Abstract: The just noticeable difference (JND) in black level (BL), white level (WL), and chroma as visible in natural images under practical viewing conditions is determined. The potential effect of cultural background on the JNDS is evaluated conducting the same experiment in China and in the Netherlands. In general, there is a big difference in JND depending on the content of the natural image, as we expected. Only for some images, we found a difference in JND between China and the Netherlands. It seems that for these images people on average were looking to a different area in the image during their assessment.
Influence of cultural background on just noticeable difference in black level, white level and chroma for natural images

Shaoling Qin a, c, Shu Ge a, Hanchun Yin a, *, Jun Xia a and Ingrid Heynderickx b, c

a Dong Fei Display R&D Centre, Southeast University, Nanjing, P.R. China
b Philips Research Laboratories, Eindhoven, the Netherlands
c Delft University of Technology, the Netherlands

Abstract

The just noticeable difference (JND) in black level (BL), white level (WL), and chroma as visible in natural images under practical viewing conditions is determined. The potential effect of cultural background on the JNDs is evaluated conducting the same experiment in China and in the Netherlands. In general, there is a big difference in JND depending on the content of the natural image, as we expected. Only for some images, we found a difference in JND between China and the Netherlands. It seems that for these images people on average were looking to a different area in the image during their assessment.

Keywords

Just noticeable difference (JND), black level, white level, chroma, cultural background

* Corresponding author. Tel.: +86-25-83792449-208; fax: +86-25-83363222. E-mail address: yhc@seu.edu.cn (Hanchun Yin).
1. Introduction

Marketing research has revealed that image quality is an important decisive factor for end-users when buying an electronic display [19]. Hence, to understand how to improve existing or develop new display systems, the relation between image quality and the technology variables of the display system is a prerequisite. In this context, the term technology variables refers to all technological characteristics related to the display design, such as the pixel size, color filter thickness, pixel aspect ratio, transmission ratio, driving voltages, etc. To find the relation between image quality and technology variables, Engeldrum [18, 19] introduced the “Image Quality Circle” model, existing of three intermediate steps. The first step of the model starts from the observation that a customer’s image quality rating is a weighted sum of perceived image quality attributes. These are sub-aspects of quality, such as sharpness, colorfulness, brightness and image uniformity, which are usually assessed unconsciously by the viewer. In the second step, the model relates each perceived image quality attribute to the physical image characteristics. These characteristics include all optical and electronical aspects of the image that can be measured instrumentally, such as e.g. the output luminance, color gamut size, display’s white point, gamma value for each color channel, noise level, etc. Finally, by fully understanding the display's physics these physical image characteristics can be linked to the technology variables of the display system in the third step of the model.

The emphasis of our research is to better understand the first step in the Image Quality Model, i.e. to investigate the relative importance of the various image quality
attributes to the overall image quality rating. To be able to determine the weighting coefficients in this relation in a reliable way, two constraints have to be fulfilled: the image quality attributes should be varied (1) independently, and (2) to a comparable extent. The four most important image quality attributes, at least for non-expert viewers when assessing image quality of high-end TVs, are stated to be brightness, contrast, color rendering and sharpness [6]. These four attributes, however, are not independent and some of these attributes are complex, which means that they actually comprise a set of so-called sub-attributes. For example, perceived sharpness is known to be a combination of perceived detail rendering and perceived contour rendering, and to be affected also by perceived contrast [16]. One way of circumventing the problems related to the first constraint is by simplifying the problem, and considering the physical characteristics involved in these four perceived image quality attributes. This resulted in a new set of four variables affecting overall image quality, namely black level (BL), white level (WL), chroma (CS) and contour rendering (CR). BL and WL both influence perceived brightness and contrast. Chroma is one of the important factors determining perceived color rendering, and contour rendering is an important aspect of perceived sharpness, related to the sharpness of edges. The second constraint, namely that the four new variables should be varied to a comparable extent, can be fulfilled using just noticeable differences (JND), i.e. using in each variable an equal amount of steps that viewers just can distinguish. So, in order to be able to pursue our general research goal, namely to better understand the relation between image quality attributes and the overall image quality rating, we first have to investigate a
sub-problem, namely to determine the JND of WL, BL, CS and CR, which is the research discussed in this paper.

