https://doi.org/10.2352/issn.2694-118X.2021.LIM-97

A study of neural network-based LCD display characterization

Joan Prats-Climent; Instituto de Instrumentación para Imagen Molecular, Universitat Politècnica de València; Valencia, Spain
Luis Gómez-Robledo; Departamento de Óptica, Universidad de Granada; Granada, Spain
Rafael Huertas; Departamento de Óptica, Universidad de Granada; Granada, Spain
Sergio García-Nieto; Instituto de Automática e Informática Industrial, Universitat Politècnica de València; Valencia, Spain
María José Rodríguez-Álvarez; Instituto de Instrumentación para Imagen Molecular, Universitat Politècnica de València; Valencia, Spain
Samuel Morillas; Instituto Universitario de Matemática Pura y Aplicada, Universitat Politècnica de València; Valencia, Spain

Abstract

In this paper we study up to what extent neural networks can be used to accurately characterize LCD displays. Using a programmable colorimeter we have taken extensive measures for a DELL Ultrasharp UP2516D *to define training and testing data sets that are used, in turn, to train and validate two neural networks: one of them using tristimulus values, XYZ, as inputs and the other one color coordinates, xyY . Both networks have the same layer structure which has been experimentally determined. The errors from both models, in terms of ΔE_{00} color difference, are analysed from a colorimetric point of view and interpreted in order to understand how both networks have learned and how is their performance in comparison with other classical models. As we will see, the comparison is in average in favor of the proposed models but it is not better in all cases and regions of the color space.*

Introduction

Display characterization has been an important topic in the field of color imaging for years, gaining interest with the recent availability increase of a variety of display technologies. Having a precise display characterization allows accurate image reproduction. This is important not only from the consumer point of view in order to optimize viewing experience, but, also, it is critical for any color imaging application, where accurate image reproduction is paramount, and for vision science research to precisely control the stimuli shown to observers when carrying out psychophysics.

In general, display characterization means to create a model able to relate device dependent *DACRGB* inputs with display outputs expressed in an appropriate device independent color space (usually tristimulus values XYZ or xyY color coordinates). A colorimeter, or even better a spectrophotometer or spectroradiometer, is used to measure outputs related to a set of *DACRGB* inputs. Some years ago, usually, the measurements needed to be taken manually. This led to the development of display models based on their physical behavior and that could be built using just a few measurements, which was convenient. Examples of these models are those in [1, 2] that are based on representing the nonlinear behaviour of inputs with some power function or look-up-table and then relate the linearized inputs with the outputs through a simple linear transformation (and so assuming constant chromaticity for them). An alternative model presented [3, 4] is based on using measurements in xyY coordinates and processing separately the Y component with a power function of the inputs and the xy chromaticity using linear interpolation between measurements. On the other hand, mathematical models have been also developed. One of the classical ones is based on trying to find the best linear application able to relate inputs (or some function of them like powers or square roots) and outputs (or some function of them) [5]. To find this linear application an overdetermined equation system is formulated using the measurements taken and the best solution is found by squared error minimization using Penrose pseudo-inverse [6], which is the reason why this method is named the Pseudo-Inverse method.

Later, the availability of programmable measurement devices has eased the access to more information about inputs-outputs relationship by measurements. This has been helpful to increase the accuracy of the physical models but it is more interesting for the mathematical models as they only use the measurements available and no more knowledge about the display. For instance, in the Pseudo-Inverse model, the more data are put into the overdetermined system, the higher the potential accuracy is (if the data is appropriately chosen). Furthermore, this also opens the possibility of using machine learning approaches to address the display characterization problem as big amounts of measurements are available.

In this paper we explore this latter possibility by building two neural network models trained to compute what *DACRGB* inputs need to be used to obtain a desired output expressed in either XYZ or xyY coordinates. This choice is the only difference between the two networks trained.

In the following section we detail the networks structure and the procedure for training them. Section 3 shows the obtained results and the comparison with state-of-the-art models and Section 4 presents some conclusions and future work.

Neural networks training

In this study we have used a LCD display *DELL Ultrasharp UP2516D 25"* and a *X-rite eye-one display* colorimeter. To build the training set we have divided each *DACRGB* input range into 22 equally spaced levels and we consider all possible combinations of them, which account for $22^3 = 10648$ different inputs so that they form an equally spaced 3D mesh of the input space. Also, to take into account the measurement device error, we have measured each input three times, which means that we have a total amount of $10648 \cdot 3 = 31944$ data in the training set. Specifically, we measured the full 3D mesh once after waiting for 2h the display to warm up and immediately after we measured two more time the set. On the other hand, for the testing dataset, we have taken 21 input levels placed right in between the 22 levels of the training set and consider all their combinations of $21^3 = 9261$.

