https://doi.org/10.2352/issn.2169-2629.2019.27.63

Quantitative Assessment of Color Tracking and Gray Tracking in Color Medical Displays

Wei-Chung Cheng, Chih-Lei Wu, Aldo Badano; Division of Imaging, Diagnostics, and Software Reliability, Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, U.S. Food and Drug Administration; Silver Spring, Maryland, USA

Abstract

The goal of this study is to develop quantitative metrics for evaluating color tracking and gray tracking in a color medical display. Color tracking is the chromaticity consistency of the red, green, or blue shades. Gray tracking is the chromaticity consistency of the gray shades. Color tracking and gray tracking are the most important colorimetric responses of a color medical display because they directly indicate the color calibration quality and can therefore be used to compare color performance between displays. Two metrics, primary purity and gray purity, are defined to measure the color shift of the primary shades and gray shades of a color display, respectively. The area under the curves of primary purity and gray purity can then represent the quality of color tracking (C_AUC) and gray tracking (G_AUC), respectively. Fifteen displays including medical, professional-grade, consumer-grade, mobile, and special displays were tested to compare their C_AUC and G_AUC. The OLED displays have the greatest C_AUC values. The medical and professional-grade displays have the greatest combinations of C_AUC and G_AUC values. Most consumer-grade displays have lower C_AUC and G_AUC values, but some show better gray tracking than color tracking. The special displays exhibit particularly poor color and gray tracking. Using C_AUC and G_AUC together can quantitatively predict and compare color performance of different displays.

Introduction

A medical display is an optoelectronics instrument that presents a two-/three-dimensional array of medical data with corresponding monochrome or color pixels/voxels arranged in the original geometrical order for the human reader to visually perceive the structured information. Grayscale medical displays are used in modalities such as mammography, computed tomography, and magnetic resonance, where the light sources are invisible to the human eye. Color medical displays are used in visible light-based modalities including pathology, dermatology, endoscopy, and ophthalmology, and pseudocolor modalities including Doppler ultrasound, positron emission tomography and single-photon emission computed tomography. Synthetic images that fuse data from multiple modalities also require color medical displays.

As the bridge between the medical data and the physicians, a medical display must faithfully convey the visual information to the human visual system via a well calibrated optical system [1]. For patient safety considerations, medical displays are universally regulated as medical devices and imposed by minimum performance requirements. Table 1 lists a subset of characteristics that need to be assessed for clearing a display as a medical device in the U.S. [2]. The American Association of Physicists in Medicine (AAPM) Task Group 18 (TG18) recommended a set of essential optical characteristics that should be rigorously assessed [3]. However, the test methods recommended by AAPM TG18 are not for color medical displays but grayscale medical displays only [4-5].

Luminance response of a display is the relationship between the input signal and the output optical stimulus. To be precise, for a *K*-bit digital display, the digital input, called *digital driving level* (DDL), is an integer *D* between 0 and 2^K-1. The output is the measured luminance of the output optical stimulus, *Y*, in cd/m^2. *Colorimetric response* of a display is the relationship between the DDL and the three-dimensional color coordinate of the output optical stimulus. Figure 1 shows the CIE XYZ measurement results of a high-end color medical display. The normalized CIE *X*, *Y*, and *Z* of the red, green, blue, and gray shades were measured from DDL 0 to 255 for the 8-bit display. In each of the red, green, and blue charts, the *X*, *Y*, and *Z* curves deviate from each other as the DDL decreases. It indicates the change in the ratio of *X*, *Y*, and *Z*, which equates to the chromaticity represented by (*x,y*). In other words, all three primary shades see chromatic shift as the normalized DDL decreases from 1 to 0. Further observation can be made about the red primary, which tends to shift toward the blue hue because the percentage of *Z* increases as the DDL decreases. Nonetheless, in the gray chart, the *X*, *Y*, and *Z* curves overlap perfectly. It implies that the gray shades have very little chromatic shift throughout the whole DDL range.

The goal of this study was to derive a set of quantitative metrics for assessing the color tracking and gray tracking of a color display. The methods are described in Section 2. The experiment results and applications are presented in Section 3 followed by discussion and conclusions.

