single technology variable to optimize, but a collection of variables including the size, spacing, and spatial luminance profile of backlight segments. For the purpose of this research, the effects of all of these variables were distilled to the practical, physical image characteristic of localized image black level luminance. Never infinitely dark even when a given backlight segment is dimmed, the black level in a display is raised incrementally by light leakage through the LC itself and by the light from neighboring backlight segments, and then incrementally further within the observer's eyes thanks to glare from the image content and surrounding environment.

Visually, the perceptual attributes related to the physically-lifted black level luminance are both a worsening of perceived blackness (i.e., lighter than black) and a lowering

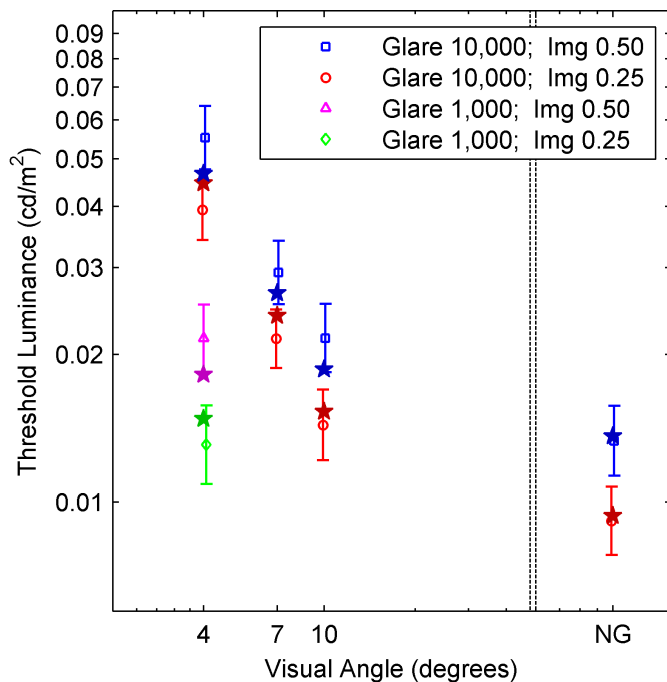


Figure 7.1: Black level visibility results. Black level luminance thresholds measured in the experiment described in Chapter 2, shown as a function of visual angle from the glare source (the no-glare condition is labeled on the x-axis as “NG”). Fitted model points are shown as stars near the corresponding experimental points. The thresholds indicate the black level luminances which are indistinguishable from much-darker blacks for each combination of glare luminance, glare distance (visual angle), and average image luminance. Luminance values listed in the legend are in cd/m^2 .

in perceived contrast. Of these two, blackness was assumed to be more intuitively described, so it was chosen as the perceptual attribute to be evaluated in the experiment. The experiment uncovered the visibility threshold of black level luminance differences over different images and physical glare intensities and proximities. The experimental results depend not only on black level, but importantly on intra-ocular glare, so a model was assembled that describes the visibility of black level differences as a function of the distance and luminance of a glare source.

The measured and modeled final relationship is repeated here as Figure 7.1 (same as Figure 2.9). The thresholds indicate the black level luminances which are indistinguishable from much-darker blacks for each combination of glare source luminance, glare source distance (visual angle), and average image luminance. Tying this back to the original question about HDR LCDs, it is apparent that sufficient dynamic range is any that keeps the black level luminance below these thresholds under all or the most likely combinations of image luminance, glare luminance, and glare distance. In fact, this visibility threshold-based relationship may be conservative. It is possible that an acceptability-based relationship, which presumably would have thresholds above those for visibility, would allow for more relaxed display specifications.

As a practical example, consider a full-HD 46" TV viewed in a dark room from 2m. According to the modeled results, if at the center of the screen there is a 100 pixel square block of full-white, 500 cd/m², then at 14 cm (4 degrees of visual angle) from the center the black level visibility threshold is about 0.18 cd/m². Because a typical LCD can only attenuate 3 log units of luminance, meaning 1000:1 contrast, a fixed backlight would produce a black level of 0.5 cd/m², which is visibly poor. To produce a sufficiently-dark black, a segmented backlight would need a spatial luminance profile which dropped to about 0.36 times its maximum over those 14cm. This sort of calculation can easily guide segment design in an HDR LCD.

7.2 RGBW OLED Displays

OLED displays employing a white-emitting material and color filter arrays can benefit from using a fourth, white primary in addition to the typical red, green, and blue. The benefit is not visual at all, but rather to reduce power consumption. However, the implementation can affect the visual image, generally in an adverse way, depending on how this technology is employed – technology variables can include how much white to use, how much RGB remains, and any color pre-processing parameters that affect these. Myriad options present themselves, some adding image luminance, some reducing color saturation, others having little visual effect. Options were assessed in two steps, effectively asking separate questions.

In Chapter 3, the essential question being asked was *how can a W primary be incorporated without degrading the visual image?* Because the goal is to not affect the image appearance at all, we must preserve all relevant perceptual attributes, importantly brightness, colorfulness, and hue. Using the solid foundation of colorimetry, it is possible to compute a metameric match, in which the colorimetry of the pixels of the RGBW image matches those of the original RGB image. The extension from colorimetric match to appearance match is a reasonable one when viewing conditions are identical. Colorimetric models of the additive RGB and RGBW systems were used to develop an algorithm that adds W to the image without visual effect. Conceptually, the subtraction of common luminance from the color channels and transferring to a neutral channel is similar to

algorithms employed in subtractive CMYK printing; however, an important generalization was made in this RGBW work to allow for colorimetric precision with an arbitrarily-colored (i.e., not pure white) W primary.

