Influence of luminance and resolution on the perceived quality of black-and-white images on soft displays.

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, 50 (4). pp. 277-289, 2002. The Imaging Science Journal is available online at:

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

Influence of luminance and resolution on the perceived quality of black-and-white images on soft displays

E Bilissi, R E Jacobson* and **G G Attridge**
Imaging Technology Research Group, School of Communication and Creative Industries, University of Westminster, Harrow, Middlesex, UK

Abstract: The perceived quality of an image displayed on a computer monitor depends on a number of different factors that can affect viewers' preferences. Two of these are the luminance of the monitor display and the resolution of the image. The effect of luminance is of interest for applications such as on-line access of images where the computer displays used for viewing these images could have different luminance settings. A relationship between cathode-ray tube (CRT) display luminance and resolution has been shown in previous studies. It was therefore interesting to investigate whether there is a relationship between image resolution and CRT display luminance on perceived image quality. Image resolution is related to the image file size, which is an important factor for applications such as on-line access of images. This work used a CRT display to study the effect of the above-mentioned factors on the perceived quality of the displayed image. Three sets of black-and-white images, each set with a different resolution, were presented to observers at three brightness settings of the computer monitor. Results are discussed regarding the effect of monitor display luminance and image resolution on perceived image quality and the interaction between them. Evaluation of results is further extended to the influence of the different backgrounds of the images. The scene content of the images was also shown to affect the viewers' judgement.

Keywords: cathode-ray tube (CRT) display, image resolution, display luminance, image quality

1 INTRODUCTION

The relationship between resolution and luminance of a cathode-ray tube (CRT) display concerning viewers' preference has previously been studied mainly for television systems. Lubin [1] referred to work carried out by Lubin *et al.* [2] who, for example, determined values for resolution and brightness at which the perceived quality was maximized. Their experiments were performed for two different ambient lighting conditions: 20 lx (home television viewing) and 300 lx (showroom conditions) and for viewing distances of 1.5 and 3.0 m. Their results showed that the viewers' preference was scene dependent and higher screen luminance was preferred for images with a greater proportion of pixels at lower grey levels compared with those with a greater proportion of pixels at higher grey levels. They also found that no single combination of the above-mentioned display parameters, resolution and luminance was subjectively optimal for all possible displayed images. A more precise estimate of the preferred combination of resolution and luminance may be obtained if many different representative images are used, for the reasons given above.

Caronna *et al.* [3] extended this approach by including the results of judgements by expert and

This paper was presented at Digital Futures 2000, The Royal Photographic Society Imaging Science Group Annual Conference, held at the University of Westminster, Harrow, UK on 11–13 September 2000. The MS was received on 15 January 2002 and was accepted after revision for publication on 21 June 2002.
* *Corresponding author: Imaging Technology Research Group, School of Communication and Creative Industries, University of Westminster, Harrow Campus, Watford Road, Northwick Park, Harrow HA1 3TP, UK.*

non-expert observers. These experiments, which also concerned television CRT displays, were aimed at determining the preferred peak luminance and resolution for colour television systems. Changes in luminance and resolution were made at the display device. The results of this work showed that, for close viewing distances, resolution was the most important parameter that affected the judgement of image quality. For longer viewing distances, a higher luminance level was preferred and this was shown to be more important than resolution. It was also shown that judgements were scene dependent.

The relationship between 'brightness' and viewing distance has also been investigated by Ardito [**4**] who found that the preferred viewing distance increases as the brightness increases. This confirmed the previous findings of Caronna *et al.* [**3**]. Ardito also pointed out that the results were influenced by home viewing habits. This has been mentioned by Lund [**5**], regarding viewing distance preferences relative to image size and image resolution. He referred to experimental work by Duncanson and Williams who found that the ratios of viewing distance preference to image height were approximately 12, 8 and 7 for monitors with image heights 0.10, 0.25 and 0.36 m respectively. He then compared these results with findings from research on the average viewing distance in peoples' homes. This distance was approximately 2.74 m which corresponds to ratios of 6–10 for images 0.43–0.27 m high. Lund's conclusion was that there was little evidence that home viewing habits affected the results on preferred viewing distance that he had mentioned earlier.

Investigations by Breneman [**6**] on the effect of the relative luminance of the surround showed that the perceived contrast of printed black-and-white images was enhanced in the middle tones and highlights and reduced in shadows when they were viewed with an illuminated surround than with a dark surround. Katoh *et al.* [**7**] have also shown that, when images were viewed on a CRT display, the human visual system was approximately 60 per cent adapted to the monitor's white point and approximately 40 per cent to the ambient light. Other experimental work carried out by Berns and Choh [**8**] gave similar results. Their experiments involved a comparison of soft copy images with hard-copy images viewed under ambient illumination.

The work presented in this paper is concerned with the relationship between display luminance and image resolution for computer CRT displays. Psychophysical experiments were designed to simulate viewing images on a computer CRT display under different luminance settings and image resolutions. Among the parameters taken into consideration were the viewing distance, duration of the observations and lighting conditions.

These experiments aimed to investigate

(a) the relationship between CRT display luminance and image resolution on the perceived image quality,
(b) the influence on image quality when the CRT display luminance is changed from the calibrated setting and
(c) the effect of different image backgrounds.

2 EXPERIMENTAL DETAILS

The concept of the experimental investigations reported here was based on the earlier work of Lubin *et al.* [**2**] as described by Lubin [**1**] and of Caronna *et al.* [**3**]. Psychophysical experiments [**9–13**] were performed with the participation of ten observers (seven male and three female) with an average age of 28 years. They were all chosen because of their familiarity with viewing images on CRT displays. Four of the observers were experienced in judging image quality. The observers were asked to view images displayed under three levels of luminance and three resolutions and to rate them according to the perceived image quality. The results were then analysed using statistical methods.

