Forward and inverse colour calibration models for OLED displays

***Maliha Ashraf*[1], *Dounia Hammou*[1], *Rafał K. Mantiuk*[1]**

[1]*Department of Computer Science and Technology, University of Cambridge, UK*

Abstract

We compare different methods of colour calibrating OLED 4-primary displays. The forward models use different colour transformation techniques to predict device-independent tristimulus values from device native RGB values. The inverse models transform the device-dependent tristimulus values to predict the native RGB values that could produce the required colour. We found that the performance of the models depended on the display characteristics and the models that performed better in terms of forward model error were not necessarily better for inverse model performance as well.

Background

OLED (Organic Light Emitting Diodes) displays are becoming increasingly popular because of their ability to achieve high luminances along with larger colour gamuts. To boost the luminance range and reduce power consumption, many OLED display manufacturers use an additional white primary to supplement the red, green, and blue channel responses. However, this makes OLED displays challenging to calibrate as the drive signals as the relation between input and emitted light becomes more complex. Moreover, OLED display manufacturers also face additional engineering constraints, such as limiting power consumption, limit on current density, pixel burn-in, which add additional complexity to the display's control and consequently its characterisation.

The existing OLED colour models in the current literature were evaluated in terms of the accuracy of the forward display model, which maps input display-encoded pixel values (e.g. sRGB or BT.2100) into colorimetric values of emitted light (XYZ or linear BT.2020 RGB) [1, 2, 3]. However, for most practical applications, we are interested in the inverse display model, which maps colorimetric values (which we want to display) into display-encoded pixel values. For that reason, in this paper, we evaluate all models in terms of the accuracy of the inverse display model, in addition to the forward display model, and also test the error introduced by computing the inverse of the model. It should be noted that display models for OLED are not directly invertible (bijective functions) because of the look-up tables, interpolation, and decision boundaries. For that reason, the performance of the inverse display model can be worse than that of the forward display model.

Display Measurements

To collect data for our model comparison, we measured the colour response of two OLED TVs: two 55-inch OLED TVs: LG OLED Evo G1 (firmware version: 03.34.80) and G2 (firmware version: 03.21.30). The Psychtoolbox software [4] running with MATLAB in a Windows 10 environment (Windows HDR enabled) was used for presenting the measurement patches with `HDR10` and 16-bit floating point precision per colour channel enabled. We collected the measurements using the same settings that were used for both displays to assure fairness in the comparison. The measurements were done in a dark room. For the colour settings, the native gamut was chosen. To avoid any unwanted alteration of the display outputs, all auto adjustment and enhancement settings were turned off. Additionally, both GSR (Global Sticky Reduction) and TPC (Temporal Peak Luminance Control) were disabled using the displays' service menus. These settings prevent image retention and pixel burn-in by dimming the brightness of the image. However, this is not desirable if we want to colour calibrate the displays accurately.

We collected 3 sets of measurements (XYZ) from each of the two OLED displays that we will call: i) ramps, ii) grid, and iii) test using the Konica Minolta CS-200 colorimeter. The measurements were done by displaying the colour in a rectangle in the middle of the screen. The test patches covered 5% of the total screen area and the rest of the screen was black.

Ramps The individual responses of each of the r, g, b, and w channels were measured using 120x4 HDR10 pixel values. These measurements are used to estimate the EOTF of the displays.

Grid A full 13x13x13 grid of RGB pixel values was measured for HDR10 pixel values. The grid measurements are used to optimise the transformation matrices and to compute 3D LUTs.

Test The performances of the models were tested using these measurements. We used the XYZ measurements of the colours of the X-Rite ColorChecker taken under Illuminant C from [5] and scaled them to five different luminance levels: 1, 4.7, 22, 106 and 500 cd/m^2 to generate a set of 120 colour coordinates and measured the display responses.

Models

Three sub-gamuts This model is based on the three-matrix method proposed in [3]. Bodner et al. divided the four-primary rgbw gamut into three-subgamuts rgw, rbw and gbw which assumes that three out of four primaries drive the display for any given colour depending on its location in the sub-gamuts, and then finds the correction matrices for reducing XYZ measurement errors between a reference spectroradiometer and test colorimeter. We use the same approach to estimate the transformation matrices between linear RGB pixel values and the corresponding RGB BT.2020 values. The transformation matrices were estimated using the grid measurements. The native RGB values were assigned to their respective sub-gamuts by finding the minimum of the three RGB linear values and the M_{RGW}, M_{RBW}, and M_{GBW} matrices were optimised by MATLAB `GlobalSearch` with CIE ΔE_{00} as the error function. We extended the model by adding an extra compensation step, in which a 3D LUT was used to correct residual errors.

PLCC-based compensation (PC) The PLCC-based compensation model is proposed in Tian et al. [2] in which the PLCC (Piece-wise Linear interpolation assuming Constant Chromaticity) [6] method is used to predict the intermediate XYZ values and then a 3D LUT is mapped to compensate for the remaining prediction errors.

