The Shifting Gamma Perception

J. Besuijen

Department of Psychonomy and Human Performance Technology, Faculty of Behavioral Sciences, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

Received 26 May 2004; revised 25 September 2004; accepted 25 March 2005

Abstract: True to life, color display and color management depend on a proper technical model of the display used. Current gamma models and fitting procedures are not accurate in modeling the lower part of the tone reproduction curve. The GOG- and GOGO-model used in color management standards tend to clip the luminance to zero for digital values were luminance can be seen and measured. Two improvements to the models are suggested. First, the models should be fitted by optimizing the root mean square error (RMSE) of the CIE lightness instead of the luminance. Second, a shifting gamma model is adopted, with gamma increasing in value for lower voltages. Results show that the adapted models correspond better with the luminance measurements. The clipping values are nearer to the measured zero luminance threshold, and the average RMSE and ΔE^*_{ab} over the whole scale are smaller. Col Res Appl, 30, 332–340, 2005; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/col.20137

Key words: gamma; color management systems

INTRODUCTION

Electronic displays are rapidly becoming part of nearly all human-machine interfaces. A substantial amount of visual inspection tasks, which used to be done on paper, on photo, or in real life, are now done on electronic displays. Some of these tasks set high demands to the reliability of the pictures shown on the displays. The visible impression caused by the pictures should be consistent for the viewers over different systems, with different graphics cards, different displays, and even different display technologies. Exchange of graphical designs, (remote) medical diagnostics, preparation and projection of presentations, and online shopping are examples of applications where visible display quality is a critical criterion. Various display standards and color management systems have been developed to preserve the correct representation of pictures. Many of these use gamma, the exponent of the power function describing the relation between input voltage and light output of a cathode ray tube (CRT) display, as the main parameter in the display's tone reproduction curve (TRC). The merits of gamma are not common knowledge: "Owing to poor understanding of TRC, and to misconceptions about nonlinear coding, gamma has acquired a terrible reputation in computer graphics and image processing," Poynton stated in one of his persistent efforts to explain gamma and clear its reputation.[1] The nonlinearity of the CRT is not a defect, but a highly desirable feature, because it is very nearly the inverse of the lightness sensitivity of human vision. The nonlinearity causes a CRT's response to be roughly perceptually uniform.

In a video camera, digital still camera, or scanner, a nonlinear transfer function—gamma correction or inverse gamma coding, similar to the CIE L^* function of color science—is imposed. This provides an optimal perceptual performance with a limited number of bits throughout recording, storage, processing, compression, transmission, decompression, and presentation.[1]

The use of gamma has several other merits:

1. It characterizes the nonlinearity of the TRC with a single meaningful parameter, unlike the parameters of cubic splines or polynomial fittings.
2. It relates to the technical characteristics of the main display technology CRT.

But with the advances in measurement methods and technologies, the number and relevance of the disadvantages that have come to light are growing:

1. The single gamma model for CRT technology is disputed. Olson[2] showed that gamma could be much higher (9.5) for low voltages than for high voltages (1.5).
2. The CRT has lost its supremacy in the market, and the liquid crystal display, the projected main technology for at least the next decennium, has a dissimilar TRC.[3,4] And so have the main other contending dis-

Correspondence to: J. Besuijen (e-mail: j.besuijen@gw.utwente.nl)

play technologies PDP and OLED. The non-CRT display technologies therefore need a lookup table correction to display the images encoded for display on CRTs, which is the actual image encoding standard.
3. Several platforms, Macintosh and SGI, use a default partial gamma correction of 1/1.4 and 1/1.7, respectively, resulting in display system gammas, that is , the gamma from the stored digital image to the image on display, of about 1.8 and 1.47. For an extensive review of the consequences see Poynton.[1]
4. Recent models for gamma fitting (GOG, GOGO) are no longer single parameter models but have two or three parameters.[5,6]
5. The gamma figure is not an absolute standard; the sRGB and IEC standard[7] has a nominal gamma of 2.2 (single parameter) for the characterization of the standard display, but uses a power of 2.4 in its formulas (with an extra gain/offset parameter).
6. The meaning of the value of the gamma is still confusing to the average user. It is unclear in which cases gamma correction is needed, and the term gamma is often used for inverse gamma as well.
7. Other curve fitting techniques like cubic splines give better fitting results.
8. The human brightness perception curve differs substantially from the gamma curve, especially for lower light levels.
9. Different standard curves based on human brightness perception are already developed (*e.g.* the Grayscale Standard Display Function for medical applications).[8]

These in turn raise questions whether gamma is still the best way to characterize a display and whether users are able to handle the gamma concept to control the brightness and contrast reproduction on their displays.

