Perceptibility and Acceptability of Gamma Differences of Displayed sRGB Images

Efthimia Bilissi, Ralph E. Jacobson, and Geoffrey G. Attridge
Imaging Technology Research Group, University of Westminster
Harrow, Middlesex, United Kingdom

Abstract

The standard RGB (sRGB) colour space was developed to ensure accurate colour reproduction of images when viewed on Cathode Ray Tube (CRT) displays under specified conditions. Typical display and viewing conditions may vary, however, especially when accessing on-line images. Previous work has been carried out on the effects of different display white points and phosphor chromaticities and on modeling gamma tolerances and display conditions for the sRGB colour space. The work described in this paper investigates the effects of different gamma values on viewing sRGB images.

Ten images with variable scene content were selected and converted to the sRGB colour space. A set of test images was generated for each scene by adjusting the display gamma of the sRGB image to a value in the range 1.8 to 2.6. Comparative judgments were conducted in which the reference sRGB image (calibrated for displays with gamma equal to 2.2) in each set was compared to each of the images adjusted to different display gammas. Each pair of images was displayed on the same monitor using software specially developed for the purpose.

In the first series, the observers' responses concerned the perceptible difference between the reference sRGB image and the images calibrated to the different gamma value. In the second test their response concerned the acceptable gamma difference. The experimental results were evaluated and discussed. Conclusions were drawn regarding the effects of gamma differences on perceived image quality while viewing on-line sRGB images.

Introduction

Accurate colour reproduction of images via the web can be a difficult task due to the variability in parameters that affect the perceived quality of the images. The sRGB colour space has been developed to ensure accurate colour reproduction of images when they are viewed under reference display and viewing conditions.[1] Typical display and viewing conditions may vary, however, while accessing images via the Internet. Research work has been carried out to investigate the effect of some of these variations on perceived image quality.

Rehak et al.[2] conducted experiments on different CRT displays with different white points. They showed that the most important factor affecting perceived image quality was the difference in the white point of the monitors and that the effect on image quality due to the variability in phosphor chromaticity was less noticeable. Bilissi et al.[3] investigated the colour differences between 24-bit and 8-bit colour depth settings of the computers' graphics card adaptors and the use of the web-safe palette. Moroney[4] has described a method whereby a simplified CIE gain, offset and gamma model could be used to estimate the gamma, offset and phosphor tolerances for an sRGB monitor. Gamma differences between computer monitors can also affect the perceived image quality. Barten[5] has discussed this effect and also described the modeling of the preferred optimum gamma. Triantaphillidou et al.[6] have also investigated gamma preferences of images displayed on CRT monitors.

A visual study conducted over the Internet by Lavin et al.[7] showed that the typical gamma value of an Apple Macintosh display was around 1.8, of a Unix display was between 2.4 and 3 and of an IBM-compatible display was about 2.2 – 2.4. This showed that the range of typical monitor gamma values may differ from the sRGB reference gamma value. The work described here investigated the effects of different gamma values on the perceived image quality of sRGB images viewed on a CRT monitor under reference viewing conditions.

Experimental Details

The experimental investigation involved the comparison of an sRGB image (calibrated for a monitor with gamma 2.2) with images that were calibrated for different monitor gamma values. Before setting up the experiment, a survey was carried out of monitor display settings and the environmental conditions under which people access the Internet. The survey also included questions that provided information related to on-line shopping preferences and the likelihood that users would be willing to calibrate their monitors on-line. The sample size for this survey was 32. In the literature a sample size of 30 is considered to be the minimum for very large unknown populations[8] (as is the case with the population of internet users). The results from the survey were used to set up the display conditions experiment (Table 1).

Table 1. Survey on Display Conditions

Screen resolution	
800x600	12.8%
1024x768	59.0%
1280x1024	20.5%
Don't know	7.7%
Colour bit-depth	
8-bit	5.1%
16-bit	7.7%
24-bit	59.0%
32-bit	20.5%
Don't know	7.7%

Although the survey was primarily intended to provide information on display settings, we were also interested in measuring the range of gamma settings over a sample of typical computer monitors. Two methods for determining monitor gamma were considered. The first, and most accurate, method was to use a device such as a colorimeter to measure luminance. The colorimeter, however, can only give accurate results when the measurements are performed in controlled conditions. The second method was to use specially developed software that can provide gamma estimations without the use of hardware devices. This method of gamma estimation is often used on web sites for monitor calibration by viewers or is bundled as part of a commercial imaging software package. The second method was chosen for our work due to the fact that the measurement conditions in our computer rooms were difficult to control. An additional advantage of using this method was that a simulation of the on-line gamma estimation method would be provided. Special software was developed in Microsoft Visual Basic 6 to determine the monitor gamma.

