

# A non–device-specific approach to display characterization based on linear, nonlinear, and hybrid search algorithms

Graduate School of Human and Environmental Studies,  
Kyoto University, Kyoto, Japan  
School of Psychology,  
University of Birmingham, Birmingham, UK  
Japan Society for the Promotion of Science, Tokyo, Japan

**Hiroshi Ban**



**Hiroki Yamamoto**

Graduate School of Human and Environmental Studies,  
Kyoto University, Kyoto, Japan



In almost all of the recent vision experiments, stimuli are controlled via computers and presented on display devices such as cathode ray tubes (CRTs). Display characterization is a necessary procedure for such computer-aided vision experiments. The standard display characterization called “gamma correction” and the following linear color transformation procedure are established for CRT displays and widely used in the current vision science field. However, the standard two-step procedure is based on the internal model of CRT display devices, and there is no guarantee as to whether the method is applicable to the other types of display devices such as liquid crystal display and digital light processing. We therefore tested the applicability of the standard method to these kinds of new devices and found that the standard method was not valid for these new devices. To overcome this problem, we provide several novel approaches for vision experiments to characterize display devices, based on linear, nonlinear, and hybrid search algorithms. These approaches never assume any internal models of display devices and will therefore be applicable to any display type. The evaluations and comparisons of chromaticity estimation accuracies based on these new methods with those of the standard procedure proved that our proposed methods largely improved the calibration efficiencies for non-CRT devices. Our proposed methods, together with the standard one, have been implemented in a MATLAB-based integrated graphical user interface software named Mcalibrator2. This software can enhance the accuracy of vision experiments and enable more efficient display characterization procedures. The software is now available publicly for free.

## Introduction

Display calibration and characterization are essential parts of experimental procedures in vision science because almost all current experiments are conducted in computer-aided environments: Visual stimuli are manipulated via a programming language and displayed on a computer display, and the observer’s responses are acquired by pressing keys connected to a computer. To ensure that visual stimuli—their luminance, color, timing, and so forth—are presented precisely in such computer-based experiments, researchers need to characterize display devices accurately in advance of actual experiments.

Cathode ray tube (CRT) displays are currently the most widely used devices for vision experiments, and the calibration procedures to characterize their luminance and chromaticities are well established with a two-stage procedure: gamma correction followed by a linear color transformation (Ban, Yamamoto, & Ejima, 2006; Berns, 1996; Brainard, Pelli, & Robson, 2002). The calibration results obtained through this standard two-step procedure have been tested (Ban, Yamamoto, & Ejima, 2006; Brainard et al., 2002), and the quality of luminance and chromatic stimuli on CRT displays satisfies the researchers’ criterion.

However, non-CRT devices, such as liquid crystal display (LCD), have come into the mainstream recently, and researchers are required to use non-CRT over CRT devices. This is because few companies still manufacture CRTs, and it has become increasingly difficult to obtain them. Furthermore, experimental

Citation: Ban, H., & Yamamoto, H. (2013). A non–device-specific approach to display characterization based on linear, nonlinear, and hybrid search algorithms. *Journal of Vision*, 13(6):20, 1–26, <http://www.journalofvision.org/content/13/6/20>, doi:10.1167/13.6.20.

setups have become more complex these days. Researchers have to select different types of display devices for different experiments. In some cases, researchers have no choice but to use the display devices already equipped and mounted in the experiment rooms. For example, researchers sometimes need to present stimuli via LCD or digital light processing (DLP) projectors to a translucent screen mounted in the shield room for functional imaging experiments.

These changes surrounding display devices could severely affect the quality of visual experiments because the current widely used standard display calibration method is established based on the internal model of CRT devices (Berns, 1996; Brainard et al., 2002) and may not be appropriate for non-CRT devices. Our previous study demonstrated that the standard method is not applicable to some LCD and DLP display devices (Ban, Yamamoto, & Ejima, 2006). We therefore proposed a new display calibration procedure based on recursive least-square estimations (Ban, Yamamoto, & Ejima, 2006) and showed its applicability to non-CRT devices in human functional imaging experiments (Ban, Yamamoto, Fukunaga, et al., 2006). In the present study, we further refined and advanced our proposed display characterization procedures applicable to non-CRT as well as CRT display devices.

The present study is especially focused on developing fairly quick and efficient methods for finding display inputs that produce specific prespecified luminance or chromaticity output values. Our new methods use a data-driven gamma-correction procedure combined with a linear/nonlinear (Nelder-Mead Simplex; Dennis & Woods, 1987; Nelder & Mead, 1965) hybrid or line search (Powell's method with Coggins constrain; Brent, 1973; Farhi, 2011; Farhi, Debab, & Willendrup, 2013; Powell, 1964; Press, Teukolsky, Vetterling, & Flannery, 2007) algorithm to get the optimal RGB video input values to produce the required luminance and chromaticities. The methods are relatively model free, assuming only (a) a monotonic increment of luminance against the increment of video input values and (b) a piecewise linearity of the system in the initial estimation step. Several pros and cons of our new methods compared with the standard procedures are summarized below. More details of the procedures will be described later.

#### Pros:

1. These new methods have a much broader application because they do not presume the internal model of the display device and can handle nonlinearity of the device. Therefore, our methods are suitable to calibrate non-CRT devices such as LCD, DLP, and even future display devices such as

the organic-EL (Electro-Luminescence) and Laser displays.

2. These methods are most useful for experiments with visual stimuli that contain only a relatively small number of luminance and chromaticity.
3. These methods are considerably robust against noise in the measured data since they estimate chromaticities based on only a small limited color spaces.
4. These methods can achieve fast and accurate calibrations of target chromaticities within one to five repetitions of measurements (1-2 min for estimating a single chromaticity value).
5. Even without an assumption of monotonicity in the display gamma functions, these methods can theoretically estimate valid video input values to display required chromaticities. This is because our methods adopt goal-seeking algorithms within a small limited color space.

#### Cons:

1. These methods cannot produce chromaticity values successively in real time, whereas standard procedures can estimate the required chromaticities very quickly based on simple linear transformations, once a global color transformation matrix has been acquired.
2. These methods cannot model RGB-phosphor cross-talks, though this is also true for the standard calibration procedure.
3. These methods are unable to model quantization effects due to bit-depth, though this is also true for the standard calibration procedure.
4. Though our methods are flexible, they may not be able to model DLP projectors when they use a RGBW, not RGB, color filter wheel.

Furthermore, all of the procedures are integrated into a graphical user interface (GUI)-based display characterization software written in MATLAB (The Mathworks, Inc., Natick, MA) language and termed "Mcalibrator2" (Figure 1). The applicability and efficiency of our software to a wide range of display devices, including LCD and DLP types of devices, were confirmed by comparing the calibration accuracies of our procedures with that of the standard two-stage method. We found that all of these new approaches improved calibration accuracies for non-CRT devices compared with the standard display characterization procedure.

In the next section, we will briefly describe the standard two-step color calibration method widely used in the current vision experiments with CRT display devices (Figure 2). We will then propose our methods and compare the efficiencies of them with those of the standard method.

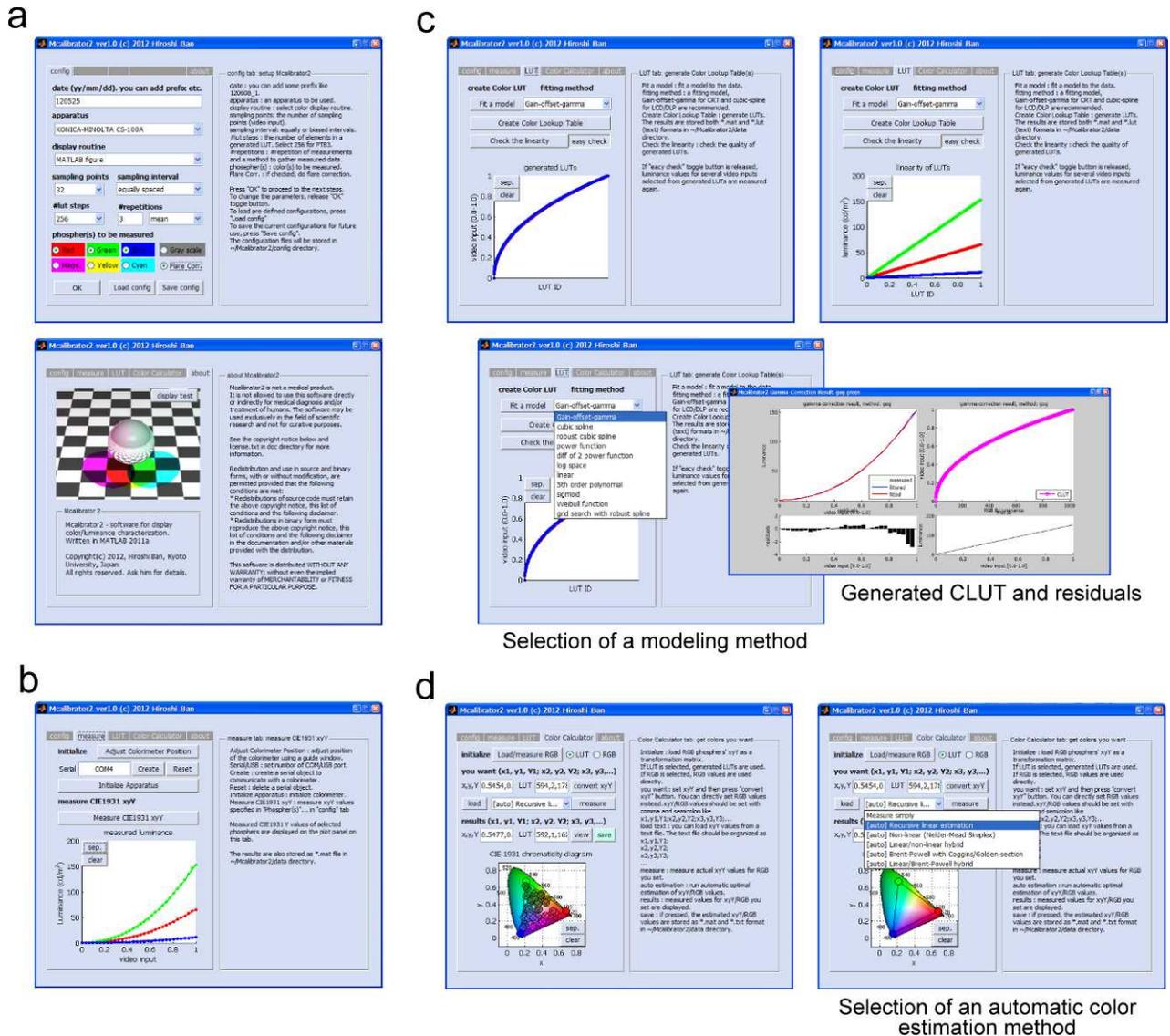


Figure 1. MATLAB-integrated display characterization software—Mcalibrator2. Mcalibrator2 consists of several tab-based GUI windows and buttons. All of the tabs are aligned along the display characterization procedures. (a) Parameter configuration tab. Calibration parameters are set through this window. (b) Measure tab. A colorimeter can be initialized and CIE1931 xyY are measured through this tab. (c) LUT tab. Color lookup tables are generated through this window. Researchers can select several fit options other than the methods described in this article. The goodness of fits and residuals are output in separate windows. (d) Color calculator tab. RGB video input values needed to display the target chromaticity are estimated based on this tab. Researchers can access all of the procedures described in this article and even customize the procedures as they like.

### Standard gamma correction: Gain-offset-gamma and gain-offset-gamma-offset models

The standard color display calibration method first tries to linearize the relationship between video input values of RGB phosphors and luminance outputs. To achieve this, we need to measure luminance values for the several video input values (generally 32 points are enough for an eight-bit phosphor (256 steps; Brainard et al., 2002) separately for each of the RGB phosphors

(Figure 2a). This method models the intact relationship between video inputs and the modeled luminance output values by Equation 1 (Figure 2b). This function is termed the “gain-offset-gamma” (GOG) model (Berns, 1996; Day, 2002).

$$L(x) = \begin{cases} \left( gain \times \frac{x - x_0}{1 - x_0} + offset \right)^\gamma, & x \geq x_0 \\ 0, & x < x_0 \end{cases} \quad (1)$$

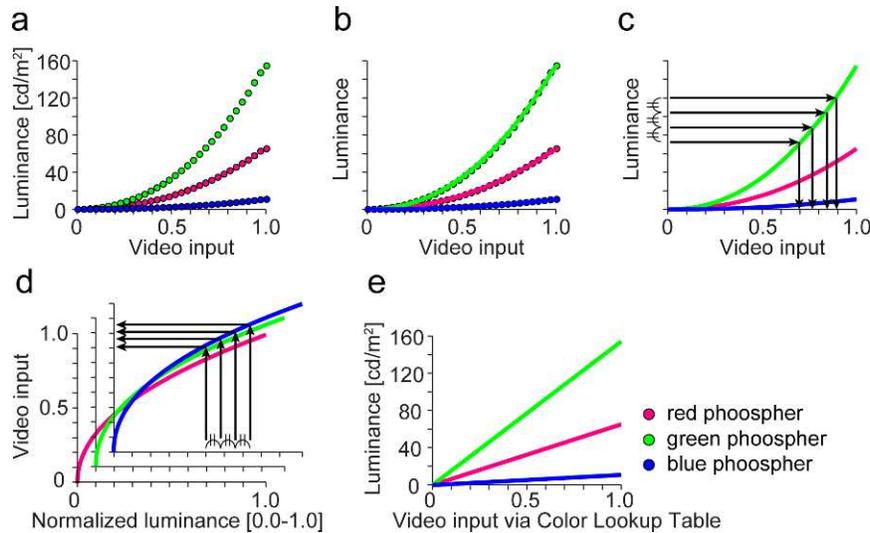


Figure 2. Standard luminance gamma-correction procedure. (a) In the standard gamma correction, display input/output properties are first characterized by measuring luminance outputs for several video input values. (b) Then, the measured points are interpolated by the GOG model (see the Introduction section) and the input/output relationships are modeled by exponential functions. (c) What we need to do after the interpolations is to get RGB video input values so that the luminance outputs against these input values are linearized. (d) To this end, the inverse functions of the exponential fits are calculated and video input values to get linearized output luminance are saved to color lookup tables (CLUTs). (e) The display input/output relationship is linearized by displaying chromaticities via the generated CLUTs.