We focus the training on addressing the problem of finding out what *DACRGB* inputs need to be set for the display to show a desired color specified in XYZ or xyY coordinates. This is known as the inverse model for the display (the direct model aims to predict what output will be obtained for a given input). Using tristimulus values has the advantage of being closely re-

Model	Average ΔE_{00}	Standard dev.
RIT [1]	5.1	1.8
Pseudo-Inverse [5]	4.3	2.0
NN_{XYZ}	2.6	1.2
NN_{xyY}	4.2	1.5

Table 1: Average and standard deviation ΔE_{00} error per model

lated to the *DACRGB* inputs, while the *xyY* coordinates separate the chromatic information *xy* from the luminance information *Y* and so divides the problem in two. Both options have potential learning benefits so we decided to study both.

For the sake of fair comparison, we have used the same structure for both neural networks. As the structure should be set related to the complexity of the whole problem rather than inputs format, we think a common appropriate structure should not limit the accuracy of either network. Both networks have been defined and trained using libraries available in Python [7, 8]. After extensive experimental simulations we concluded that a structure with two hidden layers of 256 and 64 units each, respectively, was used. This structure is depicted in Figure 1. In each unit a Rectified Linear Unit (ReLU) activation function is used. For training we used Adam optimizer [9] with a learning rate of 10^{-3} and the Mean Squared Error (MSE) as loss function. As commented above, we train two different networks using the same training data set one of them with XYZ inputs and the other with *xyY* ones, which we name NN_{XYZ} and NN_{xyY}, respectively. NN_{XYZ} needed 300 epochs to converge while NN_{xyY} convergence was found after 500 epochs. Batch size was 64 in both cases.

Results and discussion

We provided the desired colors in the testing set to each trained network to compute the corresponding input *RGB*. Then, we have measured with the *X-rite eye-one display* the real color obtained for the computed *RGB* inputs and computed the perceptual error using the ΔE_{00} formula [10, 11]. An analogous procedure has been followed to assess the performance of two state-of-the-art methods which we compare our proposals with: RIT model [1, 2] and Pseudo-Inverse model [5]. For this latter model we consider the following linear function to compute each one of the *RGB* inputs:

$$F_k(X,Y,Z) = a_k\sqrt{X} + b_k\sqrt{Y} + c_k\sqrt{Z} + d_k\sqrt{X}\sqrt{Y} + e_k\sqrt{X}\sqrt{Z} + f_k\sqrt{Y}\sqrt{Z} + g_kXY + h_kXZ + i_kYZ, \quad (1)$$

where $k = R,G,B$ and $a_k,b_k,c_k,d_k,e_k,f_k,g_k,h_k,i_k$ are the parameters to compute by least-squares minimization for each k using the training dataset.

We have to take into account that device measurement error is about $\Delta E_{00} = 0.1$ on average, which we have computed by looking at the ΔE_{00} between the three measurements taken for each *DACRGB* input in the training set.

Table 1 shows the average error for the testing dataset for each of the models in the comparison along with its standard deviation. We can see that both figures of merit are in favor of the NN_{XYZ} network. This better performance can also be seen in Figure 2, which shows histograms for the ΔE_{00} errors for each model. NN_{xyY} and Pseudo-Inverse show a similar average error but the latter has a higher standard deviation, which means a higher variability for different color regions, as it is expected for a linear model. The RIT model shows a slightly worse average performance but with a not too large standard deviation.