2.1 Apparatus

The computer monitor was a Hewlett–Packard Ultra VGA 1280 17 in CRT display with dot pitch 0.28 mm. The display setting of 1024×768 pixels at 75 Hz was used for the experimental work. The dimensions of the displayable area were $0.30 \text{ m} \times 0.23 \text{ m}$ ($12 \text{ in} \times 9 \text{ in}$). The resolution for the chosen setting was therefore set at approximately 86 pixels/in (ppi). The host computer was a Hewlett–Packard Vectra HV IBM-compatible personal computer (PC) with an S3 Trio64V2 (DX/GX) graphics card adapter working under the Microsoft Windows 95 operating system.

The images used for the experiment were captured with a Nikon F3 HP 35 mm single-lens reflex (SLR) camera with a Nikkor 50 mm *f*/1.4 lens. All pictures were captured under natural daylight. Agfa Scala ISO

200 35 mm black-and-white transparency film was used, since in this work only the relationship between image resolution and display luminance was investigated and considerations of colour were excluded. The images were then digitized using a Nikon LS-1000 35 mm film scanner and saved in TIFF format at 1350 pixels/in resolution.

The measurements of display luminance and chromaticity were performed using a Minolta Chromameter II incident colorimeter with Commission Internationale de l'Éclavage (CIE) *Yxy* output with a hood attached. All patches for the measurements covered a larger area than the area read by the colorimeter in order to avoid interreflection errors [**14**] and flare [**15**].

2.2 Calibration and characterization of the CRT display

The CRT display was calibrated under room lighting conditions. The Human Factors Society recommends illumination levels between 200 and 500 lx in offices and surveys of illumination levels in offices with computer equipment have shown that they range between 300 and 500 lx [**15, 16**]. Reflections on the screen reduce the contrast and this was likely to affect the viewers' judgement in the experiment. Since low illuminance values reduce the level of reflections, the room ambient light was set to an illuminance level of 300 lx on the screen. Typical conditions of image viewing were thus simulated [**17**]. Reflections of the light sources from the screen were also reduced by positioning the display parallel to luminaires and tilting the screen so that the reflections would not be included in the observers' field of view [**15, 16**]. Attachment of a hood on the screen to reduce reflections was shown to cause shadows on the screen. A polarizing filter was not used to eliminate surface reflection because it would have reduced the monitor display luminance and its effectiveness would have depended on the angles of illumination and viewing.

The warm-up time of the monitor and the spatial variation in the CRT faceplate were investigated before the experimental procedure. This was an essential step for the design of the experiment since it gave information on the area in which the images should be presented, on the warm-up time needed for the display to be stabilized and also on the evaluation of the final results.

The time that a CRT display needs to be stabilized depends on the type of device. Estimations of the warm-up times vary from 15 min to over 3 h [**18**]. Variation with warm-up time for the display used for this work was measured according to the method described by Ford *et al.* [**19**]. The results showed that the warm-up time of the monitor for stable performance was at least 60 min.

Uniformity of luminance of the CRT display was investigated by measuring white patches forming a 4×6 grid across the faceplate. The results showed that there was luminance variation, with the highest values measured in the centre and the lowest on the top left and the bottom right areas. The central area was also shown to have more uniform luminance distribution (Fig. 1).

The possibility of power supply overload, which

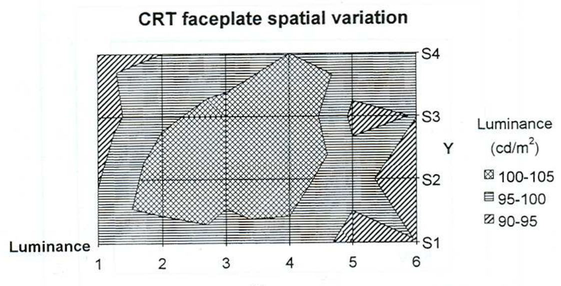

Fig. 1 Experimental investigation on the CRT faceplate spatial variation showed that the central area of the faceplate had the highest and most uniform distribution of luminance. Each interaction of the grid shown represents the centre of one of the 24 measured areas

causes spatial interdependence and instability of the monitor, was also investigated before final calibration. A small maximum white patch in a black background was displayed in the centre of the screen. The luminance was found to be 97 cd/m^2. Then the background was filled with the maximum white and the patch was measured again. The luminance of the patch was then 88 cd/m^2. This supported the view that there was an unequal power drain on the supply under these different conditions [**15**].

The white point of the monitor was set to a simulated D65 (CIE chromaticity coordinates $x = 0.3127$, $y = 0.3290$) using the hardware adjustments of the device. Spatial display targets were used to set up the brightness and contrast of the display [**20**, **21**] (Figs 2 and 3). The targets had 3 pixel value (PV) intervals (approximately from 0 to 6 per cent) and were displayed with a black background. One target was used to set the brightness of the monitor and had intervals from 0 to 15 PV and one for contrast with intervals from 240 to 255 PV. The settings were achieved by using the controls of the CRT for brightness and contrast. The brightness control was set to its mid-point and the contrast control to the minimum point. Firstly, the target for setting the brightness (Fig. 2) was displayed and the brightness was adjusted until the outer (black) step could not be distinguished from the black background of the CRT monitor and each step of the image could be distinguished. Then the target for setting the contrast (Fig. 3) was displayed and the contrast was changed until all the steps of the test chart could be distinguished and the maximum white step had the maximum luminance. The target for setting brightness was displayed again to check whether the black step (0 PV) was still indistinguishable from the black background of the monitor. It was shown that no further adjustments were needed.

The tone reproduction of the scanner, computer and display system (scanner–display) was measured [**22**] for the calibrated settings and also for the luminance settings used for the experiment after the pilot tests were completed. The results were expressed as γ values. In this work, γ was obtained by applying the following equation [**15**, **23**, **24**] and was the power parameter of the obtained transfer function after fitting an appropriate power function (Table 1):

$$\frac{L}{L_{\max}} = k_g \left(\frac{V}{V_{\max}} \right)^{\gamma} \qquad (1)$$

where L is the luminance, $L_{\max}$ is the maximum luminance of the display, V is the PV, $V_{\max}$ is 255 for the bit depth of the system used and k_g is a constant referring to gain (contrast).