Because the underlying colorimetry is well-understood and because OLED displays may be straightforwardly additively modeled, a perceptual user study was not conducted – the only example in this thesis in which it was not necessary – but indeed a series of visual verifications were made using simulated images on CRT displays and digitally-written transparency films. Prototype OLED displays confirmed that the W-RGBW system with the proposed algorithm yielded no adverse impact on image quality relative to a reference RGB OLED [1]. The degradation-free method that was developed is still used in W-RGBW OLED displays currently in the market.

In Chapter 4 the RGBW implementation was pushed toward more power savings (as well as electronic design freedom) by using less of the red, green, and blue primaries relative to white, and a second question was asked, *how degraded does the image become if the use of W is increased further?* The technology variables of interest build on the list from the preceding chapter, adding two important image processing algorithm parameters: the magnitude of RGB intensity limitation and amount of color desaturation.

These parameters result in visual image distortion that may be described as a loss of colorfulness and color contrast, but identifying a single important perceptual attribute was difficult. However, because the goal was to understand potential image degradation at a higher level, the experiment focused on overall image quality. This means that while the immediate question of quality optimization was answered without a problem, we have not built an easily-reusable basic understanding of how the different perceptual attributes contribute to the measured quality degradation.

The experimental results shown in Chapter 4 are repeated here as Figure 7.2 (same as Figure 4.4). This diagram, which shows the image quality degradation from the original image (at top right) as a contour plot over the two parameters of desaturation and RGB limit, makes it clear that there is a substantial plateau of high quality surrounded by declines in quality as either parameter is reduced. The pictured measured relationship between these physical image characteristics and overall quality informs the design decision on the quality half of the quality-power tradeoff.

7.3 Wide Gamut Displays

The potential for color gamut enlargement via a variety of technologies is almost unlimited; however, wider gamut generally comes with increased cost and power consumption. Thus, an obvious question which was addressed in Chapter 5 and Chapter 6 is *how much color gamut is necessary for pleasing reproduction of natural images?* The technology variables of interest are the color primaries chosen through the combination of light source (i.e., backlight) and color filter array materials. These define the size and shape of the three-

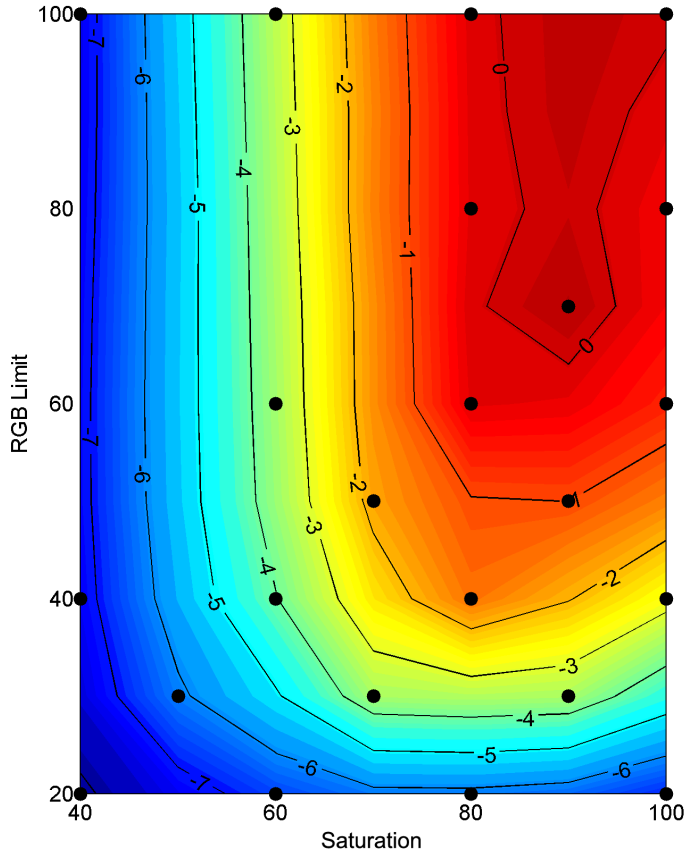


Figure 7.2: RGB limitation image quality results. Contour plot of image quality, represented as JNDs of image-quality loss relative to the original, due to parameters of desaturation (percent) and RGB intensity limit (percent). A large plateau of image quality nearly matching the original image at position (100, 100) is seen, with eventual falloff as either parameter is reduced.

dimensional color gamut, which are physical image characteristics contributing to the perceptual attributes of colorfulness and hue.

Within the gamut boundaries provided by the display, some enhancement of chroma or colorfulness of the standard-gamut input image is required. Early work on this enhancement showed that preference for colorfulness is hue-dependent, which pointed toward studying the physical image characteristic of gamut size per hue. This is interesting regardless of whether the gamut is implemented with a RGB or multiprimary architecture. Especially with multiprimary displays, the luminance dimension of the color gamut is very important, because the relative luminance of colors to white affects perceived colorfulness strongly. In the experiments conducted, this known effect was controlled and avoided, which means caution should be given to applying these results to a multiprimary system without similar control of luminance.

The primary perceptual effect of gamut is on colorfulness of the image, but it is well known that observer preference (which is assumed to be driven by overall image quality) is not monotonic with colorfulness – preference increases with colorfulness to a point, then levels off and declines. Because of this lack of monotonicity and the general interest in maximizing overall image quality, preference and overall quality were studied in the set of experiments described.

