Just noticeable gamma differences and acceptability of sRGB images displayed on a CRT monitor.

Efthimia Bilissi
Ralph E. Jacobson
Geoffrey G. Attridge

Harrow School of Computer Science

This is an electronic version of an article published in the Imaging Science Journal, 56 (4). pp. 189-200, August 2008.

The Imaging Science Journal is available online at:

http:// www.ingentaconnect.com/content/maney

Just noticeable gamma differences and acceptability of sRGB images displayed on a CRT monitor

E. Bilissi*, **R. E. Jacobson** and **G. G. Attridge**
Imaging Technology Research Group, University of Westminster, Harrow, UK

Abstract: The standard RGB colour space (sRGB) has been proposed as a means for obtaining accurate reproduction of colour and tone for images displayed across the Internet, provided that they are viewed under the reference display and viewing conditions defined in the standard. It has been found, however, that typical display and viewing conditions when accessing online images vary and therefore deviate from the reference sRGB conditions. One of the parameters that may affect the perceived quality of online images is the gamma setting of the display. In this work psychophysical experiments were conducted to determine the imperceptibility and acceptability of gamma differences of sRGB images when they are viewed on cathode ray tube displays. These experiments were carried out under both controlled and uncontrolled display and viewing conditions. The results of these experiments are presented and discussed, including the estimated points of subjective equality and the just noticeable difference of gamma values.

Keywords: cathode ray tube, sRGB, gamma, perceptibility, acceptability, just noticeable difference

1 INTRODUCTION

Accurate colour and tone reproduction of online images viewed on a computer display is difficult to achieve. The reason is that the parameters that can affect image quality such as hardware, software, display and viewing conditions that are used by the viewer are many and their combinations even more.

The standard RGB colour space (sRGB)[1] was proposed to enable more accurate colour reproduction of online images, provided that the online sRGB images are viewed in display and viewing conditions that are in agreement with the standard's reference conditions (Tables 1 and 2). When accessing the Internet, however, the typical display and viewing conditions may vary from the reference conditions of the standard.

The MS was accepted for publication on 9 January 2008.
* *Corresponding author: Imaging Technology Research Group, School of Media, Arts and Design, University of Westminster, Watford Road, Harrow, Middlesex, HAI 3TP, UK; email: E.Bilissi02@wmin.ac.uk*

The display system gamma setting is a parameter that affects the perceived image quality. This parameter may deviate from the reference sRGB conditions while viewing online images. Lavin *et al.*[2] conducted a visual study over the Internet which revealed that the typical display gamma value of an Apple Macintosh cathode ray tube (CRT) was ~1.8, of a Unix was between 2.4 and 3 and of an IBM compatible computer was between 2.2 and 2.4.

Moroney[3] has investigated viewers' tolerances to gamma deviations of sRGB images using objective measures. In his work he described a method whereby a simplified Commission Internationale de L' Eclairage (CIE) gain, offset and gamma model could be used to estimate the gamma, offset and phosphor tolerances for a sRGB monitor. Colorimetric errors due to the changes in the CRT display transfer function that result from variations in the gain and offset settings of the display have been estimated by Bodrogi *et al.*[4] In their work it was shown that the use of the sRGB transfer function with variable gain and offset settings could lead to

Table 1 Standard RGB reference display conditions

Condition	sRGB
Display luminance level	80 cd m^{-2}
Display white point	x=0.3127, y=0.3290 (D65)
Display model offset (R, G and B)	0.0
Display gun/phosphor gamma	2.2

ΔE^*_{ab} colour differences up to 33. Mitsubayashi *et al.*[5] and Washio *et al.*[6] conducted psychophysical experiments on gamma preference and acceptable gamma differences for liquid crystal displays (LCDs). In their experiments they used CRT displays set to a gamma value equal to 2.2, which is the reference sRGB gamma setting. Their experiments, however, were not conducted with the reference display and viewing conditions of the sRGB colour space and they did not report the use of sRGB images.

In this work we conducted psychophysical experiments that aimed to investigate the imperceptibility and acceptability of gamma differences of sRGB images viewed on a CRT display:

(i) using reference sRGB display and viewing conditions
(ii) using variable display and viewing conditions.

2 EXPERIMENTAL METHOD

2.1 Survey on typical display and viewing conditions

Before proceeding to experimental work, a survey of typical display settings and viewing conditions under which people access the Internet was carried out. This enabled us to determine trends on viewing online images across the Internet. Results from the survey, shown in Table 3, were also used to set up the display conditions for the experiment. The survey was conducted with a questionnaire that included a total of twenty-five questions on display settings, ambient lighting conditions, Internet access and preferences regarding online shopping. Thirty-nine completed questionnaires were collected and this formed the sample size of the survey.