THE SIMPLE GAMMA MODEL

The basic gamma model relates the electron beam current I of a CRT gun to the input voltage E with a simple power function with exponent γ:

$$I = E^{\gamma} \tag{1}$$

The relation between the beam current and the phosphor luminance is nearly linear,[2,6] leading to the one-parameter gamma function for the relation between digital code value d (or gray level) and the display luminance Y:

$$Y = \left(E_{\max}\frac{d}{d_{\max}}\right)^{\gamma} = Y_{\max}\left(\frac{d}{d_{\max}}\right)^{\gamma} \tag{2}$$

if the digital to analog conversion form (digital code value d to voltage) is linear. The simplest solution to determine the exponent γ is by fitting a straight line through the measurement points in the log-luminance–log-voltage domain.[9] Roberts[10] and Berns[5] showed that small offsets in voltage or luminance result in curves at the ends of the straight line that have a severe impact on gamma found with this fitting method.

A positive voltage offset results in a luminance offset, a luminance for digital code value 0 and a low estimate for gamma in the log–log domain. A negative voltage offset results for low gray levels in too low voltages for the CRT to produce electrons and therefore luminance, with the result that all digital code values below the threshold are mapped to luminance 0. If this offset is not compensated the gamma fitted will be too high. In practice optimal black level adjustment is nearly impossible and small offsets will remain. Even small offsets can have considerable impact on the fitted gamma.[10,5]

Not all luminance offsets result from voltage offsets; they can also be caused by internal flare or veiling glare also called external flare (see below). Berns[11] also argued that differences for unreliable low luminance measurements are overrated in the log–log domain.

The Dark Side of the Scale

Correctly measuring the low light levels linked with the lower digital levels on the graphics card has proven to be a problem for much research on display characterization. In color measurement with spectroradiometer, the incoming light is divided into many spectral bands reducing the sensitivity for low light levels considerably. In colorimeters, where the light is divided over three or more colored filters on as many sensors, the light level per sensor is substantially higher, but deviations in the spectral transmission curves can introduce considerable errors, especially for narrow band primaries, like the red primary of most CRTs. If luminance is measured with a photometer, measurements for the green primary will generally be accurate and sensitive, but the blue and red primaries fall on the sides of the spectral photometric band and their light is considerably attenuated by the photometric filter. The original objective of the research project was to develop a fast low cost display characterization maintenance method based on a combination of the merits of a contact photometer and a spectroradiometer. The spectroradiometer, a Photo Research PR-650, would be used to measure color coordinates and peak luminances. The contact photometer, a Macam L203, would be used to measure primary luminance profiles with a precision of 0.001 cd/m^2. The reliable spectroradiometer peak luminance measurements should be used to correct the photometer measurements, especially for the red and blue primary, with a scale factor.

THE GAIN-OFFSET-GAMMA MODEL (GOG)

Together with accurate luminance measurements for low light levels, the inclusion of the voltage offset in the gamma model should provide a better display characterization. Berns[5] developed a model for the amount of spectral radiant exitance $M_{\lambda,R}$ generated by a CRT gun, which incorporates an offset voltage and a gain parameter. The model derivation is limited to the red channel denoted by the subscript R.

$$M_{\lambda,R} = k_{\lambda,R}\left\{a_R\left[(v_{\max} - v_{\min})\left(\frac{d}{d_{\max}}\right) + v_{\min}\right] + b_R - v_{C,R}\right\}^{\gamma_R} \quad (3)$$

The amount of spectral radiant exitance of a computer-controlled display depends on the digital counts in the digital to analog converter (DAC) (d, d_{max}), the video generator (v_{min}, v_{max}), the video amplifier (a_R, b_R), the CRT ($v_{C,R}$, γ_R), and the properties of the faceplate and phosphor materials ($k_{\lambda,R}$). Note that a_R and b_R are supposed to correspond with the "contrast" and "brightness" settings of the monitor, and that γ_R is assumed constant under all conditions. By normalization the number of parameters can be reduced, leaving a normalized system gain term $k_{g,R}$ and offset term $k_{o,R}$, and producing a normalized tristimulus output:

$$\begin{aligned} R &= \left(k_{g,R}\frac{d}{d_{\max}} + k_{0,R}\right)^{\gamma_R} & \left(k_{g,R}\frac{d}{d_{\max}} + k_{0,R}\right) &\geq 0 \\ R &= 0 & \left(k_{g,R}\frac{d}{d_{\max}} + k_{0,R}\right) &< 0 \end{aligned} \quad (4)$$

$$k_{g,R} = \frac{a_R(v_{\max} - v_{\min})}{a_R v_{\max} + b_R - v_{C,R}} \quad (5)$$

$$k_{o,R} = \frac{a_R v_{\min} + b_R - v_{C,R}}{a_R v_{\max} + b_R - v_{C,R}} \qquad k_{g,R} = 1 - k_{0,R} \quad (6)$$

Notice that the gain is dependent on the monitor offset setting ("brightness"), and the offset is dependent on the monitor gain setting ("contrast"), which could explain the problems faced by the users to correctly set their display configuration.

