A Statistical Model of Tristimulus Measurements Within and Between OLED Displays

Matti Raitoharju

Department of Automation Science and Engineering
Tampere University of Technology
matti.raitoharju@tut.fi

Samu Kallio and Matti Pellikka

Microsoft Mobile Oy

Abstract—We present an empirical model for noises in color measurements from OLED displays. According to measured data the noise is not isotropic in the XYZ space, instead most of the noise is along an axis that is parallel to a vector from origin to measured XYZ vector. The presented empirical model is simple and depends only on the measured XYZ values. Our tests show that the variations between multiple panels of the same type have similar distribution as the temporal noise in measurements from a single panel, but a larger magnitude.

Index Terms—displays, measurement uncertainty, noise measurement, calibration, mathematical model

I. INTRODUCTION

Tristimulus measurements from displays can be used for determining color reproduction characteristics and calibration. These measurements are commonly made in Commission internationale de l'éclairage (CIE) 1931 XYZ color space. The components of XYZ color space are defined as integrals of product of a spectral radiance and color matching function over the visible light spectrum. The color matching functions for X, Y, and Z are positive and no tristimulus value is negative [1]. In practice, tristimulus measurements contain noise.

In [2] the angular optical properties of active matrix organic light emitting diode (AMOLED), organic light emitting diode (OLED)-TV, and polymer OLED displays are evaluated by making accurate measurements from various angles. In [3] the variation of displayed colors in different parts of a display is evaluated.

Our work differs from earlier work in two main ways. First, we measure the each color multiple times from multiple panels. Our measurements provide insights to the variation of displayed colors in time and between panels, while related work on OLED panels [2], [3] is concentrated on measuring a single or a couple of displays from various locations and angles.

Second, we build a model that is based only on the statistical properties of the tristimulus measurements measured from OLED displays. Using a physics-based approach as in [4] requires a better knowledge of the technical details of the measured panel, while the empirical model can be applied and fitted using only the measurements as the source of information and, thus, the model can be used and verified without access to the underlying hardware details. The statistical modeling takes the combination of all noise sources observed into account.

The measured XYZ values are usually converted into another color space, which is selected according to the application. The conversion from XYZ space to a RGB space, such as sRGB, can be done using a linear transformation [5]. The XYZ values can also be converted to a uniform color space. The purpose of a uniform color space is to present the observed color differences [6]. In calibration of the devices, it is usually desired to minimize the observed color difference. Also uncertainties in the XYZ space can be transformed into other color spaces [7]. This allows the use of the proposed noise model also with other color spaces.

For color calibration there are several different models [8], [9].The color reproduction of OLED panels is not linear. With uncalibrated panels in our test the luminance of pure white at maximum brightness was about 93% of the luminance of the sum of luminances of the pure red, green, and blue. The proposed noise model for a single panel depends only on the measured XYZ values, which makes it independent of the colors sent to the panel and, thus, the calibration of the panel.

In this paper, we build models for measurements from a single region of a single panel and between multiple panels. In measurements from a single panel the noise is temporal. In the measurements between panels there is also variation that is caused by the different static properties of the different panels. We will show that even these noises are caused by different sources they have similar statistical properties.

Rest of this paper is organized as follows. In Section II the noise model is developed. In Section III we show an example of the use of the model. In Section IV we compare our results with results presented in [3]. Section V concludes the paper.

II. NOISE MODEL FOR TRISTIMULUS MEASUREMENTS

A. Background

In tests with cathode ray tube (CRT) displays [10], it is proposed that most of the noise is along the luminance component (Y). In [7] it is stated that the source noise in a spectral measurement contributes to uncertainties in each of the tristimulus values, but not to chromaticity values because of the correlations. The chromaticity values x and y are defined as

$$x = \frac{X}{X + Y + Z} \text{ and } y = \frac{Y}{X + Y + Z}. \qquad (1)$$

If the noise does not affect x and y values, but affects X, Y, and Z values, the noise vector has to be in the null space of Jacobian of the mapping (1). The null space in this case is aligned with $\begin{bmatrix} X & Y & Z \end{bmatrix}^T$ vector. This model would be valid if the noise would be such that the spectral radiance of a display is multiplied by a wavelength independent noise. In next subsections, we test how the measured variation from displays vary along different directions to see if the largest variation is along $\begin{bmatrix} X & Y & Z \end{bmatrix}^T$ and what is its relation to noise along other directions.