Literature already contains information on the JNDs of some image quality aspects. The JND in BL and WL, for example, can in principle be deduced from the DICOM standard [5]. This standard gives the JND in luminance as a function of luminance, determined for a side-by-side comparison of homogeneous test patterns under the condition that the human eye is fully adapted to the reference luminance level. T. Tamura [26] determined the JND in luminance non-uniformity as a function of background luminance, by investigating so-called “Mura” in liquid crystal displays (LCDs). He found that the Weber-Fechner’s law [10], stating that the visible luminance increase as a function of luminance is constant, was broken at dark backgrounds. For dark backgrounds, absolute luminance rather than relative luminance determined the visibility threshold. For chromaticity, the experiments reported in [12-14] resulted in the same order of magnitude for the perceptibility threshold in natural images, i.e. between 1 and 3 units color difference in CIE1976 (ΔE76). In professional or electronic prints, the uncertainty in color reproduction is typically of the order of several units of ΔE76. On the other hand, color difference thresholds below a ΔE76 of 1 are required in some fields of industrial coloring, when comparing, for example, colors of cars or textiles [22].

Literature on JNDs, so far, mainly focused on visibility thresholds determined with specific patterns under well-controlled viewing conditions. Assessing image quality, however, is usually done on natural images, viewed under more practical viewing
conditions. As a consequence, the results obtained in literature are not expected to be directly applicable to our research problem. First support to this hypothesis is already given in [8, 24], in which visibility thresholds in BL, WL, CS and CR are reported for critical, though realistic natural image content, viewed under practical viewing conditions. The JNDs found were, in general, higher than the ones predicted in literature. Moreover, they were strongly affected by image content. The JND of BL steadily increased with the averaged image luminance. The JND of WL was substantially lower for a mainly white image than for images with a more intermediate averaged luminance, but in general, this trend was less obvious. The JND of CS seemed to be more affected by the degree of details in the image than by its averaged chroma. Finally, the JND of CR was found to be \( \sigma = 1.3' \) when expressed in terms of the width (in visual angle) of a Gaussian convolution filter [11]. Its image content dependency was small. The results obtained so far already gave some rough indications on the JNDs, but were – apart from the JND in CR – insufficiently accurate to be used for establishing the relation between image quality aspects and overall image quality. Therefore, this paper reports on the JND of BL, WL and CS determined with a more accurate methodology, and for a larger set of images, in order to establish the image content dependency.

Additionally, a potential effect of cultural background on JND is investigated by performing the same experiment in China and the Netherlands. Lu Liu et al. already addressed the effect of cultural background on the JND of CR [11]. They found the same visibility threshold in sharpness between Chinese and Dutch people.
Nonetheless, differences in preference for contrast and color among different cultural backgrounds are known to exist. E.g. Fernandez and Fairchild [21] found a small, though statistical difference in preference for the image characteristics gamma, contrast and color reproduction, between different cultural backgrounds, i.e. Americans, Chinese, Europeans, and Japanese. Japanese in general prefer lighter images in comparison to all other sub-groups. Chinese prefer a higher contrast compared to Americans and Japanese, which is consistent with the results reported in [9] for Chinese and European people. For color reproduction, the Eastern people prefer more chroma compared to the Americans. Japanese demonstrate a preference towards redder or warmer colors compared to Americans, while Chinese demonstrate a shift towards bluer or cooler colors. Whether apart from these differences in preference also the JND in BL, WL and CS depends on cultural background is so far unknown, and is in this paper evaluated for Chinese and Dutch people.

2. Perception experiment

Two perception experiments, one in China and one in the Netherlands, were conducted with the aim to accurately determine the JND of BL, WL and CS using representative natural images.

2.1 Stimuli

Nine images with a spatial resolution of 620*700 were chosen as a representative mix of image content, covering a broad distribution of averaged brightness, colorfulness and level of detail. They are shown in Figure 1. All nine images were used to
determine the JND of BL and WL, whereas only the first seven images were used for
the JND in CS, since the last two images were already previously used for the JND in
CS [24].