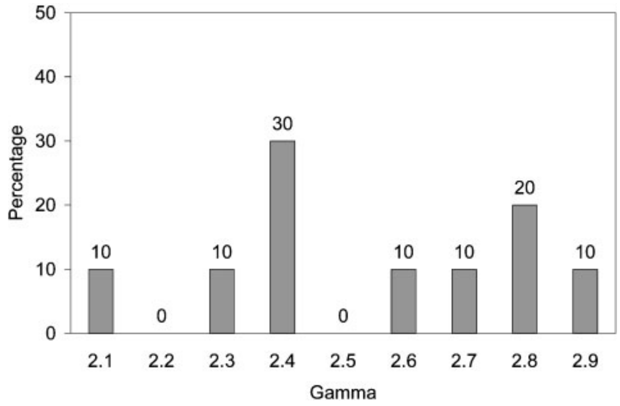
Table 3 Results from survey on typical display settings

Screen diagonal	Percentage	Screen resolution	Percentage
14″	15	800 × 600	13
15″	28	1024 × 768	59
17″	39	1280 × 1024	21
Other	18	Other	0
Do not know	0	Do not know	7.7
Display white point	**Percentage**	**Bit depth setting**	**Percentage**
5000 K	3	8 bit	5
6500 K	33	16 bit	8
9300 K	10	24 bit	59
Other	3	32 bit	21
Do not know	51	Do not know	8

The experimental investigation was to be conducted over a range of simulated display system gamma settings, and for this reason it was important to determine a range of typical settings that should be used for the experiment. A survey taken over a sample of typical computer monitors was conducted for this purpose, using special software for measuring the display system gamma setting. The software was developed in Microsoft Visual Basic. Display gamma was measured following the method described by Lavin *et al.*[2] The user viewed a pair of patches: an inner patch which was initially set to black 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 simultaneously by scrolling a bar over a predetermined range of 0 to 255. The bar was moved by the user until the inner, solid grey, patch matched the outer patch. The display system gamma γ_d was

Table 2 Standard RGB reference viewing conditions

Condition	sRGB
Background	For the background as part of the display screen, the background is 20% of the reference display luminance level (16 cd m^{-2}); the chromaticity should average to x=0.3127, y=0.3290 (D65)
Surround	20% reflectance of the reference ambient luminance level (4.1 cd m^{-2}); the chromaticity should average to x=0.3457, y=0.3585 (D50)
Proximal field	20% of the reference display luminance level (16 cd m^{-2}); the chromaticity should average to x=0.3127, y=0.3290 (D65)
Ambient illuminance level	64 lx
Ambient white point	x=0.3457, y=0.3585 (D50)
Veiling glare	0.2 cd m^{-2}



1 Cathode ray tube display gamma estimation

calculated using the equation

$$\gamma_{\mathrm{d}}=\frac{\log 0{\cdot}5}{\log\left(\frac{PV}{255}\right)} \quad \text{for } PV<255 \qquad (1)$$

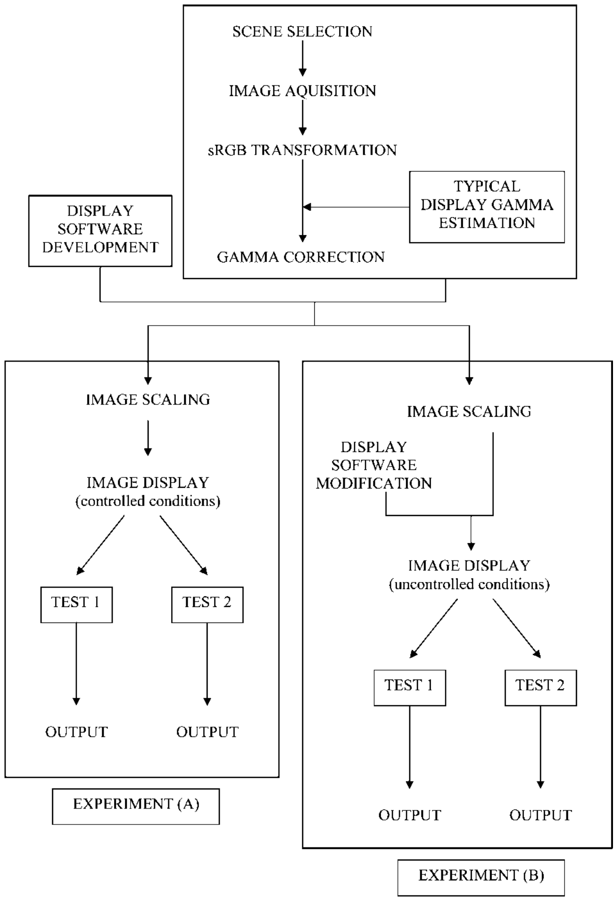
where PV is the RGB pixel value of the inner patch when it has been matched to the grid pattern.

Ten CRT displays, driven by IBM compatible computers at the University of Westminster were tested. Each gamma measurement was performed ten times and the mean for each display system was calculated. Results from the gamma estimation survey showed that the gamma values of the tested CRT display systems ranged from 2.1 to 2.9, with the highest percentage of displays, 30%, set to a gamma value equal to 2.4, as shown in Fig. 1. It should be noted that with equation (1), errors in matching the two patches have a greater effect as the calculated gamma value increases. For example, a change of 4 RGB pixel values is needed for a change in calculated gamma from 1.7 to 1.8, but a change of 2 RGB pixel values for a change in gamma from 2.5 to 2.6.

2.2 Psychophysical experiments

Two psychophysical experiments were conducted. These experiments involved the comparison of a reference sRGB image (calibrated for a display system with gamma equal to 2.2) with images that were calibrated for different display gamma settings. The first experiment, Experiment (A), was conducted in a laboratory under reference sRGB display and viewing conditions, using a characterised CRT display. The imperceptible and acceptable gamma differences of sRGB images when displayed under sRGB reference viewing conditions were determined.

