

A Display Simulation Toolbox for Image Quality Evaluation

Joyce Farrell^a, Gregory Ng^a, Xiaowei Ding, Kevin Larson^b and Brian Wandell^{ac}

^aDepartment of Electrical Engineering, Stanford University, Stanford, CA 94305

^bAdvanced Reading Technologies, Microsoft Corporation, Redmond, WA 98052

^cDepartment of Psychology, Stanford University, Stanford, CA 94305

Keywords: display linearity, display image quality, display simulation

ABSTRACT

The output of image coding and rendering algorithms are presented on a diverse array of display devices. To evaluate these algorithms, image quality metrics should include more information about the spatial and chromatic properties of displays. To understand how to best incorporate such display information, we need a computational and empirical framework to characterize displays. Here we describe a set of principles and an integrated suite of software tools that provide such a framework. The Display Simulation Toolbox (DST) is an integrated suite of software tools that help the user characterize the key properties of display devices and predict the radiance of displayed images. Assuming that pixel emissions are independent, the DST uses the sub-pixel point spread functions, spectral power distributions, and gamma curves to calculate display image radiance. We tested the assumption of pixel independence for two LCD displays and two CRT displays. For the LCD displays, the independence assumption is reasonably accurate. For the CRT displays it is not. The simulations and measurements agree well for displays that meet the model assumptions and provide information about the nature of the failures for displays that do not meet these assumptions.

Introduction

The vast majority of image and video compression and coding tools are designed on the assumption that consumers view the images on one type of display – a conventional CRT. For example, digital cameras are typically designed to produce output for display on a standard sRGB display, which is a model of a CRT. Yet, most laptop and workstation computers are equipped with LCD displays whose color properties differ significantly from CRTs. Further, standard models of displays do not account for the spatial structure of the display pixels or their sub-pixel color components. The effect of such display features can be quite significant for the proper display of fonts and fine detail in images. This diversity of devices requires an expansion in the scope of image processing algorithms and metrics: such algorithms and metrics should include more information about display spatial and chromatic properties.

To understand how to best incorporate display information, we need a computational and empirical framework to characterize displays. Here we describe an integrated suite of software tools that provide such a framework. These tools help the user (a) characterize the key properties of display devices, (b) build images that contain controlled amounts of specific digital coding artifacts, and (c) predict the visibility of these artifacts given the display properties.

In this paper, we describe a series of measurements and analyses that characterize the spatial, spectral and chromatic properties of two LCD displays and two CRT displays. For these displays, measurements of the pixel point spread, the spectral power distribution and the gamma functions for the red, green and blue pixel components of the display enable us to predict the spectral radiance for any image bitmap rendered on a linear display. We describe a display simulation toolbox that accepts digital data as input and produces an estimate of the image radiance emitted by the display. And finally, we discuss how the display simulation toolbox can be incorporated to measure image quality when developing image coding and compression algorithms.

Display Analysis

The display simulation toolbox (DST) provides a framework that guides the estimation and simulation of the spatial-spectral radiance emitted from a display by any image. Calculating the spatial-spectral radiance is the useful input because, unlike the digital image values usually used for compression and coding, this is the stimulus that actually reaches the eye.

The DST uses three functions to predict the spatial-spectral radiance emitted by a display. First, the DST converts digital values into a measure of the linear intensity (display gamma). Second, the DST models the spatial spread of light using a point spread function for each color component (sub-pixel point spread function). Third, the DST uses the spectral power distributions of the display color primaries to calculate the spectral composition of the displayed image. These three functions – the display gamma, the sub-pixel point spread functions and the spectral power distributions – are sufficient to characterize the performance of linear displays with independent pixels. Simplifying the process of modeling the radiance distribution makes it possible to use the radiance field as the input to objective image quality metrics.

The display simulation toolbox relies upon the assumption that displays can be modeled as linear elements controlled by digital values coupled to the display by a static nonlinearity (display gamma). We assume, for example, that the light emitted by each sub-pixel component (red, green and blue) can be described by a spectral power distribution and a spatial point spread function. The spectral power distribution describes how much light is emitted as a function of wavelength. The spatial point spread function describes how the light is distributed over space. The DST model is valid if the display pixel components have the same spectral power distribution and the same spatial point spread function except for a scale factor that accounts for intensity differences and a spatial translation that accounts for differences in position. Hence, we set out to measure how well – or poorly – this assumption is met in several conventional displays.

The display simulation toolbox models the pixel components as (a) space-wavelength separable functions and (b) sub-pixel independent. Separable means that the spectral power distribution emitted by a component is the same across the entire pixel. Formally, the spatial-spectral distribution of light emitted from the i^{th} sub-pixel can be defined by the product of two functions:

$$p_i(x, y, \lambda) = s_i(x, y)w_i(\lambda)$$

The function $w_i(\lambda)$ describes the spectral power distribution of the i^{th} sub-pixel. The function $s_i(x, y)$ is the spatial spread of the light from that sub-pixel. Sub-pixel independence means that the light emitted from a pixel is the sum of the light emitted by the pixel color components

$$p(x, y, \lambda) = \sum_i p_i(x, y, \lambda)$$

This additivity assumption means that the light emitted from the i^{th} sub-pixel does not depend on the intensity of the other sub-pixels.

We introduce notation to describe the usual static nonlinearity between the digital value, $\mathbf{v} = (R,G,B)$, and the emitted light, referred to as the display gamma. We describe the static nonlinearity for the i^{th} sub-component as $g_i(v)$. The i^{th} gamma function converts the digital controller values, $\mathbf{v} = (R,G,B)$, into the intensity of each sub-pixel. Taking all of these assumptions together, we expect the spatial-chromatic image from a pixel, given a digital input, (R,G,B) , to be

$$\begin{aligned} p(x, y, \lambda) &= \sum_i g_i(v) p_i(x, y, \lambda) \\ &= \sum_i g_i(v) s_i(x, y) w_i(\lambda) \end{aligned}$$

These equations apply to the light emitted from a single pixel. We create the full display image by repeating this process across the array of display pixels. In so doing, we assume that the light emitted from a pixel is independent of the values at adjacent pixels. We refer to the spatial-spectral independence between pixels as display-independence (Brainard, 1989).