(Figure 1)

Starting from the original images, series of stimuli were generated adapting the level
of BL, WL and CS separately. To change the BL and WL of an image, the luminance
component of each pixel was adapted, while keeping its chromaticity constant. This
was done in the linear xyY color space. The general approach is shown in Figure 2a.
First, the RGB-values of each pixel in the original image were scaled from the range
[0, 256] (for 8-bit images) to the range [0, 1]. Additionally, they were transformed to
the linear light domain by applying a gamma (i.e. power) of 2.2, as specified in the
EBU standard [7]. The RGB-values were then transformed to XYZ-values using the
primaries and white point specified in the EBU standard. These XYZ-values were
translated into xyY-values, of which the Y-component was used to change the BL and
WL of the image. As shown in Figure 2b, for the WL, the Y-value of each pixel in the
image was scaled down according to $Y_o = \alpha \cdot Y_i$, where the scaling factor $\alpha$ was
varied in the range [0.60, 1]. $Y_i$ and $Y_o$ are the input and output value of the
Y-component per pixel. For BL, the minimum luminance was first increased to a
certain value indicated as $1 - \alpha_b$, and then the Y-component per pixel was scaled to fit
the new range $[1 - \alpha_b, 1]$. This can be formulated as $Y_o = \alpha_y \cdot Y_i + 1 - \alpha, \quad$ where the
scale factor $\alpha_b$ was varied in the range [0.981, 1], and $Y_i$ and $Y_o$ are again the input
and output value of the Y-component per pixel. The resulting new xyY-values were
transformed back to XYZ-values, and consequently to RGB-values, using the inverse characterization model of the display panel applied in the experiment. In the Netherlands the white point of the display panel differed from D65, and as a consequence, a white-point shift was implemented in software by changing the relative contribution of the maximal luminance of R, G and B in the linear light domain.

Since natural images were used in the experiment, the actual distortion for all pixels in an image having a different gray level was not the same. In order to make the criteria comparable over all images in the experiment, the possible maximum distortion was taken as a measure for the JND in the natural images. For example, the (corresponding) distortion for full white (R=G=B=255) was used as a measure for the JND of WL, whereas the (corresponding) distortion for full black (R=G=B=0) was used as a measure for the JND of BL. Therefore, the calculated JNDs were larger than the actual distortions in the natural images observed by the subjects. The relation between the calculated JNDs and the actual distortions in the natural images was linear according to the principle shown in Figure 2b.

For changing the level of CS, a similar approach was used (see again figure 2a). The RGB-values per pixel were transformed (after scaling and applying gamma) into XYZ-values, then into LAB-values, and then into LCH-values. The CIELAB color space was selected to adapt the chroma-value (C-component), while keeping the lightness (L-component) and hue (H-component) constant. The step size in chroma change was expressed in $\Delta E_{94}$ instead of in $\Delta E_{76}$, as the former is perceptually more
uniform than the latter [15, 27]. The relevant equations to calculate a step size in $\Delta E_{94}$ are given in appendix 1, and were implemented in Matlab according to [25]. For $K_1$ a value of 0.045 (as prescribed for graphics) was used, yielding a relation between $\Delta C$ and $\Delta E_{94}$ of $\Delta E_{94} = \Delta C/(1 + 0.045C)$. The C-component per pixel was decreased with an amount of $\Delta E_{94}$, ranging between [0.1, 4.9].

(Figure 2)

It should be noted that the ranges for changing BL, WL and CS were defined such that the most distorted image in each variable was obviously different from the original image. For each variable, the whole range in $\alpha_s$, $\alpha_w$ and $\Delta E_{94}$ was divided in 49 intervals, in order to create sufficient stimuli to determine the JNDs with a high accuracy.

2.2 Experimental conditions

The experimental settings in China and in the Netherlands were the same except for the display used. In the Netherlands, a 40" LCD TV (Sony Qualia KDX40Q005) with a spatial resolution of 1368*768 was used. Its maximum and minimum luminance was 296.0 and 0.40 cd/m$^2$, respectively. In China, a 19" LCD monitor (Samsung 943N) with a spatial resolution of 1280*1024 was used. Its maximum and minimum luminance was 292.4 and 0.30 cd/m$^2$, respectively. The white point of both displays was adjusted (by hardware or software) to D65. Since the size of the stimuli was 620*700 pixels, two images could be displayed side by side simultaneously without scaling or clipping. The remaining pixels around the two images were set to a grey
background (R=G=B=128). The viewing distance was 4 times the screen height, which resulted in a viewing distance of 1.2m for the 19" Samsung monitor, and a viewing distance of 2.0m for the 40" Sony TV. The ambient illumination was 20 lx measured on the screen in the direction to the viewer, and about 18 lx measured on the background. The settings for the experiment in China and in the Netherlands are summarized in Table 1.