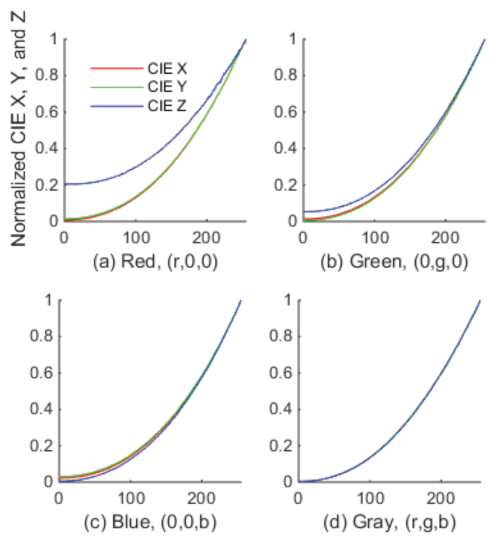

Figure 1. *Normalized CIE XYZ measurement results of the red, green, blue, and gray shades of a high-end color medical display (RadiForce R31, Eizo Corporation, Ishikawa, Japan). The overlapped XYZ curves of the gray channel in (d) indicate good gray tracking. The separation between the XYZ curves in the red (a), green (b), and blue (c) channels indicates imperfect color tracking.*

Methods and Materials

Definitions

The definitions and notations used in this paper are summarized in Table 2.

Table 2: Summary of notations

DDL	Digital driving level. Actual DDL D is an integer between 0 and 2K-1 in a K-bit display. Normalized DDL d is D/(2K-1).
(r,g,b)	DDL for the red, green, and blue channels, and its resulting shade
w_r:w_g:w_b	The ratio between the red, green, and blue primaries to synthesize shade (r,g,b)
X(r,g,b) *Y(r,g,b)* *Z(r,g,b)*	CIEXYZ tristimulus of shade (r,g,b)
CIE (x, y)	Two-dimensional coordinate of chromaticity
$s(d,0,0)$ $s(0,d,0)$ $s(0,0,d)$	Primary purity (curves) for the red, green, and blue channels as functions of normalized DDL
$t(d,d,d)$	Gray purity for the gray shades
C_AUC	Area under the color tracking curves
G_AUC	Area under the gray tracking curve

Measurement Data

The methodology is based on the colorimetric data collected by measuring the tristimulus of the red, green, blue, and gray shades with a colorimeter or spectroradiometer. The following colorimetric data need to be collected.

Table 1: Measurable Characteristics of a Color Display

Spatial resolution	Filling factor, pixel pitch, pixel count, MTF
Temporal resolution	Gray-to-gray response, refresh rate
Uniformity	Pixel-level noise, Mura, angular response
Luminance response	L_{min}, gamma, L_{max}
Colorimetric response	Gray tracking, color tracking, color gamut

$[X(d,0,0)\ Y(d,0,0)\ Z(d,0,0)],\ 0 \le d \le 1$ … red

$[X(0,d,0)\ Y(0,d,0)\ Z(0,d,0)],\ 0 \le d \le 1$ … green

$[X(0,0,d)\ Y(0,0,d)\ Z(0,0,d)],\ 0 \le d \le 1$ … blue

$[X(d,d,d)\ Y(d,d,d)\ Z(d,d,d)],\ 0 \le d \le 1$ … gray

Since d is the normalized DDL, the actual DDL used to drive a K-bit display is $\lfloor (2^K - 1)d \rfloor$. Ideally, all shades between DDL 0 and $(2^K - 1)$ should be measured unless the data transits smoothly such that a simple interpolation of subsampled data is sufficient.

Ratio between the Primary Colors

Consider a red shade generated from DDL triplet $(d,0,0)$ on a display resulting colorimetric measurement $[X(d,0,0)\ Y(d,0,0)\ Z(d,0,0)]$. This red shade can be synthesized by mixing the primary colors (1,0,0), (0,1,0), and (0,0,1) with a specific ratio w_r:w_g:w_b that satisfies the following matrix equation.

$$\begin{bmatrix} X(d,0,0) \\ Y(d,0,0) \\ Z(d,0,0) \end{bmatrix} = \begin{bmatrix} X(1,0,0) & X(0,1,0) & X(0,0,1) \\ Y(1,0,0) & Y(0,1,0) & Y(0,0,1) \\ Z(1,0,0) & Z(0,1,0) & Z(0,0,1) \end{bmatrix} * \begin{bmatrix} w_R(d,0,0) \\ w_G(d,0,0) \\ w_B(d,0,0) \end{bmatrix} \quad (1)$$

where w_r, w_g, and w_b can be found by solving the matrix equation by left-multiplied by the inverse of the 3×3 matrix.