Here,  $L$  is the modeled luminance, *gain* and *offset* are constant variables to scale the magnitude and the section of the fitting curve,  $x$  is video input value ( $0 \leq x \leq 1$ , digital),  $x_0$  is the starting point from which luminance is above zero, and  $\gamma$  describes a nonlinear form of the typical gamma function. We then calculate the inverse function  $x = f^{-1}(y)$  of the fit for each of the RGB phosphors and get the adjusted RGB video input values as color lookup tables (CLUT) so that luminance increments follow a linear function against the adjusted video inputs (Figure 2c-e).

To characterize display luminance output more precisely, this GOG model can be extended to the gain-offset-gamma-offset (GOGO) model in which display flare and glare are more explicitly modeled in the formula described below (Besuijen, 2005; Deguchi & Katoh, 1998; Katoh, 2002).

$$L(x) = \begin{cases} (L_{max} - offset) \\ \times \left( gain \times \frac{x - x_0}{1 - x_0} + 1 - gain \right)^\gamma \\ + offset, & x \geq x_0 \\ 0, & x < x_0 \end{cases} \quad (2)$$

Here,  $L_{max}$  is maximum luminance value of the target phosphor.

## Standard linear color transformation

By applying linear transformations with CLUTs, we can estimate the required RGB video inputs to produce target chromaticities. To perform the estimation, we first need to measure tristimulus values ( $XYZ$ ) for maximum RGB video inputs and create an array consisting of these values, such as,

$$pXYZ = [rXYZ, gXYZ, bXYZ] \quad (3)$$

Here,  $rXYZ$ ,  $gXYZ$ , and  $bXYZ$  are row vectors of  $XYZ$  values of each of the RGB phosphors, and  $pXYZ$  is a  $3 \times 3$  matrix. Here, we call this  $pXYZ$  transformation matrix a “global” transformation matrix to differentiate our “local” linear transformation procedures described later. Then, the tristimulus values  $XYZ$  for the  $rgb$  video inputs are calculated as a linear sum of these values by

$$XYZ = pXYZ \cdot rgb \quad (4)$$

From Equation 2, the required  $rgb$  video input values to produce the desired chromaticity ( $xyY$ ) can be acquired by

$$rgb = pYXZ^{-1} \cdot xyY_{to\_XYZ}(xyY) \quad (5)$$

Here,  $xyY_{to\_XYZ}(\ )$  is a function to convert CIE1931  $xyY$  values to tristimulus values,  $XYZ$ , and defined as

$$XYZ = \left[ Y \times \frac{x}{y}, \quad Y, \quad Y \times \left( \frac{1-x-y}{y} \right) \right]^T \quad (6)$$

where,  $x$ ,  $y$ , and  $Y$  correspond to CIE1931  $xyY$  values and  $T$  is the transposition of the matrix. We subtract zero-level  $XYZ$  values (flare) in advance of calculation if the flare cannot be ignored (Figure 3; Equation 2).

This linear transformation procedure after gamma correction can generally estimate required chromaticities precisely for CRT devices (Ban, Yamamoto, & Ejima, 2006; Berns, 1996; Day, 2002). The two-step display calibration procedure, however, was established following the internal mechanisms of CRT display devices (Berns, 1996; Brainard et al., 2002). Specifically, the procedures assume that (a) the relationship between video inputs and output luminance can be described clearly by Equation 1 or 2, (b) the device can be described as a linear system, (c) color constancy is held against video inputs (Figure 3), and (d) there is no interaction between RGB phosphors (that is, they are independent of each other).

In contrast, there is no evidence that other types of display devices also follow the same mechanisms with CRTs. We have actually reported that the standard two-step procedure is not applicable to some LCD display devices (Ban, Yamamoto, & Ejima, 2006). We therefore propose alternative display characterization procedures suitable for a wide range of devices. The procedures consist of data-driven gamma correction and direct color search algorithms. The entire procedure is illustrated as a flowchart in Figure 4a and b, with sample MATLAB codes presented in Figure 5. In the following, we will describe the details of this procedure.

### Cubic spline–based gamma correction

To make more accurate and efficient chromaticity estimations, the present study adopted a data-driven gamma-correction approach in the first step. Specifically, we characterized the relationship between video inputs and the corresponding luminance outputs based on a modified version of the cubic spline interpolation technique. The standard GOG or GOGO model can realize fairly accurate linearization of input/output relationships of CRT displays. However, it is also well known that display gamma does not always follow an extended power function, especially for non-CRT displays. Some vision experiment software (e.g., the Psychtoolbox MATLAB tools; Brainard, 1997; Pelli, 1997) therefore allows choice of many different functional forms (Weibull function, etc.) and also contains a more data-driven gamma-correction approach using a cubic-spline interpolation.

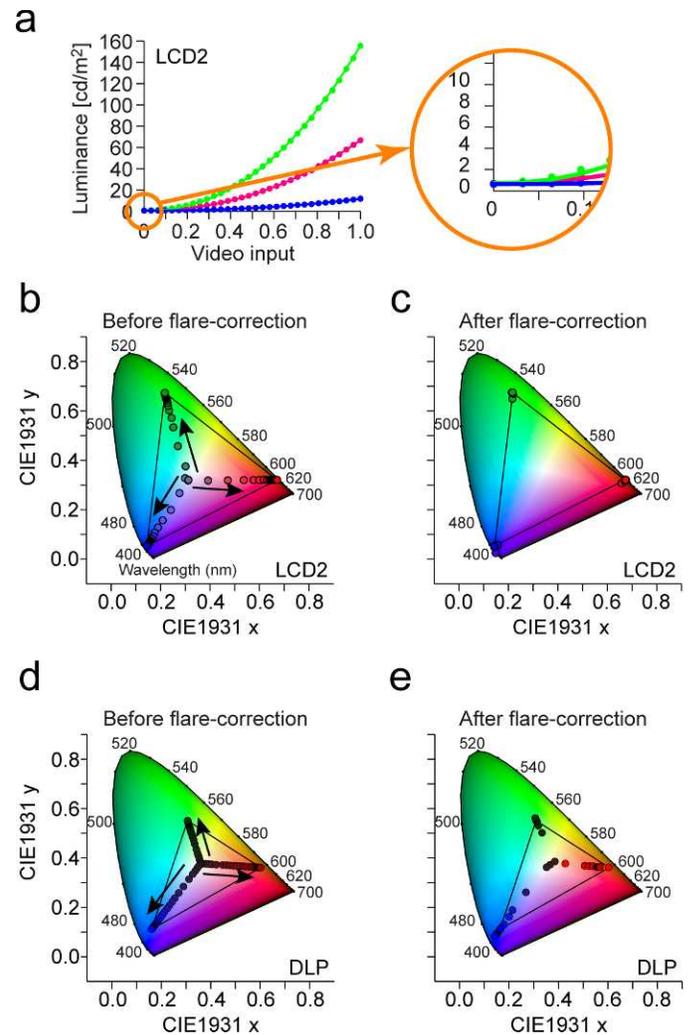


Figure 3. Flare correction and color constancy against the video input values. (a) In some display devices, leaking light at the zero-level video input (flare) can be observed, although displays should be completely black in principle. We need to subtract the flares in advance for such display devices. (b) Color constancy of the tested LCD without flare correction. When the flare effect is not ignorable, the constancy rule is broken. Then, the standard simple linear color transformation can not be applicable. (c) After the flare correction, color constancies were preserved for LCD. (d) Color constancy of the tested DLP. Inconsistency is larger than the LCD. (e) For DLP, the constancy is not realized even after the flare correction, which means that some nonlinear property is included in this display device. We need to calibrate such display devices in a different way to take the nonlinearity into account.

The present study also applied a cubic-spline approach because it can be potentially applicable to a wider range of display devices, including non-CRTs, as it interpolates data without any assumption of the system. However, a standard cubic spline interpolation has the potential problem of overfitting when it is

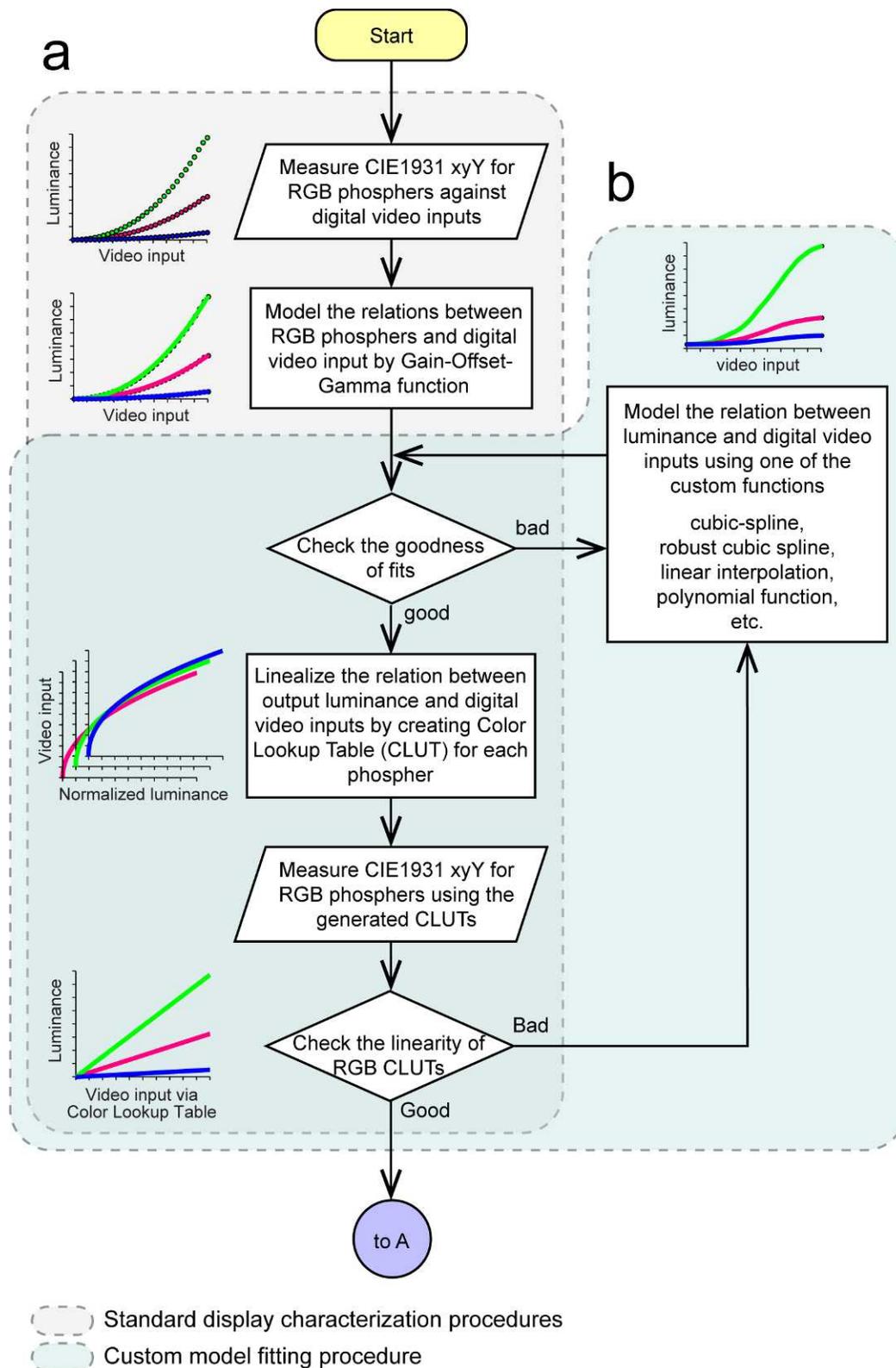


Figure 4a. Flowchart of our data-driven gamma-correction methods. By preparing several methods (some of them follow data-driven fitting procedures) to describe the luminance profiles of the target displays, Mcalibrator2 can perform more accurate gamma correction for a wide range of display devices. (a) A standard gamma-correction procedure. (b) Our customized gamma-correction procedure can be used when the standard procedure cannot describe input/output relationships of the target displays. For details, see the corresponding section of the Introduction (continued in Figure 4b).