Other models for the transfer function of the display system have been proposed [**14**, **25**] which take into account both gain and offset. In this work the simpler model was applied including constant k_g for gain.

2.3 Images

Three images were used for the experiment. This small number was chosen primarily to minimize observer

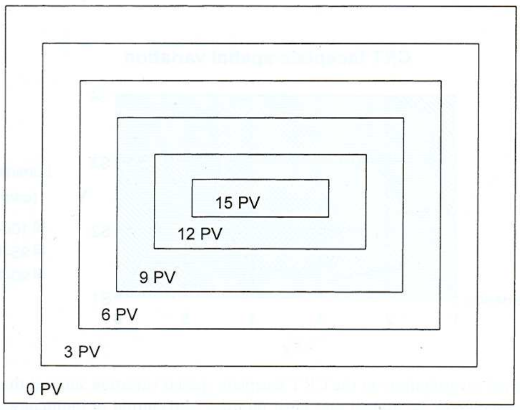

Fig. 2 Display target for setting the brightness of the CRT monitor display (PV, pixel value)

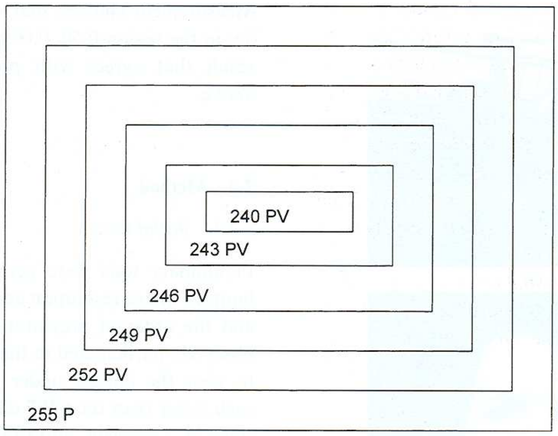

Fig. 3 Display target for setting the contrast of the CRT monitor display (PV, pixel value)

Table 1 The display system tone reproduction in terms of γ values

Luminance level	γ
Low (83 cd/m^2)	2.03
Calibrated (90 cd/m^2)	1.81
High (97 cd/m^2)	1.61

fatigue, which could adversely affect the results [**26**]. The scene content of the images was varied, in order to determine where there was any scene dependence (Fig. 4).

The displayed images had a width of 0.12 m and a height of 0.09 m and a maximum resolution of 86 pixels/in (display screen resolution). These dimensions were obtained by reducing the size of the original scanned images using bicubic interpolation [**27**]. This method of interpolation results in sharper images than bilinear interpolation [**28**, **29**]. The images were presented on a black background in the central area of the screen which was the brightest and most uniform area of the display. As the test on power drain showed, this arrangement ensured that images would be displayed with maximum luminance (Fig. 5). This agrees with suggestions of Berns *et al.* [**18**] where, for highest accuracy, the dimensions of the image should be smaller than those of the display and the background of the image should have a uniform colour. In this way, changes in the image will have a small effect on the electrical loading [**15**, **18**].

The display of images with lower resolution on the computer screen, however, had to be simulated. The image dimensions in pixels were decreased and then the images were displayed with the same physical size using special display software. The pixel size on the monitor was increased, giving the effect of an image with lower resolution. The interpolation method applied in this case was bilinear. Tests were conducted to determine whether the differences that resulted when the selected images were resized to the same physical size using bicubic and bilinear interpolation were visible. The results indicated that the differences between the two methods were not perceptible. The selected image resolutions that are referred to in this experiment were calculated by dividing the physical dimensions of the images by their pixel dimensions. Changes were made to the monitor luminance by using the hardware adjustments for brightness. The monitor was also fitted with a black cardboard screen frame so that the observer's field of view would be filled only with the monitor screen and that black surround [**17**].

It has been shown that viewing distance has a significant role in the judgement of image quality [**5**]. Viewers attempt to optimize the image quality by adjusting their distance from the screen. For example, viewers move closer to the screen when they want to see important detail and move away from it when they want to achieve a crisp image. Experiments on the preference of viewing distance, however, have been carried out mainly for video images. The working distance from a computer monitor can

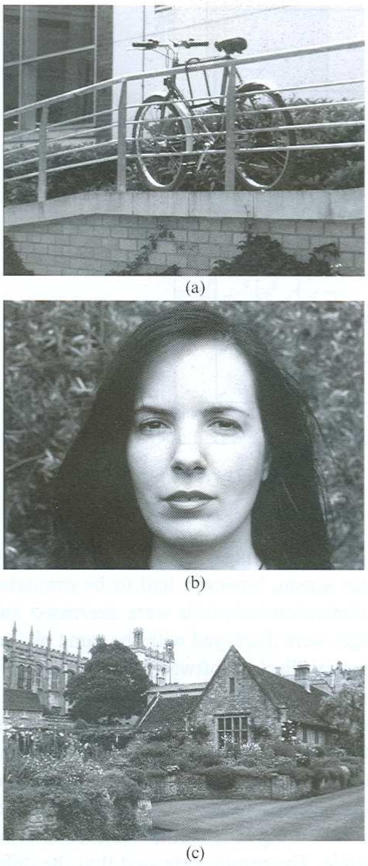

Fig. 4 Images of three selected subjects used for the experiment: (a) a bicycle behind rails, which is lit by natural daylight and where the scenery includes many lines and fine detail; (b) outdoor portrait, which is lit by natural daylight; (c) outdoor scenery which includes a house, trees and flowers and where the scene contains a large area of high-frequency information

normally have only small adjustments. Research by Harpster [**30**] on the preferred viewing distance from such equipment has shown that the viewer's distance from the display is about 0.40 m. Other research [**31**] has shown that the shortest preferred viewing distance is about 0.45 m and the longest is about 0.60 m. One of the reasons affecting the preferred viewing distance was shown to be the positioning of the keyboard and the mouse [**32**]. In the present work the viewers were asked to select the optimal viewing distance. Measurement showed that optimal viewing distance lay in the region 0.50–0.60 m for all the observers, a result that agrees with previous work mentioned above.