$$\begin{bmatrix} w_R(d,0,0) \\ w_G(d,0,0) \\ w_B(d,0,0) \end{bmatrix} = \begin{bmatrix} X(1,0,0) & X(0,1,0) & X(0,0,1) \\ Y(1,0,0) & Y(0,1,0) & Y(0,0,1) \\ Z(1,0,0) & Z(0,1,0) & Z(0,0,1) \end{bmatrix}^{-1} * \begin{bmatrix} X(d,0,0) \\ Y(d,0,0) \\ Z(d,0,0) \end{bmatrix} \quad (2)$$

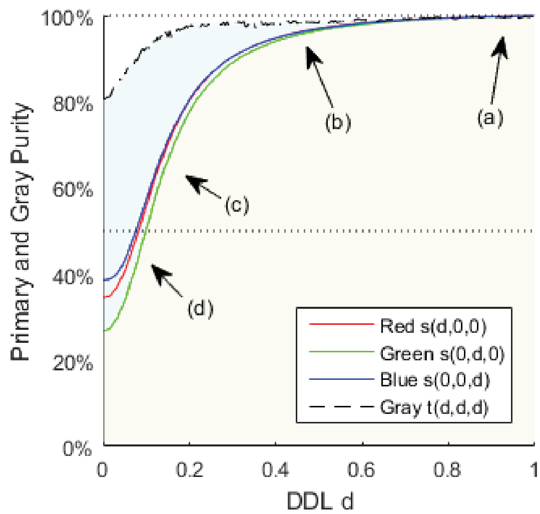

Figure 2. *Color tracking is indicated by the three solid curves in red, green, and blue. Gray tracking is indicated by the black dashed curve. Both gray tracking and color tracking are perfect in region (a). Gray tracking remains near perfect while color tracking starts to degrade in region (b). In region (c), gray tracking becomes poor because the three primary purity curves split. The gray shades shift toward magenta because the green curve is lower than the other two. Region (d) annotates where the green primary purity fails to meet 50% primary purity. The yellow-shaded area under the green curve expresses the degree of color tracking (the larger, the better). The cyan-shaded area under the black curve expresses the degree of gray tracking (the larger, the better).*

Color Tracking

Define *primary purity* of the red shade *(d,0,0)* as

$$s(d,0,0) \equiv \frac{w_R(d,0,0)}{w_R(d,0,0)+w_G(d,0,0)+w_B(d,0,0)} \quad (3)$$

$s(d,0,0)$ represents the contribution of the primary red (1,0,0) to the red shade (d,0,0) as a percentage. The primary purity is a unitless measure, so 1.0 and 100% are equivalent and can be used interchangeably. If there is no chromaticity difference between (d,0,0) and (1,0,0), then both w_g and w_b are zero while w_r is 100%, the maximum "purity" of the red primary. Otherwise, (d,0,0) and (1,0,0) differ in chromaticity and (d,0,0) is considered "contaminated" by green and blue, so its purity of red is reduced from 100% to $s(d,0,0)$.

The primary purity of green and blue shades are defined similarly:

$$s(0,d,0) \equiv \frac{w_G(0,d,0)}{w_R(0,d,0)+w_G(0,d,0)+w_B(0,d,0)} \quad (4)$$

$$s(0,0,d) \equiv \frac{w_B(0,0,d)}{w_R(0,0,d)+w_G(0,0,d)+w_B(0,0,d)} \quad (5)$$

Figure 2 shows the *primary purity curves* of the red, green, and blue channels of a color display. The *color-tracking area-under-curve* (C_AUC) function is defined as the area under the lowest primary purity curves.

$$C\_AUC \equiv \int_{x=0}^{1} MIN\big(s(x,0,0), s(0,x,0), s(0,0,x)\big)\, dx \quad (6)$$

A greater C_AUC value indicates a higher quality of color tracking. The C_AUC of an ideal display is 100%.