It is important, specially when using machine learning approaches, to analyze the performance in detail in order to fully understand the behaviour of the different methods. We approach this analysis from a colorimetric point of view. To do so, we have analyzed ΔE_{00} errors in terms of luminance and chrominance for each model using *xyY* representations. In Figure 3 we show plots of ΔE_{00} versus luminance *Y* and chrominance *xy* for each of the models in the comparison. By closely looking at this plots we can point out the following:

- NN_{XYZ} has a better average performance for low *Y* than when *Y* gets higher, where we can see that the model can make both low and high errors, most probably depending on chromaticity. In terms of the latter, we see the error is worse for medium *x* and *y* (achromatic colors), and a bit higher for high *y* and low *x* (greenish colors). On the other hand, performance is quite consistent in the rest of the chromaticities and specially good for high *x* and low *y* (purplish colors). As a consequence, we can see that the network performs worse in the sides and vertices of the 3D training mesh. This is due to the mesh having less data in these regions and the network trying to minimize the average error tends to optimize performance in the core of the data set.
- NN_{xyY} and RIT model perform quite similar when analysing the errors for different luminances and chromaticities: both perform better when *Y* is high. That is explained for the RIT model because in these cases the constant chromaticity assumption holds better than for lower *Y*. Also, both models perform better for achromatic colors and when *x* is high (reddish), *y* is high (greenish), both *x* and *y* are medium-high (yelowish) or both *x* and *y* are low (bluish). For the RIT model, same explanation about chromaticity constancy holds. For the NN_{xyY} it seems that the learning process has maken it perform in an analogous way and higher luminance colors are better managed.
- The Pseudo-Inverse model also works better for higher *Y* but with a higher variability than the rest. In terms of chrominance, we see a variety of areas with very different behaviour, which means that the model only adjust well in some regions of the color space and most depending on chromaticity regions that are better represented by the correlations included in the model in Eq. 1.

Overall, we can conclude that the networks are better in learning $XYZ - DACRGB$ relations than $xyY - DACRGB$. The former makes the network learn better in general. We think this is becasuse of a more direct individual correlation between inputs and outputs. Conversely, it means that $xyY - DACRGB$ relation is more complex to learn and maybe the network architecture and/or the training set need to be extended for the training. Also, this latter option has some advantages with respect to the learning acquired by the network as, for instance, when processing high luminance colors. Moreover, in the NN_{XYZ} case it is easy to see how important is the design of the training dataset and the influence it has on performance. This suggests also that increasing the number or points in the regions of lower performance can be beneficial.

Conclusions

In this paper we have studied the use of neural networks for LCD display characterization. Using a programmable colorimeter we have taken extensive measures for a *DELL Ultrasharp UP2516D* to define training and testing data sets that are used, in turn, to train and validate two neural networks: one of them using tristimulus XYZ inputs and the other xyY color coordinates. Both networks have the same layer structure which has been experimentally determined. Average performance in