This GOG model was expanded by Katoh[12] to the GOGO-model with a luminance offset parameter $Y_{0,R}$ for extra luminance caused by flare or glare.

$$Y_R = (Y_{\max,R} - Y_{0,R})\left(k_{g,R}\frac{d}{d_{\max}} + 1 - k_{g,R}\right)^{\gamma_R} + Y_{0,R} \quad (7)$$

Veiling Glare and Flare

Veiling glare and internal flare can disturb the light measurements. Veiling glare or external flare is caused by ambient illumination reflecting in the front glass plate of the display. When spot measurements are used, external flare can be removed by darkening the room or it can be neutralized by subtraction, but only if the source is constant and at the expense of an extra error source. With contact measurements, covering the immediate surround of the measuring probe on the glass plate will suffice to eliminate external flare. Internal flare is harder to counter, but because the sources are internal they should be included in the monitor model. Internal flare can have three sources as follows[11]:

1. Reflections in the glass face plate from light generated elsewhere on the display surface, also called neighboring pixel interreflections.
2. Reflections from light generated by electrons deflected from their original destiny, also called "secondary emissions."
3. Black level luminance caused by one of the other primaries, also called cross-channel emissions or unwanted emissions.[13]

If the background around the measurement stimulus is (nearly) black the internally reflected light will have the same color for a primary (R, G, or B) measurement stimulus. The amount of reflected light does depend on the size of the stimulus: the highest contribution will come from inside and just outside the measurement field of the light meter. In practice, it is impossible to match exactly the measurement stimulus to the measurement field of the instrument used. The risk of a stimulus being too small is far greater than the accuracy gained by an exact match.

The secondary emissions could add light from other primaries to the stimulus, but if these emissions are linearly proportional to the amount of electrons reaching the shadow mask, the effect on the measured color of a primary should be the same for all input levels. As the flare from these sources is presumably small and proportional to the electron current of the gun under measurement, for the purpose of display characterization, these forms of internal flare should be regarded as an inseparable part of the model.

Neither of the above sources will have any effect on the black level luminance ((R, G, B) = (0, 0, 0)). The luminance offset at black level is obviously the same for all the three primaries. And as most instruments are not able to measure color at these low luminances, it will be difficult if not impossible to determine the source. Most monitors do not support external black level ("brightness") adjustment for separate primaries. Just subtracting the measured luminance at (0, 0, 0) from all measurements is not the solution, because the offset could be partly resulting from the primary under measurement, that is, if $k_{o,R}$ is positive. If displays are used in brightly illuminated places, a higher black level setting might be desired to be able to distinguish dark colors.

In the GOGO-model, Y_{max} is a fixed value that can be measured and the luminance for the input $d = d_{max}$. Y_o is a model parameter, and for a characterization with separate primaries it typically does not equal the luminance for $d = 0$, because part of the luminance is caused by the voltage offset (if $k_g < 1$) of the target primary.

The model parameters are typically estimated by measuring the luminance for 9 (0, 32, 64, . . . , 224, 255) or 17 equi-stepped code values* (*e.g.*, 0, 16, 32, . . . , 240, 255) and performing a nonlinear optimization on the mean square

* The last steps, 31 and 15, in both these code value sequences are of course smaller than the others, 32 and 16. A more elegant solution for an 8-bit system, suggested by Charles Poynton, would be to use 18 really equi-stepped code values with all steps equal to 15. However, this is not a general solution as it cannot be expanded to 10- and 12-bit display systems. For the measurements in this article, 17 steps starting with the smaller step: 0, 15, 31, . . . , 239, 255, were used. This way more bits are involved in the measurements, thereby minimizing the influence of calibration errors in individual bit conversions of the DAC.

TABLE I. Results for different model fits to the measured data of the *R, G,* and *B* primaries of the delta-gun CRT.

	γ	Range	k_g	L_0	Lum RMSE	CIEL RMSE	Clip value
Delta gun R, method:							
GOGO Lum	1.938		1.151	0.036	0.0181	6.604	33.45
GOGO CIEL	2.128		1.083	0.006	0.0272	1.634	19.54
GLID Lum	2.491	0.693	0.978	0.000	0.0072	2.613	−5.74
GLID CIEL	2.245	0.323	1.045	0.002	0.0118	0.704	10.98
Delta gun G, method:							
GOGO Lum	2.015		1.140	0.052	0.0074	2.940	31.32
GOGO CIEL	2.127		1.101	0.012	0.0139	1.065	23.39
GLID Lum	2.240	0.276	1.066	0.000	0.0021	0.498	15.79
GLID CIEL	2.205	0.230	1.077	0.005	0.0024	0.277	18.23
Delta gun B, method:							
GOGO Lum	2.088		1.157	0.007	0.0084	3.646	34.60
GOGO CIEL	2.209		1.118	0.003	0.0157	1.190	26.91
GLID Lum	2.378	0.345	1.065	0.001	0.0031	1.087	15.56
GLID CIEL	2.282	0.219	1.092	0.002	0.0042	0.458	21.48

error (MSE) in the luminance domain. The resulting display characterization has an adequate color reproduction with average errors reported of about $\Delta E^*_{ab} = 0.4$[5] and $\Delta E^*_{ab} = 0.8$[11] depending on the color set used for testing.