Gray Tracking

The *gray purity* is defined as follows.

$$t(d,d,d) \equiv MIN \begin{cases} \dfrac{w_R(d,d,d)}{w_R(d,d,d)+w_G(d,d,d)+w_B(d,d,d)} \Big/ \dfrac{w_R(1,1,1)}{w_R(1,1,1)+w_G(1,1,1)+w_B(1,1,1)} \\ \dfrac{w_G(d,d,d)}{w_R(d,d,d)+w_G(d,d,d)+w_B(d,d,d)} \Big/ \dfrac{w_G(1,1,1)}{w_R(1,1,1)+w_G(1,1,1)+w_B(1,1,1)} \\ \dfrac{w_B(d,d,d)}{w_R(d,d,d)+w_G(d,d,d)+w_B(d,d,d)} \Big/ \dfrac{w_B(1,1,1)}{w_R(1,1,1)+w_G(1,1,1)+w_B(1,1,1)} \end{cases} \quad (7)$$

The *gray tracking area-under-curve* (G_AUC) function is defined as the area under the gray purity curve.

$$G\_AUC \equiv \int_{x=0}^{1} t(x,x,x)\, dx \quad (8)$$

The range of the G_AUC value is between 0 and 100%. A greater G_AUC value indicates a better quality of gray tracking. The G_AUC of an ideal display is 100%.

Experiments

In this study, a spectroradiometer (CS2000, Konica Minolta Sensing Americas, Inc., New Jersey, USA) was used to measure the spectra between 380 and 780 nm at 1 nm interval. Fifteen displays were measured and tested with the abovementioned method. The manufacturer and model information of the displays are listed in Table 3. The cohort of displays comprised LCD and OLED monitors that were used in medical, professional, office, entertainment, and mobile applications. The displays are categorized based on their intended use and target markets. Mobile OLED and mobile LCD are hand-held displays embedded in mobile phones, so the adjustable display hardware settings are limited. Consumer-grade, in this paper, is used to describe the class of computer monitors that provide basic brightness, contrast, and color temperature adjustment, and support only the sRGB color space. Professional-grade, in this paper, is used to qualify high-end desktop monitors that have internal look-up tables for color calibration guided by external measurement instruments and support multiple color spaces with wider color gamut. Medical-grade displays must support the DICOM grayscale standard display function (GSDF) mode and usually carry out the calibration process with built-in sensors automatically. Two special displays were intentionally added to demonstrate subpar color performance. Display *#n* is one of the pioneering shutter-based stereoscopic displays with a fast response time designed for computer gamers. Display *#o* is a 7-inch portable LCD that is intended to be used as an accessory of cameras, camcorders, or DVD players.

All displays were put in their default settings via the on-screen display (OSD) menu. The sRGB color space was selected for the professional-grade displays. The mobile devices were set to their maximum luminance. No color profile was used in the driving computer to alter the pixel data. No external color calibration kit, optional for *#f* and *#h*, was used to calibrate the displays. The medical-grade LCD *#d* had a built-in sensor for maintaining luminance consistency. All desktop displays were driven with 8 bits via a DVI/HDMI cable (though *#d* and *#e* are capable of 10 bits). Color/image enhancement features in the nVidia graphics card driver were disabled. Each unit was warmed up for at least two hours before the measurement.

Table 3: Displays under test

a	HTC Nexus 1	Mobile OLED
b	Samsung Galaxy S	Mobile OLED
c	Sony PVM-2551MD	Consumer-grade OLED
d	Barco MDCC 6130	Medical-grade LCD
e	Eizo CG243W	Professional-grade LCD
f	HP Z24X	Professional-grade LCD
g	Apple Cinema	Consumer-grade LCD
h	NEC PA271W	Professional-grade LCD
i	HP ZR2240W	Consumer-grade LCD
j	HP ZR2440W	Consumer-grade LCD
k	Dell 1907FP	Consumer-grade LCD
l	Apple iPhone4	Mobile LCD
m	Dell P2212Hb	Consumer-grade LCD
n	Samsung SM2233RZ	Gaming 3D LCD
o	Pyle PLCMHD70	Portable LCD