(Table 1)

2.3 Protocol

In both experiments, the JND was measured with the one-up-two-down two alternative forced-choice (2AFC) staircase method [3]. Subjects were requested to select the distorted image out of a pair of images, containing the distorted as well as the original one. The specific task for the subjects depended on the JND to be measured. For measuring the JND of WL, subjects had to indicate the darkest image, while for the JND of BL, they had to indicate the brightest image, and for the JND in CS they had to indicate the least colorful image. If the answer was right, the two images were displayed once more. If a subject gave two consecutive correct responses (i.e. indicating the distorted image), the distortion was decreased for the next pair of images. In case of a wrong response, the distortion was increased. This procedure of increasing and decreasing the distortion was repeated until the subject had nine reversals (i.e. changing from increasing to decreasing the distortion, or vice versa). The level of distortion at the last six reversal points was averaged to calculate the JND
value. The step size of increasing and decreasing the distortion changed at different reversals. The starting step size was eight intervals (of the maximum of 49 intervals). After each reversal, the step size was halved, until a minimum step size of one interval.

The user interface used in the experiment is shown in Figure 3. Both images, i.e. the distorted and the original one, were displayed simultaneously side-by-side with the position of the original image varying randomly between left and right for each presentation. The influence of a slight non-uniformity in the display was circumvented by horizontally flipping the right image with respect to the left image.

(Figure 3)

Before the experiment, each subject was asked to carefully read the written instructions about the purpose of the experiment and the tasks to finish. There were three tasks in total, i.e. the JND for BL, WL and CS separately. After each trial, each subject was asked to mark for that specific image and task, the area they focused on during the assessment. The order, in which the images were shown per task, and the order of the tasks was randomized over the subjects in order to avoid time effects. Since finishing all trials per task took about half an hour, subjects were free to select either one or two tasks per experimental session in order to make sure that the experiment was finished before they got exhausted.

2.4 Participants

The distribution of subjects is shown in table 2 for both experiments. In China, 11
males and 9 females with their age ranging from 22 to 27 performed all three tasks. In the Netherlands, 14 males and 6 females evaluated the JND of WL, while the JND for BL was assessed by 14 males and 7 females, and the JND for CS by 14 males and 8 females. The age of the participants in the Netherlands ranged from 22 to 60. Thus, in total, we had more than 20 subjects for each of the JNDs, and they evaluated at least 7 images per JND, which was substantially more than the minimum requirement of ten observers and three scenes as prescribed in the ISO 20462 standard [1]. All volunteers had (corrected to) normal vision ($\geq 1.0$ on a Landolt chart) and no color deficiencies (tested with the Ishihara color vision test).

(Table 2)

2.5 Statistical analysis

The raw data of the JNDs obtained from the staircase method were imported in SPSS version 13. These data were first explored on their reliability computing various descriptive statistics and by means of graphs [17]. The box plot of the JNDs obtained in China or the Netherlands was plotted per image, and outliers were excluded. Then the box plot was plotted again with the filtered data to exclude the remaining outliers. This procedure was repeated until there were no outliers any more. This is not a standard procedure; but, nonetheless, we did so to reduce the spread in our data in order to make our results more accurate in the determination of the JND. The amount of data remaining after this procedure is given in Table 3 per attribute.

(Table 3)
To get an impression of how this data exclusion procedure affected the JND values of China and the Netherlands, the box plot for the JND of BL based on the raw data and after filtering is plotted in Figure 4. For this comparison, the JND of BL was chosen, since it was the attribute for which most data were excluded by the procedure (see Table 3). Figure 4 shows that the filtering shrank the body of the box plot by removing the less reliable results; it, however, did not decrease the relative difference between China and the Netherlands in most cases.