The user viewed a pair of patches: an inner patch which was white and an outer patch which had a 50% luminance formed by an alternating black and white grating. The 8-bit RGB pixel values of the inner patch were changed by scrolling a bar over a predetermined range. The bar was moved by the user until the inner, solid grey, patch matched to the outer patch. It has been pointed out[7] that when this method is used over the Internet the web browser may not always be able to assign the frame buffer values requested for display. In this case it is recommended to display a set of patch pairs where the solid grey patch brightness level of each patch pair is preset. The range of brightness levels is selected to accommodate most typical monitor gamma value settings. For the purposes of our work, gamma estimation was not performed on-line and the use of the scroll bar provided more accurate results.

Ten CRT displays driven by IBM-compatible computers at the University of Westminster were tested. The computers selected were used by students for accessing the Internet and other applications and were located in rooms situated in different areas of the University campus. An estimate of whether the monitors were calibrated for the room lighting conditions was made. The areas selected had dim lighting. Each measurement was performed ten times and the average for each display was calculated. It was shown that 60% of the monitors were set to a gamma value within the range 2.1 to 2.6, while 10 % were set to gamma 2.7 and the rest to gamma 2.8 or higher. It should be taken under consideration, however, that these computers were in open access areas where users were able to alter the brightness and/or contrast settings to match their personal preference. In the survey conducted, 82.1% of the participants had replied that they adjusted the brightness and contrast controls of their monitor. These adjustments could be due to personal preference or by following an on-line calibration method. The survey therefore provided statistical data for the spread of gamma values of typical consumer CRT monitor displays.

Apparatus

The experimental investigation was performed by conducting a psychophysical experiment where observers viewed a pair of images (a reference and a sample) on a CRT display. The host computer was a Hewlett-Packard Vectra VA with a Matrox Graphics MGA Millennium display interface card running Microsoft Windows 95. This operating system does not have colour management facilities. Measurements on the CRT faceplate were conducted using a Minolta Chromameter II incident colorimeter with CIE Yxy output. The measuring range was 5.1 – 32000lux. The units were converted from lux (illuminance) to cd/m^2 (luminance) by using the following equation[9]:

$$Luminance = Illuminance \times luminance\ factor / \pi \qquad (1)$$

where *luminance factor* is the ratio of the luminance of a colour to that of a perfect diffuser identically illuminated.

In this work we made the assumption that the display was a perfect diffuser so the luminance factor was equal to 1.

A Hewlett-Packard 6100C scanner was used to scan photographic prints. The scanner was characterised using a Macbeth ColorChecker colour rendition chart and an X-Rite Digital Swatchbook. The images were displayed on an NEC MultiSync M500 15" CRT display with dot pitch of 0.25mm. The screen resolution was 1024x768 pixels and the colour bit-depth setting was 24-bit. These settings were chosen from the results obtained from the survey described previously (Table 1).

Device Characterisation and Calibration

Device characterisation enabled the accurate conversion of scanned images to sRGB and also meant that the experiments could be run in a controlled environment.

Scanner colorimetric characterisation was performed to determine the transformation matrix from scanner RGB colour space values to device independent D65 CIEXYZ. This procedure was also necessary for the transformation of the images to the sRGB colour space.

Methods for scanner colorimetric characterisation have been described in detail in the literature and researchers have also worked on scanner characterisation procedures for converting images to the sRGB colour space.[10-11] The method applied here has been described in previous work.[3] The RGB to XYZ 3x3 transfer matrix was derived and was used for the transformation of the images to the sRGB colour space.

Characterisation of the CRT display was conducted using special test targets and display software.[12-13] Focus, convergence and geometry of the monitor were tested. The results showed that good performance was obtained for the central area of the screen.

The spatial variation of the CRT was investigated by measuring a white patch situated on 12 different areas of the screen. CIELAB Colour Difference, ΔE^*_{ab}, values were computed across the screen. The plotted results showed that colour difference was more uniform in the central area of the screen than at the edges. The highest differences occurred at the right side of the faceplate. The stabilisation time of the CRT was found to be approximately 80 minutes. The display system was set up using special targets and the brightness and contrast controls were adjusted according to the sRGB reference viewing conditions.

The CRT display system transfer function was found by measuring the individual and combined red, green and blue channel values of a displayed 8-bit per channel patch. The difference between consecutive patches in the target was 15 code levels and measurements were performed using the Minolta Chromameter II.