The second experiment, Experiment (B), was conducted using the same display method as in Experiment (A). This experiment, however, aimed to investigate the effect of non-standard gamma settings on displayed image quality in ‘real’ Internet viewing conditions. For this reason the observers that participated in Experiment (B) were asked to carry out the experiment at the computer they mainly used to access the Internet, under the usual display and viewing conditions.



2 Design of Experiments (A) and (B)

Each experiment consisted of two tests. Test 1 investigated the imperceptibility of gamma differences of the displayed sRGB images, and Test 2 investigated the acceptability of these differences. A flowchart that describes the design of Experiments (A) and (B) is illustrated in Fig. 2. It should be noted that the tests were the same for both experiments.

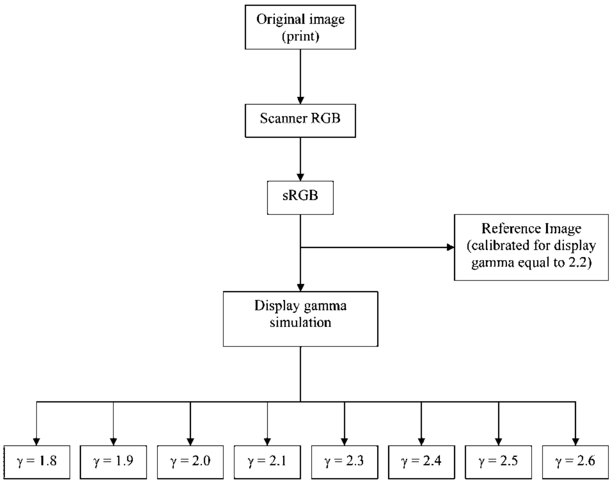
Ten images with random scenes were selected for this experimental investigation. They included scenes with people, nature, buildings and different levels of detail. The selection of the scenes aimed to cover a wide range of subjects taking into account observations mentioned in the literature.[7–9] They were also visually evaluated for contrast and colour. Some examples are illustrated in Fig. 3.

3 Some of the scenes used for the psychophysical experiment

The images were generated from scans of 100×150 mm photographic prints made using a Hewlett-Packard 6100C flatbed scanner set to its optical resolution of 600 dpi. Colorimetric characterisation of the scanner was previously conducted to derive a 3×3 transformation matrix from scanner RGB values to device independent CIEXYZ values. This matrix, shown in equation (2) was used to convert the images to the sRGB colour space using the method described in the sRGB standard.[1]

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0{\cdot}381 & 0{\cdot}432 & 0{\cdot}107 \\ 0{\cdot}231 & 0{\cdot}697 & 0{\cdot}040 \\ 0{\cdot}007 & 0{\cdot}207 & 0{\cdot}904 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2)$$

The range of selected gamma values for this experiment simulated display system gamma settings of 1.8, 1.9, 2.0, 2.1, 2.3, 2.4, 2.5 and 2.6. The selection was based on findings in Ref. 2 and on the survey on typical gamma settings of CRT display systems described in the section on 'Survey on typical display and viewing conditions'. A set of nine test images was



4 Creation of reference and test images

created for each of these scenes. Each set consisted of the reference image, which was the sRGB image calibrated for a CRT display gamma of 2.2, and eight images with modified gamma. A flowchart of how the reference and test images were created is illustrated in Fig. 4.

Simulation of a specific display system gamma γ_{ds} was achieved by applying gamma correction to the sRGB images. The gamma correction factor γ_c for simulating γ_{ds} on a CRT display system set at the reference gamma γ_r was calculated using equation (3) and was applied to the images using Matlab

$$\gamma_c = \frac{\gamma_{ds}}{\gamma_r} \quad (3)$$

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

The images were displayed on the CRT monitor using custom software developed in Microsoft Visual Basic. One pair of images was displayed at one time in the central area of the screen. The image surround was set to grey, with a luminance equal to 16 cd m^{-2}. This was equivalent to 20% of the sRGB reference display luminance.

Previous research work[7] has shown that the order in which images are displayed can affect observer judgment. To minimise this effect, each observer viewed each image pair three times in random order, thereby resulting in 30 observations for each gamma value. The total number of displayed image pairs was 240. The position of the reference and test images on the monitor display (i.e. left–right order) was selected at random, thus minimising possible errors due to spatial variation of the screen.

The distance between the two images on the display remained constant, since all the images had the same spatial dimensions. There was, therefore, no effect of either observer visual angle, or the gap, on the results obtained for each scene.

The method of 'Yes/No' choice was used for both Tests 1 and 2. In Test 1 of both experiments the imperceptibility of gamma differences was investigated. The observers, after being informed that the images might differ in contrast, responded to the question:

'Do the images appear to be the same?'

The hypothesis in this test was that observers do not observe differences over a wide range of display gamma deviations from the reference sRGB gamma value. They responded by answering 'Yes' via the corresponding option button of the software if the images appeared to be the same and 'No' if the images did not appear to be the same.

The acceptability of gamma differences was investigated in Test 2 of both experiments. Observers viewed the same pair of images as in Test 1 but responded to the question:

'Are the images an acceptable match?'

Once again the observers were informed that the images might differ in contrast. If the images appeared to match acceptably then the observers chose the option button labelled 'Acceptable'. If the images did not match acceptably then the observers chose the button labelled 'Unacceptable'.