These assumptions are a practical starting point for display simulation, but they will not be sufficient in many cases. For example, the pixel independence fails for some CRTs because they are designed with underpowered components that are unable to adequately drive rapid, full-swing signal changes in the electron beam. Some LCDs are designed so that the differential path taken by photons through the liquid crystal element causes the spectral composition of each sub-pixel to differ as a function of position. The practical model we use, however, is a good first-order approximation that has many desirable features: in general, one would prefer to build displays with these simple properties. Moreover, some displays do satisfy the first-order model well. Hence, we believe that the model forms a good basis for an initial toolbox and this model is a good foundation that can be extended to include more complex display properties.

Results

In this section we describe measurements of four displays: two LCD monitors (a Dell Model Number 1907FPc and a Dell Model Number 1905FP) and two CRT monitors (a Dell Model Number P1130 and an HP Model D2845).

We used a PhotoResearch PR650 spectrophotometer to measure the gamma functions and the spectral power distributions (SPD) of the color primaries for the 4 different displays.

***** Insert Figure 1 about here *****

We used a calibrated Nikon D100 digital camera with a customized lens and light baffle (Figure 1) to measure the spatial distribution of light intensity produced by red, green, blue and white pixels. The lens is a 20 mm focal length objective placed in the reversed direction.

In this position, the lens magnifies the pixels by a factor of 10 on the camera sensor. In raw mode, the Nikon D100 produces digital values that are linear with the display radiance. The spatial image, comprising 200 x 200 camera samples per display pixel, measures the spread functions at a spatial resolution of 1.5 microns per sample. This sampling rate is adequate to measure the spread of sub-pixels in all conventional displays. Figure 2 shows camera images of a white pixel illuminated on the four different displays.

***** Insert Figure 2 about here *****

We developed tools for testing the various properties of the spatial-chromatic distribution emitted by pixels. First, we determine if the relative spectral power distribution of the display color primaries is invariant as digital value increases (spectral homogeneity). Second, we test whether the spectral power distribution of any combination of pixel components can be predicted by the sum of the spectral power distribution of the individual pixel components measured separately (spectral additivity). Third, we determine whether the relative spatial spread of each pixel component is unchanged as digital values increase (spatial homogeneity). Finally, we test whether the spatial distribution of light emitted by any combination of pixels is predicted by the sum of the spatial light distribution of the individual pixels (spatial additivity).

Spectral Homogeneity

For each color pixel component, we measured the SPD at several different digital values, d . We then found the scale factor, a_d , that best (in the mean squared error sense) scales the SPD at maximum digital value into the measured SPD. The value a_d is always between 0 and 1.

Figure 3 compares the scale factor for each color pixel component to the gamma function measured by luminance (a weighted sum of the three sub-pixels) as a function of digital value. In all cases, the scale factor, a_d , and the luminance-based gamma function (normalized so that the peak luminance is scaled to one) agree to within 0.35%.

***** Insert Figure 3 about here *****

Spectral Additivity

We tested the spectral additivity of the displays using the method described by Brainard [1,2]. We measured the spectral power distributions (SPD) for red, green, and blue display primaries for each display and the SPD when all three primaries are turned on simultaneously (i.e., a white signal). We then determined whether the SPD measured for the white signal can be accurately predicted by the sum of the SPDs measured for the individual red, green and blue pixel primaries.

***** Insert Figure 4 about here *****

Figure 4 plots the SPD measured for the white signal and the SPD calculated by adding the SPDs for the individual red, green and blue signal components measured separately. The two SPD functions are indistinguishable for the four displays we tested. It is safe to assume, therefore, that these displays are additive in the spectral domain.

Spatial Homogeneity

We tested the spatial homogeneity of the displays by comparing linear camera images of pixels that were displayed with different digital values. If the spatial properties of the display obey the principle of homogeneity, the relative spatial spread of each pixel component should remain unchanged as digital values increases.

For each display, we calculated the scale factor that maps images of a pixel displayed at one intensity into the image of the same pixel displayed at maximum intensity. Figure 5 superimposes scale factors for each pixel intensity on the gamma functions measured for each of the display color primaries. This figure shows that the scale factors derived from the images of a single red, green or blue pixel displayed on the LCD monitors can be predicted by the normalized display gamma for the red, green and blue color channels, respectively (panels A, B). The mean square error between observed and predicted scale factors is less than 0.5% for both LCD displays. The mean square error between CRT scale factors is

1.35% and 1.94% for the HP and Dell CRTs, respectively. The CRT errors can be traced to the fact that the measured intensity of a single red CRT pixel is higher than predicted.

***** Insert Figure 5 about here *****

Spatial Additivity

We tested the spatial additivity of each of the displays by determining whether the spatial distribution of light emitted by any combination of pixels is predicted by the sum of the spatial light distribution of the individual pixels. To test the intra-pixel spatial additivity of color pixel components, we compared the RGB values in the linear camera image of a white pixel (with all three color components simultaneously displayed) to the RGB values predicted by the sum of the camera images of the three color pixel components displayed separately.

The differences between the camera image of a single white pixel and the composite image (created by adding camera images of the red, green and blue pixel components displayed separately) is shown in the left column in Figure 6. The graphs plot the red, green and blue pixel values of the camera image of the white pixel against the red, green and blue pixel values of the composite camera image. Since both images are RGB camera images, we plot the data separately for the R, G and B camera values. If the color pixel components add linearly, then the data should fall along the identity line. The data show that the color pixel components add linearly for the LCD displays but not for the CRT displays. The CRT displays both have an interaction between the color components within the pixel (panels C, D). The deviations of the data from the identity line show that the spatial structure of the light spread from one component depends on the illumination of another color component. These measurements are made for a single fully illuminated pixel on an otherwise dark screen. These conditions place strong demands on the amplifier to switch on and off, rapidly. Hence the spatial interactions may represent imperfections in the ability of the amplifiers to respond to the demand under these conditions (slew rate limitations). Manufacturers respond to these demands in a variety of ways, including the introduction of special circuitry to detect and manage these slew rate limitations on the amplifier.

