Predict LCDs' Real-World Color Performance Based on Generic Image Statistics and Gamut Mapping Rules

Pei-Li Sun, Dept. of Information Management, Shih Hsin University, Taiwan;
Chia-Yun Lee, Dept. of Graphic Communications & Digital Publishing, Shih Hsin University, Taiwan

Abstract

We proposed a patch-based metric to predict LCDs' real-world color performance. This metric takes 125 RGB patches for display and measurement. Generic image statistics, tone characteristics, weighted deltaE summary and the correlation to the general gamut mapping rules are taken into account in our metric. A visual test has been done for using its results to optimize the proposed metric. The metric is capable of predicting LCDs' color fidelity on 9 image types individually. It would be useful for LCD manufacturers to improve their produces.

Introduction

The color fidelity of LCDs was commonly estimated by standard color patches. However, the correlations between the patches and real-word color images are uncertain. Therefore, simply measuring the mean color differences between the displayed color patches and their standard values could be misleading. Previous studies on color gamut mapping suggested that the best transformations were typically image-type dependent.[1] If we regard the color variations of un-calibrated LCDs as a "standard-to-un-calibrated" LCD gamut mapping problem, the performance should be depended on image types as well. However, the conventional patch-based estimations were unable to report a display's color performances based on image types. Thus, the LCD manufacturers miss a great opportunity to make good use of the information to improve their products.

The aim of this study is to derive a metric to predict the real-world color performance of a LCD. The basic idea is to use RGB based 125 (5x5x5) uniform patches together with generic image statistics and tone characteristics to predict LCDs' overall performance on various image types. To this end, we have to simulate a calibrated LCD and several un-calibrated LCDs for optimizing weighted ΔE metrics. The whole process for the optimization was conducted by means of the following steps: (1) calibrating a LCD monitor, (2) collecting image-type-dependent color statistics based on the 125 color coordinates, (3) choosing different classes of images as references, (4) manipulating their colors intentionally to simulate the color variations of various un-calibrated LCDs, (5) asking observers to scale the image differences (ΔV) between calibrated reference images and their color-perturbed counterparts, (6) simulating the color errors on the 125 patches and then measuring the 125 sets of ΔE . (7) optimizing the weighted ΔE metrics to predict the ΔV based on the ΔE s, generic image statistics and the correlation to generic gamut mapping rules. The details will be introduced in the following sections.

Generic Image Statistics

Image Type

Color statistics for 9 image types were collected. The first type of images, Portrait, were obtained by scanning Japanese vogue magazines with a GretagMacbeth ProfileMaker Pro (Version 4) calibrated UMAX Astra 2500 scanner. The reason of scanning printed portrait rather than shooting real persons by digital camera is that we think preferred reproduction is more important than accurate reproduction for skin tones.

Four classes of images, Scenery, Night Scenes, Buildings, Flowers-and-plants, were selected from a series of royalty free image bank CDs. Originally, we tried to summarize image statistics of six categories, including Portrait, Scenery, Night Scenes, Buildings, Flowers-and-plants, and Computer Graphic (CG). We collected 1024 CG image from the 3D image gallery of http://www.3dshop.com. The 1024 images in its Fantastic category are more correlated to computer game and unrealistic images. However, we realized that the CG images must be subdivided due to the color variations across the type of images are too great. Therefore, we selected the mean and standard deviation of L* and C*, pixel frequencies of Red, Yellow, Green and Blue domains (separated by 45, 135, 225 and 315 hue angles with chroma no less than 15) as variables. Converting the 8 statistics of 1024 CG images into z scale, then performed K-means classifier1 to group four classes of CG images. The dominated colors of the four classes, notated as CG1, CG2, CG3 and CG4, were gray, blue, red-and-yellow and black respectively (see Fig.1). To save the computation cost, image statistics were firstly summarized as 51x51x51 bins 3D histograms. The histograms quantized images' L*/a*/b* values in 2/4/4 unit intervals respectively. Each image normalized its color pixel frequencies into probability before putting into the histograms (so the total number is 1). After averaging the values of 3D histograms for each type of images, finally we obtained nine 51x51x51 bins 3D histograms for nine image types respectively.