The software output was given as follows: for sample x_i, where x refers to the scene and i the index to image gamma, a '1' was scored for 'Yes' (or 'Acceptable') and a '0' for 'No' (or 'Unacceptable'). Detailed instructions on how to proceed with the experiment were made available to the observers, who were also notified of the progress of the experiment.

2.2.1 *Experiment (A)*

As mentioned earlier, the investigation in Experiment (A) was concerned with the imperceptible and acceptable gamma differences of sRGB images viewed under reference sRGB display and viewing

5 Display software for image comparison

conditions. It was conducted in a laboratory, using as host computer a Hewlett-Packard Vectra VA with a Matrox Graphics MGA Millennium display interface card running Microsoft Windows 95. This operating system did not have colour management facilities enabled.

The images were displayed on an NEC MultiSync M500 15″ CRT display. The display was characterised and set up to a gamma value equal to 2.2. The screen resolution was 1024×768 pixels and the colour bit depth setting was 24 bits. The choice of these settings was based on the results obtained from the survey previously conducted. The monitor was degaussed and allowed to warm up for approximately 90 min before each run of the experiment, according to the results obtained from the CRT display characterisation that was carried out previously.

The spatial dimensions of the displayed images on the monitor were 75×112 mm. Figure 5 illustrates the layout of the display software and the spatial dimensions of the displayed image pairs. These dimensions were selected by taking into account the findings on spatial uniformity of the screen provided by the display characterisation. Reduction of the image size from the original scanned images was performed using bicubic interpolation. The observer distance from the display was set to $\sim$0.50 m, based on previous research work.[10] It should be noted, however, that the observer distance was not strictly controlled.

Table 4 Viewing conditions for Experiment (A)

Illumination CIE x,y chromaticity coordinates	x=0.3270, y=0.3681
Illuminance	63 lx
Observer distance from display	$\sim$0.50 m

The computer and the calibrated CRT display were placed on a desk, in a room with controlled lighting conditions. The ambient light sources consisted of two sets of overhead fluorescent light tubes, each of which was filtered with a white plastic diffuser. A Minolta Chromameter II was used to measure the illumination and the colour temperature of the light sources. The measurements were conducted by placing the colorimeter in front of the screen with the sensor facing away from the display. The viewing conditions of the experiment are presented in Table 4, and it can be seen that the illumination chromaticity coordinates approximated CIE D50. This was the best approximation to D50 that could be obtained with the existing light sources.

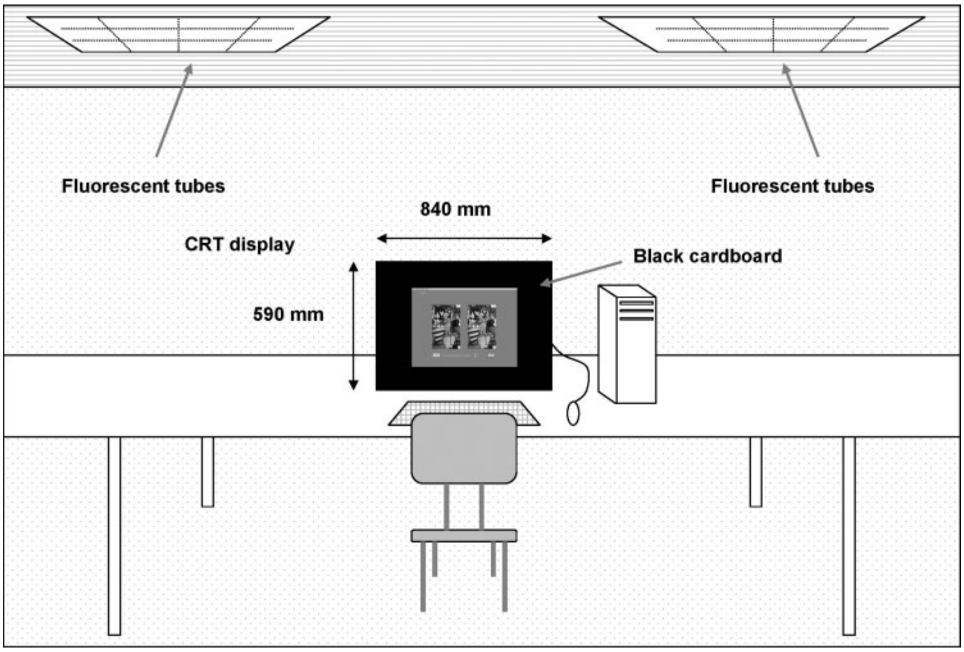
For even illumination on both sides of the screen the display was placed between the two overhead light sources as shown in Fig. 6. A black cardboard screen frame was fitted to the monitor. In this way the observers' field of view would be filled only with the monitor screen and the black surround and they would be, therefore, adapted only to the white point of the CRT display.

Before proceeding to the psychophysical experiment, the display was checked for possible screen reflections.[11] The reflections were minimised by positioning the display parallel to the source of light and tilting the screen. A hood attached to the monitor to minimise any screen reflections was not used because it was shown to cast a shadow over the screen.

Ten volunteer observers participated in the experiment, five male and five female. All had normal or corrected visual acuity, and normal colour vision. The observers were experts or were familiar with digital imaging. Instructions on conducting the experiment and on the use of the software were given to the observers. During this time the observers adapted to the ambient lighting conditions in the viewing room. Throughout the course of the experiment an observer was allowed to view an image pair for as long as was necessary for him or her to proceed to a judgment. Test 1 was conducted initially, followed by Test 2. The time each observer needed for the completion of each test was $\sim$40 min.