The two chapters on this topic provide a first look and then a more refined answer to the basic question of gamut requirement. Utilizing first a wide-gamut RGB and second a multiprimary RGBCY display, the preferred image colorfulness as a function of hue was probed with highly-saturated natural image content by a large number of observers. A clear hue-dependent result emerged, within which the main inter-observer preference variation is in overall magnitude. The powerful result of this is a relationship showing the added value of enlarging color gamut into any given portion of color space, presented in terms of observer percentiles, illustrated in Figure 7.3 (same as Figure 6.12). Such a relationship can guide the choice of number and chromaticity coordinates of display primaries with a clear understanding of how satisfied a population will be with the result. Such preferential predictions can be weighed against other, non-imaging constraints such as cost and power consumption.

7.4 Comments, Connections, & Extrapolations

The display system and advanced technologies described herein were not chosen with a global view on what is most important or most academically rigorous. To the contrary, they were real examples from industrial research, tackled as they came, with all the practical and industrial constraints that that implies. That being said, if the most important perceptual attributes of displayed images include brightness, contrast, colorfulness, sharpness, temporal stability, and motion smoothness, it is no surprise that industrial research (whose goal is feature-based market leadership) led to research on three of these topics. Overall image quality is generally dominated by the worst contributor in

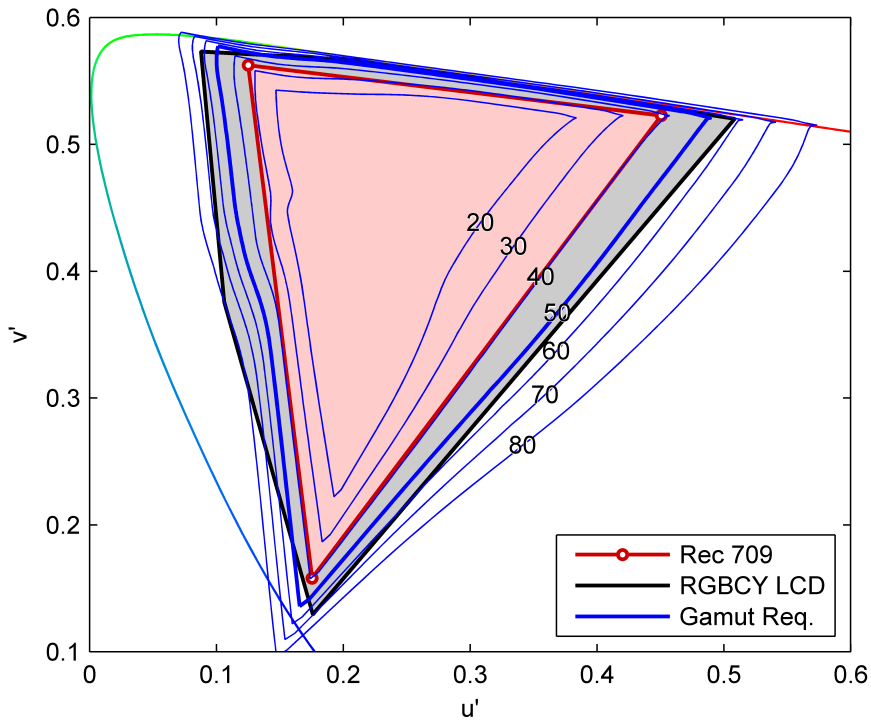


Figure 7.3: Hue-dependent gamut preference results. Gamut boundary requirements shown as a contour plot of the percent of observers satisfied. The underlying red triangle indicates the standard Rec. 709 gamut, and the blue curves show expanded (and contracted) versions of the standard using the mean-based hue-dependent chroma multiplier curves corresponding to the observer percentiles labeled.

the multidimensional sum of perceptual attributes, which has the practical result that industrial research tends to address the “currently-worst” characteristic of a system in its series of improvements.

This is certainly an applied, pragmatic thesis, which is why there is strong emphasis on efficiently uncovering the relationships between technology and perception and effectively communicating those relationships to inform design and development choices. The realistic pragmatism gives further validation to the approach of identifying key attributes and optimizing technological advances: it has been used successfully in real-life application and not only as an exercise. In the next sections, some retrospective connections and future directions are revealed.

7.4.1 Perception & Preference

In perceptual research in general it is useful to keep the difference between perception and preference clear. As described in this thesis, perception relates to detection thresholds and similar measures, such as the black level differences studied in Chapter 2. Thresholds at the level of detectability have more to do with fidelity than expectations. Preference is by definition a suprathreshold task, meaning that a difference must be perceptible in order to get a meaningful response as to whether it is preferred. The studies described for RGBW OLED image degradation in Chapter 4 and wide-gamut color in Chapter 6 all have to do with preference. They support a holistic and human-centered definition of image quality as a concept tempered by observer expectations. This distinction also implies that perception results are more robust over population differences, because they are more physiological in nature, while preference results may vary over populations due to things like cultural biases, or vary over time as improvements raise expectations. For this reason, use caution in applying the OLED degradation preference results obtained in North America, or the gamut preference results obtained with a primarily northern European population, to a global display market.

Logic says in some cases perception is the appropriate thing to measure; in others, preference is more important. In certain studies it may not matter much which is studied – for example in the perception of black level differences, if indeed people generally prefer darker blacks and higher dynamic range, then a similar study done with preference-based pairs may have resulted in similar trends though perhaps with higher thresholds. However, using a perception task in place of preference doesn't work in most cases. For example, it would not make sense to ask observers to choose the more colorful image in the gamut experiments rather than the preferred level of colorfulness, because they would presumably just find the physical boundary of the proxy display used in the experiment.