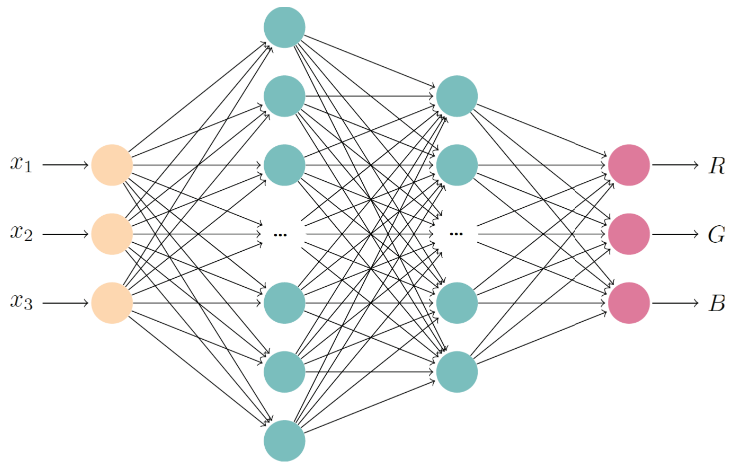

Figure 1. *Neural network structure*

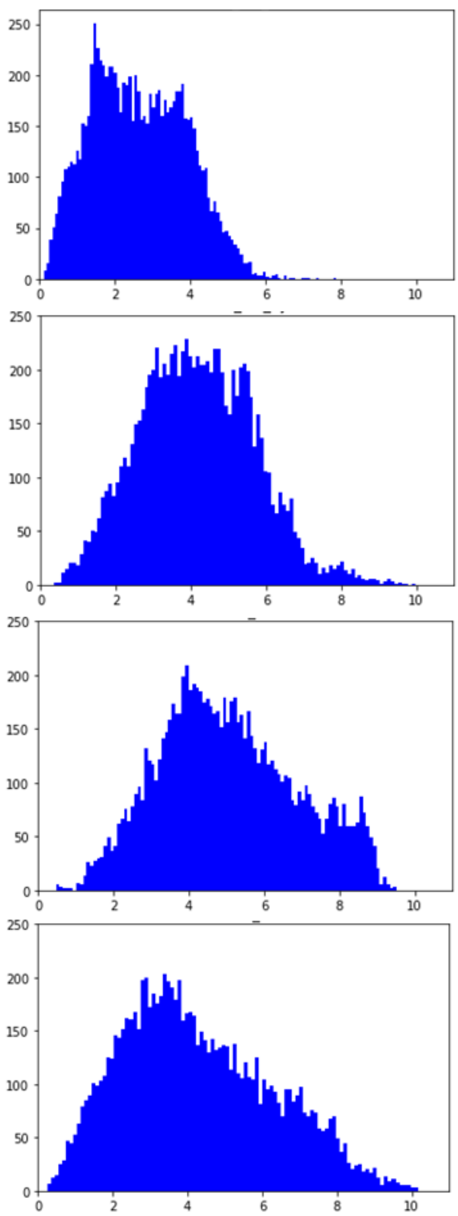

Figure 2. *ΔE_{00} error histograms for: (a) NN_{XYZ}, (b) NN_{xyY}, (c) RIT physical model, (d) Pseudo-Inverse data model.*

terms of ΔE_{00} color difference favors the neural network trained with XYZ inputs in front of the network trained with xyY data and two state-of-the-art methods. A more detailed analysis reveals that performance is far from being better in all cases for the XYZ trained network and suggests some ways to improve performance. Furthermore, we point out that the networks are better in learning $XYZ-RGB$ relations than $xyY-RGB$. The former has a more direct individual correlation between inputs and outputs.

Future interesting work could include extending the study to other LCD display models and other display technologies.

Acknowledgments

The authors acknowledge the support of Generalitat Valenciana under grant AICO2020-136, Spanish Ministry of Science under grants FIS2017-89258-P and PID2019-107790RB-C22, and thank Professor Michael J. Murdoch for providing the code for the RIT model.

References

[1] M. Fairchild, D. Wyble, Colorimetric characterization of the Apple studio display (Flat panel LCD), RIT 1998: http://scholarworks.rit.edu/article/920/

[2] Berns, Methods for Characterizing CRT Displays, Displays (1996): https://doi.org/10.1016/0141-9382(96)01011-6

[3] J. Malo, M.J. Luque. ColorLab: the Matlab toolbox for Colorimetry and Color Vision. Univ. Valencia 2002, http://isp.uv.es/code/visioncolor/colorlab.html

[4] P. Capilla, M. Diez, M.J. Luque, J. Malo, Corresponding-pair procedure: a new approach to simulation of dichromatic color perception. JOSA A 21(2): 176-186 (2004)

[5] Westland, S., Ripamonti, C., & Cheung, V. (2012). Computational colour science using MATLAB. John Wiley & Sons.

[6] R. Penrose, On best approximate solution of linear matrix equations, Proc. Camb. Philos. Soc. 52, 17–19 (1956)

[7] Pedregosa F. et al., 'Scikit-learn: Machine Learning in Python', J. Mach. Learn. Res., vol. 12, pp. 2825-2830, 2011

[8] Chollet F. et al. 'Keras', GitHub. Retrieved from https://github.com/fchollet/keras (2015).

[9] D. P. Kingma, J. L. Ba, 'Adam: A Method for Stochastic Optimization',Proceedings on the 3rd International Conference for Learning Representations, 2015. https://arxiv.org/abs/1412.6980

[10] CIE Publication 142-2001. Improvement to industrial colour-difference evaluation. CIE Central Bureau, Vienna, 2001.

[11] G. Sharma, and W. Wu, The CIEDE2000 color-difference formula:

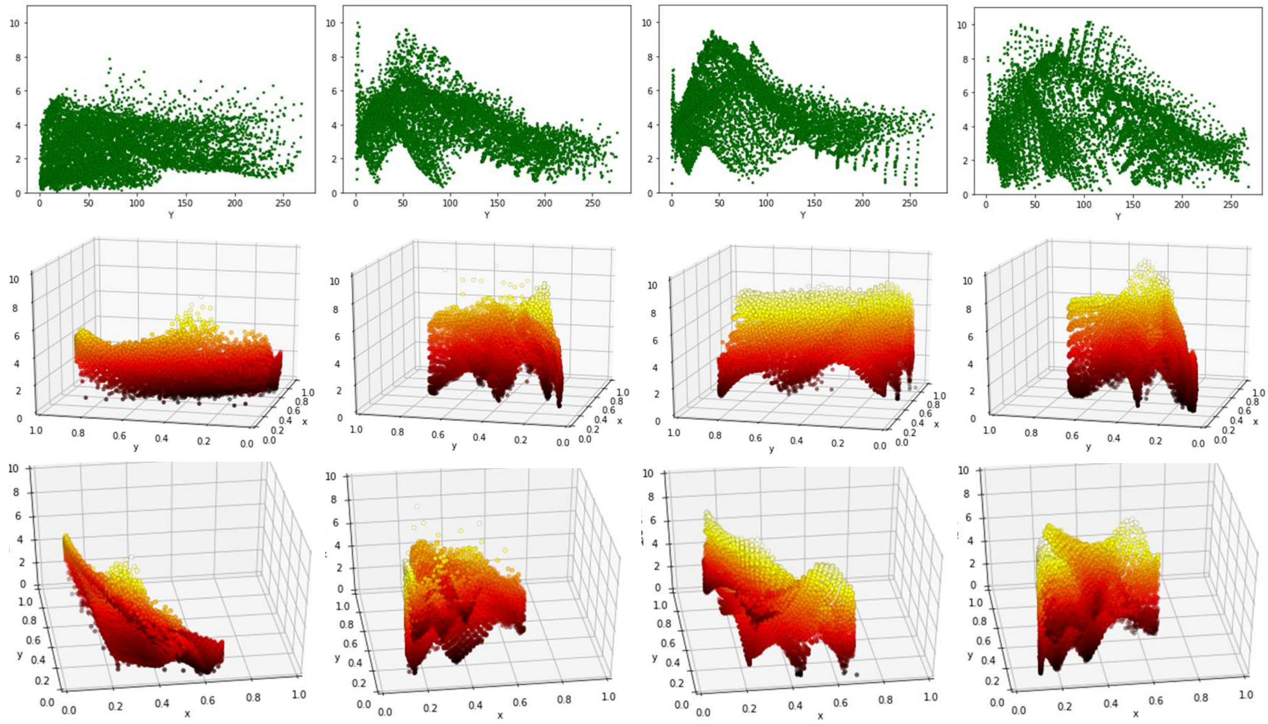

Figure 3. *ΔE_{00} error analysis plots. First row: ΔE_{00} versus Y; Second and third row: ΔE_{00} versus xy; First column: NN_{XYZ} inputs; Second: NN_{xyY} inputs; Third: RIT physical model; Fourth: Pseudo-Inverse data model.*

Implementation notes, supplementary test data, and mathematical observations, Color Research and Applications 30 1 21-30 (2015)

Author Biography

Joan Prats Climent received his BSc in Physics (2017) and his MSc in Advanced Physics with specialization on Theoretical Physics (2018) from the University of València. Since 2018 he is working as research technician at Instituto de Instrumentación para Imagen Molecular (i3M) of Universitat Politècnica de València. His work is focused on research of machine learning methods and its applications to medical image reconstruction in positron emission tomography.

Luis Gómez-Robledo received his PhD in physics from Granada University (2010). He studied Physics (2006) and Optometry (2009). His research has been focused in color measurement and colour vision. At this moment he is working in color perception of colorblind people, measurement of thresholds and its application for screen simulations.

Rafael Huertas earned a BSc in Physics in 1997 and a BSc in Optometry in 1998 form the University of Granada. In 2004 he received his PhD in Physics from UGR. Since 2011 he is an Associate Professor at the Department of Optics of UGR. His research interests include basic and applied colorimetry.

Sergio García-Nieto received his BSc (2004) and PhD (2010) degrees in control and automation from Universitat Politècnica de València. Since 2007, he has been an associate professor at the Systems Engineering and Control Department at UPV. His research interests cover non-linear modeling, predictive control, fuzzy logic, robust control and genetic algorithms.

María José Rodríguez-Alvarez received her PhD in Physics from Universidad de València, Spain. She is currently a full professor in the Applied Math Department at Universitat Politècnica de València. Her research interest includes medical image reconstruction, image processing, and deep learning.

Samuel Morillas received a degree on Computer Science by Universidad de Granada (2002) and a PhD on Applied Math from Universitat Politècnica de València (2007). Since 2019 he is a full professor in the Applied Math Department at UPV leading research on fuzzy logic systems and their applications, image processing, color science and vision science.