2.2.2 Experiment (B)

Experiment (B) was conducted under uncontrolled display and viewing conditions. Thirteen volunteer observers participated in this experiment. They were



6 Set-up of Experiment (A)

given a CD-ROM containing the display software, together with the images. The observers conducted the two tests on the computer they usually use for Internet access. They were asked to retain the display and viewing conditions they use when they normally access the Internet. The participants had normal or corrected visual acuity and all reported that they had normal colour vision. Owing to the nature of the experiment, however, the colour vision of all the observers could not be checked. The display software developed for Experiment (A) was extended for Experiment (B) to handle the three most common screen resolutions. These were determined from the survey and were shown to be: 800×600 pixels, 1024×768 pixels, and 1280×1024 pixels.

3 RESULTS

3.1 Observer responses

The proportion of Yes' (or 'Acceptable') responses was calculated using the following equation

$$p_{x,i} = \frac{1}{N}\sum_{k=1}^{N} f_{x,i,k} \tag{4}$$

where $p_{x,i}$ is the proportion of 'Yes' responses for scene x and image gamma index i. N is the number of observations and $f_{x,i,k}$ the result of observation k, to scene x with image gamma index i, where:

$f_{x,i,k}=1$ if a '*Yes*' response was obtained
$f_{x,i,k}=0$ if a '*No*' response was obtained

The number of observations N was equal to the number of observers multiplied by 3, since each observer viewed each image pair three times. The date and time of the experiment were recorded as well as the time for each image pair judgment.

Graphs of the average proportion of observer responses for every scene and for all the scenes together were plotted as a function of gamma value, for both tests. This was performed for both experiments. A normal distribution (or Gaussian) function[12] was hypothesised to fit the distribution of data from the two tests in both experiments. The normal distribution function is given in equation (5) shown below

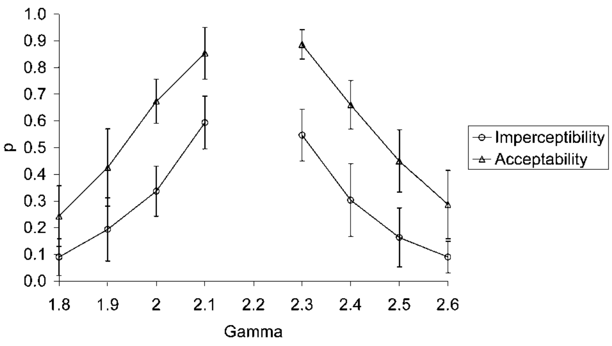
$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{5}$$

where $p(x)$ is the proportion of 'Yes' (or 'Acceptable') responses for a stimulus x, μ the mean and σ the standard deviation of the distribution.

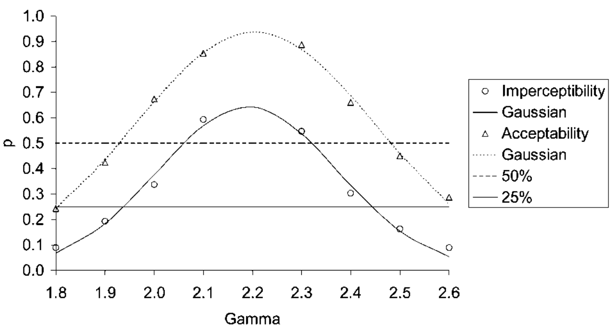
The fit was applied using curve fitting software.[13] The function applied for the curve fitting was the following

$$p'(x) = \alpha e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{6}$$

where x corresponds to image gamma value, $p'(x)$ is the proportion of 'Yes' responses, μ the mean of the



7 Average p values from all scenes plotted as function of gamma value, in Experiment (A)



8 Normal distribution curve fitting to p values obtained in Tests 1 (imperceptibility) and 2 (acceptability), for average of all scenes in Experiment (A)

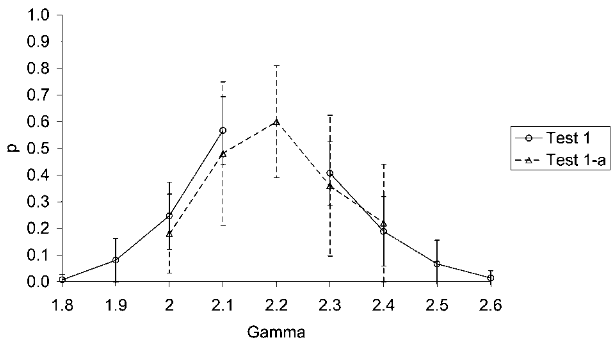
distribution, α a parameter related to the function and σ the standard deviation of the distribution. The correlation coefficient R^2 between the observed p values and the predicted values from the fitted normal distribution was calculated for each scene and the average of all the scenes.