Measurements

The R, G, and B primary luminance profiles of a 21″ Eizo Flexscan 780i-W delta gun and several Iiyama 17″ Trinitron CRTs were measured with the MACAM L203 photometer. Ambient illumination was excluded by a dark grey foam cover. The linearity of the photometer was checked by comparisons with measurements through neutral density filters: no trend was visible. The displays had at least 1 hour warm-up time. The time between stimulus change and luminance measurement was empirically optimized to 5 seconds.

Again the Dark Side

In using the GOG(O)-model for display characterization, the results are not completely satisfying: the model seems to work well for the high end of the digital code scale, but has serious defects on the low end. Even if luminance and luminance change can be perceived and measured in the lowest tenth of the digital code scale for a particular CRT gun, the fitted model tends to map these densities to zero luminance. Especially for displays were the black level is set correctly, with minimal luminance at digital code 0 and measurable luminance increase near digital code 10, the fitted solution could map about one tenth of the digital code scale to zero.

Measurements on a delta-gun CRT display with these characteristics were used to produce the model fitting results in Table I: the rows marked “GOGO lum” show the results for the GOGO-model, with clipping thresholds above 30. For this display and graphics card, this would mean a reduction of the amount of possible colors from 16.7 million to 11 million. This is not an isolated case; results for other displays showed the same discrepancy between observed luminance and fitted clipping values. The effect can also be found in the data reported by Berns[14] where luminance was measured for low digital counts (<10), but clipping values, d_c, of 24, 29, and 18 for R, G, and B are fitted by the GOGO model (see Table II). This could lead to considerable color differences, especially if one or more of the primary densities is near threshold. These densities are rarely used in tests. Berns tested his display characterization either on color sets with factorial designs with digital counts of 0, 559, 755, 903, and 1023 (10 bits)[5] or with digital counts of 0, 85, and 255 (8 bits).[13]

The same pattern can be observed for the Trinitron display tested (Table III): the black level setting is nearly optimal, but nearly one tenth of the values is mapped to zero.

Minimizing the Error in Perceptual Space

An obvious cause for the failure of the model fitting could be that the space in which the error is minimized does not represent the perceptual reality. For luminance, only the relative weight of the spectral components is defined according to perception, but the luminance scale is not linearly proportional to brightness perception. The log–log domain might overrate the importance of small luminance values; the luminance domain underrates it. The small luminance values are at least equally important in modeling the gamma function, and maybe even more because the maximum digital value with zero luminance determines the size of the voltage offset. In the GOG(O)-model an error of 0.1 cd/m^2 is equally important for luminances of 100 cd/m^2 and 0.1 cd/m^2. If the available measurement equipment does not provide reliable low luminance measurements, then the measurements should not be used in the model-fitting procedure or if possible the reliability should be improved by averaging repeated measurements.

TABLE II. Results for different model fits to the data of the *R, G,* and *B* primaries of the CRT measured by Berns.[14]

	γ	Range	k_g	L_0	Lum RMSE	CIEL RMSE	Clip value
Berns R, method:							
GOGO Lum	1.618		1.097	0.000	0.0140	5.074	22.55
GOGO CIEL	1.667		1.086	0.002	0.0186	1.562	20.19
GLID Lum	1.671	0.116	1.080	0.002	0.0117	1.038	18.89
GLID CIEL	1.695	0.179	1.071	0.002	0.0123	0.820	16.90
Berns G, method:							
GOGO Lum	1.546		1.136	0.005	0.0120	5.325	30.53
GOGO CIEL	1.661		1.091	0.002	0.0242	1.857	21.27
GLID Lum	1.688	0.239	1.075	0.002	0.0032	0.398	17.79
GLID CIEL	1.698	0.263	1.072	0.002	0.0035	0.304	17.13
Berns B, method:							
GOGO Lum	1.539		1.099	0.005	0.0123	5.848	22.97
GOGO CIEL	1.633		1.067	0.003	0.0245	2.125	16.01
GLID Lum	1.675	0.231	1.042	0.002	0.0042	0.840	10.28
GLID CIEL	1.668	0.227	1.044	0.002	0.0046	0.812	10.75

Minimization of the error in the CIE lightness domain would be preferable as the eventual goodness of fit is determined by color differences ΔE_{uv} or ΔE_{ab} in the CIELUV or CIELAB space, where lightness is one of the three dimensions. Another option is the Digital Imaging and Communications in Medicine (DICOM) Grayscale Standard Display Function (GSDF)[8] based on a cubic spline fit of a vision model developed by Barten[15] (see Appendix). This vision model was made to fit data of just-noticeable differences in luminance modulation. The GSDF does not pass through the origin of the grayscale-luminance plane, which is not a helpful feature for an error function in an optimization procedure. Moreover the ten eight-digit constants in this standard, and the eight/nine polynomial terms of log luminance in the inverse formula make this function too complex for an iterative procedure.