2.4 Method

2.4.1 Initial tests

Preliminary tests were performed to determine the limits of image resolution and CRT display luminance and the order of presentation of the images. Three observers participated in these tests. They were asked to view the images under six different resolutions, each lower than the CRT display resolution, in equal steps of 2 pixels/in, and to choose the images, which were not only acceptable, but which were considered to be of good quality. The observers gave consistent results. From these chosen images the lowest resolution for the image set of the experiment was determined. It was shown that the three highest resolution settings were judged as acceptable. Images with lower resolutions had been judged as unacceptable. The three luminance levels were chosen in a similar way. The medium level of luminance was that of the device after calibration. This was designed to investigate the influence on image quality when the luminance settings of the monitor have been changed relative to the calibrated setting. The observers were asked to view the images with different levels of luminance and to select the lowest and highest level of luminance where the images would have good quality. The results were, again, consistent and the two levels of luminance were approximately ± 8 per cent from the luminance level after calibration.

In addition, the preliminary tests gave information on the order of image presentation, which could be chosen. It was decided then that changes in resolution should be applied in random order, and the luminance level was changed after a change in image set. An image set was a set of images of the same subject under three different resolutions.

2.4.2 Experiment

The investigation used a 3×3 factorial procedure with a within-sample design (Table 2). This meant that the factors were varied within the set viewed by a participant so that each participant experienced all levels of the factor [**32**]. The effects of two factors (display luminance and image resolution) were

Fig. 5 Images were presented with a black background using special display software

Table 2 Factorial design of the experiment for within-subject factors

Factor 1	Level (pixels/in)	Factor 2	Level (cd/m^2)
Resolution	82	Luminance	83 90 97
Resolution	84	Luminance	83 90 97
Resolution	86	Luminance	83 90 97

examined for each variable (images). Each factor had three levels. The three levels of luminance were as follows: the calibrated level (90 cd/m^2), the calibrated level + 8 per cent (97 cd/m^2), the calibrated − 8 per cent (83 cd/m^2). The three levels of image resolution were 82 pixels/in (image dimensions, 385 × 288 pixels), 84 pixels/in (image dimensions, 395 × 295 pixels) and 86 pixels/in, which was the same as the monitor resolution (image dimensions 401 × 300 pixels).

The investigation was performed by displaying each of the three images under three resolutions and each resolution set under three levels of luminance as described above. The experiment was performed without a reference image, in order to simulate everyday image display viewing conditions (e.g. on a home PC via the Internet). The observers were asked to rate the images according to the perceived image quality. The rating was from 0 (poor quality) to 10 (maximum quality). Observers, however, were permitted to rate an image with a quality rating greater than 10, if they considered it appropriate during the course of the experiment. If an observer rated an image greater than 10, then the ratings provided by the observer were scaled to a range from 0 to 10. In practice, no ratings greater than 10 were received. To eliminate judging errors due to the lack of a reference image (or images) in the test set, each test image within a set occurred three times and the images were displayed in random order. The three observer ratings received for each image were averaged and this average was considered as the final rating of the image.

The effect of different backgrounds on viewer judgement was investigated by displaying one set of images ('bicycle') against a grey background. The total number of images presented to each of the observers was 108. Each participant viewed four images (three with a black background and one with grey) at three resolutions and each resolution at three levels of luminance.

Observers were permitted to view an image for as long as they needed in order to provide a rating. This procedure was adopted because the experiment involved viewing images on a computer monitor where typical image viewing times can be long. Each observer needed approximately 20 min to complete the entire experiment.

3 RESULTS AND DISCUSSION

3.1 Luminance–resolution preference

The data were interpreted by examining statistical summaries (sample standard deviation and average rating) for each individual image and for all the images combined [**13**]. Graphs were plotted for the average rating of all images and for each image individually. They included:

(a) a box plot of resolution against rating for the three levels of luminance,
(b) a graph of resolution against mean rating ±one standard deviation for the three levels of luminance and
(c) a line graph of resolution against the average rating for the three levels of luminance.

In the graphs the levels of luminance are described as low (83 cd/m^2), calibrated (90 cd/m^2) and high (97 cd/m^2). Figure 6 refers to the average ratings of all images. Figures 7 and 8 refer to the image 'bicycle' displayed with black and grey backgrounds respectively. Figure 9 refers to the image 'house' and Fig. 10 to the image 'portrait'.

Box plots were used to compare several distributions of rating (Figs 6a, 7a, 8a, 9a and 10a). This method is based on the median and the quartiles [**33, 34**]. The data were divided into four parts and each part contained 25 per cent of the observations. The division points are called first quartile (or twenty-fifth percentile), second quartile (or fiftieth percentile) which is the median, and third quartile (or seventy-fifth percentile). A box was drawn with the ends located at the first and third quartile, including the median and the 50 per cent of the ratings. In this experiment, the box plots were used for an initial assessment of variability from one box to the other and data dispersion.