7.4.2 Optimizing Uni- versus Multi-Variate Quality

An important assumption made in this work is that new technologies and their perceptual implications may be studied separately, sequentially, rather than together in a multivariate

way. This is of course debatable. For example, it is unlikely that simply combining wide-gamut and RGB-limiting using separately-derived parameters would result in an optimal implementation. However, this situation may be unlikely to arise, practically speaking. Each of the advanced display technologies studied in this thesis is a separate advancement, and history shows that such advances happen in series with the baseline evolving along with progress. Fortunately, the tools used in image quality research are effective even in the absence of a full model of overall image quality as a function of relevant contributing attributes.

Engeldrum's Image Quality Circle (see Figure 1.11) provides a useful guide regardless of whether the entire circle is ever completely described. Even after decades of research, a generalizable model of image quality has not emerged; a recent review article by Chandler [2] explores some of the reasons why, including the difficulty of modeling the human interpretation of natural images, the effects of multiple simultaneous distortions, and the effects of geometric or image-enhancement changes. Yet, despite the incomplete model, the structure of the IQC is valuable. Simply being conscious of whether a new technology affects primarily a single perceptual attribute or multiple perceptual attributes such that they cannot be easily disentangled can guide experimentation and model-building in a positive way. In a case such as the RGB-limiting example, studying overall quality directly is more efficient than studying the mix of attributes such as colorfulness, hue, and lightness, but it does not build as much basic understanding. In that example at the time of the research, the high-level understanding gained was sufficient because it confirmed that a moderate amount of RGB-limiting was tolerable, a fact that opened up design freedom in the display electronics. If the situation had been different and repeated re-use of the results had been foreseen, then the benefits of a larger-scale image quality model like those discussed in Section 1.3.2 would be valuable. And, the extra work that building them up entails would be worthwhile.

That being said, one real obstacle to creating a multivariate quality model for advanced display technologies is that in order to evaluate the combined impact of a multitude of new technologies, a "super-prototype" display must exist with capabilities at least as great – higher brightness, better contrast, larger gamut, etc. – as expected in real implementations. The two chapters on wide color gamut build knowledge in two steps, relying on prototypes with wide- and then even wider-gamut capabilities, nevermind one that can simultaneously provide ultra-high contrast. Such super-prototypes are few and far between, which also explains why, for example, Keelan's approach to imaging system quality focused on consumer-grade cameras and utilized an impairment approach to image quality. The reference system was already in existence, and the cheaper versions in "real life" could be assessed against it. In contrast, the coming advanced display technologies can (and should) make displays ever better, with the practical reality that a new prototype may only be able to showcase the effect of one or a few of these new capabilities. This is the reality that was explored in this thesis: a series of qualitative studies on non-independent attributes, approached with the hope that the serially-optimized system will be nearly as good as the hypothetical, globally-optimized version.

Another potential problem with a multivariate image quality model is one of unknown unknowns: the completeness of a model can never be verified, and a newly-discovered imaging artifact can render a model obsolete. For example, imagine an image quality model which accurately predicts overall quality based on brightness, colorfulness, hue accuracy, contrast, noise, and sharpness, all of which are interesting perceptual attributes. An image flaw that doesn't directly affect one of those attributes will cause unknown problems: for example JPEG artifacts may affect noise, sharpness, and possibly other attributes, but if such artifacts were not included in the fitting of the model, it is very unlikely that the model will accurately predict the effect of JPEG artifacts on overall quality. The impossibility of incorporating all possible physical image characteristics makes a full image quality model a fragile concept.

It is interesting to note that the thesis approach lies somewhere between highly-structured image quality modeling approaches and a pragmatic full-system approach that is typically known as a "beauty contest:" simply lining up competing displays and voting on a winner based on overall image quality (a good example is provided by Heynderickx & Langendijk [3]). Such beauty contests are not mentioned to point at their weaknesses; in fact they are the closest to the real-life experience of a customer choosing a new display in a showroom. And, as described in that paper, much insight may be gained about the perceptual attributes affecting observers' choices by asking them not only which display they prefer, but why. However, what the beauty contest can never uncover are the effects of single technology variables as discussed in this thesis. Each in this range of approaches has its merits and should be used when appropriate.

7.4.3 Extrapolating these Results & Methods

Extrapolating image quality results is always a difficult task. Image quality models always seem to be constrained to similar systems; our understanding of the human visual system is never sufficient; effects of image content, saliency, and emotion cannot be excluded or easily predicted; and something new always seems to appear. For these reasons, some big questions remain difficult or impossible to answer. For example, it would be great to know *what would a truly optimal display look like?* Perhaps the impossibility of an answer is forgivable without a very wise (and very high image-quality) crystal ball.

But, more practically (and admittedly somewhat disappointingly), a question like, *is it more beneficial to provide the viewer with additional color gamut or additional dynamic range?* remains difficult. Here, if anywhere, the limitations of serial univariate study of new technologies manifest themselves. As long as each perceptual attribute is studied separately, it remains impossible to weigh them against one another. Keelan's approach solves this by comparing every attribute directly to a common, understandable image quality attribute, sharpness. This anchoring technique could certainly be applied, even without building a multivariate final model, if the anticipated value of answering such cross-attribute questions were high. In this work it was not, thanks to time constraints and disconnected industrial projects, but in future work it could be valuable.

While the different technology variables studied herein relate to non-independent perceptual attributes, combinations may still be discussed. For example, it is reasonable to assume that black-level visibility limits discussed in Chapter 2 would remain the same with or without the kinds of color enhancement and distortion processing presented in later chapters. While the perceived benefit of the combination is not specifically studied, we remain confident that the optimum for each variable is stable regardless of the presence of the other. A perceptual interaction would be expected, however, between color enhancement and overall brightness, because luminance is known to have a strong effect on perceived colorfulness. The different color processing approaches presented would be expected to interact: at first glance, the hue-dependent color gamut expansion shown in Chapter 6 and the color distortion and desaturation processing in Chapter 4 appear to counteract one another; but, a nuanced combination may be possible – perhaps the RGB color limit would be preferentially applied in a hue-dependent way – an idea that is neither studied nor adequately predicted by the present work, but which seems a logical step.