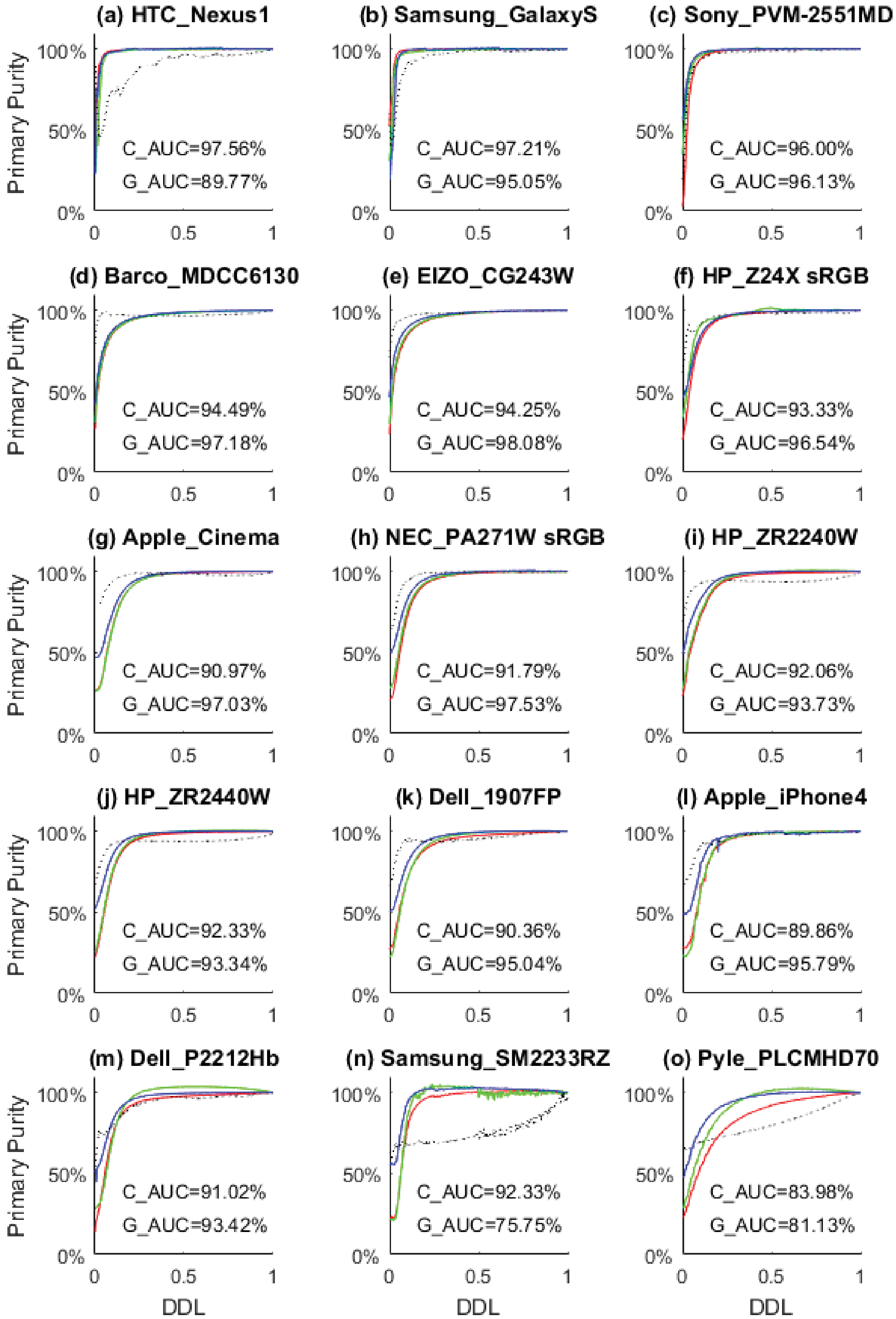

Figure 3. *The primary purity and gray purity curves of 15 displays.*

Results

The primary purity and gray purity curves of the 15 displays are shown in Figure 3 with their C_AUC and G_AUC annotated. Recall that an ideal display has flat curves and 100% C_AUC and G_AUC, so flatter curves enclosing larger areas indicate better display performance.

In the first row for the three OLED displays (*#a*, *#b*, and *#c*), the primary purity curves (in red, green, and blue solid lines) are virtually flat between DDL 0.2 and 1, and their C_AUC values are between 96% and 97.56%. Thus, the OLED displays have the best color tracking among the 15 displays. Although their color tracking is comparable, they differ in gray tracking, which is indicated by the gray purity curves (in black dashed lines). Display *#a*, a pioneering OLED mobile phone, has worse gray tracking (89.77%) than *#b*, a more modern OLED mobile phone (95.05%), and *#c*, a desktop OLED display (96.13%).

The rest 12 displays are all LCD-based. Their primary purity curves are not as flat as the first three, so their C_AUC values are lower. Displays *#d*~*#f* are professional-grade displays having C_AUC between 93.33% and 94.49%. None of the remaining 9 LCDs has C_AUC greater than 92.33%.