(Figure 4)

For better understanding the image content dependency of the JNDs, a saliency map per image and JND variable (i.e. BL, WL or CS) was constructed from the areas marked by the viewers during the perception experiment. Based on the result, the similarity in saliency between China and the Netherlands was evaluated. The creation of the saliency map was similar as proposed in [4]. It started from a blank image having the same dimensions in pixels as the original image. First, the center of mass of each marked area was determined and considered as a position of fixation. For each fixated location, an identical 3D Gaussian of unit height was projected onto the blank image. The parameter sigma of the Gaussian was chosen to be 2 degrees width at half-height, which corresponded to 61 pixels and 46 pixels in China and in the Netherlands, respectively. This value was chosen to approximate the size of the fovea [20]. Then the fixations selected by the various subjects were superimposed and taken as the saliency map. As an illustration, the saliency map of the image “balloon” assessed in China for the JND of BL is shown in Figure 5. Figure 5a) shows the map
of fixations accumulated over all subjects and 5b) shows the corresponding projected contour plot. The redrawn image is shown in Figure 5c) with brighter areas corresponding to areas with a higher averaged degree of fixation. The resemblance between two saliency maps, e.g. for the same image assessed in China and in the Netherlands, is calculated as the Pearson correlation of the weights over all pixels. This resemblance value is then compared to the difference in JND between the two countries. This difference is calculated as:

\[ \Delta \text{JND} = \frac{\text{abs}(\text{JND}_{1} - \text{JND}_{2})}{[(\text{JND}_{1} + \text{JND}_{2})/2]} \]  

(1).

(Figure 5)

3. Results

To get a first impression of the JND of each image attribute as well as of the difference between China and the Netherlands, the average with 95% confidence interval of the JND of BL, WL and CS in both countries and for each original image is shown in Figure 6. Figure 6a shows that the JND of BL changes in the range from 0 to 1 cd/m² for all images in China, while in the Netherlands the JND is larger than 1.5 cd/m² for four images (i.e. “balloon”, “flower”, ”parrot” and ”toy”). These are also the images with the broadest error bars. For WL, the JND is roughly in the range from 10 to 50 cd/m² both in China and in the Netherlands. The JND strongly depends on image content, as shown in Figure 6b. It is smaller for those images with white areas, e.g. “rabbit”, “parrot” and “balloon”, and this trend is observed both in China and in the Netherlands. The JND of CS in China and the Netherlands is around 0.7 expressed
in $\triangle E94$ (as can be seen in Figure 6c) except for the image “carriage” in China, whose JND has an apparently large confidence interval.

(Figure 6)

To investigate the effect of image content and cultural background on JND further, a full ANOVA is performed on the JND of BL, WL and CS separately, taking into account image content and country as independent factors and subjects as a random factor, and including the two-way interaction between image content and country.

For the JND of BL, the effect of image content ($p<0.001$, $F=14.9$, $df=8$), country ($p<0.001$, $F=90.5$, $df=1$) and their interaction ($p<0.001$, $F=15.6$, $df=8$) are all significant. The latter indicates that the image content dependency is different between China and the Netherlands. Hence, a post-hoc analysis over image content is performed per country; the results are shown in Table 4, and clearly confirm the difference in image content dependency between both countries. The main effect of country on the JND of BL results from the fact that the JND on average is lower in China than in the Netherlands. This effect is mainly dominated by the JND of the four images, scored considerably higher in the Netherlands than in China.

(Table 4)

For the JND of WL, the effect of image content ($p<0.001$, $F=18.5$, $df=8$) and country ($p<0.001$, $F=14.9$, $df=1$) are both significant. The interaction between image content and country ($p=0.100$, $F=1.7$, $df=8$) tends to be significant, as can also be observed in Figure 6b, where the JND of WL is clearly different between China and the
Netherlands only for the images “balloon” and “house”. The results of a post-hoc analysis over image content per country are shown in Table 4. The main effect of country on the JND of WL results from the higher values found on average in China than in the Netherlands.

For the JND of CS, the effect of image content (p<0.001, F=12.8, df=6), country (p<0.001, F=12.3, df=1) and their interaction (p<0.001, F=6.4, df=8) are all significant. Since the image content dependency is not the same for both countries, a post hoc analysis over image content was done per country; the results are shown in Table 4. There is no significant difference between images in the Netherlands. In China, only the image “carriage” is found to be significantly different from the other images. It has been checked that when excluding the image “carriage” from the data set, the new ANOVA revealed that the effect of image content (p=0.06, F=2.2, df=5), country (p=0.21, F=1.6, df=1) and their interaction (p=0.27, F=1.3, df=5) was no longer significant.

4. Discussion

The JND of BL, WL and CS as well as their image content dependency are studied in this paper with several natural images in China and in the Netherlands.