Extrapolating the presented results to different display modalities and even lighting applications is eminently doable. As a simple example, the visibility model based on contrast perception explained in Chapter 2, which was built up in reference to HDR displays, could be used to predict whether differences between the lowest-reflectance bit values of a reflective electrophoretic display would be visible under specific lighting conditions when considering screen-surface specular reflections. A simple extrapolation of this is to the lighting requirements in a projection theater. The visibility of image degradation due to unwanted spill light from a lighted “EXIT” sign, given theater specifications and projected image content, could be computed with the visibility model described. As another example, the quality implications of primary luminance, studied in Chapter 4 with W-RGBW OLED displays, could be applied to multiprimary (often RGBW or RGBCY) projection displays, in which the importance of efficiency, generally in terms of screen brightness, must be balanced with image quality. Beyond image quality, the perceptual insight derived from this work can be applied to other visual experiences; for example, Vogels proposed using an analogue of the Image Quality Circle to describe the experience of atmospheric lighting [4]. The structural way of thinking about physical characteristics, perceptual attributes, and overall visual experience is analogously useful. Even without further experiments, the results described in this thesis inform requirements for LED lighting, for example with respect to intra-ocular glare, the design and operation of RGBW LED luminaires, as well as the preferential enhancement of object colorfulness with appropriate spectral content in the light.

The approach promoted by this thesis is a way to make the best of technology trends as they arise, finding a balance between system-level and attribute-specific optimization as a result of practical considerations and efficiency. Even better than following technology trends, perhaps, would be to lead the technology advancements directly based on perceptual understanding. This does not mean obvious, empty direction like, *let's make a holographic display whose visual experience is imperceptibly different from reality*. Rather, it

might include an insight such as, *based on the HVS' ability to discern smaller variations in green and yellow than in other hues, a display design would be more visually efficient if it distributed its bit-depth and addressability accordingly.*

Next steps beyond this thesis could take several interesting paths. One, of course, is to continue to address the sequence of technology advancements in the display industry. Many additional chapters could be added to a book like this about new display technologies such as active or passive 3D stereo, lenticular autostereoscopic structures, active and adaptive image processing algorithms, motion portrayal, ultra-high resolution for TV, monitor, and mobile use, and yet-unknown advancements that are sure to come. Another is to apply these techniques to a different, visually-rich industry, such as lighting. This has already begun, of course, with an analog of the Image Quality Circle theorized to describe quality of light. This is an interesting development because both the methodology and the perceptual insights on visibility and preference may be transferred to this adjacent field and extended. Yet another path would be to formalize the study of an approach to perception in industrial research, addressing the efficiency in terms of benefit per effort invested and the effectiveness in terms of informed decisions or product and market impact. Whether creating technological innovations with perceptual insights, using perception to steer innovations, or studying the role of perception research in innovation itself, this thesis provides a departure point based on practical examples in the display industry.

7.5 Concluding Remarks

Through this thesis, a practical approach to technology choice guided by visual perception insights for display design has been presented. Examples of the most important advanced display technology applications were provided, with key perceptual attributes identified, measured, and modeled relative to the technologies in question. In some cases, this means pushing a new technology only as far as is necessary for visibility or preference, and not wastefully further; in other cases, this means trading off image quality for a different benefit, understanding exactly the perceptual or preferential impact of such a decision. In each case, the goal of making optimal human-centered use of these new technologies was achieved.

HDR LCDs may be constructed with knowledge of the visibility limits of black level differences, which means this advanced technology may be implemented with maximal visual impact at minimal cost and complexity. RGBW OLED displays may be designed that take full advantage of the highly efficient white subpixel with no image quality penalty, thanks to thorough modeling of color perception; or, with a clear understanding of the image quality loss that is caused by a cost- and power-saving technology choice to limit the current to the display's subpixels. Wide-gamut displays, RGB or multiprimary, may be crafted cleverly, with the added value of the additional color gamut area directly computable based on an understanding of observer preference for hue-dependent color gamut limits.

These examples are real, practical examples from industrial research, and the relationships between technology and perception that were uncovered and modeled in every case were efficiently found and effectively expressed. The efficiency of this approach means not expending unneeded effort to construct a complete model of system image quality as a function of all system components because of the time involved and the potential for obsolescence. Rather, full focus is given to each technology variable in order to understand clearly its impact on a most-relevant image quality attribute, revealing relationships through studies that are relatively narrow but which still take advantage of the structure of the Image Quality Circle. The effectiveness of the approach is shown by the use of these relationships while weighing the other factors such as cost and lifetime that together determine how and which technologies make their way into new displays. Performing industrial research in this way means that quantified perceptual insights can be used to positively impact decisions in the technology and product development cycle.

With this approach, each of the still-unforeseen ideas that will come as display technology continues its unyielding progress can be evaluated efficiently on its image quality merits, ensuring a properly human-centered implementation that balances technology and perception.