Interaction between the factors was investigated by plotting the average ratings for all combinations of resolution and luminance for each image in the test set [**35**]. Interaction would mean that the level of one factor affected the levels of the other factors. If there was no interaction between the factors, the lines would be parallel, while in the case of interaction, the lines would cross each other. Typically sampling error

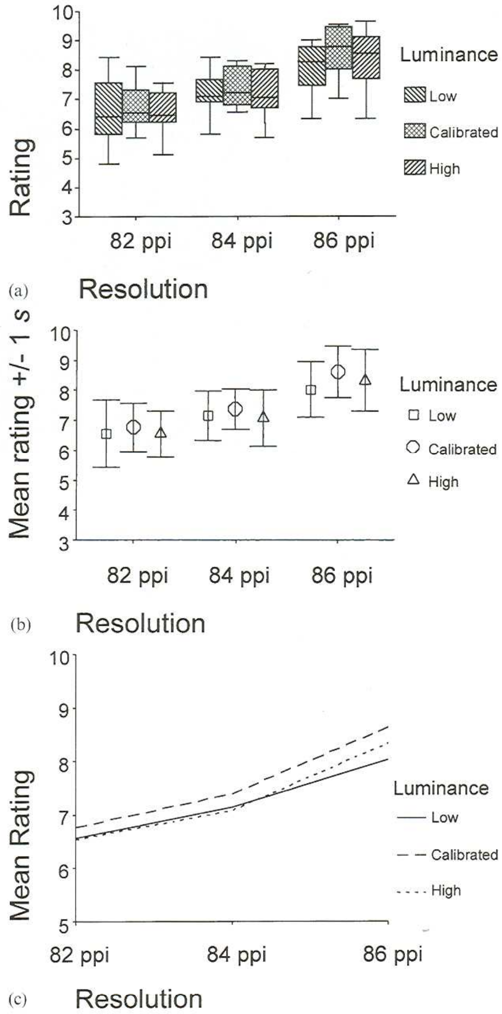

Fig. 6 Graphs for the ratings of all images (ppi, pixels/in): (a) box plots showing the dispersion of the data; (b) standard deviation *s*; (c) average ratings

may cause lack of parallelism in the lines even when there is no interaction [**26**]. In this experiment, the graphs showed that there could be interaction between the factors since the lines were crossing each other in most cases (Figs 6c, 7c, 8c, 9c and 10c).

The significance of the interaction between luminance and resolution was investigated by applying an analysis of variance (ANOVA) to the data. ANOVA is a method for statistical hypothesis testing based on

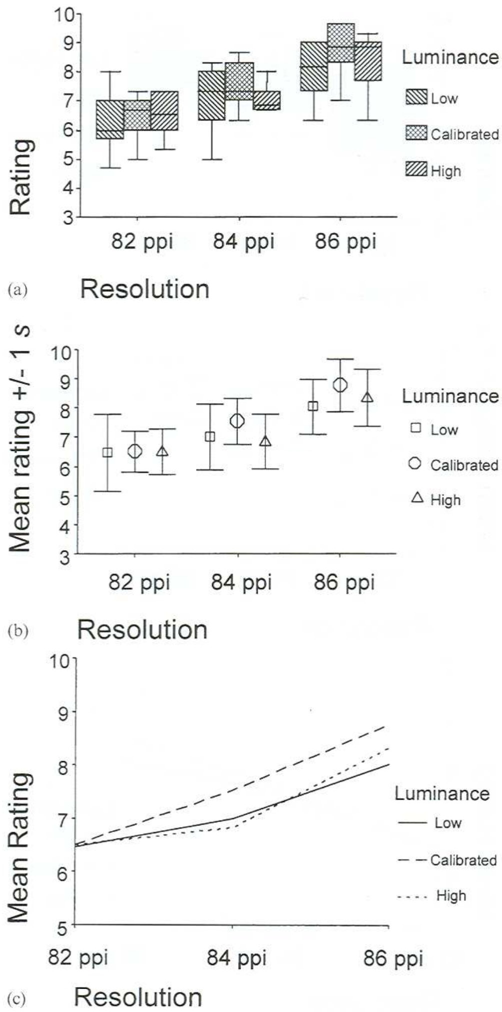

Fig. 7 Graphs for the ratings of the image 'bicycle' in a black background (ppi, pixels/in): (a) box plots showing the dispersion of the data; (b) standard deviation s; (c) average ratings

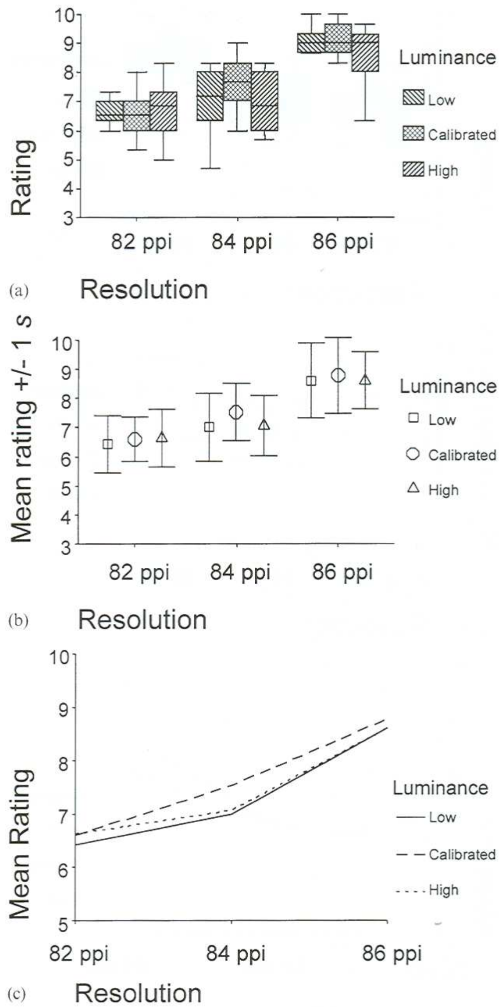

Fig. 8 Graphs for the ratings of the image 'bicycle' in a grey background (ppi, pixels/in): (a) box plots showing the dispersion of the data; (b) standard deviation s; (c) average ratings

examination of different sources of variability in a complex situation [**35**]. One of the assumptions in ANOVA is that the variabilities are equal among the population. One of the methods of measuring variability is standard deviation [**33**]. By observing the standard deviation for each variable and factor it was shown that the variability was reasonably constant from one to another [**35**]. An example for the average ratings of all images is given in Table 3.