Colour mixing In the model proposed by Sun and Luo [1], two transformation matrices are used — one for high and one for low chroma colours. The high chroma matrix (M_{rgb}) is estimated using the XYZ values of the rgb primaries. The low chroma matrix (M_{gray}) is estimated by scaling (M_{rgb}) to match the response of the w primary. The final XYZ values are the weighted sum of the XYZ values from the high and low chroma matrices.

Polynomial regression We tested the 3rd order polynomial regression model proposed in Sun and Luo [1]. The XYZ values are predicted by a 14x3 matrix with the polynomial terms: $(R, G, B, R^2, G^2, B^2, R^3, G^3, B^3, RG, GB, BR, RGB, 1)$. The coefficients of the polynomial matrix are optimised using MATLAB `GlobalSearch` with CIE ΔE_{00} as the error function to be minimised. Our 13x13x13 grid measurements are used for this optimisation.

RGBW gamut We tested a model which works on a simple assumption of independent RGBW primaries where linear combinations of each of the four pixel responses produce the combined colour response. The final display colour output is a non-linear function of the four sub-pixel responses. We tested the model to see whether this simple assumption could compete with the more sophisticated models outlined above. The coefficients of the XYZ to RGBW linear transformation matrix are optimised with CIE ΔE_{00} as the minimising function.

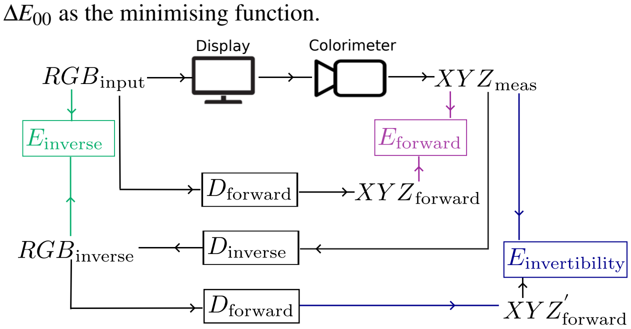

Figure 1. *Model validation flowchart. $D_{forward}$ represents the forward models and $D_{inverse}$ represents the inverse models. $E_{forward}$ is the forward model error, $E_{inverse}$ the reverse model error, and $E_{invertibility}$ is the invertibility test error*

Validation

We use 3 metrics to test the performance of the different models. All the testing was done using the 120 *test* measurements data set. The inputs of the three error metrics are shown in Figure 1.

Forward model The forward model error (E_{forward}) is the CIE ΔE_{00} colour difference between the output of the display forward model and the measured XYZ colour values. The XYZ values are converted to $L^*a^*b^*$ values with the highest value of the w primary as the white point.

Inverse model The inverse model error (E_{inverse}) is the difference between the PQ-encoded RGB values in the native colourspace of the display. The measured XYZ values are transformed to the native RGB pixel values via the respective inverse models.

Invertibility test For this test, we pass the XYZ measurements through the inverse model to predict the native RGB values and then pass those values through the forward model to predict the XYZ values again. This test tells us the extent of the invertibility of the model. A perfectly invertible model should give the error of 0. However, due to the use of interpolation, look-up tables, and decision boundaries, such error is non-zero for more complex models.

Results

Figure 2 shows the results of our three error metrics for the G1 and G2 displays. In terms of forward model performance, the 3-gamut model and the PLCC-based compensation model (PC) produce the smallest CIE ΔE_{00} for both G1 and G2 displays. The forward model errors for both the models are comparable for the G1 display, but PC shows a slightly lower error distribution than the 3-gamut model for the G2 display. We believe that the source of performance differences between the two displays is the difference in their internal EOTF curves.

The Colour Mixing (CM) model showed the smallest mean and median inverse model errors for the G1 display. For G2 display, the 3-gamut model along with the CM model showed the best performances. All the models show quite large maximum errors as compared to the forward model errors.

The results from the invertibility test show that the performances of the models are almost reversed for the two displays. The models that performed better for G1 are the worst performing ones for G2. Note that in this model, we apply both the inverse and forward models successively and this is not a measure of accurate colour transformation but of the invertibility of the models. The Polynomial and RGBW-gamut method use linear transformation matrices and show high invertibility for G1, but the opposite for G2. This could be because G2 (the more recent model) uses its white primary in a more non-linear way compared to G1.