B. Noise of measurements from an OLED display

For our first test, we measure several colors from a 5.7" AMOLED panel, each color 12 times. The measurements are made with 4 calibrated Ocean Optics USB2000+ spectrometers (3 measurements with each spectrometer).

We investigate the noise by computing standard deviations (SDs) along a different directions. First compute the sample mean

$$\mu = \frac{\sum_{i=1}^n c_i}{m}, \tag{2}$$

where c_i is a tristimulus measurement and m is the number of samples. The SD along a unit vector v is

$$\sigma_v = \sqrt{\sum_{i=1}^m \frac{(v^T(c_i - \mu))^2}{m}}. \tag{3}$$

Fig. 1 shows the SDs along the main axis of the XYZ colorspace, and along unit vectors:

$$v_1 = \frac{\begin{bmatrix} X & Y & Z \end{bmatrix}}{\sqrt{X^2 + Y^2 + Z^2}}, v_2 = \frac{\begin{bmatrix} 0 & -Z & Y \end{bmatrix}^T}{\sqrt{Y^2 + Z^2}}, v_3 = \frac{v_1 \times v_2}{\|v_1 \times v_2\|}. \tag{4}$$

Markers in the figure show the SDs for different colors, and the marker color presents the color drawn on the display. Some greys were measured in multiple brightness levels. Furthermore, the figure shows a linear fit that approximates the SDs as a function of $X + Y + Z$:

$$\sigma_j \approx k(X_j + Y_j + Z_j). \tag{5}$$

For this problem, the linear least squares estimate for k is

$$k = \frac{\sum_{j=1}^n (X_j + Y_j + Z_j)\sigma_j}{\sum_{j=1}^n (X_j + Y_j + Z_j)^2}, \tag{6}$$

where n is the number of SDs.

The SD along v_1 is clearly larger than along the other dimensions. To see if this is a general behavior we repeat the same test for 12 other panels. Each color was measured from panels from 9 to 15 times. Table I shows the values of $k \times 1000$, the average μ of $k \times 1000$ of all panels, and the SD σ of these. The SD along v_1 is larger than the SD along axis normal to it although there is some variation in the factor. This is caused by differences in displays and the relatively small sample size. To verify the accuracy of the choice of $\begin{bmatrix} X & Y & Z \end{bmatrix}^T$ direction we compute the axis of maximum variation using principal
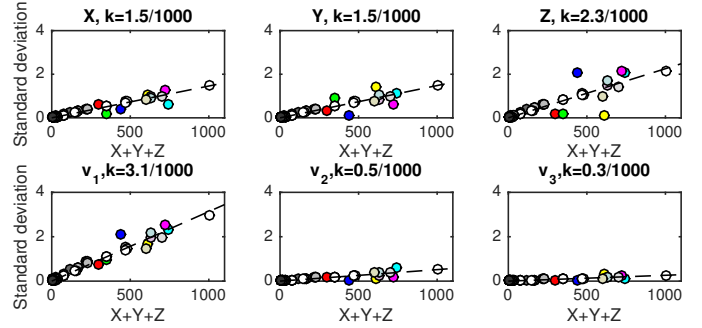