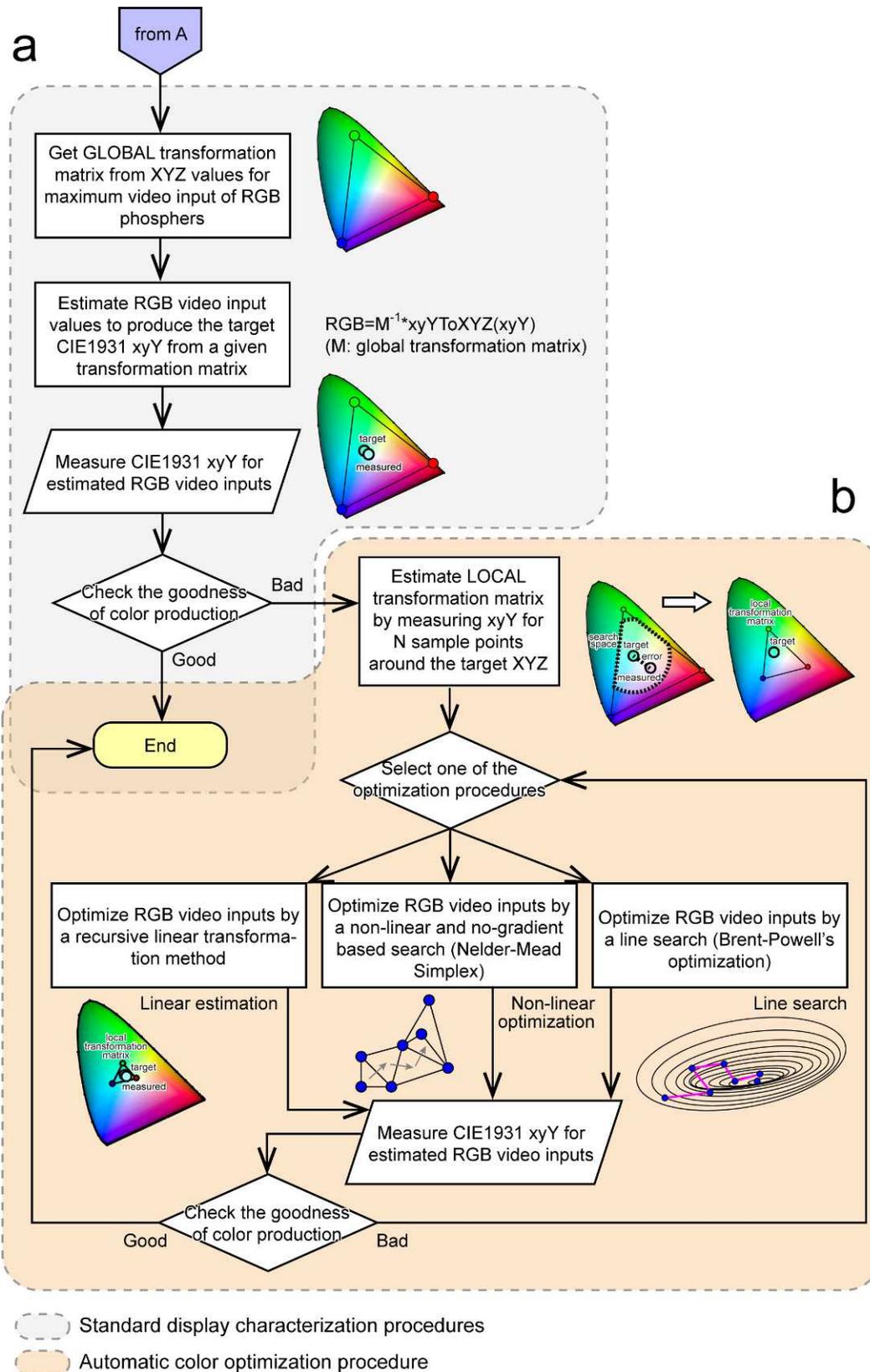


Figure 4b. Flowchart of our data-driven chromaticity calibration procedures (continued from Figure 4a). Our proposed color estimation procedures are implemented in Mcalibrator2 so that they work as optionally when the results of the standard procedure are not good. (a) A standard global color estimation procedure. (b) Our goal-seeking color reproduction procedures. The actual fine-scale estimations are performed by goal-seeking approaches in which the final values are estimated recursively following the standard numerical optimization procedures. For details, see each section of the Introduction.

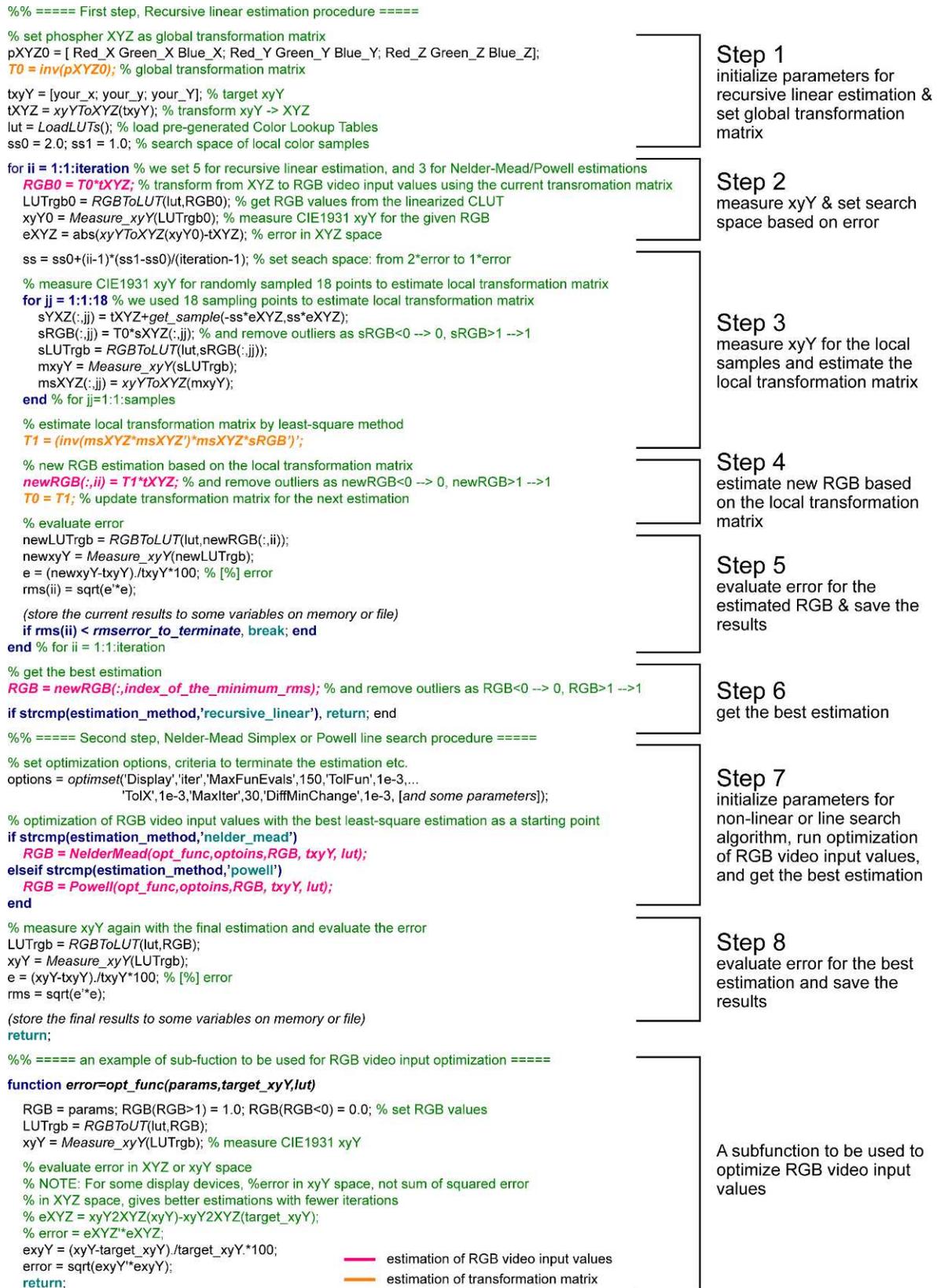


Figure 5. Sample MATLAB code of our proposed color estimation algorithms. In this sample code, each of the numbered blocks on the right side corresponds to each step of our proposed methods described in the Methods section of the main text. The irrelevant portions (i.e., variable declarations) are omitted from the sample.

applied on noisy data (Figure 6). Therefore, we first filter the measured luminance outputs for the 32 video input values (equally spaced from 0.0 to 1.0) by a least-square minimization fit so that they increase monotonically. Then, we used a modified cubic spline interpolation (Garcia, 2010) in which input data points are filtered before interpolation by a discrete-cosine-transformation (DCT)–based smoothing (Figure 6).

Note that we can alternatively use the standard low-pass filter (e.g., Butterworth filter) for smoothing the measured data. There is nothing special about DCT-based smoothing except that (a) it can perform smoothing robustly without a careful pre-estimation step to determine an adequate cutoff threshold that is required in the standard low-pass filtering procedure and (b) it can avoid any phase shifts that occurs when a low-pass filter is applied to the data. Mcalibrator2 also has the standard low-pass filtering procedure as an option. In some cases, the process for filtering out the noise may rather emphasize quantization step errors. In that case, we can avoid any filtering procedures by changing the parameters.

The cubic spline–based procedure improves results for non-CRT devices in comparison to the standard gamma correction, as shown in the Results section. Even when the input/output relationship is described as an S-shaped function, this procedure can linearize the relationship between video inputs and the luminance outputs correctly. The latter color estimation procedure would also be improved by using this data-driven gamma correction instead of using the standard GOG model.

## Recursive linear color estimation

After the gamma correction, we estimate chromaticities by applying linear transformations recursively in a local color space using a least-squares method. Specifically, we estimate color transformation matrices over several iterations while gradually decreasing search spaces that are defined based on the errors between target chromaticity and the actual measurements. We therefore call this a local color transformation matrix. The advantage of this recursive method for chromaticity estimations is that it never assumes global linearity of the system. Instead, it assumes only piecewise linearity in the limited search space. Therefore, even when a display device is not characterized by a linear system, this method can be applied by assuming approximate linearity in the limited neighboring regions of the target chromaticity. The detailed procedures of this method are described below. Each step corresponds to its equivalent block indexed in Figure 5. Illustrations

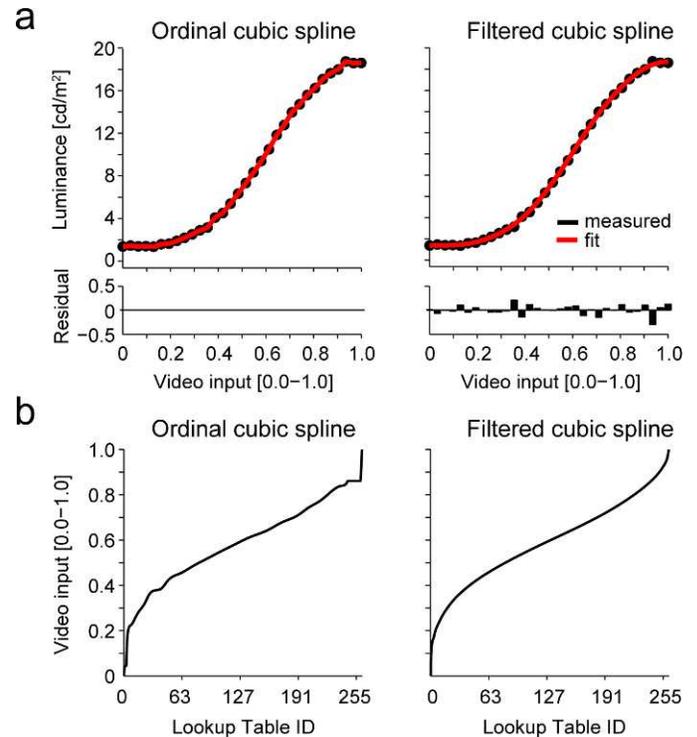


Figure 6. Filtered/unfiltered cubic spline–based gamma correction. (a) Cubic splines with/without the filtering procedures. When a cubic spline is directly applied to the raw data, it is sometimes overfitted to the raw data or distorts the data because it uses third-order polynomials (left panels). In our procedure, these problems are avoided by applying two filters (right panels) before fits with preservation of the shape of the raw data (Garcia, 2010; see the Introduction for details). Although our procedure shows larger residuals, the fit can omit any biases. (b) Generated color lookup tables (CLUTs). An ordinal cubic spline procedure results in distortion of CLUTs (left panel), whereas our filtering procedure can characterize display input/output relationship accurately (right panel).

explaining the estimation procedure are presented in Figure 7.

### Algorithm of recursive linear color estimation

$i$  is an iterator,  $N$  is the maximum number of iterations, and the estimations will stop when  $i$  reaches the maximum number of iterations.

*Step 1:* Set an initial local color transformation matrix,  $T_1$ , as  $T_1 = pXYZ^{-1}$ . Here,  $pXYZ^{-1}$  is a global transformation matrix calculated by Equation 3.

*Step 2:* Calculate the video input values,  $rgb$ , required to produce the target chromaticity  $wXYZ$ , using  $T_i$  and Equation 5. Then, measure CIE1931 xyY for  $rgb$  and calculate the error matrix,  $errXYZ$ , by subtracting  $wYXZ$  from the actual measurement.