***** Insert Figure 6 about here *****

To test the inter-pixel spatial additivity of adjacent pixels, we compared the camera image of two adjacent pixels displayed simultaneously to the composite camera image created by adding camera images of the two pixels displayed separately. We tested the additivity of horizontally adjacent pixels and two vertically adjacent pixels. Figure 7 plots the R, G and B values for the camera images of two vertically adjacent pixels against the composite camera image created by adding images of the two pixels displayed separately. For all four displays, spatial additivity holds for vertically adjacent pixels. Figure 7 shows that the differences between camera images of two vertically adjacent pixels and their corresponding composite camera images is small.

***** Insert Figure 7 about here *****

Figure 8 compares the R, G and B values for the camera images of two horizontally adjacent pixels to the composite camera image of the same two pixels. For both LCD displays, the linear camera RGB values in the two images are the same (within measurement noise). This is not true for the two CRT displays. For these displays, the R, G and B values for images of two horizontally adjacent pixels are higher than predicted by the composite image. The failure of spatial additivity for horizontally adjacent pixels illuminated on the CRT displays is illustrated by the difference images shown in Figure 8.

The success of additivity for vertically adjacent pixels, and the failure of spatial additivity for horizontally adjacent pixels, can be explained by sample and hold circuitry that presumably exists in the CRT displays. CRT manufacturers use sample and hold circuitry to compensate for the slew rate limitations of the electron beam as it moves horizontally across the screen [3]. Vertically adjacent pixels are not affected by the slew rate limitations of the electron beam and, consequently, are both independent and additive. The failure is most extreme in the case of additivity for single illuminated pixels (Fig. 6).

***** Insert Figure 8 about here *****

Display Simulation Toolbox

We implemented a Matlab toolbox, referred to hereafter as the Display Simulation Toolbox (DST), to predict the image radiance from a digital image on a calibrated display. We used the measured display properties parameters to predict the radiance of the display image. For each sub-pixel component these parameters are 1) the spectral power distribution, 2) the spatial point spread function, and 3) the static nonlinearity (gamma function).

The DST is capable of simulating theoretical displays; it is also capable of using calibration from real displays as the basis for the simulation. In the case here, we simulate real displays, based on several parametric functions (display gamma, sub-pixel point spread and spectral power distribution). These quantities are measured using the procedures described in Methods and stored in physical units. In this way, the simulated radiance function too can be described in absolute physical units.

The DST includes a variety of functions to calculate different conventional measures of the display (e.g., color gamut, sub-pixel point spread functions) and to adjust the simulated parameters (e.g., pixel size, spatial arrangement of the sub-pixels) of both theoretical and calibrated displays. Finally, the DST contains functions to render the display radiance function at fine spatial scale for any image.

To evaluate the accuracy of the display simulations we compare simulations of displayed images to measured display images. We measure the images on the display using the Nikon D70 camera with a 20 mm lens and a lens extender ring. We captured images of a lower case 'g' (10 pt, Georgia) as rendered using the Microsoft ClearType technology on the Dell LCD Display Model 1905FP and the Dell CRT Display model P1130. These two displays had the smallest and largest departure from display linearity, respectively.

We used the DST to compute the expected image radiance for this character on the two different displays. The prediction represents the screen at a spatial sampling resolution of 20

microns. We then used the simulated image radiance and an ISET-2.0 model of a Nikon D70 camera [4], to predict the camera RGB responses. In this way, we have a comparison of the predicted and observed Nikon images.

Figure 9 compares the measured and simulated characters for the Dell LCD Display Model 1905FP. The point-by-point RMSE between the measured and simulated image is very small, 0.95%. This error is on the order of the camera measurement error. Hence, the display simulator produces a representation of the display radiance as accurate as the camera image.

***** Insert Figure 9 about here *****

Figure 10 compares the measured and simulated characters for the Dell CRT Display Model P1130. There are noticeable differences between the measured (left) and simulated (right) images, revealing the failures of the DST model for this display. The point-by-point RMSE between the measured and simulated image is 3.1%. The prediction failures are anisotropic: errors in the horizontal direction are significantly larger than those in the vertical direction. We quantified this anisotropy by comparing the RMSE for the column sums and row sums separately. When we sum the column values, predicting the row sums, the RMSE is relatively low (2.7%). When we sum the rows, predicting the column sums, the RMSE is much higher (15.3%). The large horizontal errors can be traced to the failure of spatial additivity in this direction (Figure 8). Additivity fails because the pixel intensity within a row does not switch off as rapidly as the additivity model predicts. Hence, the measured column means differ substantially from the predicted column means.

It is possible to create a more complex model for CRTs. For example, the failures of additivity can be used as an empirical basis for modeling the distortion introduced by the sample and hold circuitry. The measurements and simulation show that this added complexity is necessary to adequately characterize these CRT displays, and probably many others.

***** Insert Figure 10 about here *****

Discussion

Display simulation is an important tool for the design and evaluation of imaging systems. For example, display simulation technology has been used to 1) evaluate the design of color matrix display pixel mosaics, [5, 2]) characterize the angle-dependent color properties of LCDs [6], and 3) evaluate grayscale-resolution tradeoffs in digital typography [7]. We extend this prior work by modeling the spatial and chromatic properties of display pixels and predicting the radiance of a displayed image.

The display simulator saves considerable time and effort in predicting the radiance to a wide range of important calibration targets. Measuring the spatial-chromatic radiance from a display screen to each of these targets is a challenging and time-consuming experimental procedure. To obtain estimates of the radiance field involves the use of expensive radiometric equipment and high quality digital imagers and lenses. Making a relatively small number of measurements, and using these measurements to create a calibrated display model, permits the user to investigate how the display will represent a wide variety of test images saving the time and expense of making additional measurements. The purpose of the Display Simulation Toolbox is to assist the user in capturing the information necessary for creating a simulation of the display, managing these data, and performing the final estimates of the radiance field.

The Display Simulation Toolbox models one component of an imaging system that may include other components, such as image acquisition and processing. By modeling these other components [4], it is possible to evaluate the effect that changes in these separate components have upon the perceived quality of the final output. A controlled simulation environment, then, can provide engineers with useful guidance that improves the understanding of design considerations for the individual parts and algorithms in a complex imaging system.

Summary

We describe a simulation technology that models the display image radiance from a small number of measurements. The model incorporates the sub-pixel point spread functions, spectral power distributions, and gamma curves. Using these inputs and the assumption that pixel emissions are independent, we can calculate the anticipated display image radiance.