Fig. 1. SDs along $X, Y, Z, v_1, v_2,$ and v_3 of one panel

TABLE I
FACTOR OF SDS $\times 1000$ OF TRISTIMULUS MEASUREMENTS ALONG AND NORMAL TO $[X, Y, Z]^T$ MEASURED FROM 5.7" AMOLED PANEL

panel	1	2	3	4	5	6	7	8	9	10	11	12	13	μ	σ
v_1	3.1	1.4	1.8	1.7	2.2	2.8	2.1	1.8	2.4	2.0	1.1	1.4	2.9	2.1	0.6
v_2	0.5	0.8	1.0	0.8	0.7	0.8	0.6	0.6	1.0	0.5	0.7	0.9	1.1	0.8	0.2
v_3	0.3	0.4	0.4	0.3	0.3	0.4	0.3	0.2	0.3	0.3	0.2	0.3	0.6	0.3	0.1

component analysis (PCA). The $\begin{bmatrix} X & Y & Z \end{bmatrix}^T$ and the main axis of the noise have a difference 14 degrees on average. In comparison, assuming the main axis has a 60-degree angle with Y axis on average. This shows that the assumption of having most of the noise along Y vector in [10], does not hold for OLED displays.

The SD is in every sample larger along vector v_2 than along v_3. With this data, the factor of SD along $[X, Y, Z]$ compared to a normal direction is on average 3.7. Because this test was made with 4 different spectrometers, even though they all were of similar type and calibrated, there may be some calibration related noise in the data.

Furthermore, we tested a different type of AMOLED panels that has 5.2" diameter. In this test, we measured white at maximum brightness setting from each panel 10 times within 15 minutes and computed the SDs of the measurements. Table II shows the SDs of measurements of the white. The results are similar as with the 5.7" display. We can note that the SD along v_1 and other directions is larger at 6.9 and the variation along v_3 is larger than along v_2, which is opposite than with the results in Table I.

Fig. 2 shows the value of $X + Y + Z$ of the white as the function of the measurement time from each display. Asterisks show the measurements and each color marks a panel. This figure shows no clear trends in measurement values during 15 minute measurement sequence, which could have been caused

TABLE II
FACTOR OF SDS $\times 1000$ OF TRISTIMULUS MEASUREMENTS ALONG AND NORMAL TO $[X, Y, Z]^T$ WHEN MEASURED FROM 5.2" AMOLED PANELS

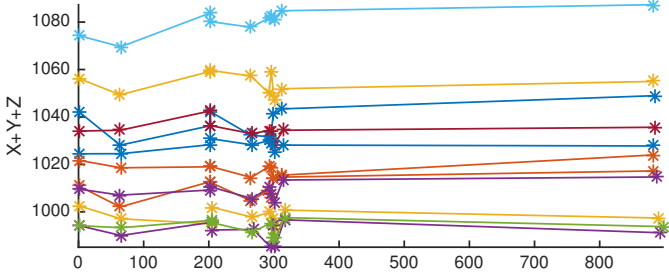| panel	1	2	3	4	5	6	7	8	9	10	11	μ	σ
v_1	3.8	2.8	1.8	2.3	1.7	2.8	2.0	1.3	1.8	2.5	1.9	2.2	0.7
v_2	0.4	0.3	0.2	0.3	0.3	0.3	0.4	0.1	0.2	0.3	0.3	0.3	0.1
v_3	0.3	0.4	0.2	0.2	0.4	0.4	0.9	0.4	0.2	0.3	0.2	0.4	0.2

Fig. 2. Value of X+Y+Z measured from 11 OLED panels showing white color at different time instances. Each line corresponds to a different panel and asterisk to a measurement time.

Fig. 3. SDs of measurements along $[X, Y, Z]^T$ (v_1) and 2 perpendicular directions (v_2, v_3) on multiple panels

be warming of the panel in use, and that measurements from each panel had approximately similar amount of noise and the variations between panels are larger than variations within panel.

C. Noise of measurements between OLED displays

Next we test how the SD of the tristimulus measurements varies between panels. For 5.7" panel we used a test sequence that contains 765 measurements in total. The measurements were done using 3 different brightness levels i.e. there are 255 colors given to panel so that the panel brightness is set to 100%, 50% and 25%. This measurement sequence was done for 22 panels. For 5.2" panels we used two test sequences. A shorter test sequence of 50 measurements was used for 702 panels and longer with 765 measurements to 11 panels. Each color was measured only once from a panel.