3.1.1 Experiment (A)

The average proportion of observer responses for all the scenes was plotted as a function of gamma value for both imperceptibility and acceptability. The graph, illustrated in Fig. 7, included standard deviation bars over variation in average p values for all scenes. Acceptability p values were shown to be

Table 5 Experiment (A) – parameters of the normal distribution function and correlation coefficient

Tests	Parameters α	μ	σ	R^2
Test 1 (imperceptibility)	0.64	2.19	0.184	0.982
Test 2 (acceptability)	0.94	2.21	0.247	0.996



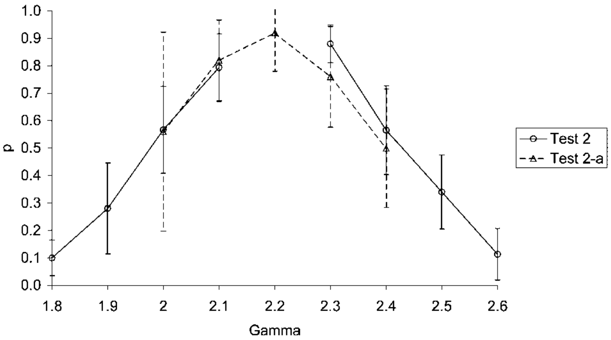
9 Comparison of p values obtained from Test 1 and control Test 1-a (imperceptibility)

higher than imperceptibility and this confirmed results from work previously conducted by Pointer *et al.*[14] and Song *et al.*[15] on colour differences of images displayed on a CRT monitor.

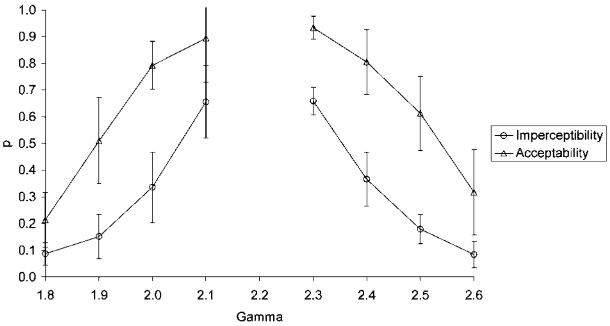
The normal distribution curve fitting to the data of the average p values from all scenes is illustrated in Fig. 8. These graphs also illustrate the 50% and 25% p value points. The parameters α, μ and σ of the fitted distribution and the correlation coefficient R^2 for each scene and average of all the scenes in Tests 1 and 2 are shown in Table 5. The curve fitting parameters were obtained by the curve fitting software. From the results it was shown that there were deviations concerning the value of the mean μ of each curve. This occurred for both Tests 1 and 2. A Student's t-test was conducted and it showed that the difference between the mean values of the normal distribution of each scene was not significant.

Evaluation of the results also showed that for all scenes in Test 1 and most of the scenes in Test 2 the proportion p for the gamma value equal to 2.2 would be less than 1. The value p for gamma equal to 2.2, however, would represent the observer's judgment when the reference sRGB image was compared to itself. It would be, therefore, expected that p should be equal to unity for each scene.

It has been shown, however, that in practice this is not the case since false responses may occur due to observer guessing. For example, Rich *et al.*[16] observed this effect in experiments they conducted on perceptibility of colour differences. In their work it was shown that a sample colourimetrically identical to the reference colour was seen as not matching the reference colour for at least 20% of the judgments. In our work we carried out control tests to investigate the observer responses when the reference image was compared to itself. These tests were conducted using the same display software with a shorter range of



10 Comparison of *p* values obtained from Test 2 and control Test 2-a (acceptability)



11 Average *p* values from all scenes plotted as function of gamma value, in Experiment (B)

gamma values, including the gamma value equal to 2.2. From the results, illustrated in Figs. 9 and 10, it was shown that the proportion *p* of positive responses when the reference image was compared to itself was very close, within the standard deviation bars, to the predicted *p* value from the normal distribution curve fitted to the data of both Tests 1 and 2.

3.1.2 Experiment (B)

The output results from Experiment (B) included information regarding the display and viewing conditions that were applied during each session. This is shown in Table 6. The highest percentage of observers, 65%, responded that they viewed the images under normal lighting conditions, and 54% of the observers reported that the light source was daylight. The display gamma setting ranged from 2.2 to 2.8.

The average proportions of 'Yes' (or 'Acceptable') observer responses for all the scenes plotted as a function of gamma value are shown in Fig. 11. Standard deviation bars regarding the different scenes were included in the graph.

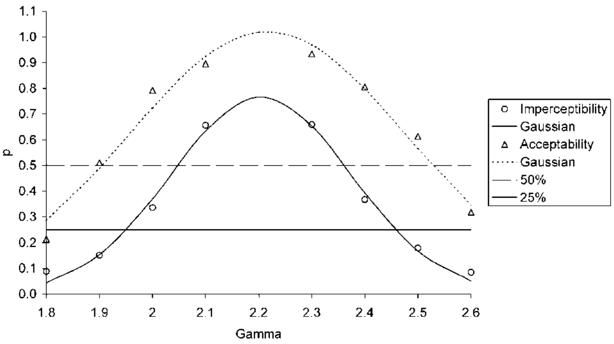
A normal distribution function was fitted to the data and graphs were plotted for each individual scene and for the averaged *p* values. The parameters of the curve fitting and the correlation coefficient R^2 between the observed *p* values and the predicted values from the normal distribution function for both tests are shown in Table 7. Figure 12 illustrates the normal distribution curve fitting for the mean *p* values of all the scenes.

3.2 Just noticeable difference

Perceptibility is related to the stimulus energy that produces a sensation. In this work it was related to the sample image that was perceived as different from the reference image. A measure of perceptibility is the just noticeable difference (JND).