The gamma (γ) value for the combination of the red, green and blue channels was equal to 2.2. The white point of the display was set to CIE Standard Illuminant D65 and the display luminance level was 80 cd/m^2 which is in agreement with the reference display conditions for the sRGB colour space. The results of the survey had also shown that 33.3% of the monitors were set to 65000 K while 10.3 % were set to 9300 K, 2.6% to 5000 K, and 2.6% to some other setting. It was interesting to note that 51.3% participants who completed the survey did not know the white point setting of their computer monitor.

Psychophysical Experiment

The perceived gamma difference of displayed sRGB images was investigated by conducting a psychophysical experiment, which consisted of two tests. Test 1 investigated the imperceptibility of gamma differences while Test 2 investigated the acceptability of gamma differences.

A database of 48 images with random scenes was initially created. The images were scanned from 10x15 photographic prints at 600 dpi (optical resolution of the scanner) in TIFF format. They were converted to the sRGB colour space following the method described in the literature[1]. The size of the images was reduced by bicubic interpolation to 414 x 275 pixels and a subset of ten scenes was selected for the experiment. A set of nine test images was created for each of these scenes. Each set consisted of the reference image, which was the sRGB image, calibrated for monitor gamma 2.2, and eight images with modified gamma.

The range of gamma values selected simulated monitor display gammas of 1.8, 1.9, 2.0, 2.1, 2.3, 2.4, 2.5 and 2.6. The simulation for a specific monitor gamma γ_s was achieved by applying gamma correction to the images. The gamma correction, γ_c, for simulating γ_s, on a display calibrated for the reference gamma, γ_r, was calculated using equation 2 and applied to the images in the Matlab environment.

$$\gamma_c = \gamma_s / \gamma_r \qquad (2)$$

The images with modified gamma were compared to the reference sRGB image in the same set and the reference image was not compared to itself. Each scene set consisted of eight pairs.

The method of constant stimuli was selected for the comparison. The dimensions of the images were selected so that the reference and test image pair could be displayed simultaneously in the central area of the monitor display. As mentioned earlier, the central area was shown to be the most uniform. The images were displayed in portrait orientation.

Image Display

Custom software developed in Visual Basic 6 was used to display the images on the monitor. Each image pair was shown three times to the observer in random order resulting in 30 observations for each gamma value. The position of the reference and test images on the monitor display (i.e. left – right order) was selected at random. Ten volunteer observers participated in the experiment and a total of 240 image pairs were presented to each observer. The background was set at 20% of the screen luminance (16 cd/m^2) and a black cardboard was fitted around the faceplate. The viewing conditions are shown in table 2.

The parametric factors that can affect colour difference evaluation have been described previously[14] and include the following: observer uncertainty (variability of colour matching and variability of judgments of constant stimuli) and physical parameters (sample size, sample separation, texture, colour of background, luminance of sample and lightness of sample).

Table 2. Experimental Conditions

Illumination colour temperature	CIE xy chromaticity coordinates x=0.3270, y=0.3681
Illumination	63 lx
Observer distance from display	Approximately 50 cm

The experiment, which consisted of two tests, used the method of Yes/No choice. In Test 1 the imperceptibility of gamma differences was investigated. Observers responded to the question: "Do the images appear to be the same?". The hypothesis in this experiment was that the observers do not observe differences over a wide range of display gamma deviations from the reference sRGB gamma value. Although the typical viewing conditions may differ from the reference sRGB conditions, the observers would not perceive that difference.

The acceptability of gamma differences was investigated in Test 2. Observers viewed the same pair of images as in Test 1 but responded to the question: "Are the images an acceptable match?". The output was given as follows[15]: For sample x_i a "1" was scored for "Yes" (or "Acceptable") and a "0" for "No" (or "Unacceptable). The proportion p_i of "yes" responses for the sample x_i was calculated using the equation:

$$p_i = \frac{f_i}{N} \quad (3)$$

where f_i is the sum of the responses and N is the number of observations.

The time at each judgment was also recorded and this gave us the average time for the completion of the experiment, which was approximately 40 minutes.

Discussion of Results

The proportion of "Yes" (or "Acceptable") responses from the observers was calculated using equation 3 and the results from the observations were plotted for each image. Figure 1 compares imperceptibility with acceptability and shows a graph for the average proportions of all the images plotted as a function of gamma value. Standard deviation error bars were also included in the graph. The proportions for the average of all the images and the respective standard deviation values are given in tables 3 and 4.