Fig. 3 shows the SDs computed from measurements with multiple panels. For all test sequences the SD along v_1 is larger than along the two other dimensions. The results are similar to the result achieved by measuring only one display, but the amount of SD is approximately sixfold compared to variations within display. Figure show also that between panels there is more variation in some colors. Specifically, green has relatively large variation along v_3, although it is still smaller than the variation along v_1.

D. Noise model

We construct a statistical model for noise from OLED displays based on the experiments in the two previous subsections. The covariance matrix uses 5-fold measurement noise along $[X \ Y \ Z]^T$ compared to other directions. It can be written as

$$P = a^2(X + Y + Z)^2 U \begin{bmatrix} 5^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} U^T, \quad (7)$$

where

$$U = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \quad (8)$$

and a is a multiplier that can be set to $\frac{1}{2000}$ for measurements within display and $\frac{1}{400}$ for measurements between displays. If we assume that the temporal noise in every panel is the same and the differences between panels does not change in time,
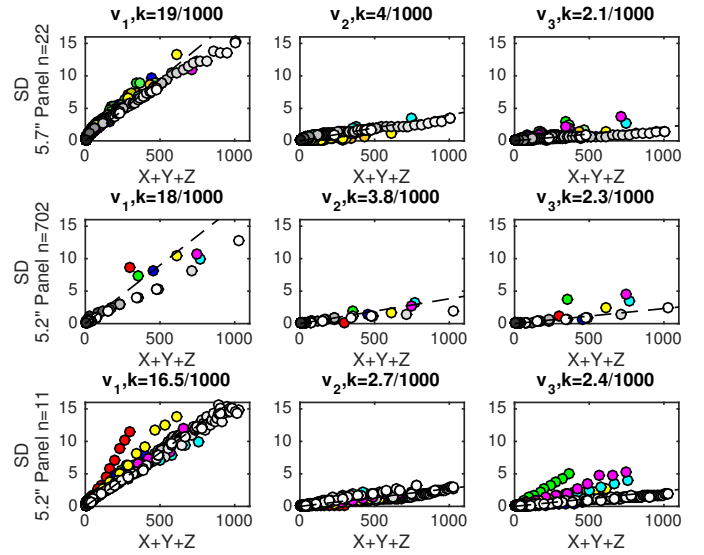
after an optimal calibration the factor a for perfectly calibrated panels should be $\frac{1}{2000}$. The temporal noise of measurements can be reduced by making multiple measurements and taking an average of them. Taking average of multiple measurements increases the measurement time.

In many applications, the value of a does not affect the result and can be omitted. In the next section, we show one such application.

III. USE OF NOISE MODEL IN CALIBRATION

In this application, we try to find a matrix M such that measurement $c_i = [X_i, Y_i, Z_i]^T$ from a display multiplied with it matches with a measurement $\hat{c}_i$ of the same color from another display

$$\hat{c}_i \approx M c_i. \quad (9)$$

To fit multiple measurements, we minimize the norm

$$\sum_i (\hat{c}_i - M c_i)^T \left(\hat{P}_i + M P_i M^T \right)^{-1} (\hat{c}_i - M c_i), \quad (10)$$

where P_i is the covariance (7) for measurement c_i and $\hat{P}_i$ is covariance for $\hat{c}_i$.