***Figure 1**. Illustrate the 9 image types used in the study. First row, left to right: Portrait, Scenery, Night Scenes, Buildings, Flowers-and-plants. Second row: CG1, CG2, CG3 and CG4.*

Color Target

The proposed study aims to predict the overall performance of a display based on weighted color differences measured from 125 color patches. We used a 15" Compaq FP5315 LCD as our test device. The device was firstly calibrated using GretagMacbeth Eye-One Pro spectroradiometer to approximate the sRGB standard (with D65 white point and 2.2 gamma).[2] According to its 24-bit RGB signals, we divided each 8-bit channel into 6-bit interval (i.e., [0 63 127 191 255]) to use all combinations to produce our 5x5x5 RGB patches. These colors were then converted into LAB space via GOG model.[3] The accuracy of forward GOG model was estimated by 3x3x3 evenly distributed RGB patches, and the resulted mean and maximum CIEDE2000 were 0.86 and 2.92 respectively. The characterization errors were acceptable for our application (i.e., comparing large color differences). We finally calculated the pixel probability (denoted as P) within 20 ΔE distances over the 51x51x51 bins 3D histograms to the corresponding (L*,a*,b*) coordinates of the 5x5x5 RGB patches. The image-type-dependent data P will be used for optimizing the metric later.

Visual Assessment

Image Preparation

As previously mentioned, images were classified into 9 types and we selected 6 images from each type as reference images. Therefore, we have 54 reference images (= 9 types x 6 images) in total. Each reference image was manipulated by 7 different transformations (see Fig.2) including: Lightness Gamma Function (LGam), Ligtness Sigmodial Function (LSigm), Chroma Gamma Function (CGam), Lightness & Chroma Gamma Functions (LCGam), Hue Shift (HShift), Ambient Flare (AFlare) and Dark Clipping (DClip). They all are common color defects in display applications. Where LGam, CGam and LCGam are relevant to gamma corrections, LSigm varies image contrast, HShift happened when 3 color primaries mis-balance, AFlare when black matrix is unable to remove the screen's surface reflections, and DClip (or gray inversion) happened when liquid crystal in LCD twisted to wrong directions. As we have 6 reference images per image type, 3 different manipulation levels were applied to the 6 reference images.

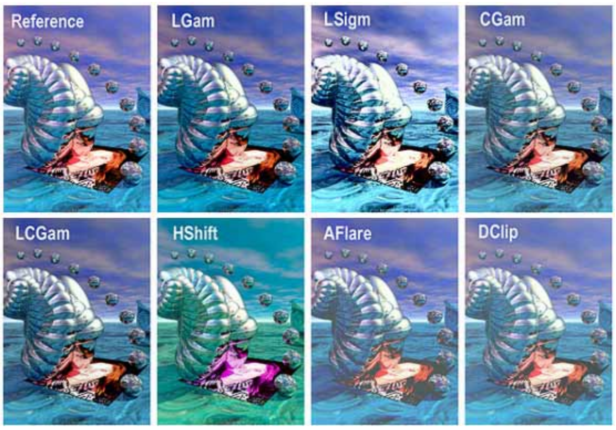

Figure 2. Color manipulation to simulate 7 types of color defects.

Psychophysical Experiment

After completing the image preparation, a psychophysical experiment was conducted to assess visual image differences between reference images to their color-perturbed counterparts (see Fig.3). A window program written by MathWork Matlab (Version 6.5) was designed for the visual assessment. The assessment was performed under dark viewing condition. 18 observers were asked to stay in the dark laboratory at least 3 minuses before starting the assessment. A pair of images including a reference image on left side and its manipulated counterpart on right side was displayed on the previous mentioned Compaq LCD in random order. A series of radio buttons indicated the image differences from level 1 to level 20 was shown on the window for choosing. Because each reference image has 7 corresponding counterparts, every observer has to evaluate 378 pairs (= 9 types x 6 ref. images x 7 manipulations) of image differences.