We tested the model assumptions using data from two LCD displays and two CRT displays. For the LCD displays, the independence assumption is reasonably accurate. For the CRT displays it is not. We developed software to use the parameter measurements needed to implement the model and create a simulated display image. The simulations and measurements agree well for displays that meet the model assumptions and depart for displays that do not.

LCD displays are rapidly replacing CRT displays, both in the home and in the office. It is fortunate that most of the LCD devices we have tested satisfy the simple model properties. This makes it easier to calibrate, control and model these displays and thus predict their effect in the imaging pipeline.

Acknowledgements

We gratefully acknowledge support from Microsoft Corporation, Nikon Corporation and Sharp Labs of America. We thank Greg Hitchcock, Tanya Matskewich, Scott Daly and Louis Silverstein for many helpful discussions. And we thank Dr. Ibrahim Sezan, Mr. Norihiro Aoki and Professor Bernd Girod for their continued support and encouragement.

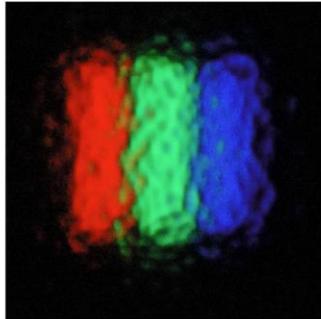
References

- [1] Brainard, D.H. (1989) Calibration of a computer controlled color monitor. *Color Research and Application*, Vol. 14, pp. 23–34.
- [2] Wandell, B. A. (1995). *Foundations of Vision*, Sinauer Press, Sunderland, MA. pp. 1-443.
- [3] Lyons, N. P. and Farrell, J. E. (1989) Linear systems analysis of CRT displays, *SID Digest 1989*, pp. 220-222.
- [4] Farrell, J. E, Xiao, F., Catrysse, P. and Wandell, B. A. (2004) A simulation tool for evaluating digital camera image quality. In *Proceedings of the SPIE Electronic Imaging '2004 Conference*, Vol. 5294, pp. 124-131.
- [5] Silverstein, L.D., Krantz, J.H., Gomer, F.E., Yei-Yu, Y. and Monty, R.W. (1990) Effects of spatial sampling and luminance quantization on the image quality of color matrix displays. *Journal of the Optical Society of America A*, 7 (10), pp. 1955–68.
- [6] Fiske, T. G. and Silverstein, L. D. (1999) “Characterizations of viewing angle-dependent colorimetric and photometric performance of color LCDs,” in *SID Int. Symp. Dig.*, May 1999, pp. 565–568.
- [7] Anthony, W. R. and Farrell, J.E. (1995) CRT-simulation of printed output, *SID Digest 1995*, pp. 209-212.

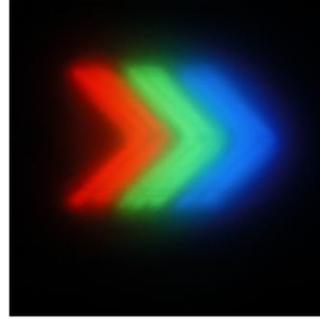


Figure 1. Nikon D100 digital camera with a customized lens and light baffle to measure the spatial distribution of light intensity produced by red, green, blue and white pixels. The lens is a 20 mm focal length objective placed in the reversed direction. In this position, the lens magnifies the pixels by a factor of 10 on the camera sensor.

LCD: Dell Model Number 1907FPc



LCD: Dell Model Number 1905FP



CRT: Dell Model Number P1130



CRT: HP Model Number D2845

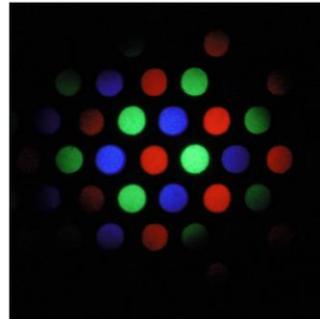


Figure 2. Camera images of a white pixel illuminated on a Dell LCD Display Model 1907FPc (top left), a Dell LCD Display Model 1905FP (top right), a Dell CRT Display Model P1130 (bottom left) and an Hewlett Packard CRT Display Model Number D2845 (bottom right).