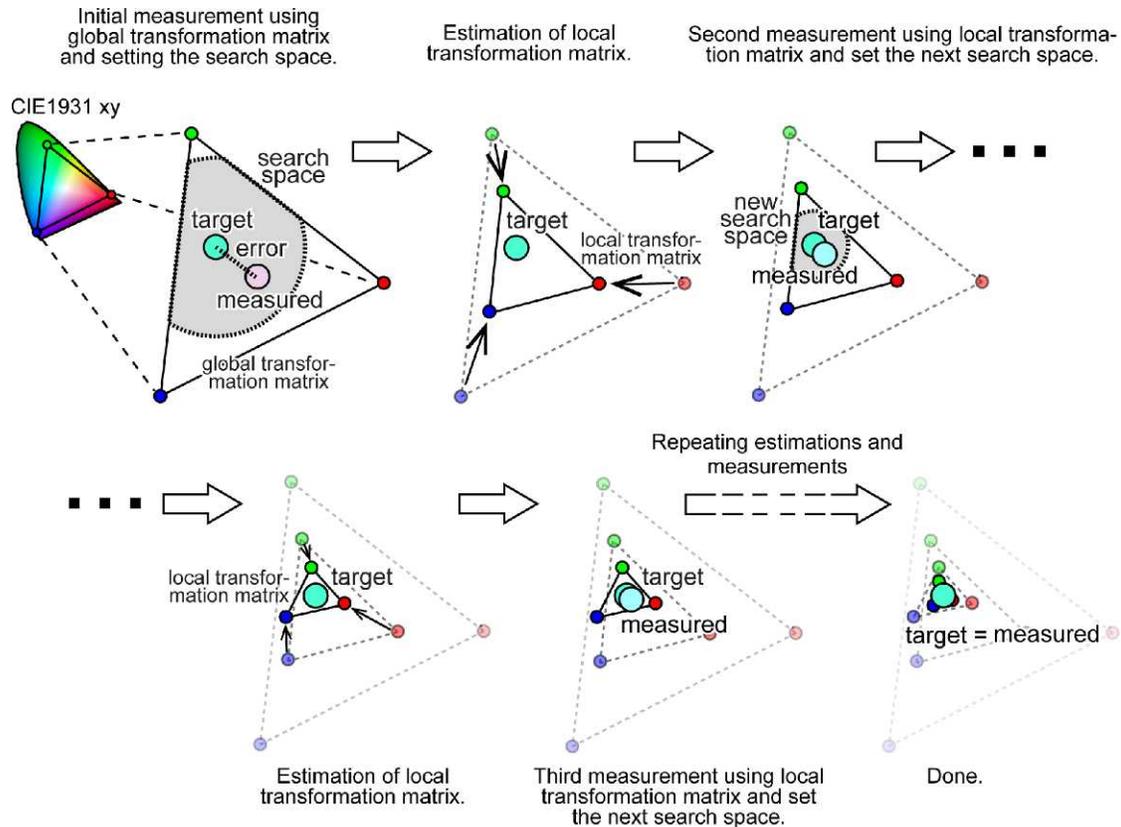


Figure 7. Recursive linear color estimation algorithm. Our proposed recursive linear color estimation method is illustrated in this figure. Here, note that the illustrations are done in CIE1931 xy space for simplicity and legibility, whereas the actual estimations are done in CIE1931 xyY (three variables) and XYZ tristimulus space. The method estimates the required RGB video inputs within local and small color spaces by recursively evaluating errors and calculating the local color transformation matrices with a least-square method.

*Step 3:* Set a local color search space,  $S$ , using  $err_{XYZ}$ , as

$$S = \left(2 - \frac{i-1}{N-1}\right) \times |err_{XYZ}| \quad (7)$$

Within this local color space  $S$ , measure CIE1931 xyY values for randomly generated 18 RGB video inputs,  $sRGB$  ( $3[RGB] \times 18$  matrix). Convert them to XYZ values,  $sXYZ$  ( $3 \times 18$  matrix). Then, using  $sXYZ$  and  $sRGB$ , estimate the next local color transformation matrix,  $T_{i+1}$ , based on the least-squares estimation,

$$T_{i+1} = [sXYZ \cdot sXYZ^T]^{-1} \cdot sXYZ \cdot sRGB^T \quad (8)$$

*Step 4:* Calculate new video input values,  $rgb$ , required to produce the target chromaticity,  $w_{XYZ}$ , using  $T_{i+1}$  and Equation 5. Then, measure CIE1931 xyY for  $rgb$ .

*Step 5:* Calculate root mean squared errors (*RMSE*) between the target chromaticity and the actual measurement. Then add 1 to  $i$ . If the error is smaller than the terminating condition that was originally set, or if  $i$  reaches the number of maximum iterations,

terminate the estimation and go to the next step. Otherwise, go to Step 2 and repeat the estimations in smaller space.

*Step 6:* Get the  $rgb$  with the best estimation from all the iterations and finish.

## Linear/nonlinear hybrid color estimation

We further propose another color estimation procedure: the linear and nonlinear hybrid color estimation method. The recursive linear color estimation described above improves the estimation accuracy even for non-CRT devices (Ban, Yamamoto, & Ejima, 2006). However, it is possible that we cannot assume a piecewise linearity even in the local space for some nonlinearity-dominated display devices. In these cases, a linear method would result in a wrong estimation or an infinite loop. To avoid this shortcoming of the recursive linear estimation procedure, we developed a hybrid method in which the recursive linear estimation was combined with a nonlinear search algorithm (Nelder-Mead Simplex direct search; Dennis &

Woods, 1987; Nelder & Mead, 1965). Specifically, in our hybrid procedure, the starting point of the nonlinear search is first estimated by recursive linear estimation. Then a nonlinear heuristic search algorithm is applied to look for the optimized video input values within a local space. Optimizations are done by minimizing a sum of squared errors between estimated and target CIE1931 xyY values.

This hybrid estimation procedure has two advantages in characterizing display devices compared with the recursive linear transformations or nonlinear search estimation alone. First, as nonlinearity is explicitly included in the algorithm without any assumption of the display system, this hybrid estimation can be applied even when a piecewise linearity in the very local space that the recursive linear estimation assumed is violated in some devices. Even when we need to characterize such completely nonlinear devices, the hybrid estimation can precisely give the accurate results in principle. Second, the linear estimation on the first iteration can prevent the latter nonlinear search algorithm from falling into a local minimum problem. The Nelder-Mead method we took here is a widely used and reliable algorithm for optimization, but it tries to optimize parameters in only the limited local space. Thus, if the estimation starts with the initial inputs estimated using the global color transformation matrix, it may result in converging to incorrect (local minimum) points or entering an infinite loop. Our proposed method avoids this problem by giving more reliable initial inputs using local transformation matrices.

The detailed procedures of this method are described below. Each step corresponds to its equivalent block indexed in Figure 5.

#### **Algorithm of linear/nonlinear hybrid color estimation**

$i$  is an iterator.

*Steps 1–6:* Repeat the same procedures of the recursive linear color estimation with  $i \leq 3$ . Obtain the  $rgb$  with the best estimation as an initial starting point of nonlinear estimation.

*Step 7:* Run the nonlinear optimization of RGB video input values to produce a target CIE1931 xyY chromaticity with Nelder-Mead Simplex direct search algorithm (Dennis & Woods, 1987; Nelder & Mead, 1965). Repeat the estimations until the function output fulfills one of or several predefined tolerance parameters.

*Step 8:* Get the output of the optimization and set it as the value of the best video input,  $rgb$ ; evaluate the error; and terminate the estimation.

Note that the nonlinear estimation here can, in principle, give valid RGB video input values even without CLUTs because they directly search the best

optimized values (Dennis & Woods, 1987; Nelder & Mead, 1965). However, after several attempts, we found that if we searched optimal video inputs without CLUTs and without giving reliable initial inputs calculated by the recursive linear estimation procedure, the estimation required extended processing time. We therefore applied gamma correction in advance of the color estimation.

#### **Line search color estimation**

Furthermore, we propose another color estimation procedure based on a line search algorithm (Brent-Powell Method, Brent, 1973; Powell, 1964) combined with the recursive linear estimations. The Brent-Powell method is useful because it can be applied without any underlying mathematical definitions for the display system. This procedure has a further advantage in that the optimization is generally converged with lesser iterations than the other optimization methods (Farhi, 2011; Farhi, Debab, & Willendrup, 2013). We can therefore accelerate the calibration procedure. This method in principle minimizes the error of the evaluation function by a type of bisection algorithm along the search direction. The new position in each search step can be expressed as a linear combination of search vectors. To efficiently search for optimal values, we used a modified version of the algorithm implemented in the iFit MATLAB toolbox (Brent-Powell's method with Coggins constrain; Farhi, 2011; Farhi et al., 2013; Press et al., 2007). The detailed procedures of this method are described below and in Figure 5.

#### **Algorithm of line search color estimation**

$i$  is an iterator in the following algorithm and iterated by  $i \leq 3$ .

*Steps 1–6:* Repeat the same procedures of the recursive linear color estimation with  $i \leq 3$ . Obtain the  $rgb$  values with the best estimation to use for the initial starting point of the line search.

*Step 7:* Run the optimization of RGB video input values to produce a target CIE1931 xyY chromaticity with a line search algorithm (Brent, 1973; Farhi, 2011; Farhi et al., 2013; Powell, 1964; Press et al., 2007).

*Step 8:* Get the output of the optimization procedure and set it as the value of the video input,  $rgb$ ; evaluate the error; and terminate the estimation.

We tested how applicable these new approaches are by characterizing different types of display devices. We also compared the efficiencies of our methods with that of the standard two-step calibration method.

Abbreviation	Display device	Photometer	Results
CRT1	CRT display GLM-20E21 (Silicon Graphics, Fremont, CA)	Konica-Minolta, CS-100A	Figure 3, Figure 6, Figure 8, Figure 9, Figure 10, Figure 11, Table 2, Table 3, Table 4, Table 5
LCD1	Thinkpad T61 laptop computer, LCD panel (Lenovo, Beijing, China)		
LCD2	MultiSync PA241W LCD display (NEC, Tokyo, Japan)		
DLP	DLP projector U2-1130 (PLUS Vision, Tokyo, Japan)		
CRT2	ViewSonic P225f (ViewSonic Corporation, Walnut, CA)	Admesy Brontes-LL	Figure 12, Figure 13, Table 6, Table 7
LCD3	SyncMaster 913N (Samsung, Seoul, Korea)		

Table 1. Display devices tested, photometers, and list of figures of the corresponding results.

## Methods

### Display devices and colorimeter

We examined the calibration accuracies of our proposed methods using six display devices selected from three different types (two CRTs, three LCDs, and one DLP; see Table 1). All of the measurements and estimations were controlled with MATLAB environment installed on a laptop computer (Thinkpad T61 with an Nvidia Quadro NVS 140 graphic driver, Lenovo, China). CRT1/2 and LCD2/3 displays and laptop were connected using a D-sub type analog display cable. For LCD1 (screen of the laptop computer), the chromaticities displayed on the computer's LCD panel were directly measured. For DLP, the images were projected on a translucent screen, and the chromaticities were measured from behind the screen. Before the measurements, all devices were warmed up for at least 45 min to ensure stability in the readings. During measurements, all display devices and colorimeter were placed into a dark cage enclosed by a black drape to avoid any effect of ambient light. For all of the devices, CIE1931 xyY values for each level of video inputs were measured using Konica-Minolta CS-100A (Konica-Minolta, Japan) or Admesy Brontes-LL (Admesy, Netherland) colorimeter. All transformations of color spaces (e.g., from xyY to XYZ space) were done using MATLAB (Mathworks) with subroutines implemented in Mcalibrator2 software. The results of our proposed methods were compared with those acquired through the standard two-step procedure.

### Software used for luminance/chromaticity measurements and estimations

All measurements and estimations were completed using our in-house software Mcalibrator2 with MAT-

LAB. The software and the related documents are publicly available and can be downloaded from the following link: <http://www.cv.jinkan.kyoto-u.ac.jp/site/mcalibrator/>.

### Standard gamma correction

To characterize the relationship between video input values and luminance outputs using the standard gamma-correction procedure, CIE1931 xyY values for 32 input values (equally spaced from 0.0 to 1.0) were measured for each of RGB phosphors separately. Measurements were repeated three times, and the results were averaged for each of the RGB phosphors before any further processing. The GOG functions (Equation 1) were then fitted to the averaged data. The goodness of fits were evaluated by plotting the actual measurements and fitted curves together and by calculating root *RMSEs* for each of RGB phosphors separately (Figure 8). The CLUTs were generated based on inverse functions of the best GOG fits and used for later chromaticity estimations. Furthermore, the linearity of the luminance outputs against the video input values after gamma correction was tested separately for each of the RGB phosphors. The tests were performed by remeasuring luminance values for 18 video input values, which were not used in initial fits (equally spaced from 0.05 to 0.95), and fitting linear functions considering the flare at zero-level video input (Figure 9). These results were evaluated by *RMSEs* (Tables 2 and 3).

### Cubic spline–based gamma correction

The same measurement procedures with the standard gamma correction were applied, except that the input/output properties were described by our cubic spline–based method (see the Cubic Spline–Based Gamma Correction section in the Introduction for

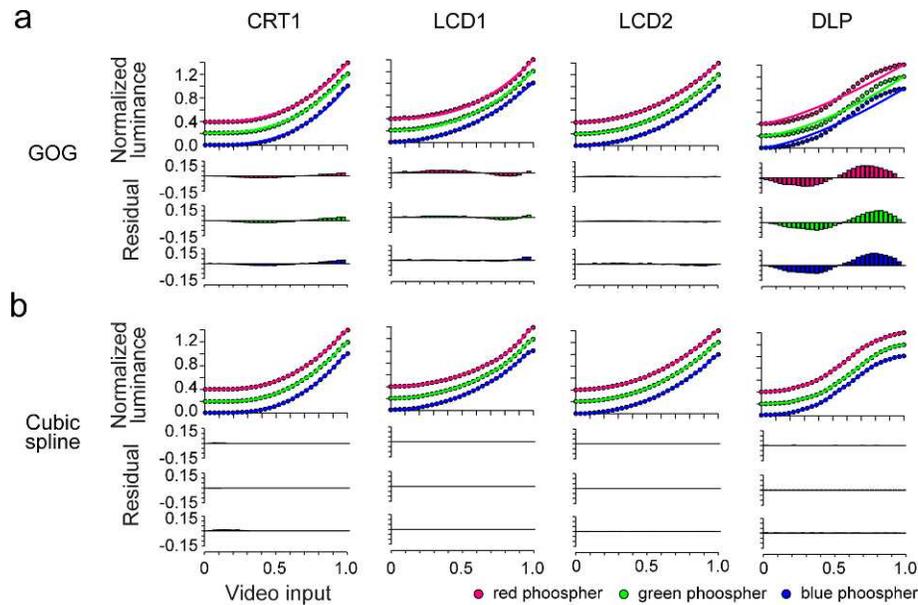


Figure 8. Measured luminance and model fits—the standard GOG versus cubic spline–based gamma correction. The fitting accuracies between the standard GOG and cubic spline–based methods were compared. (a) Fitting results of the standard gamma correction with the GOG model. (b) Fitting results of our proposed cubic spline–based procedure. In these graphs, the luminance values were normalized so that the maximum is 1. The red and green phosphors results were shifted along the y-axis to avoid overlapping of the data. The circular dots represent the measured luminance values, and lines represent the fitted curves. The residuals are also displayed as bars.

details). The accuracies of the fits were evaluated by *RMSEs* (Figures 8 and 9; Tables 2 and 3). The linearity after gamma correction was tested and evaluated using the same procedures as the standard GOG-based gamma-correction method. For all the later color estimation procedures, CLUTs generated by GOG or this cubic spline–based method were used in correcting video input values to ensure the linearity of the display input/output relationship.