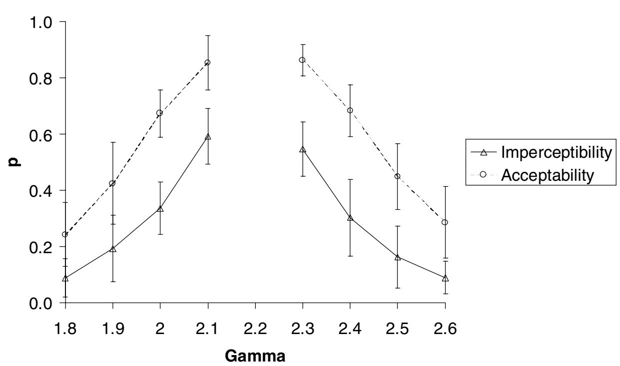

Figure 1. Imperceptibility and Acceptability vs Gamma value

A technique commonly used to evaluate the data from psychophysical experiments is to hypothesize a function for the psychometric response and to fit this function to the data[15]. The Gaussian function is given as:

$$p = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4)$$

where p is the proportion of "Yes" (or "Acceptable") responses, μ is the mean and σ is the standard deviation of the distribution.

Table 3. Average Imperceptibility Results for All Images

Gamma	Proportion (p)	Standard deviation
1.8	0.090	0.069
1.9	0.193	0.118
2.0	0.337	0.094
2.1	0.593	0.099
2.3	0.547	0.097
2.4	0.303	0.136
2.5	0.163	0.110
2.6	0.090	0.059

Table 4. Average Acceptability Results for All Images

Gamma	Proportion (p)	Standard deviation
1.8	0.243	0.113
1.9	0.425	0.145
2.0	0.673	0.083
2.1	0.853	0.097
2.3	0.887	0.055
2.4	0.660	0.091
2.5	0.450	0.117
2.6	0.287	0.128

In this work the Gaussian model in equation 4 was hypothesized to fit the distribution of data from the two tests. The fit was applied using curve-fitting software. The data are shown plotted in figures 2 and 3 and the curve fitting parameters with statistical data (standard error and correlation coefficient R) are shown in table 5. It was shown that a close fit between the data and the normal distribution function was obtained with a good correlation coefficient. These results agree with previous findings in the literature regarding psychometric data.

Table 5. Gaussian Function Parameters and Curve Fitting Statistics

Parameters	Imperceptibility	Acceptability
μ	2.190	2.205
σ	0.184	0.247
Standard error	0.032	0.020
R	0.990	0.998

Perceptibility is related to the stimulus energy that produces a sensation. In our work it was related to the sample image that was perceived as different from the reference image. We were interested in determining the range of gamma values which exceeded the just noticeable difference (JND) value between the test image and the reference sRGB image. This would represent the set of test images that would be perceptibly different from the reference.

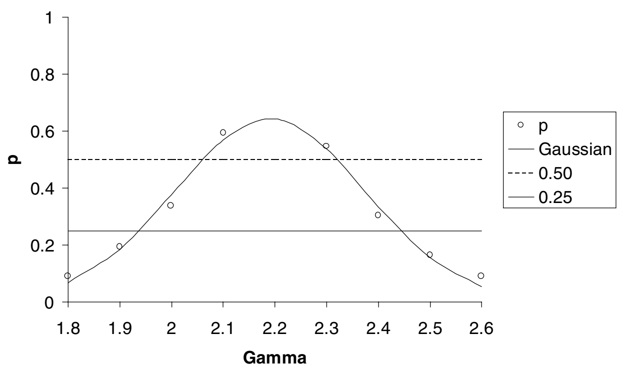

Figure 2. Imperceptibility vs Gamma value

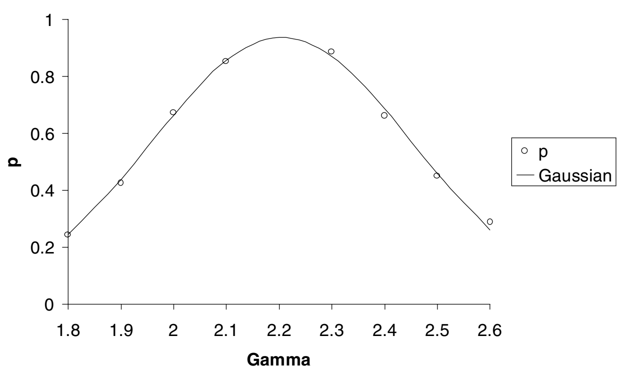

Figure 3. Acceptability vs Gamma value

The JND is considered as the difference between the reference gamma and the gamma value that corresponds to a proportion, p, equal to 0.75. The proportion $p = 0.75$ is equivalent to the value that is 50% greater than the absolute threshold[15-16] where the absolute threshold is defined as the gamma value at $p = 0.50$.