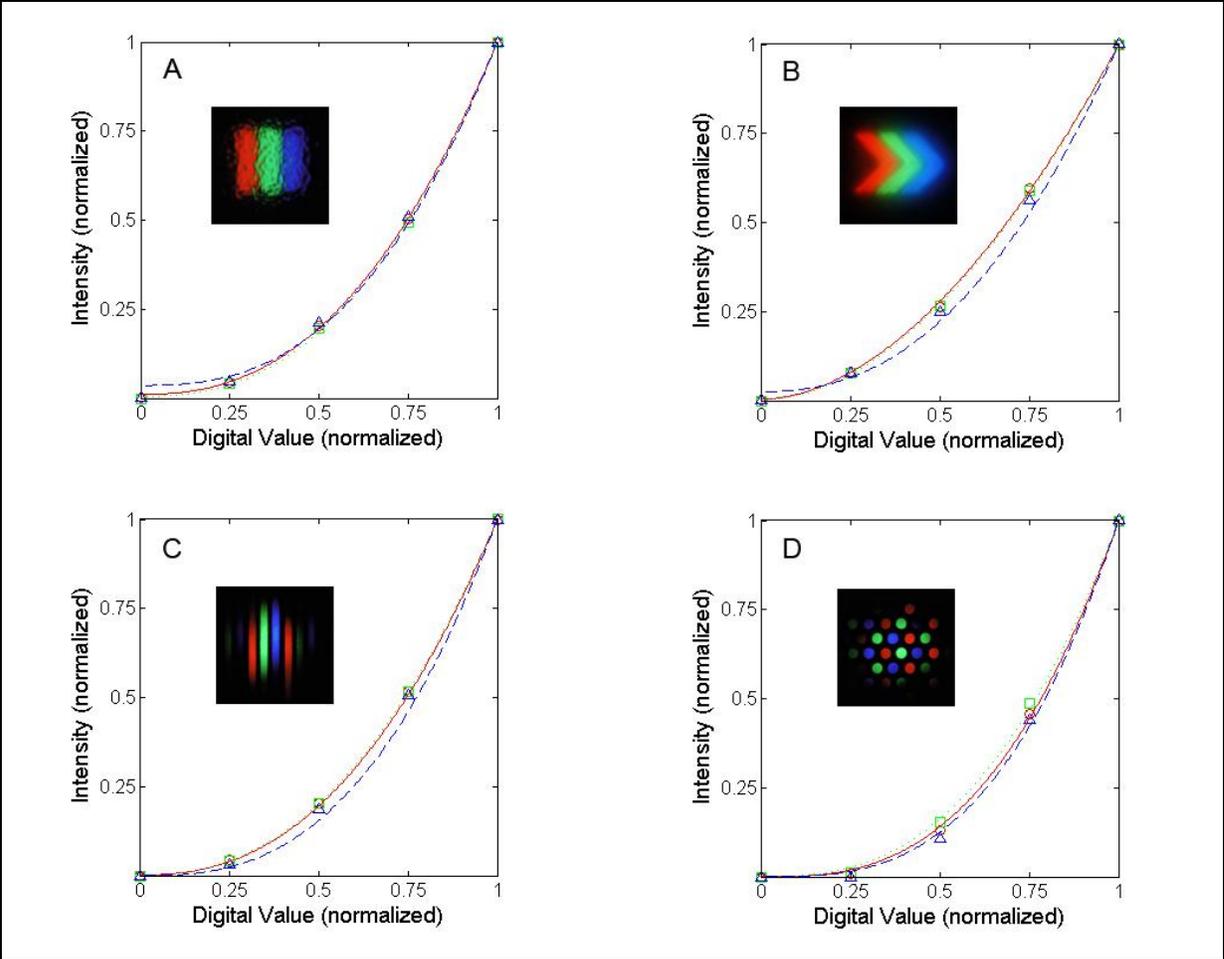


Figure 3: Scale factors that map the SPD measured at each intensity into the SPD measured at the maximum intensity for red (squares), green (circles) and blue (asterisks) pixels. Lines represent the normalized gamma functions for the red (solid line), green (dotted line) and blue (dashed line) channels. The data are plotted separately for the two different LCD displays (top) and the two different CRT displays (bottom).

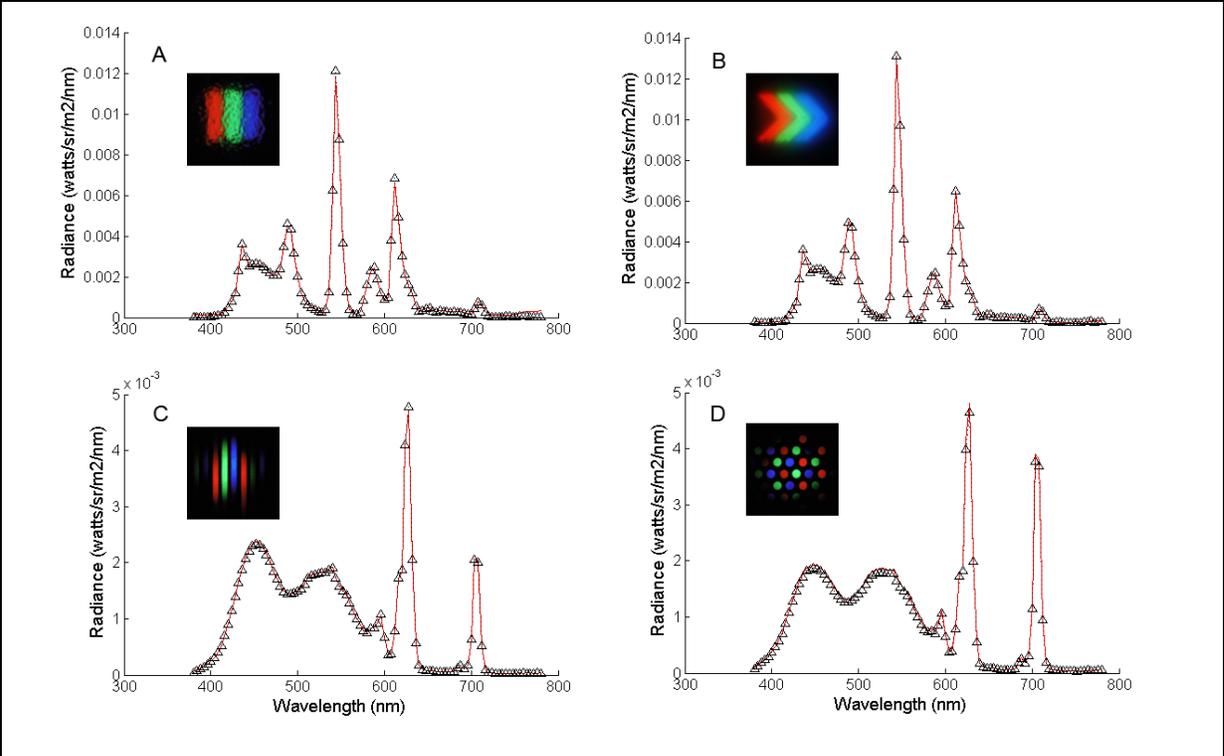


Figure 4. The SPD measured for the white signal (solid line) and the SPD calculated by adding the SPDs for the individual red, green and blue signal components measured separately (triangle symbol). The data are plotted separately for the two different LCD displays (top) and the two different CRT displays (bottom).