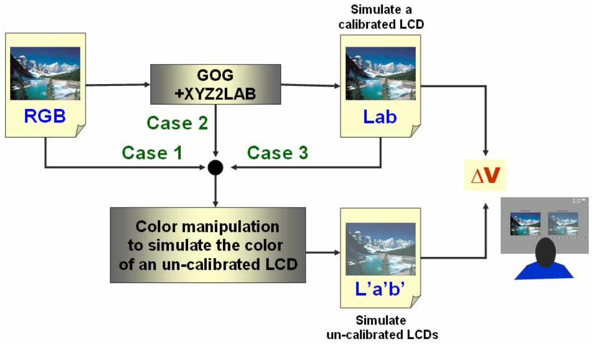

Figure 3: Color manipulation to simulate calibrated and un-calibrated LCDs. Where DClip manipulated colors in RGB space (Case 1), AFlare in XYZ space (Case 2) and LGam, LSim, CGam, LCGam and HShift in Lab space (Case 3).

Optimizing Weighted Metrics

Metric Performance

The color coordinates (L*,a*,b*) of the original 5x5x5 color patches were transferred using the 7 manipulation functions with the corresponding level indicated in the last section. Therefore each manipulated image has one corresponding 5x5x5 color patches. Because the aim of this study is to predict overall visual differences of a certain type of images, we must apply the statistics of the group of images, not the statistics of the manipulated image. We applied classified image statistics to color differences from original 5x5x5 color patches to its manipulated patches for predicting the visual differences from its corresponding reference image to its manipulated image.

The mean responses from the 18 observers were regarded as visual differences of each pair of images. The mean values are denoted as ΔV. The performance of a weighted ΔE function is determined by the coefficient of variation (CV) proposed by Alder et al.[4] Lower CV value indicates a good fitting from ΔV s to weighted ΔE s.

ΔE Metrics

ΔEab, $\Delta E94$ and $\Delta E00$ are three major color differences formulae for industrial applications. Therefore, we must compare their performances. There are 125 color differences (ΔE s) between two sets of color patches (calibrated and un-calibrated simulations). A single index would be welcome for indicating its performance. We initially reported the mean and 95 percentile of the 125 set ΔE s for comparison. As can be seen in Table 1, $\Delta E94$ is superior to ΔEab and $\Delta E00$, and the 95 percentile of the errors are more correlated to visual results compared to the mean errors.

Table 1. The overall performance of ΔE metrics.

CV	ΔEab	$\Delta E94$	$\Delta E00$
mean	67.0	61.3	67.6
95th perc.	67.1	42.7	51.0

Pixel Probability

The pixel probability, P, indicates the color distributions of a certain type of images. The fitting for $\Delta E94$ can be improved by applying a logarithmic P statistics (referring to Eqn.(1)) especially for CG4 images (see Table 2). We also tried $\log_e$, $\log_2$ and normalized P. But Eqn.(1) is the best in our tests.

$$\Delta E_{wt} = \Delta E_{94} \cdot \left(1+\log_{10}\left(1+10^5 P\right)\right) \quad (1)$$

Error Summarization

Mean and maximum ΔE s were commonly used for reporting color performance of image devices. However, referring to Table 2, we found the summation of median and maximum ΔEwt for Eqn.(1) is more correlated to visual image differences. The results are similar to previous findings.[6] We also found that applied a power of 0.7 to the weighted ΔEwt would enhance the CV performance a bit.

Table 2. The performance of different metrics on 9 image types.

CV	Mean	95th Perc.	median + max	Eqn.3	final metric
Portrait	43.3	33.6	31.1	25.5	20.3
Scenery	55.1	43.1	37.3	34.2	24.8
Night	58.7	40.7	32.7	35.9	22.5
Building	50.1	37.9	35.1	31.8	25.2
Flowers	47.1	40.8	38.4	33.0	26.5
CG1	50.9	41.7	38.6	31.7	32.1
CG2	38.1	33.8	33.7	32.5	34.5
CG3	44.7	35.3	33.0	33.0	32.4
CG4	70.2	54.8	46.3	42.1	33.9
overall	50.9	40.2	36.2	33.3	28.0

Local Color Contrast

Tone characteristics also play an important role while reproducing images. To make a better prediction on visual color fidelity, the tone characteristics for both calibrated and un-calibrated LCDs were also estimated by the 125 patches to check if the continuity and uniformity of the two devices are similar. Referring to Fig.4, the 125 colors construct 4x4x4 sub-cubes in RGB space. Each cube can be defined relatively using 3 vectors along R, G and B axes respectively. The corresponding 3 vectors in LAB space tell us the local color contrast. The length of each vector was calculated by $\Delta E94$ for correlating visual color differences. Other than the 4x4x4x3 vectors for sub-cubes, we still need 4x4x2x3 vectors to describe the local contrast of the rest parts of the gamut surface. The lengths of a RGB vector are identical for both calibrated and un-calibrated LCDs, but the $\Delta E94$ of the corresponding vector pair could be different. If the difference is considerable, we can say the two LCDs have different local color contrast. The similarity of global tone characteristics denoted as S was measured by averaging the above mentioned local differences (Eqn.2). Eqn.3 combines the above features and enhances overall CV performance to 33.3 (referring to Table 2).