It was observed, however, that there was a wide distribution of ratings between observers. This could be related to the subjective nature of judgements made without a reference image. Another reason could have been the method of rating. Training of the observers before performing the experiment [**36**] and increasing the number of images might provide more consistent ratings among the observers and more accurate results. This, however, would significantly increase the

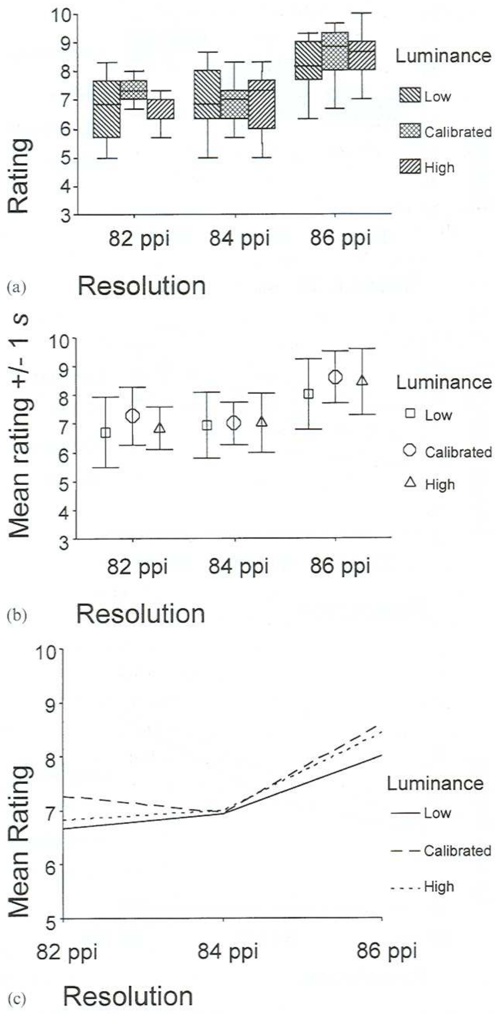

Fig. 9 Graphs for the ratings of the image 'house' (ppi, pixels/in): (a) box plots showing the dispersion of the data; (b) standard deviation *s*; (c) average ratings

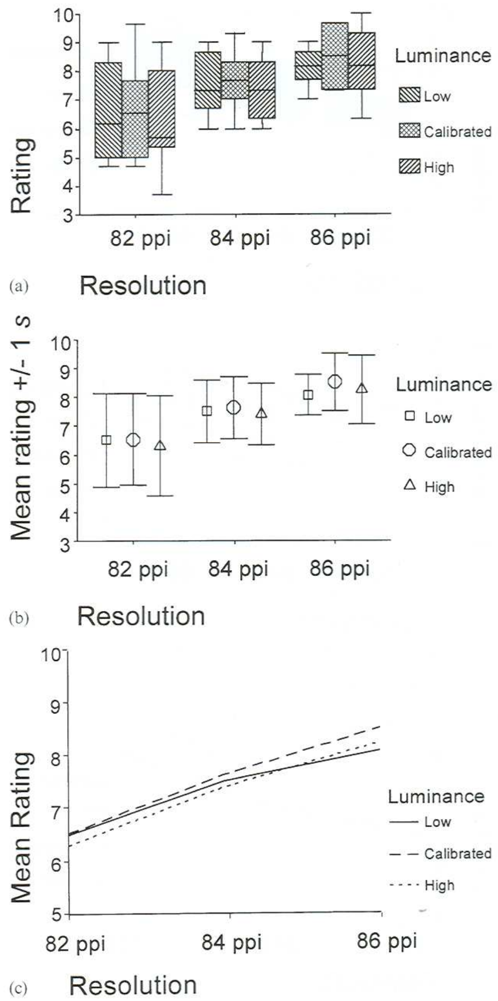

Fig. 10 Graphs for the ratings of the image 'portrait' (ppi, pixels/in): (a) box plots showing the dispersion of the data; (b) standard deviation *s*; (c) average ratings

final number of images viewed by the observers [**27**]. The effect of viewer fatigue on image rating during the experimental process would therefore have to be taken into consideration.

The experimental results were analysed using a two-way within-subjects ANOVA. The null hypotheses H_0 tested were the following:

1. The levels of each factor make no difference to perceived image quality.
2. There is no interaction between the two factors in determining perceived image quality.

The hypothesis test was performed by using an *F* test, which is based on the *F* ratio. The *F* ratio is based on two sources of variation, between-sample

Table 3 The mean rating and the standard deviation of the observers' ratings for each variable and each factor of all images, where N is the number of observations

Luminance	Resolution (pixels/in)	Mean rating	Standard deviation	N
Low	82	6.556	1.126	10
	84	7.145	0.820	10
	86	8.034	0.912	10
	Total	7.245	1.114	30
Calibrated	82	6.767	0.810	10
	84	7.388	0.684	10
	86	8.623	0.867	10
	Total	7.593	1.094	30
High	82	6.544	0.775	10
	84	7.078	0.928	10
	86	8.332	1.032	10
	Total	7.318	1.168	30
Total	82	6.623	0.891	30
	84	7.204	0.800	30
	86	8.330	0.939	30
	Total	7.385	1.123	90

variability and within-sample variability:

$$F = \frac{\text{between-sample variability}}{\text{within-sample variability}}$$

Statistical computations were performed using the SPSS v.9 software package. If the F ratio was less than 1, the effect was not significant. If the F ratio was over 1, then the p value was checked. The p value, which is the probability that the differences occur by chance, has to be less than 0.05 for the F ratio to be considered as significant [**32**].