The CIE lightness L^* is a relative measure of the luminance of a color Y to the luminance of a reference white Y_n. For a display the obvious choice for Y_n is the luminance of white on the display.

$$L^* = 116\left(\frac{Y}{Y_n}\right)^{1/3} - 16 \quad \text{if } Y/Y_n > 0.008856 \text{ and}$$
$$L^* = 903.3\left(\frac{Y}{Y_n}\right) \quad \text{otherwise.} \tag{8}$$

As shown in the rows marked "GOGO CIEL" in Tables I–III, an optimization in the lightness domain lowers the clipping threshold, whereas the decrease in lightness RMSE outweighs the increase in luminance RMSE. Because the voltage offset is lowered, the gain is decreased and gamma is increased. The fit for the luminance offset is more compatible over the three primaries.

Correcting the Physical Model

So the fitting in the lightness domain yields an improvement in the quality of display characterization, but on closer inspection several effects remain unexplained.

For simulated GOGO-model data with different voltage or luminance offsets and a little added noise, the fitting

TABLE III. Results for different model fits to the measured data of the *R, G,* and *B* primaries of a trinitron CRT.

	γ	Range	k_g	L_0	Lum RMSE	CIEL RMSE	Clip value
Trinitron R, method:							
GOGO Lum	2.161		1.094	0.022	0.0054	2.121	21.91
GOGO CIEL	2.236		1.069	0.006	0.0088	0.666	16.46
Glid Lum	2.344	0.207	1.037	0.000	0.0020	0.709	9.10
Glid CIEL	2.290	0.139	1.052	0.003	0.0026	0.251	12.60
Trinitron G, method:							
GOGO Lum	2.034		1.104	0.000	0.0074	3.164	24.02
GOGO CIEL	2.133		1.074	0.000	0.0139	1.126	17.57
Glid Lum	2.220	0.246	1.047	0.000	0.0018	0.428	11.45
Glid CIEL	2.204	0.229	1.052	0.000	0.0019	0.303	12.60
Trinitron B, method:							
GOGO Lum	2.078		1.071	0.000	0.0079	3.965	16.90
GOGO CIEL	2.196		1.039	0.000	0.0177	1.371	9.57
Glid Lum	2.293	0.274	1.006	0.000	0.0014	0.519	1.52
Glid CIEL	2.260	0.227	1.015	0.000	0.0020	0.262	3.77

procedure never failed to find the model parameters. This leads to the conclusion that the GOGO-model might not correctly describe the relation between input voltage and output luminance of a CRT. The addition of gain and offset parameters has undoubtedly improved the gamma model, but it cannot correctly predict the luminance output over the whole input voltage range.

A possible explanation is that the century old single-gamma model for CRTs is too simple to model the physical reality, and not all the differences in gamma measurement results can be attributed to differences in voltage offsets. This is exactly what Olson[2] established in extensive tests on CRTs over the whole input voltage and electron beam current domain. He found that gamma varied over the voltage domain from values as high as 9.5 on the low end to a minimum of 1.5 on the high end of the scale, with a range of about linear decrease in the middle. His measurements far exceeded the range of normal display operation and in his paper the normal range of voltage operation of the measured displays is not specified, and luminance was not measured, but it seems reasonable to assume the normal display operating range to be in the nearly linear decreasing gamma region. Berns[11] argues that changes in gamma only occur at very low voltages and luminances, and the effects could hardly be measured and would be too small to be perceived. But Olson's findings place the normal gamma of about 2.2 in an area where gamma is rapidly decreasing with increasing voltage, and Olson concludes that a simple gamma does not suffice for the purpose of high-resolution film recording.

An indicative test for the changing gamma hypothesis can be performed by splitting the data samples in an upper 128–255 and a lower 0–128 half. For all the tested displays the gamma fitted in a GOGO-model for the lower half was invariably higher than the gamma fitted for the upper half.

The change in gamma can also be visualized by plotting the apparent gamma, a differentïal measure characterizing the slope at one point of the curve.

$$\gamma_{app} = \frac{\log(Y_{i+1} - Y_0) - \log(Y_i - Y_0)}{\log(d_{i+1} - d_c) - \log(d_i - d_c)} \quad (9)$$

The apparent gamma is extremely sensitive to noise in the measurements and to a correct estimation of the voltage offset d_c. In Fig. 1, a moving average over 9 samples of a full scale (0, 1, 2, . . . , 255) measurement of a delta-gun CRT is plotted. Even then the small peaks caused by slight misalignments in the graphics card DAC (at 64, 128, and 192) and between different luminance meter scales (at 197) are clearly visible. It also shows the influence of different voltage offset estimations on the apparent gamma. If the GOG(O)-model is correct, then with the right voltage offset the apparent gamma would be a straight horizontal line at gamma level. In the figure, the lines are either curved or show a decreasing gamma for increasing digital counts in the range 60–255. These curves can be compared with the apparent gamma for simulated data shown in Fig. 2. Comparison shows that the measured and offset corrected curves in Fig. 1 have shapes similar to the curves with a descending gamma in Fig. 2. Instead of using Olson's complex formula, with voltage constants that should be measured inside the display housing, these curves are based on the GOGO-model function with a linearly decreasing gamma over the luminance range of the CRT. This linear shift can be characterized by the parameter A_R denoting the range of gamma variation in the gliding gamma model.