7.6 Bibliography

- [1] A. D. Arnold, P. E. Castro, T. K. Hatwar, M. V. Hettel, P. J. Kane, J. E. Ludwicki, M. E. Miller, M. J. Murdoch, J. P. Spindler, S. A. Van Slyke, K. Mameno, R. Nishikawa, T. Omura, and S. Matsumoto. Full-Color AMOLED with RGBW Pixel Pattern. *Journal of the Society for Information Display*, 13(6):525–535, 2005.
- [2] Damon M. Chandler. Seven Challenges in Image Quality Assessment: Past, Present, and Future Research. *ISRN Signal Processing*, 2013(905685):1–53, 2013.
- [3] I. Heynderickx and E. H. A. Langendijk. 47.1: Invited Paper: Image Quality Comparison of PDP, LCD, CRT and LCoS Projection. In *SID Symposium*, pages 1502–1505. SID, 2005.
- [4] Ingrid M.L.C. Vogels. How to Make Life More Colorful: from Image Quality to Atmosphere Experience. In *17th Color Imaging Conference*, pages 123–128, 2009.

Summary

Human-Centered Display Design: Balancing Technology & Perception

Displays for entertainment, such as televisions and cinema screens, and displays for information, such as control panels and digital signage, have become ubiquitous in the human environment. Both the range of devices with displays and the number of applications for which they are used continue to increase, so it is important to take a close look at the human-centered aspects of display design. Humans perceive displays with their visual sense, the human visual system (HVS), which is both wonderfully effective and complex. Designing displays is best done at the meeting point between the details of advanced display technology and an understanding of the performance and limits of human visual perception. The goal of this thesis is to emphasize the intersection of these themes and to present a general approach with concrete examples of the implementation of advanced display technologies in a human-centered way.

The first chapter lays a conceptual foundation of visual perception, display technology, and how they meet. We perceive light with our eyes, seeing patterns of brighter and darker, the difference between which is contrast. Adaptation allows the HVS to function over a huge luminance range, but intra-ocular glare limits perceivable contrast. We see color thanks to the three differently sensitive types of cone cells in the retina of the eye. Because of this, color displays can present a wide range of colors to the HVS with as few as three different color primaries: this is a realization of metamerism, meaning physically different light sources can provide the same color stimulus to the observer. In images, we perceive attributes such as brightness, contrast, sharpness, and colorfulness. These perceptual attributes contribute to an overall impression of image quality, which is also tempered by the expectations of the observer. Overall quality, as well as its component perceptual attributes, can be quantified through experiments presenting images to human observers using psychometric methods. This in turn allows the quantification of the perceptual effect of the underlying display technology choices.

The approach espoused by the thesis to implement new display technologies in a human-centered way is to focus on the perceptual attribute which is most important to the technology in question. In the framework of an image quality circle, two relationships are paramount: that between the technology and the physical image, which explains how that specific technology affects what is physically and measurably on the display screen; and the relationship between the physical image characteristics and the relevant perceptual attribute, which should remain general and independent of specific technologies. For a series of examples, such relationships are uncovered and analyzed, and they can subsequently be used to guide implementation and to weigh perceptual quality with other

non-visual characteristics such as cost. The advanced display technologies discussed are designed to improve contrast, efficiency, and color reproduction over a baseline display system, the television-sized LCD technology prevalent around 2007.

In Chapter 2, the requirements for increased display contrast are studied. Replacing an LCD's fixed backlight with addressable backlight segments enables the creation of a high dynamic range display capable of a much wider contrast range than was possible previously. But, the design of a segmented backlight may only be optimized with an understanding of the spatial limits of human contrast sensitivity considering the glare induced by nearby bright regions which are inevitable in displayed content and in normal viewing environments. A psychometric experiment was conducted to find the visibility limits of black-level differences in proximity to a bright glare source, varying the intensity and distance of the glare. These results were successfully modeled using the CIE model of intra-ocular glare, the DICOM contrast-visibility model, and a new estimate of adaptation luminance which depends on the surrounding visual field. The result is a model which can be used to assess the design of backlight segment size and spatial luminance profile so that perceived contrast is maximized.

The implementation of OLED display technology, which offers an alternative to LCD with lower system complexity and potentially better color and contrast characteristics, is addressed in Chapters 3 and 4. For manufacturability reasons, OLED displays are better implemented with a white-emitting OLED material and color filters, but this implies a low efficiency due to light absorption by the filters. Leaving some subpixels unfiltered results in a more efficient RGBW system, but care must be taken with such an architecture to avoid color degradation. Based on colorimetry, which is a robust model of color matching in the HVS, a basic method is presented for making the best use of the efficiency of RGBW without disturbing color reproduction at all. Further, an advanced method is explained which gives even more efficiency and design freedom at the expense of limiting the maximum displayed RGB luminance relative to white. Because this method results in color degradation, a perceptual experiment was done in which the image quality loss due to a combination of limited maximum RGB and intentional color desaturation was quantified. The results enable the implementation of RGBW OLED over a wide range of color accuracy and power efficiency and make the image quality implications of that range clear. Currently the first OLED televisions are reaching the consumer market, and they utilize this RGBW OLED architecture.

The implementation of wide color gamut is explored in Chapters 5 and 6. As a first step, experiments were conducted using a wide-gamut RGB display to determine the preferred level of color saturation of natural images presented at a variety of single hues. A strong hue dependence for maximal and preferred color saturation was found which was notably different from the per-hue capability of the RGB display. This led to a hue-dependent chroma boost algorithm which was then verified using multi-hue natural images. Because the capabilities of the wide-gamut RGB display were not sufficient, especially due to limitations in yellow hues, a second study was undertaken using a

prototype 5-primary (RGBCY) display which offered additional color gamut where needed. This experiment verified the previously-measured hue dependence and confirmed the expected need for increased yellow gamut. Further, observers significantly preferred images transformed with the hue-dependent chroma boost algorithm over those with hue-independent chroma boost and those presented colorimetrically. Throughout the gamut experiments, the primary inter-observer variation relates to an overall chroma preference, meaning that some people prefer higher color saturation than others, but the hue dependence of that preference remains similar for all observers. Thus, based on these results, a display may be designed with color gamut boundaries wide enough to satisfy the observer population and a single control parameter given to the end user for overall chroma.