### Linear color estimation based on tristimulus values

The standard color estimations based on RGB phosphors' tristimulus values (*global* color transformation matrix) were done following Equations 3 to 6 for 50 chromaticities. These 50 chromaticities were randomly generated in CIE1931 xyY space once for each display device, with a restriction that they were within a triangle enclosed by x and y values of RGB phosphors (see Figure 10). This restriction means that the corresponding video input values for the target chromaticity fall from 0.0 to 1.0 and that the target chromaticity can theoretically be reproduced based on Equations 3 to 6. The actual CIE1931 xyY values for the estimated RGB video inputs based on Equation 5 were measured. The estimation accuracies were

evaluated by *RMSEs* and  $\Delta E_{Lab}$  errors. Here, *RMSEs* were calculated by the formula below after converting the measured error in the CIE1931 xyY space to percentage residuals because the scale of Y is relatively larger than x and y in the raw data.

$$error = \begin{bmatrix} (measured\ x - target\ x)/target\ x \times 100 \\ (measured\ y - target\ y)/target\ y \times 100 \\ (measured\ Y - target\ Y)/target\ Y \times 100 \end{bmatrix} \quad (9)$$

$$RMSE = \sqrt{error^T \cdot error} \quad (10)$$

The tests were repeated three to five times on different days to ensure repeatability. We are convinced of the robustness of the estimation results and errors. The repeatability tests were done for all estimations described below, except for those described otherwise. Here, note that we must carefully compare the absolute differences of estimation accuracy among devices. This is because the target 50 chromaticities were generated randomly for each display device separately so that they could be reproduced theoretically based on Equation 5. These separate selections were required because there were large differences in the available color space of display devices (compare the triangles enclosed by RGB phosphor CIE1931 x,y values of different display devices). It therefore might be possible that some

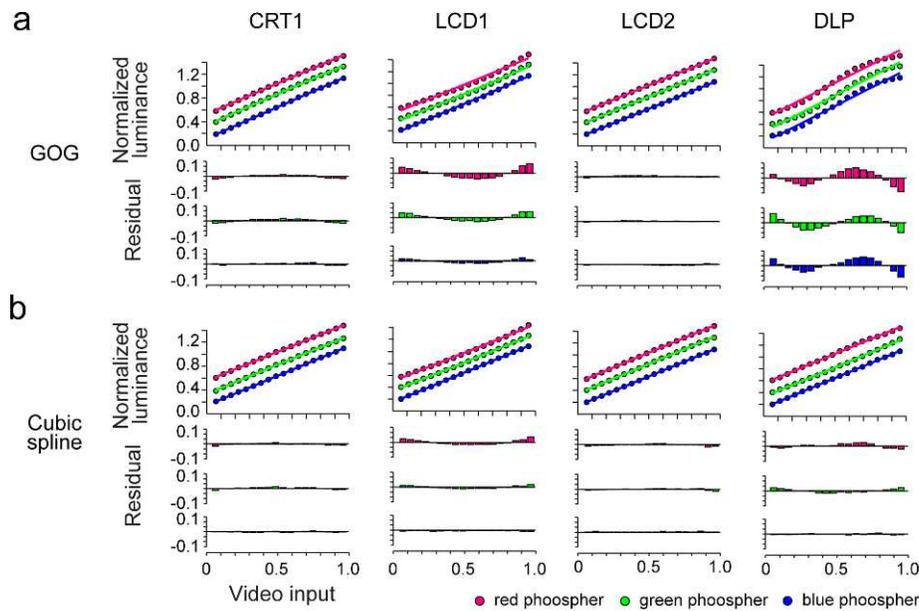


Figure 9. Linearity of generated color lookup tables—the standard GOG versus cubic spline-based procedures. The linearity of the input/output relationships after gamma corrections was tested by remeasuring luminance against 20 input values. (a) Linear curve fittings to the input/output relationships after correcting by the standard GOG model. (b) Linear curve fittings to the input/output relationships after correcting by cubic spline-based procedures. In these graphs, the luminance values were normalized so that the maximum is 1. The red and green phosphors results were shifted along the y-axis to avoid the overlapping of the data. The circular dots represent the measured luminance values, and lines represent the fitted curves. The residuals are also displayed as bars in the panels below.

chromaticities selected for a display device were much easier to reproduce than the others for another display, which would have affected the final absolute accuracies, although we can confirm there was no such bias in the results. The comparison among estimation methods within a display can be done directly.

### Recursive linear color estimation

Using our recursive linear estimation method, we estimated video input values to produce the same 50 randomly generated chromaticities used in the stan-

Method	Phosphor	CRT1	LCD1	LCD2	DLP
GOG	Red	0.0781	0.1069	0.0178	0.4064
	Green	0.0900	0.0747	0.0218	0.3606
	Blue	0.0867	0.0576	0.0466	0.3925
	Mean	0.0849	0.0797	0.0287	0.3865
Cubic spline	Red	0.0072	0.0002	0.0001	0.0067
	Green	0.0034	0.0001	0.0002	0.0001
	Blue	0.0244	0.0001	0.0029	0.0086
	Mean	0.0117	0.0001	0.0011	0.0051
(RMSE)					

Table 2. Averaged RMSEs of model fits to the measured luminance.

dard linear estimation. The starting points of the recursive computations were set as the values calculated by the standard global linear transformation. From these starting points, the local color transformation matrices were estimated at most five times. The local transformation matrices were estimated by a least-squared method from the locally measured 18 chromaticities chosen randomly within a small search space defined by Equation 7. Search spaces were set so that they gradually decreased from  $2.0\times$  errors (RMSEs) to  $1.0\times$  errors during the estimation iterations. We set two termination conditions: when the RMSEs between the estimation and the target came to less than 1 or the number of repetitions

Method	Phosphor	CRT1	LCD1	LCD2	DLP
GOG	Red	0.0386	0.1289	0.0168	0.1867
	Green	0.0428	0.0936	0.0096	0.1544
	Blue	0.0232	0.0473	0.0130	0.1587
	Mean	0.0349	0.0899	0.0131	0.1666
Cubic spline	Red	0.0237	0.0657	0.0243	0.0525
	Green	0.0270	0.0369	0.0195	0.0467
	Blue	0.0149	0.0116	0.0117	0.0161
	Mean	0.0219	0.0381	0.0185	0.0384
(RMSE)					

Table 3. RMSEs in linearizing input/output relationship.

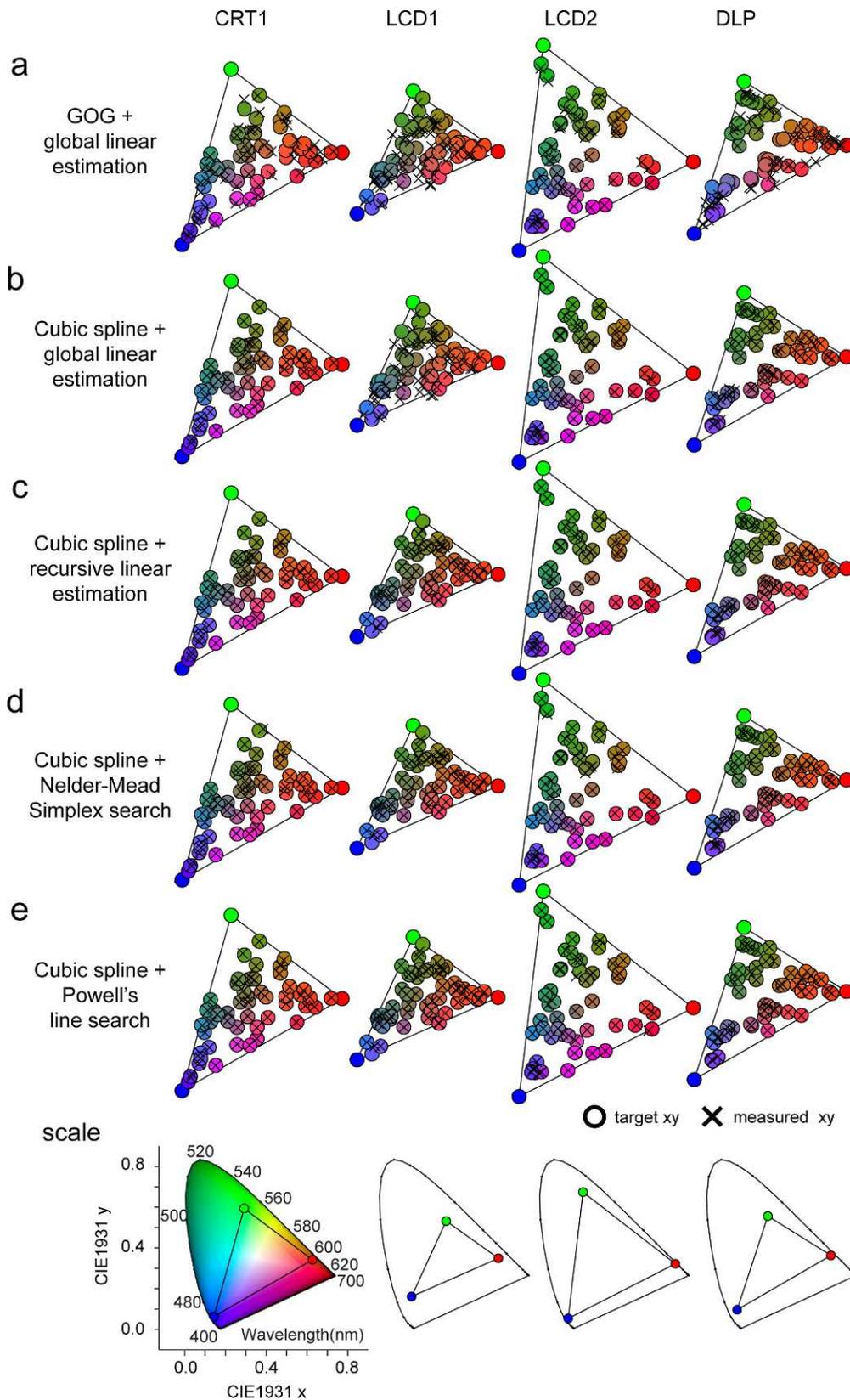


Figure 10. Chromaticity estimation results for the standard global linear transformation and our proposed estimation procedures. Color estimation results of all the procedures tested are visualized in this figure. Along the columns from left to right, the results of CRT, LCD1, LCD2, and DLP are displayed. Along the rows, results of different estimation methods are displayed: (a) GOG model + the standard global color transformations, (b) cubic spline + the standard global color transformations, (c) cubic spline + recursive linear estimations, (d) cubic spline + linear/nonlinear hybrid color estimations, (e) cubic spline + line search color estimations.

reached the maximum (five iterations). *RMSEs* were calculated based on Equations 9 and 10. Thus, when a percentage error in each of the  $xyY$  values is 1.0 (%), the *RMSE* is 1.73. The goal of the present study is to provide accuracy to within 1 of *RMSE*. The criterion is relatively demanding, but our methods actually achieved this for non-CRT devices, as shown in the Results section.

In this estimation procedure, at most 96 measurements (the initial measurement +  $5 \times [18 \text{ local measurements} + 1 \text{ remeasurement to test error}]$ ) were required for each chromaticity estimation. However, in the actual tests, more than half of all the estimations finished within four or fewer repetitions. We intentionally set these parameters to ensure that the estimations would finish within a reasonable period (1–5 min) and to show that results were still reliable and accurate even in these limited estimations. Although accuracy could be improved with more rigid parameters, these restrictions were important to ensure vision experiments could be conducted after daily quick display calibrations.

### Linear/nonlinear hybrid color estimation

We also estimated 50 chromaticities using our linear/nonlinear hybrid estimation method. The initial starting point of the line search was set based on recursive linear estimations to prevent a local minimum problem. All parameters remained the same as for the recursive linear color estimation described above, except that the number of maximum iterations was set to three. A nonlinear search algorithm (Nelder-Mead Simplex direct search; Dennis & Woods, 1987; Nelder & Mead, 1965) was run to find the optimal video input values, with the following termination conditions: the number of iterations of the whole estimations reached 50, the number of function evaluations (calculated by the sum of squared errors) reached 150, or tolerance of the function's output reached 0.5. With these parameters, each nonlinear estimation generally converged to the optimal point within 30 iterations of function evaluations. Estimation accuracies were evaluated by percentage errors and *RMSEs*.

### Line search color estimation

We also estimated the 50 chromaticities using a line search algorithm implemented in the iFit MATLAB toolbox (<http://ifit.mccode.org/>; Brent-Powell's method with Coggins constrain; Brent, 1973; Farhi, 2011; Farhi et al., 2013; Powell, 1964; Press et al., 2007) combined with recursive linear estimations. The initial starting

point for the line search was based on the recursive linear estimations with at most three iterations. Termination occurred when either the number of iterations of the whole estimations reached 80, the number of function evaluations reached 200, or the tolerance of the function's output reached 0.5. With these parameters, each line search optimization generally terminated within 20 iterations of function evaluations.