Line graphs of quality rating against resolution for the average ratings of each image and for all the images together showed that there was some interaction between luminance and resolution, as mentioned before (Figs 6c, 7c, 8c, 9c and 10c). This was indicated by the crossing of the lines. The significance of this interaction was tested and shown in the results of the ANOVA table (Table 4). For each image the individual effect of luminance was shown not to be significant whereas the individual effect of resolution was significant. The interaction between luminance and resolution was also shown to be not significant.

ANOVA for the averaged results of all images showed that the effects of both display luminance and image resolution were significant when taken into account individually. This indicated that changes in display luminance setting affected the perceived image quality. The interaction between them, however, was not significant. It should also be mentioned that although the individual effects (resolution and luminance) were shown to be significant, the interaction between them was not significant because the individual effect was based on the average for each case while the interaction was based on all the combinations of the two factors [**35**].

All the previous results obtained with ANOVA showed that hypothesis 1 was rejected regarding the individual effect of resolution. For the individual effect of luminance on image quality, hypothesis 1 was accepted in all cases with the exception of the case where the individual effect of luminance was considered for all images combined. In this case, hypothesis 1 was rejected. It was also shown that hypothesis 2 was accepted since there was no significant interaction between the two factors: image resolution and display luminance.

Graphs of resolution versus mean rating $\pm$standard deviation (Figs 6b, 7b, 8b, 9b and 10b) gave information on data dispersion and on scene dependence of the ratings. Observations on scene dependence were also performed by examining the average rating graphs and the box plots.

Concerning the effect of image background (black and grey) on rated image quality, ANOVA showed that the effect of image background did not alter the significance of luminance and resolution (for the cases of individual effect and interaction) on perceived image quality (Table 4). In practice, differences in rating between the two cases were observed from the graphs. In Figs 7c and 8c it was shown that luminance differences affected the perceived quality of the image on a black background more than on a grey background. Box plots (Figs 7a and 8a) also showed that the observers gave more consistent results when the image was viewed on a grey background, although the standard deviation was still high (Figs 7b and 8b).

In the image 'portrait' there was no significant preference among luminance levels but there was clear preference for high resolution (Fig. 10c). The low-resolution image had the widest distribution of the ratings (Fig. 10a) and, hence, also the highest standard deviation (Fig. 10b). It also included the lowest rating from the observers. This agreed with comments from observers that images displayed at low resolution decreased the perceived image quality because the characteristics of the face were noticeably affected.

Table 4 ANOVA results for each image and all images together, where the subscripts in parentheses to the F ratio include the degrees of freedom (first number) and the degrees of freedom for the error associated with each main effect or interaction (second number)

	F ratio	p	Effect
Bicycle (black background)			
Luminance	$F_{(2,18)} = 1.494 > 1$	$0.251 > 0.05$	Not significant
Resolution	$F_{(2,18)} = 40.134 > 1$	$0.000 < 0.05$	Significant
Luminance × resolution	$F_{(4,36)} = 1.201 > 1$	$0.327 > 0.05$	Not significant
Bicycle (grey background)			
Luminance	$F_{(2,18)} = 0.792 < 1$	$0.561 > 0.05$	Not significant
Resolution	$F_{(2,18)} = 32.578 > 1$	$0.000 < 0.05$	Significant
Luminance × resolution	$F_{(4,36)} = 1.203 > 1$	$0.326 > 0.05$	Not significant
House			
Luminance	$F_{(2,18)} = 1.898 > 1$	$0.179 > 0.05$	Not significant
Resolution	$F_{(2,18)} = 19.556 > 1$	$0.000 < 0.05$	Significant
Luminance × resolution	$F_{(4,36)} = 1.048 > 1$	$0.396 > 0.05$	Not significant
Portrait			
Luminance	$F_{(2,18)} = 0.845 > 1$	$0.446 > 0.05$	Not significant
Resolution	$F_{(2,18)} = 8.779 > 1$	$0.002 < 0.05$	Significant
Luminance × resolution	$F_{(4,36)} = 0.538 < 1$	$0.709 > 0.05$	Not significant
All images			
Luminance	$F_{(2,58)} = 4.005 > 1$	$0.023 < 0.05$	Significant
Resolution	$F_{(2,58)} = 43.498 > 1$	$0.000 < 0.05$	Significant
Luminance × resolution	$F_{(4,116)} = 1.352 > 1$	$0.255 > 0.05$	Not significant

4 CONCLUSIONS

The experimental work showed the following:

1. Resolution affected the perceived image quality to a greater degree than luminance for typical computer monitor viewing distances.
2. The interaction between luminance and resolution was not significant in all cases.
3. Luminance did not have a significant effect on the perceived image quality for each image but had a significant effect when the ratings of all images were averaged.
4. For small changes in image resolution that did not significantly degrade image quality, as indicated by the preliminary tests (Section 2.4.1), the results showed that observers were more critical when judging images of different resolutions than different levels of luminance. This was an interesting result because the three luminance levels were chosen with the same criteria as resolution in the preliminary tests.
5. The distribution of observer ratings was wide. This indicated that a larger number of images with suitable subjects and a larger number of observers having initial training on rating before proceeding to the experiment might give more precise results. A different method of rating could also be a subject of consideration.

REFERENCES

1 **Lubin, J.** The use of psychophysical data and models in the analysis of display system performance. In *Digital Images and Human Vision* (Ed. A. B. Watson), 1993, pp. 163–178 (MIT Press, Boston, Massachusetts).

2 **Lubin, J., Pica, A. P.** and **Barbin, R. L.** Resolution vs. brightness trade-offs in CRT image quality, 1992 (unpublished, cited in reference [**1**]).

3 **Caronna, G., Santoro, G.** and **Testa, P.** Preference study of peak luminance vs. resolution for color television tubes. *J. Soc. Motion Pict. Telev. Engrs*, August 1998, 538–543.