Conclusions

We compared the performances of different state-of-the-art OLED colour calibration models for both forward and reverse transformations. We recommend the use of 3-gamut or PLCC-based compensation model method for the forward display char-

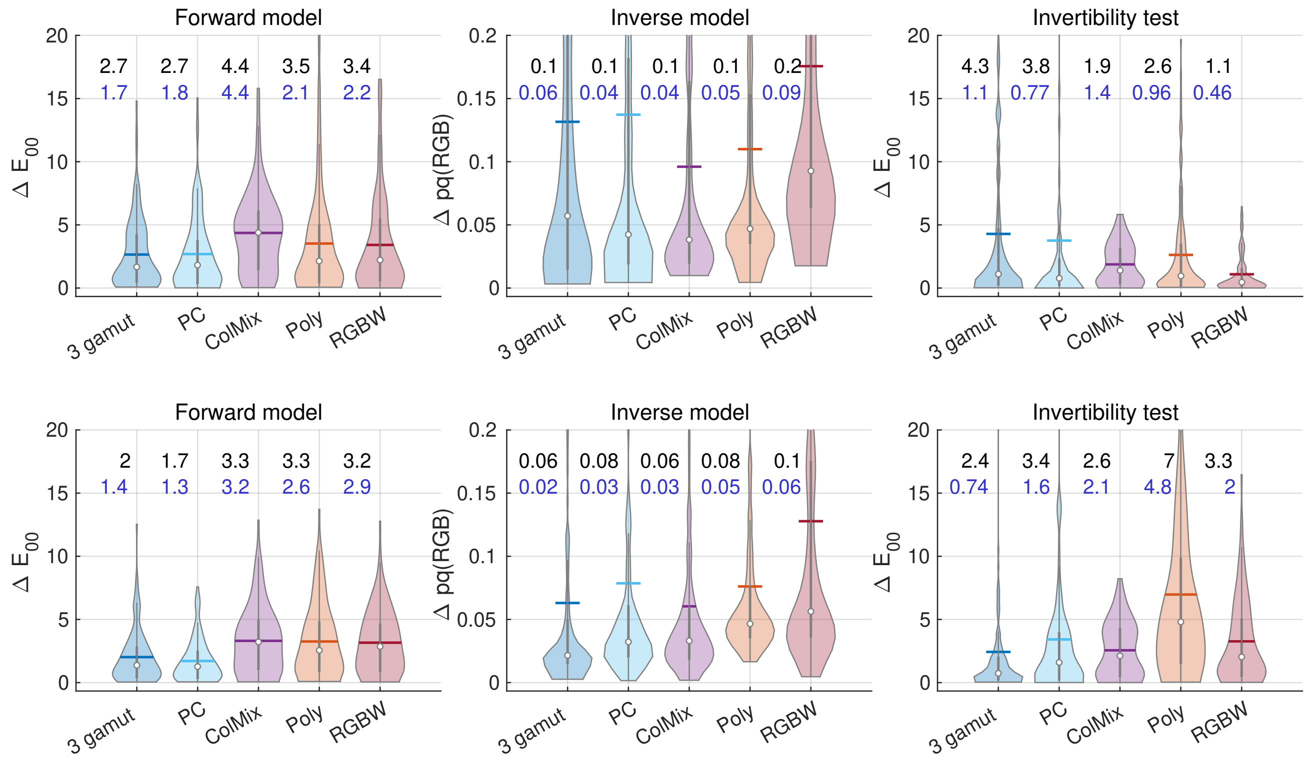

Figure 2. *The distribution of errors for the G1 (top) and G2 display (bottom). The models are: 3-gamut, PLCC-based compensation (PC), Colour mixing (ColMix), Polynomial regression (Poly) and RGBW gamut. The white circle in each violin plots are the medians and the horizontal dash are the means. The black text at the top left of each violin is the mean, the blue text are the medians.*

acterisation and either 3-gamut or Colour Mixing model depending on the luminance response characteristics of the display for inverse display characterisation. The comparisons demonstrate the need for further work in developing better colour characterisation models for OLED displays.

Acknowledgements

This research received funding from Meta.

References

[1] P.-L. Sun and R. M. Luo, “Color Characterization Models for OLED Displays,” *SID Symposium Digest of Technical Papers*, vol. 44, no. 1, pp. 1453–1456, 2013.

[2] D. Tian, L. Xu, and M. R. Luo, “The Characterization of HDR OLED Display,” *Electronic Imaging*, vol. 31, pp. 326–1–326–6, Jan. 2019.

[3] B. Bodner and N. Robinson, “83-3: Calibration of Colorimeters for RGBW Displays,” *SID Symposium Digest of Technical Papers*, vol. 50, no. 1, pp. 1206–1209, 2019.

[4] M. Kleiner, D. Brainard, and D. Pelli, “What’s new in psychtoolbox-3?,” 2007.

[5] “ColorChecker (Macbeth) Chart.” https://poynton.ca/notes/color/GretagMacbeth-ColorChecker.html.

[6] D. L. Post and C. S. Calhoun, “An evaluation of methods for producing desired colors on CRT monitors,” *Color Research & Application*, vol. 14, no. 4, pp. 172–186, 1989.