## Results

### Gamma correction: Standard versus cubic spline-based methods

We compared the results of the gamma correction between the standard GOG-based procedure and our cubic spline-based method. The results indicate that our method can be applied to a wide range of display devices, including non-CRTs, and that this method improves calibration accuracy. As shown in Figure 8 and Table 2, our method can describe the relationship between video inputs and luminance outputs more accurately for all devices than the standard GOG model.

Even for a CRT device, the cubic spline-based method could describe the input/output property more accurately than the GOG model. This result is important in that although the GOG model is established based on the internal model of CRT devices, some CRTs may not necessarily follow the model, and rather the data-driven cubic spline-based method is suitable for such CRTs.

Further, our method was more accurate for LCDs and DLPs than the standard GOG model. The input/output relationships of the tested DLP (and also LCD1 slightly) followed an S-shaped function (the rightmost panel in Figure 8). This S-shaped relationship was impossible to interpolate by a GOG function correctly, as it was based on an exponential function. For this kind of S-shaped input/output display system, the GOG model will be problematic. The residuals of the GOG model for the DLP were periodic against video input values. This will next lead to severe periodic biases when trying to estimate chromaticities. In contrast, our cubic spline-based method accurately described the measured values irrelevant to the shape of the input/output property, with smaller and randomly distributed residuals.

Using the CLUTs generated by the GOG and our methods, we compared accuracies of the linearization of the display inputs and outputs (Figure 9; Table 3). The plots were generated by remeasuring luminance

values for 18 video inputs for each RGB phosphors that were not measured in the gamma-correction procedure. These were then fit with linear functions. For three of the four measured devices, our cubic spline-based method achieved better linearization. The means of *RMSEs* for RGB phosphors were 1.59 times smaller for the CRT, 2.36 times for the LCD1, and 4.34 times for the DLP, whereas the standard GOG was 1.4 times better for the LCD2 (Table 3). One reason that we may have observed a reduction in linearization for the LCD2, despite better initial fittings, may be due to an overfitting of the cubic spline curves. However, as the absolute values of *RMSEs* for the LCD2 were smaller than those for the other devices, we can conclude that both methods were successful for the LCD2.

For LCD1 and DLP, the residuals of the linear fits after gamma correction with the GOG model were periodic against the video input increments. These results would be easily predictable, as the initial fitting with the standard GOG model for these devices could not correctly describe the input/output relationships. In contrast, S-shaped functions in these devices could be linealized correctly by our cubic spline-based method.

### Chromaticity estimation: Standard linear transformation versus our search methods

We compared chromaticity reproduction accuracies between the standard two-step method and several of our described methods. Specifically, we compared accuracies of the following five color estimation procedures:

- Method 1: GOG model-based gamma correction + linear color transformation
- Method 2: cubic spline-based gamma correction + linear color transformation
- Method 3: cubic spline-based gamma correction + recursive linear color estimation
- Method 4: cubic spline-based gamma correction + linear/nonlinear hybrid color estimation
- Method 5: cubic spline-based gamma correction + line search color estimation

For all devices, our methods reproduced the required chromaticities in the CIE9131 xyY space more accurately than the standard global transformation procedure. Furthermore, the results suggest that our methods can be applied to a wider range of devices than the standard method. Table 4 shows average *RMSEs* ( $RMSE \leq 1.0$  was the preset termination criterion) and  $\Delta E_{Lab}$  errors of chromaticity estimations for 50 randomly generated CIE1931 xyY values. A mixed-design analysis of variance (ANOVA) on *RMSEs* for

50 reproduced chromaticities found significant differences between estimation methods ( $F(4, 196) = 1134.862, p < 0.00001$ ). A significant interaction was also observed for Display Device  $\times$  Method ( $F(12, 784) = 133.364, p < 0.00001$ ), but the interactions were only for Display  $\times$  The Standard Global Estimation Procedures ( $F(3, 980) = 285.911, p < 0.00001$ , for Method 1 and  $F(3, 980) = 127.101, p < 0.00001$ , for Method 2). Further multiple comparisons (corrected with Ryan's method,  $p < 0.05$ ) showed that all three of our methods (Methods 3, 4, and 5) significantly improved the estimation accuracies for all devices when compared with the standard method (Methods 1 and 2).

Figure 10 plots the target and estimated 50 chromaticities in CIE1931 xy space for each display device and estimation method. These plots affirm that the errors were distributed randomly and there was no bias in estimations. Namely, the estimations were achieved correctly in all available CIE1931 color spaces of target display devices.

Comparing the standard deviations (*SDs*) and minimum/maximum errors across methods should be also important as well as averaged errors. As shown in Table 5, the smallest error for each device was obtained using a different method. Notably, our proposed methods could give better estimations with smaller errors, especially for non-CRT displays compared with the standard procedure. The different estimation accuracies observed for different displays likely derive from differences in the profiles of the display devices. Some of our methods such as recursive linear transformations are based only on linear transformation and are not suitable for displays with a nonlinear profile. In contrast, although the nonlinear or direct search methods can handle nonlinearity, they are not necessarily suitable for some linear display devices because they may overfit to the local values. We therefore need to select different methods for different display devices. Our software Mcalibrator2 can overcome this linearity/nonlinearity problem by preparing both linear and nonlinear direct search calibration algorithms.

The numbers of incremental measurements required until the estimations converged to the termination condition are also important. Figure 11 plots the *RMSEs* against the number of iterations of Method 3 and Method 4 for LCD1 and DLP devices. For LCD1, although all the chromaticities estimated by a global transformation method were above our predefined criteria ( $RMSE > 1.0$ , see Equations 9 and 10), 43 of 50 chromaticities converged to the termination criterion within three iterations. For DLP, after two iterations, 47 of 50 chromaticities reached the termination. When linear (Method 3, red line in Figure 11) and nonlinear (Method 4, blue line in Figure 11)

Method	RMSE				Delta *E Lab			
	CRT1	LCD1	LCD2	DLP	CRT1	LCD1	LCD2	DLP
Method 1	2.7443	3.2804	9.6667	11.1526	4.7555	10.3887	2.6495	12.0693
Method 2	0.7411	1.4808	4.4507	4.9502	2.0338	10.7232	1.6921	1.9483
Method 3	0.3764	0.5988	0.9600	1.2907	0.9993	0.8304	0.6645	0.7955
Method 4	0.8005	1.2398	1.0467	1.9792	1.7840	0.8878	1.6584	0.8217
Method 5	0.5788	0.8360	0.7880	1.4404	1.3705	0.9237	1.1532	0.5180

Table 4. Average percentage and *RMSEs* for chromaticity estimations.

search methods are compared, the linear method converged faster than the nonlinear procedures. Notably, the nonlinear method was more unstable; it could give better estimations for some chromaticities, as shown in Figure 11, whereas it also gave worse estimations for some chromaticities (for DLP, only 42 of 50 colors reached the termination conditions when the nonlinear method was applied). Thus, the best strategy would be that we should try both procedures and take the better one, if we can spend enough time for daily calibrations.

The required time for calibration differed largely between the methods. For the gamma-correction step, both the standard GOG and our cubic spline-based methods finished within 10 s for all display types. We therefore suggest that the cubic spline-based method can be easily implemented in a regular display calibration procedure. In the color estimation step, however, we found that our methods took 50 to 100 min for 50 chromaticity estimations. In our three search algorithms, the calibration speed was faster for the recursive linear estimation, followed by the line search estimation and then the linear/nonlinear hybrid estimation procedure. In contrast, global linear color transformation finished within a second, as it estimated only RGB video input values by linear conversions. In contrast, our methods search the best video input values by recursively measuring nearby points and sometimes reached maximum iterations without satisfying any other terminating conditions (note that even without

convergence, the final accuracies of our methods were better than the standard procedure; Table 4). However, when the estimation succeeded within preset iterations, it generally took 1 to 2 min for a single chromaticity.

### Control measurements and tests

We further performed additional control measurements to test the validity of our proposed methods. First, we tested the efficiency of our random sampling procedure in recursive linear estimations. In recursive linear transformation procedures (Method 3 and the initial steps of Methods 4 and 5), we randomly sampled 18 chromaticity values within small color spaces to estimate local color transformation matrices. However, it is unknown whether a considerable efficiency can be obtained from a random sampling procedure. It may be better to select chromaticity values in a structured way, such as a grid sampling within a local color space. We thus need to test the efficiency of random samplings. To this end, we compared chromaticity reproduction accuracy versus number of measurements between random and grid-sampling procedures (Figure 12). In the random sampling procedure, 18 chromaticity values were selected randomly within the error space (for details, see the Methods section). In the grid sampling, 18 chromaticity values were selected as below (also see Figure 12a).

Method	Delta *E Lab											
	CRT1			LCD1			LCD2			DLP		
	<i>SD</i>	Min	Max	<i>SD</i>	Min	Max	<i>SD</i>	Min	Max	<i>SD</i>	Min	Max
Method 1	2.0948	2.0371	14.6091	1.9862	5.8364	14.8696	1.9971	1.1320	15.5537	3.9332	6.7672	21.6162
Method 2	3.0835	0.5183	20.9798	2.4852	5.5208	15.2245	2.0760	0.5057	15.5822	0.8778	0.5560	3.8892
Method 3	2.8602	0.0591	19.2288	2.0356	0.0868	11.9246	1.8185	0.1027	13.1948	2.0172	0.0810	10.6824
Method 4	3.3935	0.1217	22.2754	2.5548	0.0629	17.2483	2.1303	0.0636	13.6405	1.7606	0.0827	10.0818
Method 5	3.2580	0.0511	22.0912	2.7172	0.0764	17.9090	2.1463	0.0818	14.4913	0.3629	0.0570	1.8337

Table 5. Standard deviations and min/max errors in delta \*E Lab for chromaticity estimations.

grid\_samples

$$= \begin{bmatrix} eX & -eX & 0 & 0 & 0 & 0 & \frac{eX}{2} & \frac{-eX}{2} & \frac{eX}{2} & \frac{-eX}{2} & \frac{eX}{2} & \frac{-eX}{2} & 0 & 0 & \frac{eX}{2} & \frac{-eX}{2} \\ 0 & 0 & eY & -eY & 0 & 0 & \frac{eY}{2} & \frac{eY}{2} & \frac{-eY}{2} & \frac{-eY}{2} & 0 & 0 & \frac{eY}{2} & \frac{-eY}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & eZ & -eZ & 0 & 0 & 0 & 0 & \frac{eZ}{2} & \frac{eZ}{2} & \frac{eZ}{2} & \frac{eZ}{2} & \frac{-eZ}{2} & \frac{-eZ}{2} \end{bmatrix} \times ss \quad (11)$$

Here,  $eX$ ,  $eY$ , and  $eZ$  are residuals between the target CIE1931  $xyY$  values and the actual measurements in the corresponding estimation step, and  $ss$  is a coefficient that gradually decreases against the repetitions of estimations (from 2.0 [the first estimation] to 1.0 [fifth estimation] in the 0.2 step). Also note that luminance values of 50 CIE1931  $xyY$  chromaticities estimated in this test were limited within 5 to 15  $\text{cd}/\text{m}^2$ . This is because we already found, through several previous tests, that chromaticity estimations with higher luminance values were considerably stable and good enough, and we could not see any clear differences between methods. We thus had to test efficiency in noisy conditions.

Figure 12b plots  $RMSE$  errors against the iterations of recursive linear transformations for 50 CIE1931  $xyY$  values. Table 6 shows average  $RMSE$  errors obtained from the random and grid-sampling procedures,

together with Method 2 as a comparison. We found that both the random and grid-sampling procedures improved estimation accuracy compared with the standard global linear transformation method. Furthermore, although we performed the comparisons in noisy conditions, we did not find any differences between the random (red lines) and grid (blue lines) sampling procedures. A mixed-design ANOVA (Device [CRT2 and LCD3]  $\times$  Method (random vs. grid)  $\times$  Iteration [one to five steps]) with 50  $RMSE$  samples showed significant differences between devices ( $F(1, 98) = 9.338$ ,  $p < 0.029$ ) and iteration ( $F(4, 392) = 94.568$ ,  $p < 0.00001$ ) and interactions of Device  $\times$  Iteration ( $F(4, 392) = 58.271$ ,  $p < 0.00001$ ) but never found differences between sampling methods ( $F(1, 98) = 0.045$ ,  $p = 0.83$ ). The reason that both random and grid samplings worked effectively may be that 18 chromaticity values are large enough to estimate a  $3 \times 3$  local color transformation matrix, even when it is selected randomly. It may be also possible that random and grid selections may never bias pooled data when they are performed in a fairly local and small color space. We can therefore conclude that the random sampling taken in our methods is efficient enough for estimating a local color transformation matrix.