The JND of BL as found in this paper for natural images mainly changes in the range from 0 to 1 cd/m\(^2\). For comparing these values to the ones predicted by the DICOM standard [5], we determined the luminance value averaged over the saliency map for the BL. These luminance values vary between 2.7-60 cd/m\(^2\) depending on the image.
content. The corresponding JND given in the DICOM standard [5] ranges between 0.05-0.5 cd/m². As expected, the JND of BL found for natural images under common viewing conditions is larger than that obtained from the DICOM standard using specific test patterns under critical viewing conditions.

For WL, the JND is roughly in the range from 10 to 50 cd/m² in both countries. If we compare these values to the DICOM standard [5], we again find much larger values than predicted by the standard. Indeed, the luminance averaged over the saliency map for the WL varies between 4.2-100 cd/m², depending on the image (the low averaged luminance level belongs to the image “Carriage”). The corresponding JNDs in the DICOM standard [5] are between 0.06-0.7 cd/m². Similar as for the JND of the BL, the JND for the WL was expected to be larger in case of natural image content measured under common viewing conditions than in case of specific test patterns measured under critical conditions. But, in the case of the JND for the WL the difference between our results and the values found in the DICOM standard is very large. This may be explained partly by the fact that the actual changes as implemented in the images are smaller than the indicated JND (giving the change on the WL). This is especially true for natural images with low to intermediate intensity levels, for which the actual change in the image linearly scales with the JND (see also section 2.1). The latter is confirmed by the observation that the JND of WL for those images with white areas, e.g. “rabbit”, “parrot” and “balloon”, is in the range of 10-25 cd/m², i.e. smaller compared to the JND for the other images. These results are also consistent with results reported in literature ([22] and [24]), indicating that people are
most sensitive, i.e. the JND value is lower, for images at higher luminance level than for images with intermediate luminance level.

The JND of CS is about 0.7 expressed in $\Delta E_{94}$, which corresponds to about 2 to 4 units in $\Delta E_{76}$, depending on the hue and for intermediate lightness levels. Thus, our result is comparable to results found in literature [12-14], where a perceptibility threshold for color differences in natural images between 1 to 3 units in $\Delta E_{76}$ is reported.

In general, we found similarities in where participants were looking at in an image, while assessing a given JND. For the JND of BL, the subjects were mainly focusing on the relatively dark areas in the images. The more brighter and colorful areas in the images were used to assess the JND of WL, while the most colorful areas were also used to assess the JND of CS.

The differences in JND found between China and the Netherlands are not systematic. For most images, the JND is equal between both countries, but for some images, there is a difference between both countries. The possible reason for the difference between the two countries may lie in the difference in experimental setup or in possible differences in assessing the images.

The difference in experimental set-up was limited to a different display, which, nonetheless, may have had some effect on the JND values obtained. The two panels were roughly the same in dynamic luminance range, but different in panel size and spatial resolution. To alleviate the effect caused by the difference in panel size, the
viewing distance was always kept at 4 times the screen height in order to keep the visual angle covered by the displays the same. However, the spatial resolution of both displays was different. As a consequence, the displayed images have a visual angle of about 9-10 degrees in China and about 12-13 degrees in the Netherlands. However, we do not expect that this difference in visual angle affected the JND values of luminance and chromaticity, mainly because the differences found were not systematic, but rather related to some particular image content.

To explain the potential difference between China and the Netherlands in assessing the images, the corresponding saliency maps are studied. The relation between the JND found and the saliency map is shown in Figure 7. The ΔJND is calculated according to Equation (1), and the resemblance between the two saliency maps (Sim) for the same image assessed in China and in the Netherlands is taken as the Pearson correlation of the weights over all pixels. Figure 7a gives an example for BL, WL and CS for which the distribution of the JNDs is similar between China and the Netherlands. In those cases, participants also clearly have been looking to the same areas in the images during the assessment, as indicated by the relatively high value of Sim. To assess the JND of BL, the subjects focused on the dark green carriage or on the bunch of flowers in the image “carriage”. For the JND of WL they focused on the bright green leaf or red flower in the image “flower”, and for the JND of CS they mainly looked at the yellow coat or red skirt in the image “children”. The examples given in Figure 7b correspond to images for which the distribution of the JNDs is dissimilar between China and the Netherlands. In these cases, the Chinese and Dutch
participants focused on different areas in the image during their assessment, which is reflected by the relatively low value of Sim for these images. For example when assessing the JND of BL for the image “toy”, the Chinese participants focused more on the blocks, while the Dutch participants focused more on the shadow lines between the blocks. When judging the JND of WL for the image “house”, the Dutch participants also focused on the blue sky or blue roof, while this was less the case for the Chinese participants. Finally, when assessing the JND of CS the Dutch participants focused more on the green carriage in the image “carriage” as compared to the Chinese participants. Hence, the difference in JND between both countries tends to be the result of looking to a different area in the image. This hypothesis will be further explored in our future research related to modeling image content dependency of JND.