Table 6 Display and viewing conditions for Experiment (B)

Screen diagonal	Percentage	Screen resolution	Percentage
14″	23	800 × 600	31
15″	0	1024 × 768	62
17″	77	1280 × 1024	15
Other	0	Other	0
		Do not know	0
Display white point	**Percentage**	**Bit depth setting**	**Percentage**
5000 K	0	8 bit	0
6500 K	38	16 bit	0
9300 K	23	24 bit	62
Do not know	38	32 bit	31
Other	0	Do not know	8
Lighting conditions	**Percentage**	**Lighting source**	**Percentage**
Bright	15	Daylight	54
Normal	65	Tungsten	27
Dim	19	Fluorescent	19
		Other	0



12 Normal distribution curve fitting to p values obtained in Tests 1 (imperceptibility) and 2 (acceptability), for the average of all scenes in Experiment (B)

For perceptibility the following terms were defined as follows:

(i) the 50% proportion of positive responses was defined as the point of subjective equality (PSE), called the perceptibility points of subjective equality (PPSE)
(ii) the perceptibility JND (PJND) value was defined as the difference between the gamma value, at which 75% proportion of responses were positive, and the PPSE.

We also determined corresponding JND values for Test 2, that investigated acceptability of gamma differences. They were computed in the same way as that defined for perceptibility. For acceptability, therefore, the following terms were defined as:

(i) the 50% proportion of positive responses as the point of subjective equality called the unacceptability PSE (UPSE)
(ii) the unacceptability JND (UJND) value was defined as the difference between the gamma value at which 75% proportion of responses were positive and the UPSE.

The PSE was computed by solving the equation for the value of gamma corresponding to p=0.50. In the case of Test 1 in both experiments, however, the 'Yes' responses corresponded to imperceptibility. For this reason it was necessary to modify the JND computation method for our data to determine the JND value between the test image and the reference sRGB image. Therefore, the JND values for each scene were computed from equation (6), by solving the equation for the value of gamma corresponding to p=0.25 for both sides of the distribution. The UPSE and UJND were calculated accordingly.

The computed PSE and JND values for Experiment (A) are shown in Table 8, for both perceptibility and unacceptability. The subscript L refers to the calculated gamma values γ, for γ<2.2, and the subscript R refers to the calculated gamma values γ, for γ>2.2. The gamma value γ, refers to the simulated display gamma for each scene. For Experiment (B) results from Tests 1 and 2 are shown in Table 9.

Table 7 Experiment (B) – parameters of normal distribution function and correlation coefficient

Tests	Parameters α	μ	σ	R^2
Test 1 (imperceptibility)	0.77	2.20	0.169	0.988
Test 2 (acceptability)	1.02	2.22	0.261	0.970

4 DISCUSSION

In Experiment (A) the $PPSE_L$ value for Test 1 was found to be 2.06 and the $PPSE_R$ was 2.32. The range of imperceptible differences in gamma was therefore equal to 0.26. In Test 2, the $UPSE_L$ value was equal to 1.93 and the $UPSE_R$ value was 2.48, resulting in a range of imperceptibility difference equal to 0.55, approximately double the range for Test 1. Experiment (B) revealed similar results, indicating that although the ambient lighting conditions may have an effect on the display gamma preference, they did not affect the perceptibility and acceptability of gamma differences by the observers. These findings

Table 8 Points of subjective equality and JND values for Experiment (A)

	$PPSE_L$	$PPSE_R$	$PJND_L$	$PJND_R$
Test 1	2.06	2.32	0.12	0.12
	$UPSE_L$	$UPSE_R$	$UJND_L$	$UJND_R$
Test 2	1.93	2.48	0.13	0.13

Table 9 Points of subjective equality and JND values for Experiment (B)

	$PPSE_L$	$PPSE_R$	$PJND_L$	$PJND_R$
Test 1	2.05	2.36	0.10	0.10
	$UPSE_L$	$UPSE_R$	$UJND_L$	$UJND_R$
Test 2	1.90	2.53	0.12	0.12

were in close agreement with results from research conducted by Gille *et al.*[17] on web safe images.

In Test 1 of Experiment (A), the $PJND_L$ and $PJND_R$ were, in both cases, found to be equal to 0.12. In Test 2, of the same experiment, the UJNDs were in both cases equal to 0.13. It was observed that the calculated JNDs in both tests were very similar. Another observation was that the JND values had approximately the same magnitude as the gamma difference between the PSE and the reference gamma value of 2.2.

In Experiment (B), the $PJND_L$ and the $PJND_R$ for Test 1 were 0.10 in both cases. The UJNDs were 0.12 for both cases in Test 2. The similarity of this result to that obtained in Experiment (A) indicated that the main differences between acceptability and imperceptibility were in the level of tolerance shown by the observer to the effects due to gamma differences. The JND values appear not to have been affected.

As mentioned in the section on 'Observer responses', the time for each image pair judgement by each observer was recorded during the experiments. In Experiment (A) it was found that the average time taken to judge imperceptibility was 8.2 s, while the average time taken to judge acceptability was 4 s. Observers, therefore, needed almost double the time to reach a decision regarding judgement of imperceptibility than acceptability of image gamma differences.