Second, we tested whether considerable efficiency could be obtained by a simpler procedure. Specifically, we investigated whether we could reproduce required chromaticities by simply adjusting phosphor intensity corresponding to the measured errors (residuals between the target and measured CIE1931  $xyY$ ). Suppose that we have our global (or local) transformation matrix  $M$  that maps  $xyY$  to  $RGB$  video input values. We have a current set of  $RGB$  values ( $RGB_i$ ), and they produce a current measured output  $xyY_i$ . Our goal is to obtain target  $xyY$  values. Then, we can compute

$$xyY_{\text{delta}} = \text{target}_{xyY} - xyY_i. \quad (12)$$

The linear model says that to correct for this error, we can add

$$RGB_{\text{delta}} = M \cdot xyY_{\text{to\_XYZ}}(xyY_{\text{delta}}) \quad (13)$$

$$RGB_{\text{corrected}} = RGB_i + RGB_{\text{delta}}. \quad (14)$$

Unfortunately, as the present study has shown, the

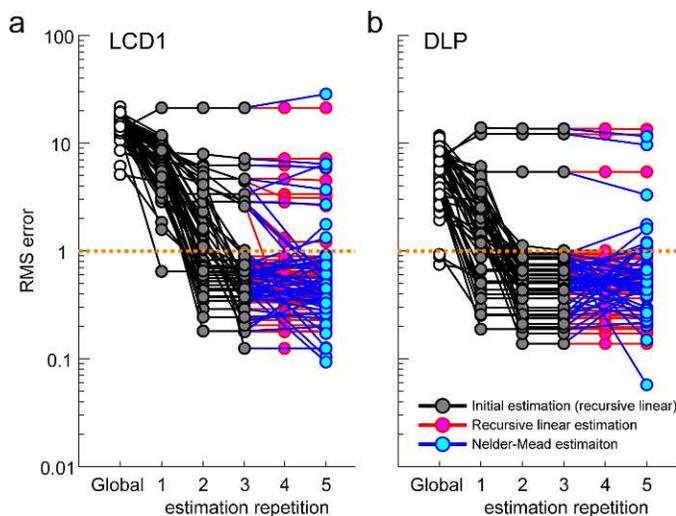


Figure 11. Convergence of recursive linear chromaticity estimations. Plots of color reproduction accuracies against the number of estimation. White dots are accuracies obtained by a global color transformation procedure. Black dots and lines are accuracies obtained by our recursive linear transformation procedures. Red dots and lines are accuracies obtained by proceeding the recursive linear estimations. Accuracies obtained by our nonlinear estimations based on the Nelder-Mead optimization procedure are interposed on these figures by blue dots and lines.

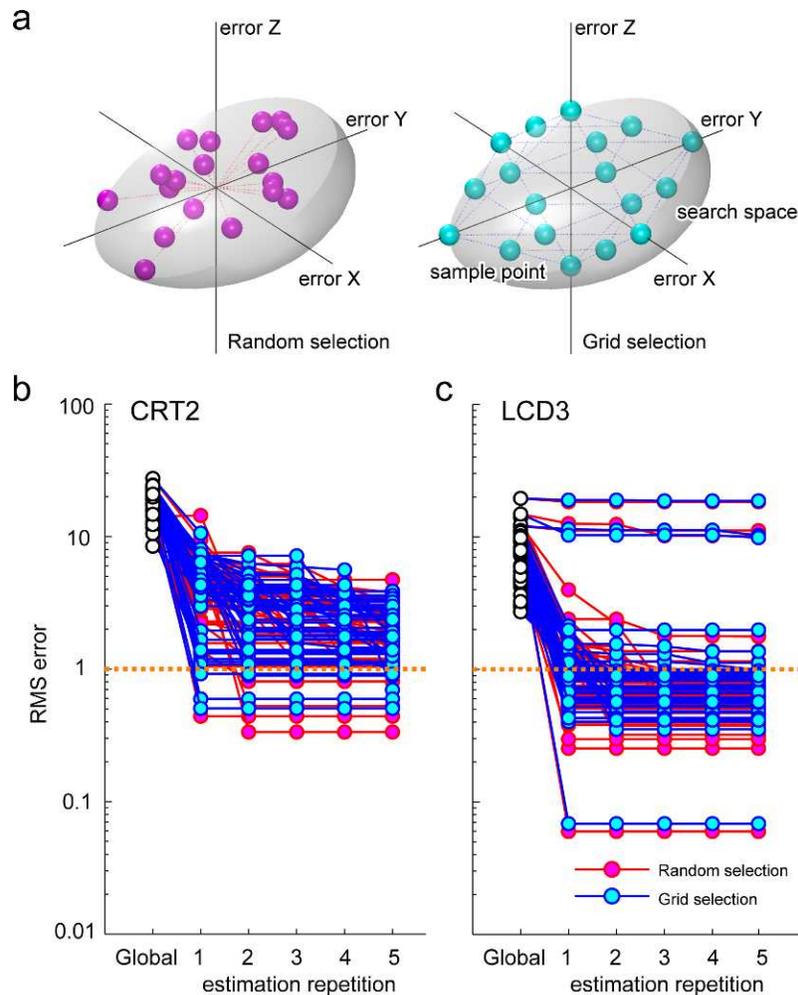


Figure 12. Random versus grid sampling: the effects on estimating local color transformation matrices. (a) Illustrations of random-sampled (left) and grid-sampled (right) 18 chromaticity values. For both methods, 18 chromaticity values were selected within a local CIE1931 xyY space. (b) Plots of color reproduction accuracies for 50 randomly generated CIE1931 xyY values against the repetitions of estimations. White dots are accuracies obtained by a global transformation method. Red dots and lines are accuracies obtained by our recursive linear estimation method with a random sampling procedure. Blue dots and lines are accuracies obtained with a grid-sampling procedure. We did not find any differences between the two sampling procedures.

linear model is not exactly right, especially for non-CRT displays. However, this sort of linear residual adjustment may be good enough to head things in the right direction, especially if it is combined with a rate parameter  $C$  to decrease and adjust  $RGB_{delta}$ . This sort of update might converge very quickly on good RGB values even based on the global transformation matrix,  $M$ . We thus investigated how accurately chromaticity values were reproduced by adjusting errors, following Equations 13 and 14 together with multiplying a coefficient  $C$  ( $-3.0 < C < 3.0$ ) on the  $RGB_{delta}$ . Namely, Equation 14 was modified as

$$RGB_{corrected} = RGB_i + C \times RGB_{delta}. \quad (15)$$

Here,  $C$  was adjusted by a standard linear optimization procedure implemented in MATLAB. We also compared color reproduction accuracies when RGB

phosphor errors were adjusted simultaneously ( $C$  was a scalar) and when they were modified independently ( $C$  was a  $1 \times 3$  vector).

The results of these estimations for CRT2 and LCD3 showed slight improvements of the estimations compared with the global transformation procedure for some chromaticity values (Figure 13). However, the overall accuracies were not comparable to our recursive linear estimation procedures (Table 7; also see the second and third rows of Table 6). Furthermore, the improvement was obtained only when RGB phosphor errors were adjusted simultaneously (Table 7; also see the first row of Table 6 for comparisons), and it was not statistically significant ( $t_{196} = 1.017$ ,  $p > 0.31$ , after Ryan's correction for multiple comparisons). The estimations became worse compared with the global transformation when RGB errors were adjusted

Method	Delta *E Lab									
	CRT2					LCD3				
	RMSE	delta *E	SD	Min	Max	RMSE	delta *E	SD	Min	Max
Method 2	18.8312	4.7055	1.3082	2.4358	7.7162	8.3665	8.0636	3.2751	2.8745	18.0133
Method 3 (random sampling)	2.0479	1.2180	0.9187	0.2226	4.2213	1.4851	1.2772	2.5188	0.0864	16.2654
Method 3 (grid sampling)	2.0072	1.1629	0.7961	0.2075	3.6344	1.4731	1.2813	2.5754	0.0960	16.8452

Table 6. Errors for chromaticity estimations: random versus grid sampling.

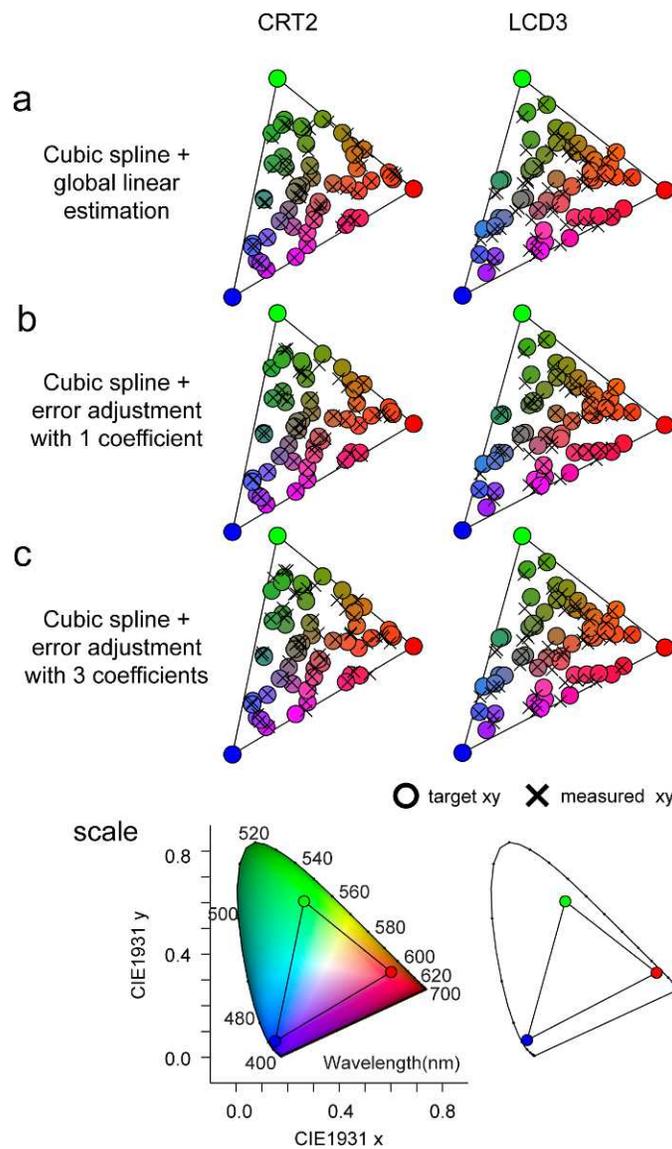


Figure 13. Chromaticity estimation results for the standard global linear transformation and error adjustment procedures. Color estimation results of a global linear estimation and error-adjustment procedures. The left column is the results of CRT2, and the right column is the results of LCD3. (a) Cubic spline + global color transformations, (b) cubic spline + error adjustment (RGB phosphor residuals were adjusted simultaneously by a scalar  $C$ ), and (c) cubic spline + error adjustment (RGB phosphor residuals were adjusted separately by a  $1 \times 3$  vector  $C$ ).

Method	Delta *E Lab									
	CRT2					LCD3				
	RMSE	delta *E	SD	Min	Max	RMSE	delta *E	SD	Min	Max
Error adjustment, RGB simultaneously	7.3554	3.6439	1.9828	0.7156	9.8965	6.9713	6.9742	3.6069	0.2723	17.6009
Error adjustment, RGB separately	18.8312	4.7055	1.3082	2.4358	7.7162	8.3665	8.0636	3.2751	2.8745	18.0133

Table 7. Chromaticity estimation results by error adjustment procedures.

independently ( $t_{196} = 14.5430$ ,  $p < 0.00001$ , with Ryan's correction).

The result may look strange considering that Olds, Cowan, and Jolicoeur (1999) could obtain accurate estimations based on a similar idea even with a single iteration. The reason why we failed to get efficient estimations may be that the present study used only a global color transformation matrix (Equation 3) to adjust errors, whereas Olds et al. (1999) used a kind of local color transformation matrix. Because global linearity is not right in precise display characterization, our estimation would be worse when it is compared with the previous local estimation procedure. However, note that our aim here is to explore quicker estimations than our recursive linear estimations. So we intentionally used a quickly obtainable global transformation matrix, even though we knew a local matrix would improve the estimations. In sum, we can conclude that just simply adjusting residuals between the required and measured chromaticities using a global transformation matrix is not good enough for precise color reproductions (also see the Discussion).

## Discussion

In this article, we proposed novel display characterization methods based on (a) a data-driven cubic spline-based procedure for gamma correction and (b) recursive linear/nonlinear hybrid and line search algorithms for estimating required RGB video input values. We have shown several advantages of our methods compared with the standard gamma correction and global linear color transformation method currently used in vision experiments. Our methods are relatively model free; they never assume any internal model of the display system and can be applied to a wider range of display devices, potentially even future devices such as organic EL and laser monitors.

In addition, we developed a software suite called Mcalibrator2 using MATLAB, which includes all the procedures described in this article as well as standard gamma correction and global linear color transformation methods. The software is user friendly, as all methods are operated through a simple graphical

interface, as shown in Figure 1. Furthermore, Mcalibrator2 has provided some additional components to communicate with any photometers that researchers are using in their daily measurements. The software has also provided simple frameworks to add alternative chromaticity optimization procedures researchers may want to use. Further, the software can automatically generate gamma-correction tables compatible with Psychtoolbox, one of the most widely used vision science tools (Brainard, 1997; Pelli, 1997). These functions will assist researchers in characterizing their display devices (for details, see a manual of Mcalibrator2 software distributed at <http://www.cv.jinkan.kyoto-u.ac.jp/site/mcalibrator>).

The gamma-correction and color estimation methods described in this article advanced our previous study (Ban, Yamamoto, & Ejima, 2006). First, by applying additional filters, the new gamma-correction procedure enables us to describe a display input/output profile more precisely and more robustly against noises. Second, our new methods can be applied to a wider range of devices as they can deal with nonlinearity and identify the best values via a direct search algorithm, whereas the previous study used only the recursive linear estimation approach assuming piecewise linearity within a local color space. Although it has been shown that the performance of the recursive linear method already improved estimation accuracy better than the standard global color transformation, our new approaches using linear/nonlinear hybrid or line search algorithms further improved the accuracies, as shown in Figure 10 and Tables 4 and 5.