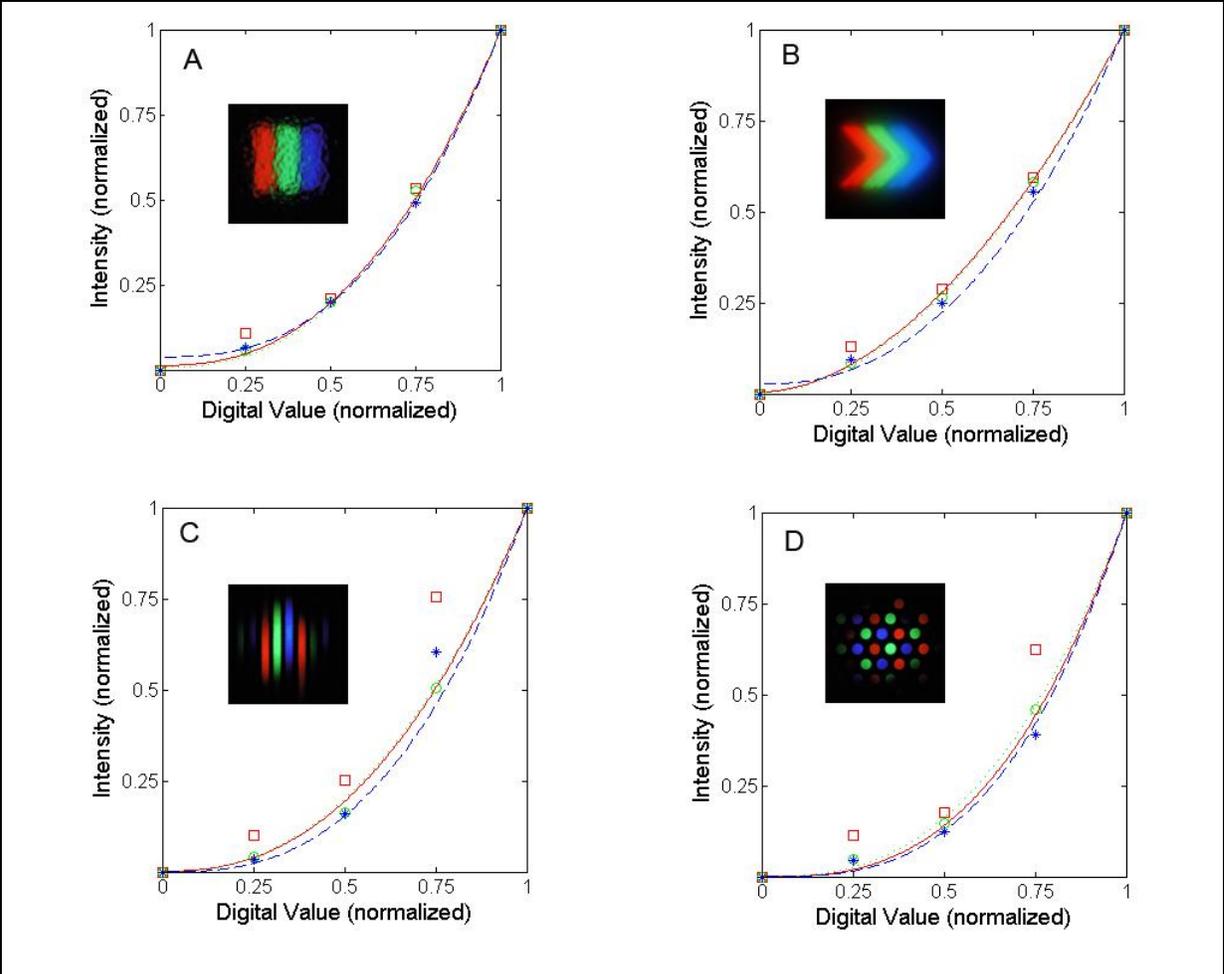


Figure 5: Scale factors that map the luminance values in images of a single pixel (displayed with different digital values) into the luminance values of the same pixel displayed with the maximum digital value. Scale factors for red (square), green (circle) and blue (square) pixels are plotted along with the normalized gamma functions for the red (solid line), green (dotted line) and blue (dashed line) channels. The data are plotted separately for the two different LCD displays (top) and the two different CRT displays (bottom).