Through these examples, a practical approach to technology selection based on quantifying the effects on visual perception has been illustrated. In each example, critical perceptual attributes were measured and modeled with respect to the technologies in question. The results are generally relational, showing the perceptual effect over a range of a technology variable or parameter. Such relationships clarify the impact of technology choices, some of which are made to improve image quality, others of which incur quality degradations while gaining other benefits such as power savings. Using this practical human-centered approach, future display advancements may be implemented by successfully balancing technology and perception.

Acknowledgements

The path to this dissertation has been long but fulfilling. Doing both MS and PhD degrees while working full time in industry requires a commitment to giving up weekends and evenings – sometimes in intense bursts, often with lulls in between, and definitely with an exhausting push to the finish line. First, I thank my wonderful wife Meredith both for tolerating all the time spent writing and importantly for pushing and encouraging me the whole way. Second, for tacitly letting me spend more than a few weekend “PhD days” in the attic instead of “Daddy days” building Lego towers and going for bike rides with her, thank you to my daughter Dina.

How can I say enough about my committee? For all of her help editing and co-authoring, her confident encouragement, her (sometimes brutal) honesty, her experience with experimental design and analysis, and her practical wisdom about everything from Dutch bureaucracy to Belgian beer, I give a big thanks to my *eerste promotor*, Ingrid Heynderickx. Your commitment to mentorship of your many students and colleagues in the Netherlands and China is truly amazing. Wijnand, thank you for your thought-provoking words and encouragement and your always-smiling demeanor, as well as previous collaborations on 3D visual comfort. Huib, I like your “big-picture” perspective and push for a more unified approach (and you would also get the award for finding the most typos in my draft!). Gerard, I enjoyed our discussions of “topical adjacencies,” relating my work to image encodings and practical implementations. Jean-Bernard, thank you for the dialog on analyses and statistical methods and for making me think further about effectiveness in research in general. And Mark, I thank you for your involvement all the way from RIT, your detailed comments and suggestions, and also the many discussions over the years of color and vision topics. I appreciate the contributions of each of you!

I have been lucky throughout my industrial work to have the opportunity for varied collaboration with experts, and I have many people to thank for their insight, inspiration, and influence. From my days at Kodak, these include John Lacey, John Nelson, Tom Madden, Paula Alessi, Geoff Woolfe, Michael Miller, Paul Kane, Ron Cok, Andy Arnold, David Long, Jim Sutton, and Kim Montz. From Philips, special thanks to Ad Leenaars, Dragan Sekulovski, Pieter Seuntiëns, Marc Lambooij, Robert de Volder, Mariska Stokkermans, Patrick Vandewalle, Karel Hinnen, Stefan Swinkels, Erno Langendijk, and Kees Teunissen. Also thanks to some of the many people not already mentioned with whom I’ve worked on other topics, discussed odd ideas, and had a bit of fun over the last few years, including Bart Salters, Marco van Boven, Fetze Pijlman, Ulrich Engelke, Ingrid Vogels, Gosia Perz, Chris Damkat, Siebe de Zwart, Ruslan Sepkhanov, Tim Dekker, Tom Bergman, Bart Kroon, Adrie de Vries, Wiebe Wagemans, and others. I look forward to further collaborations on fun and challenging topics!

List of Publications

Journal Articles

- M. Lambooij, M. J. Murdoch, W. A. IJsselsteijn, and I. Heynderickx. The Impact of Video Characteristics and Subtitles on Visual Comfort of 3D TV. *Displays*, 34(1):8, 2013.
- Michael J. Murdoch and Ingrid E. J. Heynderickx. Veiling Glare and Perceived Black in High Dynamic Range Displays. *Journal of the Optical Society of America A*, 29(4): 559–566, 2012.
- Michael J. Murdoch, Dragan Sekulovski, and Ingrid Heynderickx. Preferred Color Gamut Boundaries for Wide-Gamut and Multiprimary Displays. *Color Research & Application*, 2012. doi: 10.1002/col.21780.
- Bogumil Bartczak, Patrick Vandewalle, Oliver Grau, Gérard Briand, Jérôme Fournier, Paul Kerbiriou, Michael Murdoch, Marcus Müller, Rocco Goris, Reinhard Koch, and René van der Vleuten. Display-Independent 3D-TV Production and Delivery Using the Layered Depth Video Format. *IEEE Transactions on Broadcasting: Special Issue on 3DHorizon*, 57(2):477–490, 2011.
- Michael J. Murdoch, Dragan Sekulovski, Robert de Volder, and Ingrid Heynderickx. Preferred and Maximally-Acceptable Color Gamut for Reproducing Natural Image Content. *Journal of the Society for Information Display*, 18(12):1111–1118, 2010.
- Michael E. Miller and Michael J. Murdoch. RGB-to-RGBW Conversion with Current Limiting for OLED Displays. *Journal of the Society for Information Display*, 17(3): 195–202, 2009.
- J. P. Spindler, T. K. Hatwar, M. E. Miller, A. D. Arnold, M. J. Murdoch, P. J. Kane, J. E. Ludwicki, P. J. Alessi, and S. A. Van Slyke. System Considerations for RGBW OLED Displays. *Journal of the Society for Information Display*, 14(1):37–48, 2006.
- A. D. Arnold, P. E. Castro, T. K. Hatwar, M. V. Hettel, P. J. Kane, J. E. Ludwicki, M. E. Miller, M. J. Murdoch, J. P. Spindler, S. A. Van Slyke, K. Mameno, R. Nishikawa, T. Omura, and S. Matsumoto. Full-Color AMOLED with RGBW Pixel Pattern. *Journal of the Society for Information Display*, 13(6):525–535, 2005.