There is no simple solution for this equation, because M term is used to induce the norm and also transforming the measurement. Because the target is to transform measurements c_i to match the reference measurements we can make approximation

$$M P_i M^T \approx \hat{P}_i \quad (11)$$

Now the minimized function becomes

$$\sum_i (\hat{c}_i - M c_i)^T (2\hat{P}_i)^{-1} (\hat{c}_i - M c_i). \quad (12)$$

One measurement can be written as

$$
\overbrace{\begin{bmatrix} \hat{X}_i \\ \hat{Y}_i \\ \hat{Z}_i \end{bmatrix}}^{\hat{c}_i} = \overbrace{\begin{bmatrix} X_i & Y_i & Z_i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & X_i & Y_i & Z_i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & X_i & Y_i & Z_i \end{bmatrix}}^{H_i} \overbrace{\begin{bmatrix} m_{1,1} \\ m_{1,2} \\ m_{1,3} \\ m_{2,1} \\ m_{2,2} \\ m_{2,3} \\ m_{3,1} \\ m_{3,2} \\ m_{3,3} \end{bmatrix}}^{m} + \varepsilon_i,
$$

$$\tag{13}$$

where ε_i are independent zero-mean multivariate Gaussians with covariance $\hat{P}_i$, and $m_{i,j}$ is an element of M.

To compute the optimal estimate, the weighted linear least squares formula is used

$$
m = \left(H^T W H \right)^{-1} H^T W \hat{c}, \tag{14}
$$

where

$$
\begin{aligned}
\hat{c} &= \begin{bmatrix} \hat{c}_1 & \dots & \hat{c}_n \end{bmatrix}^T \\
H &= \begin{bmatrix} H_1^T & \dots & H_n^T \end{bmatrix}^T \\
W &= \mathrm{diag}\left((2\hat{P}_1)^{-1}, \dots, (2\hat{P}_n)^{-1} \right).
\end{aligned} \tag{15}
$$

This formulation requires at least 3 linearly independent measurements, which can be obtained e.g. by measuring pure R, G, and B colors. When using only 3 measurements the weighting with noise model does not affect to the values of matrix M, but with more measurements it has an effect on the result. Also, we can note that M is independent of the multiplier a.

This algorithm is similar to the algorithm used in [7] for colorimeter calibration, but here we calibrate outputs of two displays and use a variable covariance weighting for different dimensions in colorspace. The benefit over the four-color method proposed in [10] is that this formulation allows using any number (≥ 3) of measurements.

In our test, we measure red, green, blue and white from a pair of panels and fit matrix M. Elements of M are computed using (14). For verification of the M we measure cyan, magenta, and yellow from the displays and transform obtained tristimulus measurements from the first display with the matrix (9). The result of the transformation is compared with the measurements from the other display. Fig. 4 shows the sum of absolute values of differences of the XYZ values when device 1 is compared with other 12 devices. Uniform error model assumes that the error is distributed uniformly while the proposed model assumes most noise is along v_1. The proposed error model reduces the error compared to the uniform error model in 11 calibrations out of 12.

IV. COMPARISON TO RELATED WORK

In [3], different regions of a 6" AMOLED display are measured so that the display shows one of primary colors at the time. Then the variation of of different regions is compared
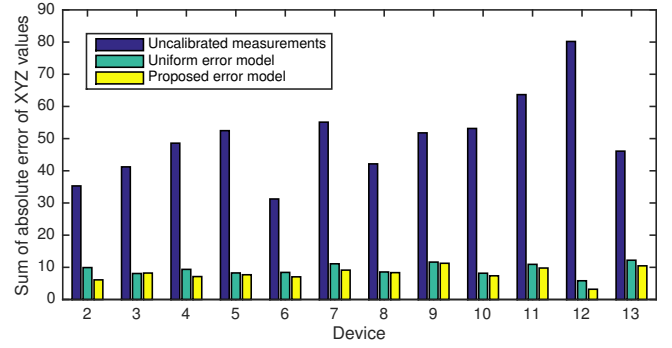

Fig. 4. Results of the calibration using the proposed error model

to the average of the whole display. Value of

$$
\Delta E = \sqrt{(L^* - L^*_{\mathrm{avg}})^2 + (a^* - a^*_{\mathrm{avg}})^2 + (b^* - b^*_{\mathrm{avg}})^2}, \tag{16}
$$

where L^*, a^*, and b^* are the color components of the CIE 76 color system and variables with subindex avg are the average values from the whole display, is used to determine the color difference.