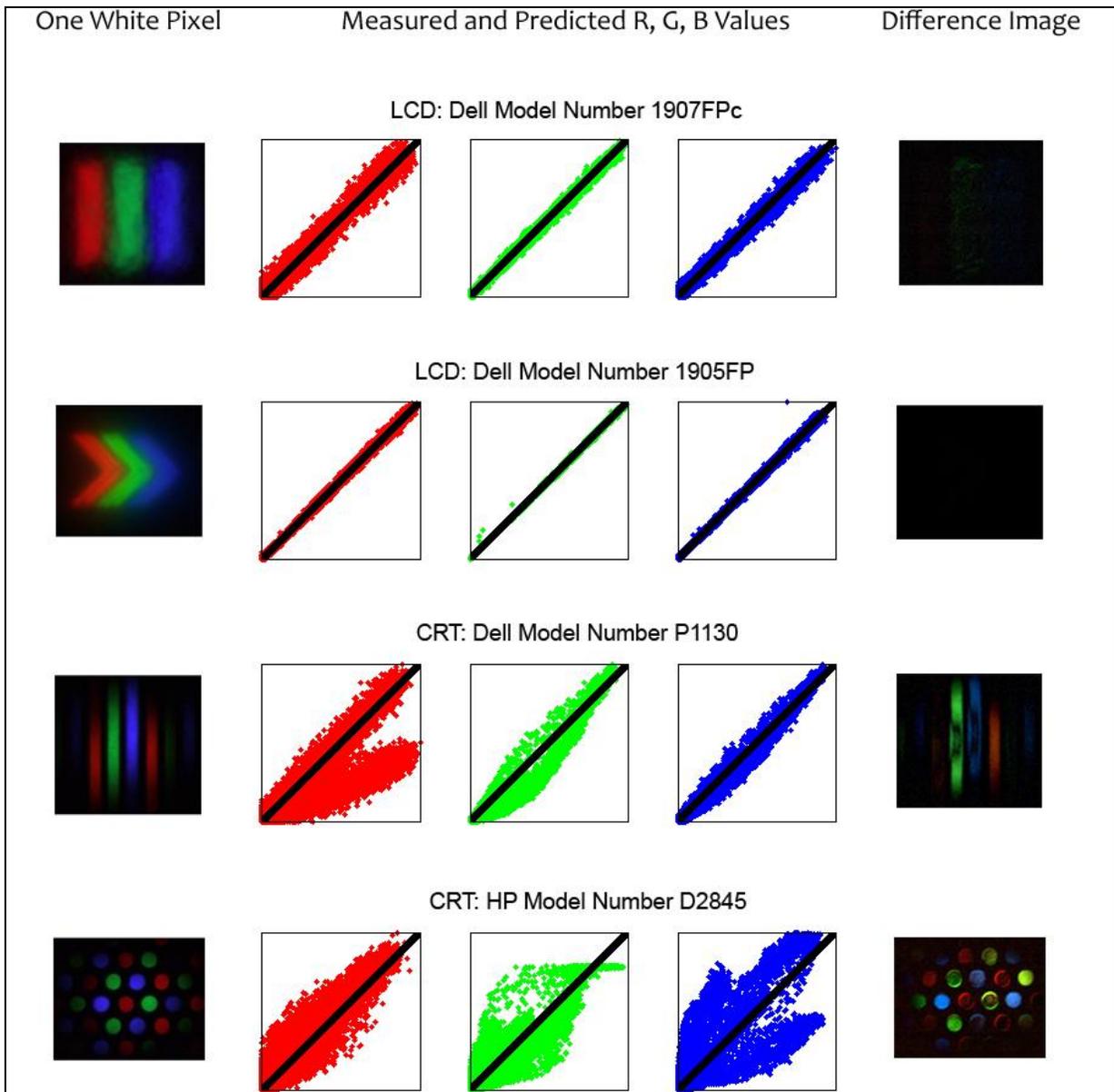


Figure 6: On the right are camera images of a single white pixel displayed on four different displays. On the left is the difference between each camera image and a composite image created by adding camera images of the red, green and blue pixel components displayed separately. The graphs plot the red, green and blue pixel values of the camera image of the white pixel against the red, green and blue pixel values of the composite camera image. If the color pixel components add linearly, then the data should fall along the identity line.

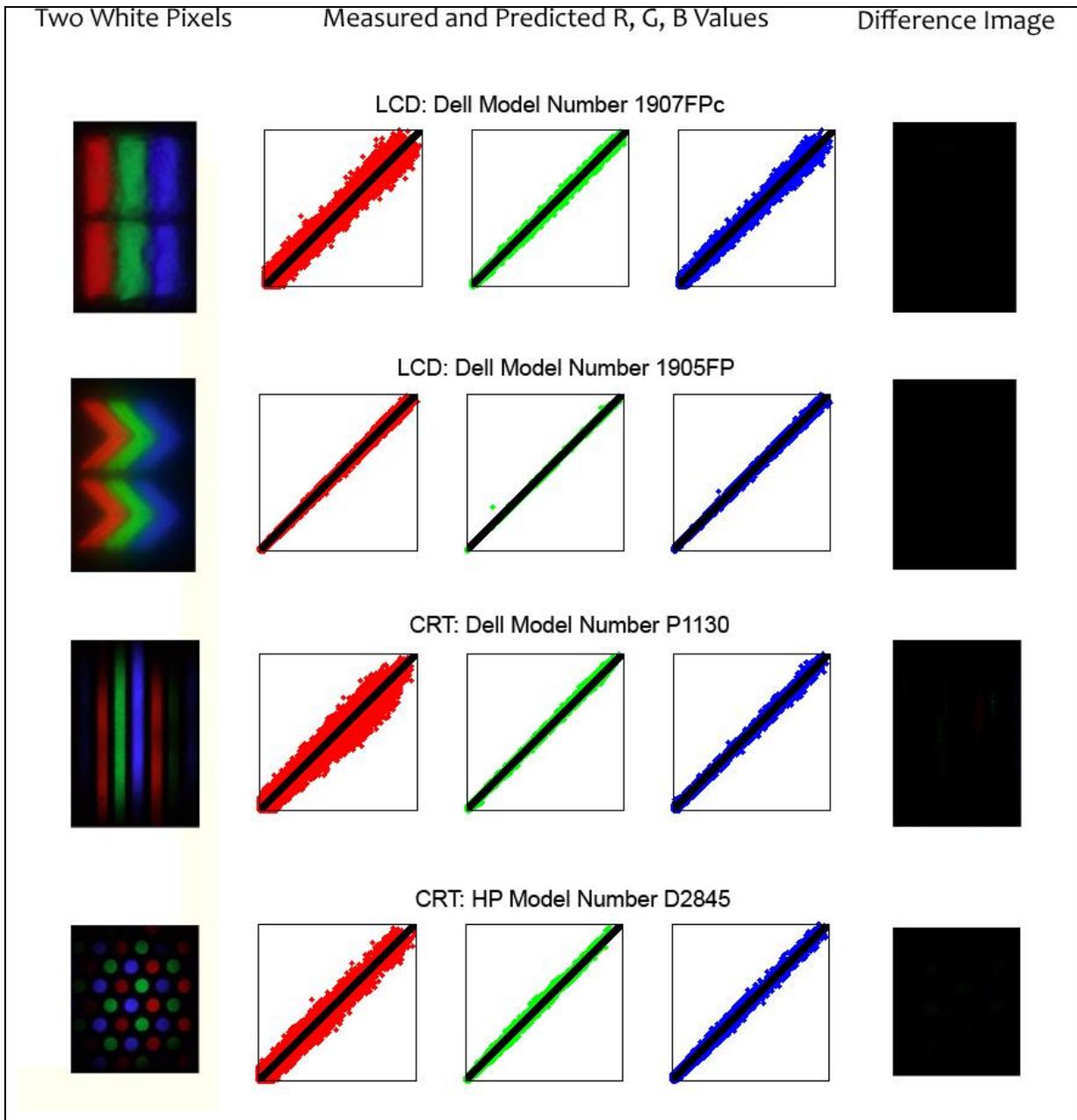


Figure 7: On the right are camera images of two vertically adjacent white pixels displayed on four different displays. On the left is the difference between each camera image and a composite image created by adding camera images of the two pixels displayed separately. The graphs plot the red, green and blue pixel values of the camera image shown on the right against the red, green and blue pixel values of the composite camera image.