The four professional-grade displays (*#d*, *#e*, *#f* and *#h*) and display *#g* show the best gray tracking among the 15 displays. Their gray purity curves are flat between DDL 0.2 and 1, and mostly above their primary purity curves. The G_AUC of these five displays are between 96.54% and 98.08%.

Displays *#i*, *#j*, *#k* and *#m* are consumer-grade desktop displays. Their color tracking (C_AUC) is in the same range as the professional-grade group. However, their gray tracking is inferior because, despite the lower G_AUC, their gray purity curves fail to maintain flat between DDL 0.3 and 1. In contrast, display *#l*, a mobile phone, maintains both color tracking and gray tracking adequately between DDL 0.3 and 1, and degrades quickly below DDL 0.3. Display *#m* is a bargain desktop display using the early LED backlighting technology. It exhibits serious color shift, which is revealed by the separation between the three primary purity curves. Notice that the green primary purity curve exceeds 100% because the green shades between DDL 0.2 and 0.9 deviate in hue and color-shift out of the original color gamut.

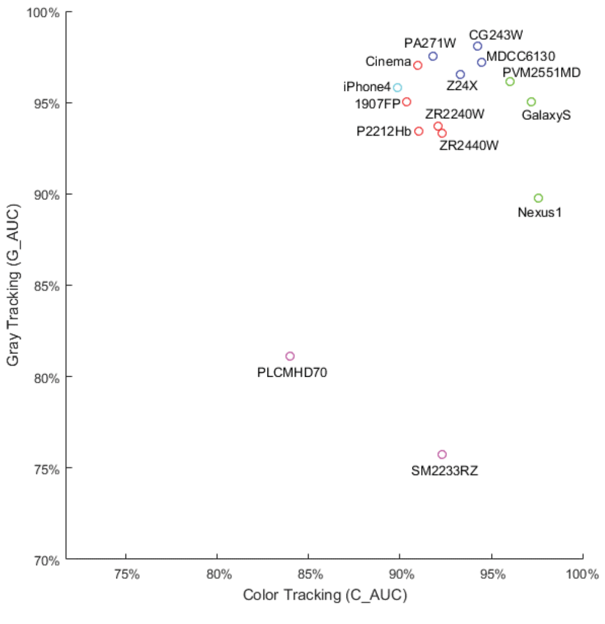

Figure 4. *The gray tracking (G_AUC) and color tracking (C_AUC) of 15 displays. The circles are color-coded by category – green: OLED, blue: professional-grade, red: LCD desktop, cyan: LCD mobile, magenta: special displays*

Displays *#n* and *#o* are special displays. Display *#n* is an early 240Hz stereoscopic display that emphasizes on minimizing transition time. Judged by its gray purity curve, calibration for gray tracking seems to be overlooked in its design. Similar problem is observed in display *#o*, a portable device that employs a 7-inch panel, which is nominally of low quality.

Color Tracking vs Gray Tracking

The 15 displays are placed on the same plane per their G_AUC and C_AUC in Figure 4. The *x*-axis is the C_AUC representing color tracking, while the *y*-axis is the G_AUC representing gray tracking.

Having both greater C_AUC and greater G_AUC indicates better color tracking and gray tracking, so better displays are located at the upper-right corner of the plot.

The displays are color-coded by their categories. The group of OLED displays (green circles) is located on the right-most side of the chart, indicating the best color tracking.

The professional-grade displays (blue circles) are in the upper-right cluster and seem to have a better balance between color tracking and gray tracking. Compared with the OLED displays, the LCD mobile display (cyan circle) is on the opposite side of the diagonal, meaning better gray tracking but worse color tracking. Similarly, display *#h* (PA271W), *#g* (Cinema), and *#k* (1907FP) are on the same side and share the same properties.

The remaining consumer-grade displays, *#i* (ZR2240W), *#j* (ZR2440W), *#m* (P2212Hb), have inferior gray tracking and color tracking to *#d* (MDCC6130), *#e* (CG243W), and *#f* (Z24X), but they also cost much less and represent the entry-level, consumer-grade displays.

The two special displays (magenta circles) do not belong to any cluster because of their poor color tracking and gray tracking.