We compute the ΔE values of our measurements for variation when measurement is repeated multiple times at one region of the display for one panel and for variation of colors of multiple panels. The measurements for variation of one location in one display is made using the 12 5.7" displays. The average values are computed for each display separately. The variation of multiple panels is computed for 702 5.2" displays.

Fig. 5 shows the histograms of ΔE for red, green, and blue computed from our data, different panels and one panel one region, and the data for different regions in a panel obtained from Fig. 4 of [3]. The histograms show that our measurements are consistent with the results presented in [3]. Measurements from one panel at one region have smaller variation than when measuring from multiple regions and variation between different panels is larger than variation within one panel. The mean of ΔE between panels is 3.6, between regions of a panel 1.8 and in one region in one panel 0.3. Just noticeable difference of ΔE is about 2.3 [11], which means that the variations of one display in one region are not detectable by human eye. The variation between different parts of a display may be just noticeable and different uncalibrated displays of same type may have color differences that are well noticeable.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we developed an empirical model for tristimulus measurements from OLED displays. In our model, most of variation is aligned along $\begin{bmatrix} X & Y & Z \end{bmatrix}^T$ vector and the measurements from OLED displays support this. In [7] was proposed that all of the variation is along this direction. We analyzed the variations and did find there is variation in directions perpendicular to $\begin{bmatrix} X & Y & Z \end{bmatrix}^T$. We found out that the variations are almost linearly dependent on the sum
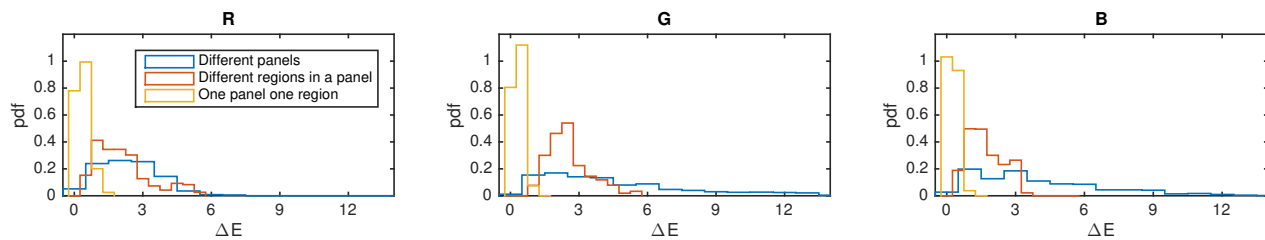
Fig. 5.   Histograms of ΔE obtained by measuring one region of one display multiple times, measuring different regions of one OLED display [3], and measuring different OLED displays of the same type

$X + Y + Z$ and the ratio between variation along $\begin{bmatrix} X & Y & Z \end{bmatrix}^T$ and directions perpendicular to it is approximately 5.

The variation of one display at one region is temporal in its nature, while most of the variation between displays is static so it cannot be explained by temporal type of noises, such as shot noise. However, results showed that the variation of tristimulus values measured from one display at different time instances is similar to variation between displays, but with a different scale. We compared to our results of amount of variation with variation between different regions of a display presented in [3]. Expectedly the variation measured from one display and one region was smaller and variation between displays was larger than the variation between regions of one display. In calibration, the within panel model can be interpreted to be the lower bound of calibration accuracy that can be achieved. However, it may be unfeasible to make the calibration separately for different parts of a display and thus the variation in different regions of a display is the lower bound in practice.

The proposed model is fitted using measurements so it covers all sources of noise that contributed to our measurement. OLEDs are current driven [12] and the source-drain current uniformity is a major factor in the short-term variations in OLED displays [13]. In [14] (as cited in [3]) variations in subpixel level radiance was determined to be caused by current fluctuations. In [12] it is stated that process variation in the electrical characteristics, such as threshold voltage, results in differences in the OLED current among pixels. Based on these references, we hypothesize that the variations are mostly caused by current. The variation of color, when measured from one location of a panel would be caused the fluctuation of current in time and the variation between displays would be caused variation of static properties of the electrical characteristics of the displays.