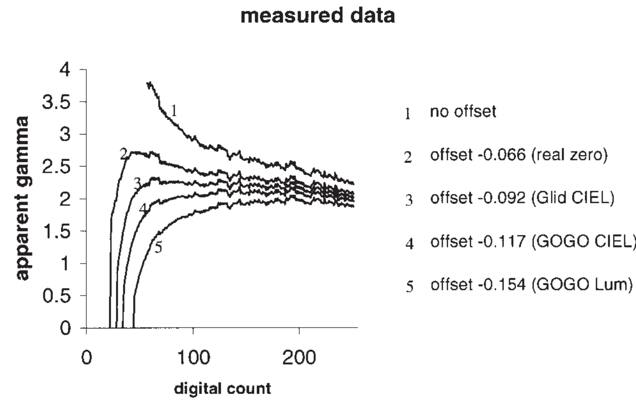

FIG. 1. Apparent gamma for measured data with different values for offset and gamma range as provided by the different model fits. The descriptions in the legend have the same top-to-bottom order as the curves at digital count 100.

$$\begin{aligned} Y_{i,R} &= Y_{0,R} && \text{if } k_{g,R}\frac{d_i}{d_{max}} + 1 - k_{g,R} \le 0 \\ Y_{i,R} &= \left(k_{g,R}\frac{d_i}{d_{max}} + 1 - k_{g,R}\right)^{\gamma + A_R(0.5 d_i/d_{max})} + Y_{0,R} && \text{otherwise} \end{aligned} \quad (10)$$

As shown in Tables I–III, the gliding gamma model is better fitted to explain the variation in the luminance measurements. The total root mean square error is smaller in the CIE lightness domain as well as in the luminance domain. The clipping value is nearer to the apparent threshold in the measurements. For the delta gun display that would mean an increase in the amount of possible colors from 11 to 13.5 out of the maximum 16.7 million.

Computational Complexity

The introduction of a fourth parameter in the optimization procedure raises the chances of divergence or finding a suboptimal solution. We used an LU-decomposition method with the following constraint: $Y_0 \ge 0$ and $Y(d_{max}) = Y_{max}$. The first is a physical constraint: luminances cannot be negative. The latter is a matter of computational efficiency

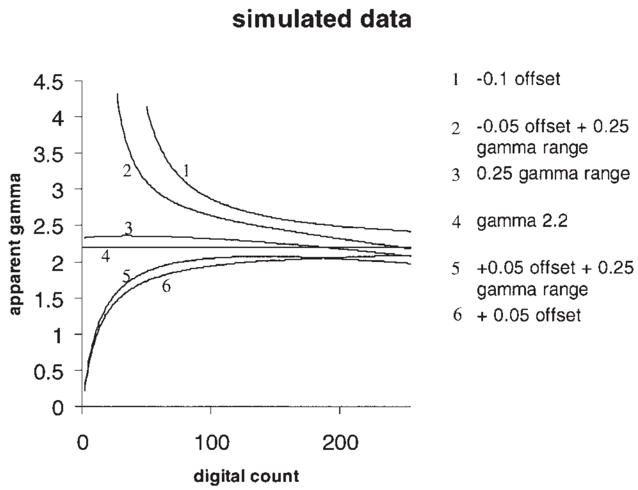

FIG. 2. Apparent gamma for simulated data with gamma 2.2 and different values for offset and gamma range. The descriptions in the legend have the same top-to-bottom order as the curves at digital count 100.

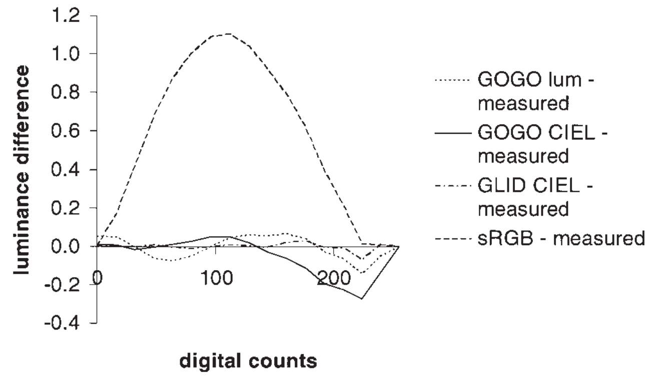

FIG. 3. Luminance differences between measured data and fits for different models for the green channel of the delta-gun CRT and the sRGB standard. The GLID lum data are very close to the GLID CIEL data and therefore left out.

in color generation: there is no need to check if the generated color exceeds the maximum luminance for a primary.