With reference to the results from the survey conducted on the sample of computer CRT monitors from the University of Westminster, shown in Fig. 1, it was found that for 50% of monitor displays the estimated gamma values lay outside $PPSE_L - 1$ PJND and $PPSE_R + 1$ PJND. It was also observed that the difference between the reference gamma and the sum of the PPSE value and 1 PJND was equal to approximately 2 PJNDs. Concerning the acceptability of gamma differences it was found that 40% of the computer monitor displays had estimated gamma values that lay outside $UPSE_L - 1$ UJND and $UPSE_R + 1$ UJND. The difference between the reference gamma and the sum of the UPSE value and 1 UJND was found to be approximately 3 UJNDs. Thus 10% more displays would be judged acceptably similar than would be judged imperceptibly different from a reference display.

The JND and point of subjective equality values used above were taken from the results obtained from Experiment (A). Since the results obtained from Experiment (B) were similar to those from Experiment (A), the observations above are expected to apply to the results from Experiment (B).

5 CONCLUSIONS

1. The range of acceptable gamma differences was 0.55, approximately double the range of imperceptible gamma differences which was shown to be 0.26. This was observed for both controlled and uncontrolled display and viewing conditions.
2. The sensitivity of the observers, expressed as JNDs, was shown to be approximately the same for imperceptibility and acceptability. It was also not dependent on whether the gamma deviation was either higher than the reference gamma or lower.
3. The results obtained from Experiment (A) were similar to the results obtained from Experiment (B). This showed that the estimated values of the perceptibility and UJNDs and UPSEs were largely independent of commonly encountered viewing conditions.
4. For applications that require accurate tone and colour reproduction of displayed images the display gamma setting should lie within the range of gamma values defined by the imperceptible gamma difference range. For other applications, the acceptable gamma difference range should be used. The range of imperceptible and acceptable gamma differences can also be used to reduce the total time allocated to gamma adjustment purposes when the number of displays that need gamma adjustment is large.

REFERENCES

1 'Multimedia systems and equipment – Colour Measurement and Management – Part 2-1: Default RGB Colour Space – sRGB (1st edition), IEC 61966-2-1', International Electrotechnical Commission (IEC), Geneva, Switzerland, 1999.

2 Lavin, Y., Silverstein, A. and Zhang, X. Visual experiment on the web. Proc. IS&T/SPIE Conference on *Human Vision and Electronic Imaging IV*, San Jose, California, USA, 1999, Vol. 3644, 278–289.

3 Moroney, N. Model based color tolerances. Proc. IS&T/SID 7th *Color Imaging* Conf., Scottsdale, Arizona, USA, 1999, 93–96.

4 Bodrogi, P., Sinka, B., Borbely, A., Geiger, N., and Schanda, J,. On the use of the sRGB colour space: the 'gamma' problem', *Displays*, 2002, **23**, 165–170.

5 Mitsubayashi, T., Shimodaira, Y., Washio, H., Ikeda, H., Muraoka, T., and Mizushina, S. Gamma range of reproduced pictures having the quality above acceptable limits. Proc. 16th Int. Display Research Conf., Birmingham, UK, October 1996, 483–486.

6 Washio, H., Takahashi, M., Shimodaira, Y., Muraoka, T., Yamamoto, Y., and Ishii Y. Effect of gamma change on picture quality tested by subjective evaluation. Proc. Int. Display Research Conf., Monterey, CA, USA, October, 1994, 333–336.

7 Farnand, S. P. The Effect of image content on color difference perceptibility. Proc. IS&T/SID 4th Color Imaging Conference, Scottsdale, Arizona, USA, 1995, 101–104.

8 Babcock, J. S., Pelz, J. B. and Fairchild M. D. Eye tracking observers during rank order, paired comparison, and graphical rating tasks. Proc. IS&T's PICS Conference, Rochester, New York, USA, 2003, 10–15.

9 Boust, C., Chahine, H., Vienot, F., Brettel, H., Chouikha, M. B. and Alquie G. Color correction judgements of digital images by experts and naive observers', IS&T's PICS conference, Rochester, New York, USA, 2003, 4–9.

10 Bilissi, E., Jacobson, R. E. and Attridge, G. G. Influence of luminance and resolution on the perceived quality of black-and-white images on soft displays. *Imaging Sci. J.*, 2002, **50**, 277–289

11 Sanders, M. S. and McCormick E. J. Human Factors in Engineering and Design, 7th edn; 1992, New York, McGraw - Hill.

12 Kreyszig E. 'Advanced Engineering Mathematics', 5th edn; 1983, New York, John Wiley and Sons.

13 *XL-Plot 2.09*: XL-Plot is the eXtra-Light version of the freeware program Serf and is available from the URL: http://www.bram.org/xlplot/xlplot.shtml. Last accessed: 11 November 2006.

14 Pointer, M. R., Attridge, G. G. and Jacobson, R. E. Food colour appearance judged using images on a computer display', *Imaging Sci. J.,* 2002, **50**, 23–36

15 Song, T., and Luo, M. R. Testing color-difference formulae on complex images using a CRT monitor, IS&T/SID 8th Color Imaging Conference, Scottsdale, Arizona, USA, 2000, 44–48.

16 Rich, R. M., Billmeyer, F. W. and Howe, W. G. Method for deriving color-difference-perceptibility ellipses for surface-color samples. *J Opt. Soc. Am.,* 1975, **65**, 956–959

17 Gille, J., Luszcz, J. and Larimer J. Error diffusion using the 'web-safe' colors: How good is it across platforms? IS&T/SPIE Conference on *Human Vision and Electronic Imaging III*, San Jose, California, USA, 1998, Vol. 3299, 368–375.