Our data-driven and goal-seeking approach is different from recent reports, which have also tried to characterize non-CRT displays (Bastani, Cressman, & Funt, 2005; Fairchild & Wyble, 1998; Gibson, Fairchild, & Fairchild, 2000; Ikeda & Kato, 2000; Kwak & MacDonald, 2000; Tamura, Tsumura, & Miyake, 2001; Thomas, Hardeberg, Foucherot, & Gouton, 2007). In these studies, display characterization is generally performed by modeling the display input/output property explicitly. For example, some studies modeled interactions of the RGB phosphors (Bastani, Cressman, & Funt, 2005; Tamura et al., 2001). These explicit modeling approaches are most successful if the system can be described precisely, but they do not ensure that

the same approaches can be applied to different devices. In contrast, our approaches and a recent study, which used a nonlinear search method (Funt, Ghaffari, & Bastani, 2004), never assume any complex model of the display system. Therefore, our approaches are potentially applicable to any devices. Further, our methods advanced recent data-driven approaches (Ban, Yamamoto, & Ejima, 2006; Funt, Ghaffari, & Bastani, 2004) by combining the initial recursive linear estimations with a fine-tuned non-linear or line search method, which can prevent a local minimal problem and reduce estimation time. Further, these other studies have focused extensively on characterizing LCD displays, whereas we also tested DLP displays. Our results showed that for color reproduction, DLP-type projectors can be used for vision experiments provided a proper calibration procedure is performed. Finally, we are providing all the methods as an integrated, user-friendly software suite as well as describing the detailed procedures in this article. Because all of the software codes are freely available, the software itself is of benefit to the vision science community.

Olds et al. (1999) also reported a display chromaticity calibration and adjustment method that is essentially based on an idea similar to our recursive linear estimation method. Both methods assumed a piecewise linearity in a local and small color space. Notably, in their study, the estimations were considerably good even with a single iteration. In contrast, our recursive estimation method required approximately two to three times the iterations to achieve the required accuracy ( $RMSE < 1.0$ ; see Figure 11). This difference in the required number of iterations may be due to the difference in termination conditions between the two methods; the termination condition in the Olds et al. study was based on the residuals of multiple regressions, whereas our termination was set based on  $RMSEs$  (see Equations 9 and 10). In addition, the most likely reason for the difference would be the differences between the numbers of samples used in estimation step. Their method first obtained seven local chromaticity values around the target CIE1931 xyY for each of the RGB phosphors (thus,  $7 \times 7 \times 7 = 343$  chromaticity values in total). Then, the whole of these values were input to multiple regression procedures. Thus, although their method gave fairly good estimations even with a single iteration, their method required establishing relatively large samples in an initial step. In contrast, our method measured only 18 chromaticity samples for each step. We showed that even with this small population of samples, our method gave considerably good results after three iterations of the estimations. Therefore, it is not valid simply to compare the number of iterations. Finally, although the procedure is slightly different, the present study clearly extended an idea of piecewise linearity of Olds et al. (1999) by applying the

idea to non-CRT displays and exploring its efficiency in details, together with introducing novel linear/nonlinear hybrid search algorithms.

Our goal-seeking approach will be especially effective in characterizing displays with nonlinear profiles. We found that the target chromaticities could not be presented correctly with the standard global linear transformation method, even when the input/output relationship was linearized accurately (see LCD1 results in Figure 8 and Table 4). One possibility of the results will be the nonlinearity of the device; the luminance and the chromaticity values may change depending on the level of video inputs. Also, it will be possible that some display devices may work differently when RGB phosphors are lit simultaneously on (Tamura et al., 2001), whereas we characterized RGB phosphors separately in the gamma-correction step. A solution to avoid this problem will be to model nonlinearity somehow and explicitly to include that effect into the chromaticity estimations. However, nonlinearity of the system will strongly depend on the devices. Further, their internal mechanisms are generally difficult to access because of proprietary secrets. We have thus proposed alternative model-free approaches in this article; we used linear/nonlinear hybrid or line search techniques. As shown in Figure 10 and Table 4, our methods could optimize RGB video input values correctly and present chromaticities more accurately than the standard procedure without describing the internal models of display devices.

Future studies should consider the followings. First, a faster estimation procedure should be developed. Our chromaticity estimation procedures currently take 1 to 2 min to estimate a single chromaticity value, even when we set optimization parameters loosely so that the estimation could finish in a reasonable period (e.g., 50 iterations with 200 function evaluations for the Nelder-Mead Simplex search, and 30 iterations with 150 function evaluations for the Powell line search). However, if we employ a more rigid criterion, the estimations would take considerably longer, affecting the actual vision experiments conducted after the display characterization.

Second, a real-time estimation procedure should be developed. Our color estimations are performed for each chromaticity separately in advance of actual vision experiments. Therefore, our procedure will work precisely only when we have to present visual stimuli with a fixed number of chromaticities. However, we cannot present successively changing chromaticities in real time during the experiments. For example, researchers should pay special attention when they need to present visual stimuli with unrestricted chromaticity values such as staircase measurements of chromaticity thresholds. In such a case, researchers can still use our methods by estimating all the possible chromaticities

each in advance and restrict a staircase range within the pre-estimated chromaticity space. However, it will require a longer estimation period and currently is not practical. The future advanced method should overcome this shortcoming.

Finally, a fundamental question researchers want to know will be whether we can use non-CRT devices for luminance or color-related vision experiments. CRTs have been commonly and widely used in the current vision experiments. It thus would be better to use CRTs when we can use them to compare the results with those of previous studies. However, we should also keep in mind that CRT displays are not perfect devices for testing visual perception; they have superadditivity across the x-direction of pixels, and pixel-point spread functions change with luminance levels (Naiman & Farrel, 1988). Most importantly, the linearity generally assumed for CRTs is not true but just a close approximation. In addition, it is now very frequent that we need to use different types of display devices in different experiments. Therefore, we have to test the validity of non-CRT devices in detail as well as CRTs for future environments of vision science. Our results in Figure 10 and Tables 4 and 5 clearly show that non-CRT devices can be calibrated accurately for the presentation of chromaticities. Furthermore, we found that the color reproduction space of LCD2 was larger than that of the CRT tested (compare the sizes of triangles in Figure 10 between devices). This means that some non-CRT devices have a wider range of color reproduction space than CRTs, and therefore, they can present visual stimuli with a larger color saturation range. The advantages in color reproduction in LCD devices would be because the current LCDs are developed with more advanced technologies. We therefore conclude that non-CRT devices can also be used in vision experiments, provided the proper calibration procedures are applied to these devices. In addition, a recent study revealed that some LCD displays' temporal response characteristics are in fact better than a CRT (Lagroix, Yanko, & Spalek, 2012; but see also Elze, 2010). Our study additionally revealed that some non-CRT displays' color reproducibility is better than that of a CRT. Although our procedures are not without limitations, the vision experiments in the next generation may be done correctly using new types of non-CRT display devices.

## Conclusions

Vision researchers are required to develop new display characterization tools in order to handle new types of display devices such as LCD, DLP, and forthcoming organic EL and laser displays, which are

increasingly used preferentially over CRTs. The present study provides a solution to characterize these non-CRT display devices, by proposing novel display characterization procedures applicable to a wide range of devices without any assumption of the internal model of display devices. Our methods perform gamma correction using a cubic spline-based data-driven approach. The subsequent fine-scale color estimations are performed using a recursive linear estimation, a nonlinear search, or a line search algorithm. These approaches give us more accurate chromaticity estimation results for CRTs and non-CRTs alike, in comparison to the current widely used standard estimation procedure.

The procedures described in this article have been implemented into integrated GUI software, Mcalibrator2. To our knowledge, this software suite provides the first publicly available comprehensive framework to characterize a variety of display devices. The authors provide this software suite in the hope that it will benefit researchers performing calibration of their display devices efficiently and improve accuracies of stimulus displays regardless of the display types.

*Keywords:* display characterization, gamma correction, luminance, chromaticity, CIE1931, psychophysics software, imaging software

## Acknowledgments

We would like to thank M. L. Patten, Y. Ejima, and S. Takahashi for comments on this study and manuscript; O. Hashimoto and his colleagues at Namoto Trading Company Japan for their technical support; and C. Hiramatsu for testing our software and procedures in different display environments and for helpful comments. We are grateful to E. Tarajan, G. Flandin, D. Garcia, J. Lundgren, O. Salvado, and E. Farhi for their contributions on developing MATLAB mathematical subroutines and GUI-development tools. This work was supported by the Fellowships of Japan Society for the Promotion of Science to H. B. (H22.220), Grant-in-Aid for Scientific Research on Innovative Areas (23135517) from the Ministry of Education, Culture, Sports, Science and Technology of Japan and Grants-in-Aid for Scientific Research (22530793) from the Japan Society for the Promotion of Science to H. Y.

Commercial relationships: none.

Corresponding authors: Hiroshi Ban; Hiroki Yamamoto.

Email: ban.hiroshi+mcalibrator@gmail.com;

yamamoto@cv.jinkan.kyoto-u.ac.jp.

Address: Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan.

## References

- Ban, H., Yamamoto, H., & Ejima, Y. (2006). Mcalibrator: MATLAB integrated GUI software for display calibration—A proposal of a new color calibration procedure applicable to a wide range of display devices and evaluation of its efficiency. *Japanese Journal of Psychonomic Science*, *24*, 149–161.
- Ban, H., Yamamoto, H., Fukunaga, M., Nakagoshi, A., Umeda, M., Tanaka, C., et al. (2006). Toward a common circle: Interhemispheric contextual modulation in human early visual areas. *Journal of Neuroscience*, *26*, 8804–8809.
- Bastani, B., Cressman, B., & Funt, B. (2005). Calibrated color mapping between LCD and CRT displays: A case study. *Color Research & Application*, *30*, 438–447.
- Berns, R. S. (1996). Methods for characterizing CRT displays. *Displays*, *16*, 173–182.
- Besuijen, J. (2005). The shifting gamma perception. *Color Research & Application*, *30*, 332–340.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Brainard, D. H., Pelli, D. G., & Robson, T. (2002). *Display characterization. Encyclopedia of imaging science and technology*. Hoboken, NJ: Wiley.
- Brent, R. P. (1973). *Algorithms for minimization without derivatives*. Englewood Cliffs, NJ: Prentice-Hall.
- Day, E. A. (2002). *Colorimetric characterization of a computer controlled [sgi] CRT display*. Technical report of a joint research program of the National Gallery of Art (Washington), the Museum of Modern Art (New York), and Rochester Institute of Technology.
- Deguchi, T., & Katoh, N. (1998). Effect of user controls on CRT monitor characteristics. *SPIE Proceedings*, 3300. Internet site: <http://dx.doi.org/10.1117/12.298283>.
- Dennis, J. E., Jr., & Woods, D. J. (1987). *New computing environments: Microcomputers in large-scale computing*. Philadelphia: Society for Industrial and Applied Mathematics.
- Elze, T. (2010). Misspecifications of stimulus presentation durations in experimental psychology: A systematic review of the psychophysics literature. *PLoS ONE*, *5*(9), e12792.
- Fairchild, M. D., & Wyble, D. R. (1998). Colorimetric characterization of the Apple studio display (flat panel LCD) (Technical report). Rochester, NY: Munsell Color Science Laboratory.
- Farhi, E. (2011). The iFit data analysis library, ILL, computing for science. Internet site: <http://ifit.mccode.org>.
- Farhi, E., Debab, Y., & Willendrup, P. (2013). iFit: A new data analysis framework. Applications for data reduction and optimization of neutron scattering instrument simulations with McStas. *Journal of Neutron Research*, *17*(1), doi:10.3233/JNR-130001.
- Funt, B., Ghaffari, R., & Bastani, B. (2004). Optimal linear RGB-to-XYZ mapping for color display calibration. *Proceedings of 12th Color Imaging Conference, Color Science, Systems & Applications*.
- Garcia, D. (2010). Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational Statistics & Data Analysis*, *54*, 1167–1178.
- Gibson, J., Fiarchild, G., & Fairchild, M.D. (2000). *Colorimetric characterization of three computer displays (LCD and CRT)* (Technical report). Rochester, NY: Munsell Color Science Laboratory.
- Ikeda, H., & Kato, N. (2000). Part 4: Equipment using liquid crystal display panels. *Multimedia systems and equipment: Color measurement and management*. International Electrotechnical Commission, IEC 61966-4 ed1.0.
- Katoh, N. (2002). *Corresponding color reproduction from softcopy images to hardcopy images* (Unpublished Ph.D. dissertation). Chiba University, Chiba, Japan.
- Kwak, Y., & MacDonald, L. (2000). Characterization of a desktop LCD projector. *Displays*, *21*, 179–194.
- Lagroix, H. E. P., Yanko, M. R., & Spalek, T. M. (2012). LCDs are better: Psychophysical and photometric estimates of the temporal characteristic of CRT and LCD monitors. *Attention, Perception and Psychophysics*, *74*, 1033–1041.
- Naiman, A., & Farrell, J. (1988). Modeling the display and perception of grayscale characters. *SID Digest*, *26*, 424–427.
- Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, *7*, 308–313.
- Olds, E. S., Cowan, W. B., & Jolicoeur, P. (1999). Effective color CRT calibration technique for perception research. *Journal of Optical Society of America*, *16*, 1501–1505.
- Pelli, D. G. (1997). The video toolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Powell, M. J. D. (1964). An efficient method for finding the minimum of a function of several variables

- without calculating derivatives. *Computer Journal*, 7, 155–162.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). Van Wijngaarden–Dekker–Brent method. In *Numerical recipes: The art of scientific computing* (3rd ed., pp. 454–455). New York: Cambridge University Press.
- Tamura, N., Tsumura, N., & Miyake, Y. (2001). Calibration of LCD colorimetry considering channel interaction: Special section of display-special contribution from Asia display/IDW '01. *ITE Journal*, 51(10), 1–5.
- Thomas, J. B., Hardeberg, J. Y., Foucherot, I., & Gouton, P. (2007). Additivity based LC display color characterization. *Proceeding of Gjobvik Color Imaging Symposium*, 4, 50–55.