Estimating the starting values for the parameters in the minimization can speed up the procedure and raises the chances of finding the right solution, with more accurate results. If there is any offset luminance, the most likely source is a voltage offset for the other primaries. It is reasonable to assume that the voltage offset for each primary is proportional to its maximum luminance.

$$Y_{0,R,\mathrm{start}} = Y_0 \frac{Y_{\max} - Y_0 - Y_{\max,R}}{Y_{\max} - Y_0} \tag{11}$$

The clipping value and therefore the gain factor can be estimated by extrapolating the power function from the digital value d_{z1}, with the lowest luminance measurement higher than Y_0, Y_{z1}.

$$d_{c,R,\mathrm{start}} = d_{z1,R} - 255\left(\frac{Y_{z1,R} - Y_{0,R}}{Y_{\max,R} - Y_{0,R}}\right)^{1/\gamma} \tag{12}$$

$$k_{g,R,\mathrm{start}} = \frac{d_{\max}}{d_{\max} - d_{c,R}} \tag{13}$$

with γ the nominal system gamma: 2.2 for PC, 1.8 for Macintosh, etc. A starting value for γ could be computed from the measured values at about 40% and 60% of the digital code scale. For a graphics card with 256 values and 17-step sampling this would be given as follows:

$$\gamma_{R,\mathrm{start}} = \frac{\log(Y_{144,R} - Y_{0,R}) - \log(Y_{112,R} - Y_{0,R})}{\log(144 - d_{c,R}) - \log(112 - d_{c,R})} \tag{14}$$

In practice, this gamma estimation is not reliable enough due to the cumulative errors of Y_0 and $d_{c,R}$ and the noise enhancing effect of the log-difference measure. Nominal gamma values, if available, generally provide better starting values. For the gamma range a value of 0.25 appeared to be a good starting point.

These starting values provide suitable approximations of the final parameters in the optimization; they cannot guarantee an optimal solution. However, the recognition of an optimal solution is not difficult. For a 17 sample fit, the total CIE Lightness RMSE should be smaller than 1. Correct solutions from other fitting methods can be used as starting values.

DISCUSSION

The gliding gamma model with optimization in the CIE lightness domain provides a better characterization of the CRT display. Figures 3 and 4 show the error of the fits for different models in the luminance and the CIE-lightness domain, respectively. The GOGO model fit optimized on luminance error performs reasonably well in the luminance domain, but has substantial errors for low digital counts in the lightness domain. Adapting the optimization to the lightness domain provides a smaller average error, which is more

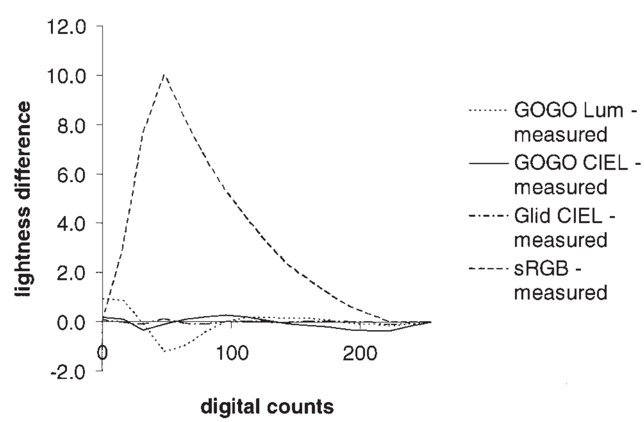

FIG. 4. CIE lightness differences between measured data and fits for different models for the green channel of the delta-gun CRT and the sRGB standard. The GLID lum data are very close to the GLID CIEL data and therefore left out.

spread over the whole domain. The error for the gliding gamma model with optimization in the lightness domain is hardly noticeable in the lightness as well as in the luminance domain. The results in the tables and in a great number of measurements and fits not reported here, show the model is consistent with the physical reality of the CRT configuration and gives good estimations of voltage offsets and luminance offsets caused by flare.

The improved model comes at the cost of an extra parameter describing the linear range over which gamma varies. The gamma shift poses problems in determining the inverse function needed to compute the digital counts for each primary (R, G, B) from the CIE-coordinates of the desired color. This can be solved by using a short iterative procedure, but a more common approach is working with lookup tables. Both methods add to the computational complexity of correct color generation.

On the whole the shifting gamma seems to make display characterization more complex, and take us farther away from the most important advantage of the gamma model methods: the single parameter display characterization. Four parameters are needed to describe the model and gamma is no longer fixed. But a closer look at the results in Tables I–III shows that the central gamma value in the gliding gamma model fits has less variation over the primaries and is closer to the nominal gamma values than the fixed gamma in the GOGO fits. So the concept of gamma as a single figure describing the TRC of a CRT still stands.

CONCLUSION

The two methods suggested clearly improve the models fitted to the CRT's TRC. The optimization in the lightness domain enhances the perceptual validity of the models. These could be further refined by making allowances for the viewing intent, that is, the intended view-surround contrast conditions (*e.g.*, office, living room television, or cinema lighting conditions).

The voltage dependency of CRT gamma, detected and physically explained by Olson,[2] can have a measurable effect on the TRCs within the luminance operating range of office CRTs. The gliding gamma model provides more accurate fits for these TRCs in the lightness and luminance domain.