Conference Proceedings

- Ulrich Engelke, Mariska G. M. Stokkermans, and Michael J. Murdoch. Visualizing Lighting with Images: Converging between the predictive value of renderings and photographs. In *Human Vision and Electronic Imaging XVIII*. SPIE, 2013.
- Bart Salters, Michael Murdoch, Dragan Sekulovski, Shih-Han Chen, and Pieter Seuntjens. An Evaluation of Different Setups for Simulating Lighting Characteristics. In *Human Vision and Electronic Imaging XVII*. SPIE, 2012.
- M. Stokkermans, M.J. Murdoch, and U. Engelke. Preference for Key Parameter of Tone Mapping Operator in Different Viewing Conditions. In *Experiencing Light*, 2012.
- Michael J. Murdoch, Dragan Sekulovski, and Pieter Seuntjens. The Influence of Speed and Amplitude on Visibility and Perceived Subtlety of Dynamic Light. In *19th Color & Imaging Conference*, pages 265–269. IS&T, 2011.
- M. Perz, D. Sekulovski, and M. Murdoch. Chromatic Flicker Perception in Human Peripheral Vision Under Mental Load. In *18th Color & Imaging Conference*. IS&T, 2010.
- M. J. Murdoch, M. van Etten, and I. E. J. Heynderickx. Effect of Environmental Factors on Image Attributes. In *Proceedings International Congress of Imaging Science*, pages 153–156. ICIS, 2010.
- I. Heynderickx, R. de Volder, D. Sekulovski, and M. Murdoch. Preferred Reproduction of Natural Image Content on Wide-Gamut Displays. In *16th International Display Workshops*. SID, 2009.
- Michael J. Murdoch and Michael E. Miller. 52.4: Distinguished Paper: RGB to RGBW Conversion for OLED Displays. In *SID Symposium Digest of Technical Papers*, volume 39, pages 791–794. SID, 2008.
- Michael E. Miller, Michael J. Murdoch, E. Ludwicki John, and D. Arnold Andrew. P-73: Determining Power Consumption for Emissive Displays. In *SID Symposium Digest of Technical Papers*, volume 37, pages 482–485. SID, 2006.
- M. J. Murdoch, M. E. Miller, and P. J. Kane. Perfecting the Color Reproduction of RGBW OLED. In *Proceedings International Congress of Imaging Science*, pages 448–51. ICIS, 2006.
- Michael E. Miller, Michael J. Murdoch, P. J. Kane, and A. D. Arnold. P-34: Image Quality Impact of Pixel Patterns and Image Processing Algorithms for RGBW OLED Displays. In *SID Symposium Digest of Technical Papers*, volume 36, pages 398–401. SID, 2005.

Master's Thesis

Michael J. Murdoch. Nonverbal Vocal Interface. Master's thesis, Rochester Institute of Technology, 2006.

United States Patents

- 8,013,817 Electronic display having improved uniformity
- 7,969,428 Color display system with improved apparent resolution
- 7,872,619 Electro-luminescent display with power line voltage compensation
- 7,764,252 Electroluminescent display brightness level adjustment
- 7,667,891 Desktop display with continuous curved surface
- 7,586,497 OLED display with improved power performance
- 7,515,122 Color display device with enhanced pixel pattern
- 7,397,485 Color OLED display system having improved performance
- 7,362,336 Four color digital cinema system with extended color gamut and copy protection
- 7,333,080 Color OLED display with improved power efficiency
- 7,230,594 Color OLED display with improved power efficiency
- 7,184,067 Color OLED display system
- 7,091,941 Color OLED display with improved power efficiency
- 7,075,242 Color OLED display system having improved performance
- 6,911,961 Method of designing an OLED display with lifetime optimized primaries
- 6,897,876 Method for transforming three color input signals to four or more output signals for a color display
- 6,885,380 Method for transforming three color input signals to four or more output signals for a color display
- 6,747,618 Color organic light emitting diode display with improved lifetime

Curriculum Vitae



Michael J. Murdoch was born in 1975 in southern Illinois, USA, lived in several states as well as the UK while growing up, and returned to Illinois to graduate from O'Fallon Township High School in 1993. He studied Chemical Engineering at Cornell University, earning his BS *cum laude* in 1997, then worked for the next 11 years at Eastman Kodak where he wisely shifted his career toward digital imaging. In his spare time he pursued a MS in Computer Science, focused on graphics and computer vision, which he completed in 2005 at Rochester Institute of Technology.

His Kodak work included system modeling and simulation for professional slide film, digital intermediate image enhancement for motion picture systems, OLED display color processing and system design, and image quality modeling and specification for commercial high-speed inkjet systems. In his first two years at Kodak, he completed the Imaging Science Career Development program, through which he learned the nuances of both chemical and digital imaging systems and found mentorship in color and imaging science.

In 2008, Michael decided to move to Europe with a job at Philips Research in Eindhoven, The Netherlands. His first few years there focused on display image quality and visual comfort for new display technologies including wide-gamut, multiprimary, ultra-high contrast, and stereoscopic 3D. Since 2011, he has led a team creating computer graphics tools for perceptually-accurate visual simulation of lighting systems which are used to develop, evaluate, and sell new LED lighting concepts and study quality of light.