4 **Ardito, M.** Studies of the influence of display size and picture brightness on the preferred viewing distance for HDTV programs. *J. Soc. Motion Pict. Telev. Engrs*, August 1994, 517–522.

5 **Lund, A. M.** The influence of video image size and resolution on viewing—distance preferences. *J. Soc. Motion Pict. Telev. Engrs*, May 1993, 406–415.

6 **Breneman, E. J.** The effect of level of illuminance and relative surround luminance on the appearance of

black-and-white photographs. *Photogr. Sci. Engng*, 1962, **6**(3), 172–179.

7 **Katoh, N., Nakabayashi, K., Ito, M.** and **Ohno, S.** Effect of ambient light on the color appearance of softcopy images: mixed chromatic adaptation for self-luminous displays. *J. Electron. Imaging*, 1998, **7**(4), 794–806.

8 **Berns, R. S.** and **Choh, H.-K.** Cathode-ray-tube to reflection-print matching under mixed chromatic adaptation using RLAB. *J. Electron Imaging*, **4**(4), 347–359.

9 **Rust, J.** and **Golombok, S.** *Modern Psychometrics: The Science of Psychological Assessment*, 1989 (Routledge, London).

10 **Guildford, J. P.** *Psychometric Methods*, 2nd edition, 1954 (McGraw-Hill, New York).

11 **Nunnally, J.** *Psychometric Theory*, 2nd edition, 1981 (Tata McGraw-Hill, New Delhi).

12 **Evans, S.** and **Attaya, W. L.** Research methods and strategies in the psychophysics of image quality. *Photogr. Sci. Engng*, 1978, **22**(2), 92–97.

13 **Thomas, W.** (Ed.) *SPSE Handbook of Photographic Science and Engineering*, 1973 (John Wiley, New York).

14 **Berns, R. S., Gorzynski, M. E.** and **Motta, R. J.** CRT colorimetry—Part I: theory and practice. *Color Res. Applic.*, 1983, **18**(5), 299–314.

15 **Travis, D.** *Effective Color Displays—Theory and Practice*, 1991 (Academic Press, London).

16 **Sanders, M. S.** and **McCormick, E. J.** *Human Factors in Engineering and Design*, 7th edition, 1992 (McGraw-Hill, New York).

17 CIE guidelines for coordinated research on evaluation of colour appearance models for reflection print and self-luminous display image comparisons. *Color Res. Applic.*, 1994, **19**(1), 48–58.

18 **Berns, R. S., Goszynski, M. E.** and **Motta, R. J.** CRT colorimetry—Part II: metrology. *Color Res. Applic.*, 1983, **18**(5), 315–325.

19 **Ford, A. M., Jacobson, R. E.** and **Attridge, G. G.** Assessment of a CRT display system. *J. Photogr. Sci.*, 1996, **44**, 147–154.

20 **Poynton, C.** 'Brightness' and 'contrast' controls, URL. www.inforamp.net/~poynton/notes/brightness_and_contrast/, 7 June 2001.

21 **Ford, A. M.** Relationships between image quality and still image compression. PhD thesis, University of Westminster, London, 1997.

22 **Trantaphillidou, S., Jacobson, R. E.** and **Ford, A. M.** Preferred tone reproduction of images on soft displays. In Proceedings of the International Congress on *Imaging Science* (*ICPS '98*), University of Antwerp, 7–11 September 1998, Vol. 2, pp. 204–209.

23 **Sproson, W. N.** *Colour Science in Television and Display Systems*, 1983 (Adam Hilger, Bristol).

24 **Pearson, D. E.** *Transmission and Display of Pictorial Information*, 1975 (Pentech Press, London).

25 **Poynton, C. A.** 'Gamma' and its disguises: the nonlinear mappings of intensity in perception, CRTs, film, and video. *SMPTE J.*, December 1993, 1099–1108.

26 **Ferguson, G. A.** and **Takane, Y.** *Statistical Analysis in Psychology and Education*, 6th edition, 1989 (McGraw-Hill, Singapore).

27 **Press, W. H., Flannery, B. P., Teukolsky, S. A.** and **Vetterling, W. T.** *Numerical Recipes, The Art of Scientific Computing*, 1986 (Cambridge University Press, Cambridge).

28 **Keys, R. G.** Cubic convolution interpolation for digital image processing. *IEEE Trans. Acousts, Speech, Signal Processing*, 1981, **29**(6), 1153–1160.

29 **Ward, J.** and **Cok, D. R.** Resampling algorithm for image resizing and rotation. *Digital Image Processing Algorithms*, Proceedings of the SPIE, Vol. 1075, 1989, pp. 260–269 (Society of Photo-optical Instrumentation Engineers, Bellingham, Washington).

30 **Harpster, J. L.** Visual performance on CRT screens and hard-copy displays. *Human Factors*, 1989, **31**(3), 247–257.

31 **Gingery, D.** and **Frusta, R.** Skimming electronic newspaper headlines: a study of typeface, point size, screen resolution and monitor size. *Inf. Processing Managmt*, 1997, **33**(5), 685–696.

32 **Brace, N., Kemp, R.** and **Snelgar, R.** *SPSS for Psychologists, A Guide to Data Analysis using SPSS for Windows*, 2000 (Macmillan, London).

33 **Anderson, D. R., Sweeney, D. J.** and **Williams, T. A.** *Essentials of Statistics for Business and Economics*, 2nd edition, 2000 (South-Western College Publishing, Mason, Ohio).

34 **Howell, D. C.** *Statistical Methods for Psychology*, 4th edition, 1997 (Duxbury Press, Belmont).

35 **Siegel, A. F.** *Practical Business Statistics*, 3rd edition, 1997 (Irwin/McGraw-Hill, Chicago, Illinois).

36 **Engeldrum, P. G.** *Psychometric Scaling: A Toolkit for Imaging Systems Development*, 2000 (Imcotek Press, Winchester, Massachusetts).