$$S = \sum \left| \Delta E_{94(un-calibratedLCD)} - \Delta E_{94(calibratedLCD)} \right| \quad (2)$$

$$\Delta E_{combo} = median(\Delta E_{wt}) + max(\Delta E_{wt}) + 3.3 \cdot S \quad (3)$$

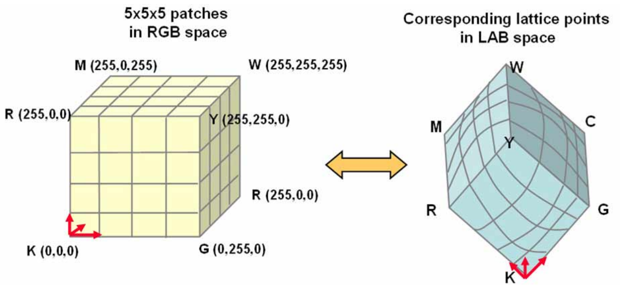

Figure 4. 3 vectors to link the color coordinates of neighbor sample patches.

Gamut Mapping Rules

Color gamut mapping refers to the transformation of an image by mapping its colors to fit the gamut of a destination medium is a hot issue that is being intensively studied. Various kinds of gamut mapping algorithms (GMAs) have been proposed in recent years [6]. General rules could be extracted by observing their visual results. For instants, color clipping is acceptable in high chroma regions [7] but is unpleased in dark regions [8]. Sigmoidal curves are favorite when compressing an image's lightness [9]. The tolerance of hue shift is limited [10] and the directions of color mapping should point to the gamut center [11]. If the LCD's "standard-to-uncalibrated" color variations follow the above rules, we could lower the final metric values as the transformations are potentially better than some other cases.

Gamut-Mapped Color Fitting

SKNEE gamut mapping algorithm (GMA) [9] first compresses lightness using a sigmoidal function and then compresses towards the hue cusp using a piece-wise-linear knee function. The SKNEE algorithm follows not only above rules but also been tested by many independent authorities as a recommended model for gamut compression. We regarded the calibrated and the un-calibrated LCD gamut boundaries as source and destination gamuts respectively. The boundaries can be constructed by their 125 patches. We first transferred the 125 patch colors from the calibrated LCD to the destination gamut using the SKNEE algorithm (image independent version), and then modified the ΔEwt using Eqn.4. Where Lmin is the minimum lightness of the un-calibrated LCD gamut and ΔG is the $\Delta E94$ color difference between color-mapped calibrated patch and its corresponding un-calibrated patch.

$$\Delta E'_{wt} = \left[1 - \frac{L_{\min}}{200}\left(1 - \frac{\Delta G}{50}\right)\right] \Delta E_{wt} \quad (4)$$

More Weights for Core Gamut Colors

Most of gamut mapping studies suggested that the color accuracy of core gamut [12] is more critical than that of near gamut boundary colors. Therefore, we gave different weights to Eqn. 2 for calculating the variations of local color contrast. We found that gave a weighting value of 2 for core gamut (i.e., all vectors comprised central 2x2x2 sub-cubes) and 0.5 for all the rest will enhance the CV performance. Adding the above two features to Eqn.3 as the final metric, the overall CV was reduced to 28.0 (referring to Table 2).

Applications

Figure 5 shows the workflow of our proposed LCD color quality assessment. First, converting the 125 RGB patch values into (L*,a*,b*) space and regarding them as calibrated colors. Second, displaying the 125 patch on an un-calibrated LCD and then measuring their (L*,a*,b*) values. Third, calculating the weighted local differences S based on the 4x4x4 cube structure and calculating the 125 sets weighted ΔE_{94} (refer to Eqn.4) using classified image statistics P. Finally, using the combo function illustrated in Eqn.3 to predict the real-world color performance of the un-calibrated LCD.