The proposed model is simple, does not require a deep knowledge of technical properties of the panel, and can be fitted using a relatively small sample size. Mmodels in Fig. 3 for 5.2" panels were almost the same with $n = 11$ and $n = 702$. The variations of some colors e.g. green may differ from the model, but our test results showed that still the proposed model improves accuracy over the uniform error model.

The developed model was tested in calibration of tristimulus

measurements from different displays. Results showed that the proposed model achieved better accuracy than a uniform error model. Due to the nonlinear nature of the OLED displays the linear calibration is not directly suitable for calibrating the input of the panel and the output, but we expect that using the proposed model with other calibration models [8], [9] would improve the calibration accuracy.

The measurements were made with two different types of panels, both from a same manufacturer. In future, the model should be verified with different types of panels. Because the model is simple and based only on measurements it is straightforward to build this kind of model and see whether other panel types exhibit similar properties.

REFERENCES

[1] G. Wyszecki and W. Stiles, *Color Science Concepts and Methods, Quantitative Data and Formulae*, 2nd ed. John Wiley and Sons, 1982.
[2] P. Boher *et al.*, "Optical characterization of OLED displays," *J. Soc. Inform. Display*, vol. 23, no. 9, pp. 429–437, 2015.
[3] ——, "Spectral imaging analysis of oled display light emission properties," in *Proc. 21st Int'l Disp. Worksh*, 2014.
[4] A. Badano, "Principles of cathode-ray tube and liquid crystal display devices," *Advances in digital radiography: RSNA categorical course in diagnostic radiology physics*, pp. 91–102, 2003.
[5] N. Ohta and A. R. Robertson, *Colorimetry : fundamentals and applications*, ser. Wiley-IS&T series in imaging science and technology. Chichester, GB, Hoboken, NJ, USA: John Wiley and Sons, 2005.
[6] M. Mahy *et al.*, "Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV," *Color Research & Application*, vol. 19, no. 2, pp. 105–121, 1994.
[7] J. Gardner, "Tristimulus colorimeter calibration matrix uncertainties," *Color Research & Application*, vol. 38, no. 4, pp. 251–258, 2013.
[8] M. Vrhel and H. Trussell, "Color device calibration: a mathematical formulation," *IEEE Trans. Image Process.*, vol. 8, no. 12, pp. 1796–1806, Dec 1999.
[9] S. Westland *et al.*, *Computational Colour Science Using Matlab*, 2nd ed. Wiley, 2004.
[10] Y. Ohno and J. E. Hardis, "Four-color matrix method for correction of tristimulus colorimeters," in *Color and Imaging Conference*, vol. 1997. Society for Imaging Science and Technology, 1997, pp. 301–305.
[11] M. Mahy *et al.*, "Evaluation of uniform color spaces developed after the adoption of cielab and cieluv," *Color Research & Application*, vol. 19, no. 2, pp. 105–121, 1994.
[12] C. L. Lin and Y. C. Chen, "A novel LTPS-TFT pixel circuit compensating for TFT threshold-voltage shift and OLED degradation for AMOLED," *IEEE Electron Device Lett.*, vol. 28, no. 2, pp. 129–131, Feb 2007.
[13] T. Tsujimura *et al.*, "4.1: A 20-inch OLED display driven by super-amorphous-silicon technology," *SID Symp. Dig. Technical Papers*, vol. 34, no. 1, pp. 6–9, 2003.
[14] D.-Y. Shin *et al.*, "Reducing image sticking in amoled displays with time-ratio gray scale by analog calibration," *J. Soc. Inform. Display*, vol. 17, no. 9, pp. 705–713, 2009.