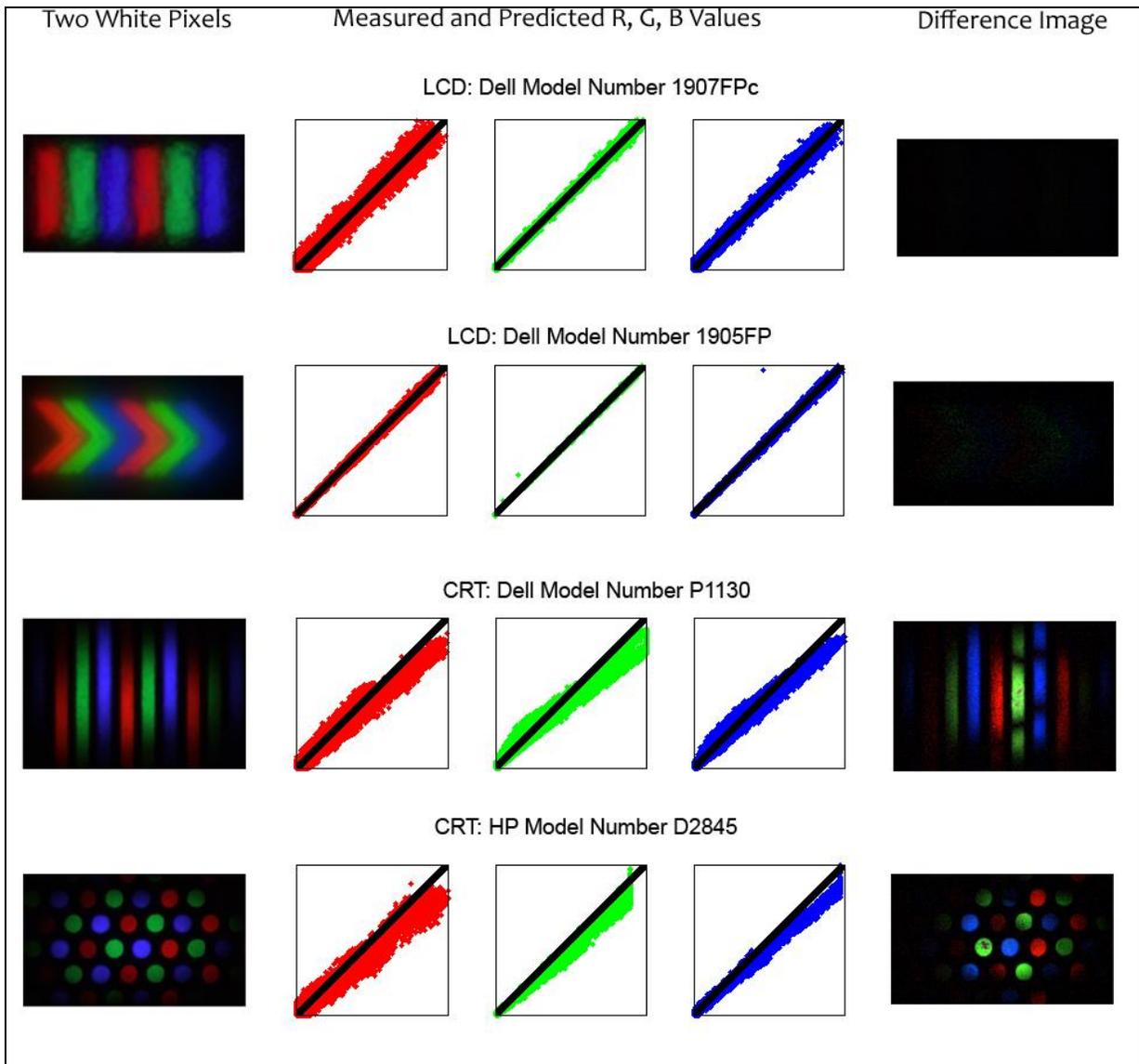


Figure 8. On the right are camera images of two horizontally adjacent white pixels displayed on four different displays. On the left is the difference between each camera image and a composite image created by adding camera images of the two pixels displayed separately. The graphs plot the red, green and blue pixel values of the camera image shown on the right against the red, green and blue pixel values of the composite camera image.

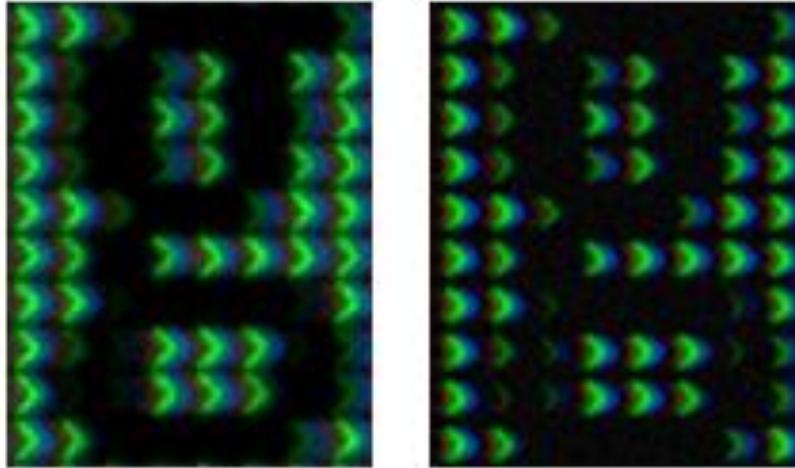


Figure 9. Comparison of measured and simulated characters. The image on the left shows the raw Nikon D70 sensor image of a character that was displayed on the Dell LCD monitor (Dell LCD Display Model 1905P). The image on the right shows the raw sensor image predicted by the DST simulation of the same character and the ISET simulation of the Nikon D70 camera. The two images appear to be green because they have not been color corrected.

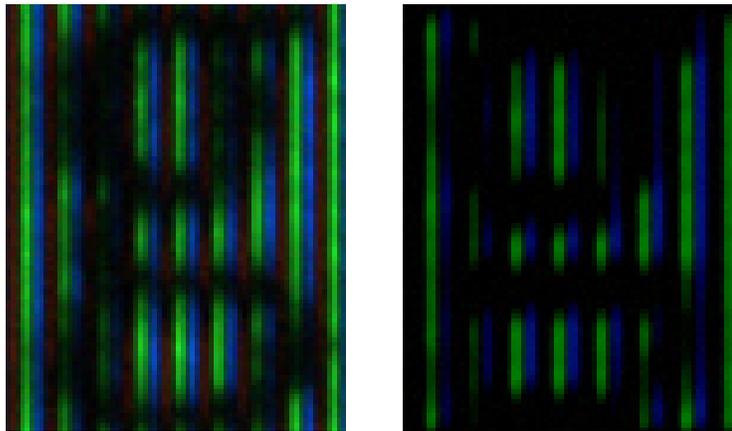


Figure 10. Comparison of measured and simulated characters. The image on the left shows the raw Nikon D70 sensor image of a character that was displayed on the Dell CRT monitor (Dell CRT Display Model P1130). The image on the right shows the raw sensor image predicted by the DST simulation of the same character and the ISET simulation of the Nikon D70 camera. The two images appear to be green because they have not been color corrected.