In this case, the "Yes" responses corresponded to "imperceptibility" and not "perceptibility". By inverting the scale (subtracting the imperceptibility datum from unity) we were able to estimate the range of gamma values that were greater than one JND. When the scale is inverted from "perceptibility" to "imperceptibility" the JND value is determined from the difference between the reference gamma and the gamma value at the proportion, p, equal to 0.25. This is shown in figure 2. From the graphs it was shown that when the gamma values were set to 2.1 and 2.3 the difference was not perceived by greater than 50% of the observers. For gamma values 2.4 and 2.0, less than 50% of the observers but greater than 25% said they did not perceive a difference. Less than 25% of the observers reported than no difference was perceived at gamma values equal to 1.8, 1.9, 2.5 and 2.6. This range represents the set of gamma values that was greater than the JND value for this experiment.

Conclusions

A survey showed that typical CRT gamma settings may differ from the reference sRGB gamma value of 2.2. The assumption that the computers measured were calibrated for the local lighting conditions was shown to be invalid. Fifty percent of the computer monitor in the survey were set to a gamma value greater than or equal to 2.5 and no monitors were measured with gamma settings less than or equal to 1.9. This would imply that an sRGB image displayed on at least 50% of the monitors in the survey would result in an image that was perceptibly different from the same image displayed on a calibrated monitor and viewed under equivalent lighting conditions.

Observers can perceive differences in the displayed image relative to the reference if the gamma value of the display monitor is at least 0.3 greater or less than the reference gamma value. The tolerance was higher for acceptability, which was an expected result.

References

1. International Electrotechnical Commission, International Standard IEC 61966-2-1 1st Edition 1999-10, Multimedia Systems and Equipment – Colour Measurement and Management, Part 2.1: Colour Management – Default RGB Colour Space – sRGB, 1999
2. Rehak, R., Bodrogi, P., and Schanda, J., *Displays,* **20,** 165 (1999)
3. Bilissi, E., Jacobson R. E., and Attridge G. G., The Effect of Reduced Colour Depth on the Colour Reproduction of Web Images, *Proc. PICS,* pg. 70 (2002)
4. Moroney, N., Model Based Color Tolerances, *Proc. IS&T/SID The 7th Color Imaging Conference: Color Science, Systems and Applications,* pg. 93 (1999)
5. Barten, P. G. J., *Contrast Sensitivity of the Human Eye and its Effects on Image Quality,* SPIE Optical Engineering Press, Bellingham, Washington, 1999, pg. 185
6. Triantaphillidou, S., Jacobson, R. E., Ford, A. M., Preferred Tone Reproduction of Images on Soft Displays, *Proc. Of International Congress on Imaging Science (ICPS '98),* University of Antwerp, pg. 204 (1998)
7. Lavin, Y., Silverstein, A., and Xuemei Z., Visual Experiment on the Web, *Proc. SPIE,* 3644, pg. 278 (1999)
8. Roscoe, J. T., *Fundamental Research Statistics for the Behavioral Sciences,* 2nd edition, Holt, Rinehart and Winston Inc., N.Y., New York, 1975, pg. 182-184
9. Hunt, R. W. G., *Measuring Colour,* 2nd edition, Ellis Horwood, Chichester, England,, 1992, pg. 261, 296

10. Kang, H. R., *Color Technology for Electronic Imaging Devices,* SPIE, Bellingham, Washington, 1997, pg. 12, 272, 342
11. Hardeberg, J. Y., Desktop Scanning to sRGB, *Proc. SPIE,* 3963, pg. 47 (2000)
12. Ford, A. M., Jacobson, R. E., Attridge, G. G., The Journal of Photographic Science, 44, pg, 147 (1996)
13. Berns, R. S., Gorzynsky, M. E., Motta, R. J., *Color Research and Application,* **18,** pg. 315 (1993)
14. Commission Internationale de l'Eclairage, Parametric Effects in Colour-Difference Evaluation, Technical Report Publication No. CIE 101, 1st edition 1993
15. Engeldrum, P. G., *Psychometric Scaling: A toolkit for imaging systems development,* Imkotek Press, Winchester, Massachusetts, 2000, pg. 64-66
16. Nunnaly, J. C., *Psychometric Theory,* 2nd edition, Tata McGraw-Hill Publishing Co. Ltd, New Delhi, 1981, pg. 63

Biography

Efthimia Bilissi received a degree in Photography from the Technological Educational Institute of Athens in 1995 and an MSc in Digital Imaging from the University of Westminster in 1999. Since then she has been a postgraduate research student in the University of Westminster's Imaging Technology Research Group, working on a project concerned with image quality on the Internet.