Discussion

Another metric for measuring gray tracking was recently recommended by the AAPM Task Group 196 (TG196) [8]. The TG196 method uses $\Delta(u',v')$ to evaluate the chromaticity deviation of gray shades from the white point, and is therefore suitable for assessing absolute errors against acceptance criteria. The TG196 method was designed for evaluating gray tracking only, but might be able to evaluate color tracking as well after some modification. The TG196 method is independent of the primary shades. It is unclear how the primary shades, or the color gamut, will influence the perception of gray tracking. In contrast, the primary purity and gray purity are tools for presenting both color tracking and gray tracking characteristics of a display. The G_AUC is a function of not only the gray shades but also the primary shades.

The primary and gray purity metrics are relative measures and do not include absolute colorimetric characteristics such as the chromaticity of the white point, the color gamut, or the maximum/minimum luminance. Although these characteristics can be retrieved from the collected measurement data, the evaluation of these characteristics needs to be done with other metrics. For this reason, the metrics are suitable for evaluating or guiding the color calibration, usually conducted by a color calibration kit, of the same display.

The correlation between the (C_AUC, G_AUC) metrics and visually perceived differences was not examined in this study. The reason was that singling out the color tracking and gray tracking characteristics from the other color characteristics of a display, including the luminance responses, color gamut, and additivity of primaries, is challenging for the study design. Furthermore, an ideal display having perfect color tracking and gray tracking is not feasible, so the instrumentation is also intractable.

Conclusions

Two metrics, primary purity and gray purity, were introduced for describing the color shift of the primary and gray shades of a display, respectively. A plot of both the primary purity curves and the gray purity curve provides a concise visual presentation of the color tracking and gray tracking properties of a display. The areas under the curves, C_AUC and G_AUC, are single-value measures representing the color tracking and gray tracking properties of a display. Based on the (C_AUC, G_AUC) dual, a binary relation inferiority can be defined to compare the color tracking and gray tracking of two displays.

References

[1] L. D. Silverstein, S. F. Hashmi, K. Lang, E. A. Krupinski, "Paradigm for Achieving Color-reproduction Accuracy in LCDs for Medical Imaging," *Journal of the Society for Information Display*, vol. 20, no. 1, pp. 53-62, January 2012.

[2] FDA. *Display Devices for Diagnostic Radiology – Guidance for Industry and Food and Drug Administration Staff*. October 2017.

[3] E. Samei, et *al*., "Assessment of Display Performance for Medical Imaging Systems: Executive summary of AAPM TG18 report," *Medical Physics*, vol. 32, no. 4, pp. 1205-1225, April 2005.

[4] W. C. Revie and P. Green, "Current Problems and Perspectives on Colour in Medical Imaging," *Electronic Imaging*, vol. 20, pp. 1-5, February 2016.

[5] A. Badano, et *al*., "Consistency and Standardization of Color in Medical Imaging: A consensus report," *Journal of Digital Imaging*, vol. 28, no. 1, pp. 42-52, February 2015.

[6] E. A. Krupinski and H. Roehrig, "The Influence of a Perceptually Linearized Display on Observer Performance and Visual Search," *Academic Radiology*, vol. 7, no. 1, pp. 8-13, January 2000.

[7] G. Wyszecki and W. S. Stiles, *Color Science*. New York: Wiley, 1982.

[8] A. Badano, et. *al*., "Gray Tracking in Medical Color Displays—A report of Task Group 196," *Medical Physics*, vol. 43, no. 7, pp. 4017-22, July 2016.

Author Biography

Wei-Chung Cheng is a research scientist at FDA. He received a PhD in electrical engineering from University of Southern California. His research interests include color imaging system, color vision, and display systems.

Chih-Lei Wu received his BS and MS in electrical engineering from I-Shou University and Chang Gung University, Taiwan in 2002 and 2004, respectively. With 13 years of experience in the display industry, he specializes in display technologies including TFT-LCD, rigid/flexible AMOLED, AMQLED, and MEMS, as well as module electronics design, EAC (Cell2), automotive testing, FPGA design, real-time image/color processing, and color performance evaluation.

Aldo Badano is a senior biomedical researcher at FDA. He currently serves as Deputy Director of the Division of Imaging, Diagnostics, and Software Reliability, Office of Science and Engineering Laboratories, Center for Devices and Radiological Health. He received an MEng degree in radiological health engineering and a PhD degree in nuclear engineering from the University of Michigan in 1999 and 1995 after obtaining a Chem Eng degree from the Universidad de la República, Montevideo, Uruguay, in 1992. His primary interests are in the characterization and modeling of medical imaging acquisition and visualization systems.