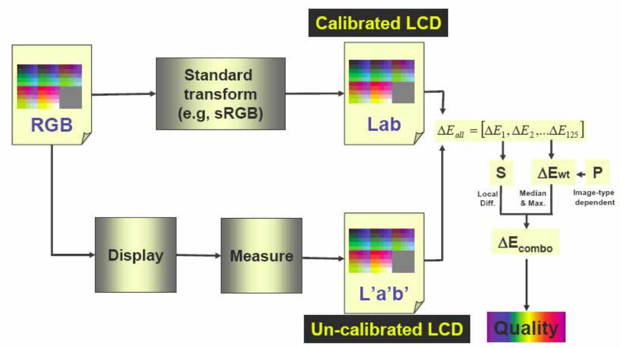

Figure 5. *The workflow of the proposed LCD color quality assessment.*

Conclusions

The objective of this study is to provide a patch-based metric to predict LCDs' real-world color performance. Generic image statistics, tone characteristics, weighted deltaE summary and the correlation to the general gamut mapping rules are taken into account in our metric. A visual test has been done for using its results to optimize the proposed metric. The metric performances have been double from simple mean measurement (overall CV=67.0) to our proposed final metric (overall CV=28.0). The metric is capable of predicting LCDs' color fidelity on 9 image types individually. It would be useful for LCD manufacturers to improve their produces. The same concept also can be used for estimating other types of display and printer media.

References

[1] J. Morovic and M. R. Luo, Verification of Gamut Mapping Algorithms in CIECAM97s Using Various Printed Media, Proceedings of the 6th IS&T/SID Color Imaging Conference, 53-56, (1998).

[2] IEC 61966-2-1: Colour Measurement and Management in Multimedia Systems and Equipment - Part 2-1: Default RGB Colour Space – sRGB, (1998).

[3] CIE, The Relationship between Digital and Colorimetric Data for Computer Controlled CRT Displays, Publication CIE 122-1996, Bureau Central de la CIE, (1998).

[4] C. Alder, K. P. Chaing, T. F. Chang, E. Coates, A. A. Khalili and B. Rigg, Uniform Chromaticity Scales –New Experimental Data, Journal of Society of Dyers and Colourists, 98:14-20, (1982).

[5] J. Morovic and P. L. Sun, Predicting Image Differences in Colour Reproduction from Their Colorimetric Correlates, Journal of Imaging Science and Technology, IS&T, Vol. 47(6), 509-516, (2003).

[6] J. Morovic and M. R. Luo, The Fundamentals of Gamut Mapping: A Survey, J. Imaging Science and Technology, vol. 45, no. 3, pp. 283-290, (2001).

[7] F. Ebner and M. D. Fairchild, Gamut Mapping from Below: Finding the Minimum Perceptual Distances for Colors Outside the Gamut Volume, Colour Research and Application, 22(6): 402–413, (1997).

[8] Hoshino T. and Berns R. S. (1993) Color Gamut Mapping Techniques for Color Hard Copy Images, SPIE Proceedings, 1909: 152-164.

[9] G. J. Braun and M. D. Fairchild, Image Lightness Rescaling Using Sigmoidal Contrast Enhancement Functions, Journal of Electronic Imaging, 8(4): 380–393, (1999).

[10] M. C. Stone, W. B. Cowan and J. C. Beaty, Color Gamut Mapping and the Printing of Color Images, ACM Transactions. on Graphics, 7(4): 246-292, (1988).

[11] L. W. MacDonald, Gamut Mapping in Perceptual Colour Space, Proceedings of the 1st IS&T/SID Color Imaging Conference, 193-196, (1993).

[12] L. W. MacDonald, J. Morovic and K. Xiao, A Topographic Gamut Compression Algorithm, Journal of Imaging Science and Technology, 46(1): 53-61, (2002).

Author Biography

Pei-Li Sun received his BS and MSc degrees from the Department of Graphic Communications of the Chinese Culture University in Taipei, Taiwan. He pursued his Ph.D. at the Colour & Imaging Institute of the University of Derby in England. At present he is working as an assistant professor at the Department of Information Management of Shih Hsin University, Taipei, Taiwan. His current interests include image quality, color management and image processing.

.