The gliding gamma model appears too complex to impose it on other display technologies. A new standard should be developed rather than trying to fit the TRCs of the new technologies to the technical oddities of the CRT. As shown in Figs. 3 and 4 the sRGB standard does not do a great job of describing the TRC of at least one sample of the CRT technology it is meant to represent. The "physical properties" of the display technologies should not be the base for color management standards, and these should be based on human lightness perception, with CIE lightness the most likely candidate. New technologies already have lookup tables to correct their technical TRCs and these could easily be filled with lightness correction curves. And with the acceptance of the digital display interface there should be no problem to integrate lightness correction LUTs in the CRT. The gliding gamma model can then be used to compute the correction tables.

APPENDIX

The DICOM Grayscale Standard Display Function (GSDF)[8] is defined by a mathematical interpolation of 1023 luminance levels derived from Barten's model.[15] The GSDF calculates the luminance, L, in candelas per square meter, as a function of the just-noticeable difference (JND) Index, j:

$$\log_{10} L(j) = \frac{a + c.\ln(j) + e.(\ln(j))^2 + g.(\ln(j))^3 + m.(\ln(j))^4}{1 + b.\ln(j) + d.(\ln(j))^2 + f.(\ln(j))^3 + h.(\ln(j))^4 + k.(\ln(j))^5} \quad \text{(A1)}$$

with ln referring to the natural logarithm, j the index (1 to 1023) of the luminance levels L(j) of the JNDs, and $a = -1.3011877$, $b = -2.5840191$ E-2, $c = 8.0242636$ E-2, $d = -1.0320229$ E-1, $e = 1.3646699$ E-1, $f = 2.8745620$ E-2, $g = -2.5468404$ E-2, $h = -3.1978977$ E-3, $k = 1.2992634$ E-4, $m = 1.3635334$ E-3.

To apply the above formula to a device with a specific range of L values, it is convenient to also have the inverse of this relationship, which is given by the following equation:

$$j(L) = A + B.\log_{10}(L) + C.(\log_{10}(L))^2 + D.(\log_{10}(L))^3 + E.(\log_{10}(L))^4 + F.(\log_{10}(L))^5 + G.(\log_{10}(L))^6 + H.(\log_{10}(L))^7 + I.(\log_{10}(L))^8 \quad \text{(A2)}$$

with $A = 71.498068$, $B = 94.593053$, $C = 41.912053$, $D = 9.8247004$, $E = 0.28175407$, $F = -1.1878455$, $G = -0.18014349$, $H = 0.14710899$, $I = -0.017046845$.

1. Poynton C. The rehabilitation of gamma. In: Rogowitz BE, Pappas TN, editors. Human vision and electronic imaging III. SPIE; 1998. p 232–249. Proceedings of SPIE, San Jose, CA, Jan. 26–30.
2. Olson T. Behind gamma's disguise. SMPTE Journal 1995;104:452–475.
3. Gibson JE, Fairchild MD. Colorimetric characterization of three computer displays (LCD and CRT). Rochester, NY: Munsell Color Science Laboratory; 2000. 40 p.
4. Glasser J. Principles of display measurement and calibration. In: MacDonald LW, Lowe AC, editors. Display systems. Design and applications. Chicester, UK: John Wiley & Sons; 1997. p 261–287.
5. Berns RS, Motta RJ, Gorzynski ME. CRT colorimetry. Part I: Theory and practice. Color Res Appl 1993;18:299–314.

6. Katoh N, Deguchi T, Berns RS. An accurate characterization of CRT monitor (I) verifications of past studies and clarifications of gamma. Opt Rev 2001;8:305–314.
7. IEC 61966–3 3CD. Colour measurement and management in multimedia systems and equipment. Part 3: Equipment using cathode ray tubes. IEC; 1998. 36 p.
8. Digital imaging and communications in medicine (DICOM). Part 14: Grayscale standard display function. Rosslyn, VA: National Electrical Manufacturers Association; 2003. Report nr PS 3.14–2003. 55 p.
9. VESA. Flat panel display measurements standard version 1.0. Milpitas, CA: Video electronics standards association; 1998. 281 p.
10. Roberts A. Measurement of display transfer characteristic (gamma, γ). EBU Tech Rev 1993;32–40.
11. Berns RS, Katoh N. The digital to radiometric transfer function for computer controlled CRT displays. Proceedings of CIE Expert Symposium '97 Color Standards for Image Technology, Vienna, Austria, 1997p 34–37.
12. Katoh N. Corresponding color reproduction from softcopy images to hardcopy images. Ph.D., Chiba University, Chiba; 2002. 150 p.
13. Deguchi T, Katoh N. The effect of user controls on CRT monitor characteristics. In: Beretta GB, Eschbach R, editors. Color imaging: device-independent color, color hardcopy, and graphic arts III. SPIE; 1998. p 219–230.
14. Berns RS. Methods for characterizing CRT displays. Displays 1996; 16:173–182.
15. Barten PGJ. Physical model for the contrast sensitivity of the human eye. Human vision, visual processing, and digital display III. SPIE; 1992. p 57–72.