(Figure 7)

5. Conclusion

The JND of BL, WL and CS, as visible in representative natural image content, were measured with the one-up-two-down 2AFC staircase method under common viewing conditions. The same experiment was conducted in China and in the Netherlands to evaluate a potential cultural difference as well.

In general, there is a big difference in JND depending on the image content, as we expected. The JND of BL changes in the range from 0 to 1 cd/m² for all images in China. In the Netherlands, the JND is similar for most images, but larger than 1.5
cd/m² for four out of the nine images. For WL, the JND is roughly in the range from 10 to 50 cd/m² both in China and in the Netherlands. Except for two images, the JND of WL is highly similar between China and the Netherlands. The JND of CS in China and in the Netherlands is around 0.7 expressed in ∆E94 for most images. Here, there is only one image, for which the JND is different between both countries.

To assess the JND, the volunteers mainly focus on darker areas in the images for the BL, on brighter and more colorful areas in the images for the WL, and on the most colorful areas in the images for the CS.

The differences in JND found between China and the Netherlands were not systematic, but rather related to particular images. For those images, for which we found a difference in JND between both countries, it seems that people on average were looking to a different area in the image during their assessment. To get a better understanding of the image content dependency in JND, a model relating JNDs to image characteristics considering the region of interest will be explored in our future research.

6. Acknowledgement

This research was supported by the China Scholarship Council. The experiment in the Netherlands was performed at the Philips Research Laboratories in Eindhoven. The authors would like to thank Arnold Van Keersop, Stefan Swinkels, Pieter Seuntiens, Justin Laird, and Ingrid Vogels from Philips Research for helping them with the setup of the perception experiment. The authors would also like to thank Hani Al-Ers for
his help with the saliency maps. In the end, the authors are grateful to all volunteers who participated into their experiment.

7. References

7. ITU-R BT.709-5: Parameter values for the HDTV standards for production and international programme exchange
8. J. Xia, S.L. Qin et al. The just noticeable difference in chromaticity, black level, white level and contour rendering in natural images, IDW/AD05, (2005), 1821-1824.


25. Stephen Westland, Caterina Ripamonti, Computational colour science using MATLAB, (Wiley; Chichester, 2004)


8. **Appendix:** Color difference calculation in CIE 1976 and CIE 1994

1. **Delta E (CIE 1976)**
   The color difference, or $\Delta E$, between two colors $L_1, a_1, b_1$ and $L_2, a_2, b_2$ is:
   
   $$\Delta E_{76} = \sqrt{(L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2}$$

2. **Delta E (CIE 1994)**
   The color difference, or $\Delta E$, between a sample color $L_2, a_2, b_2$ and a reference color $L_1, a_1, b_1$ is:
   
   $$\Delta E_{94} = \sqrt{\left(\frac{\Delta L}{K_L S_L}\right)^2 + \left(\frac{\Delta C}{K_C S_C}\right)^2 + \left(\frac{\Delta H}{K_H S_H}\right)^2}$$

   Where
   
   $\Delta L = L_1 - L_2$
   $\Delta C = C_1 - C_2$
   $\Delta H = \sqrt{\Delta a^2 + \Delta b^2 - \Delta C^2}$
   
   $$C_1 = \sqrt{a_1^2 + b_1^2}$$
   $$C_2 = \sqrt{a_2^2 + b_2^2}$$
   
   $\Delta a = a_1 - a_2$
   $\Delta b = b_1 - b_2$
   
   $S_L = 1$
   $S_C = 1 + K_L C_1$
   $S_H = 1 + K_C C_1$

   $K_L = \begin{cases} 1 & \text{default} \\ 2 & \text{textiles} \end{cases}$
   $K_C = 1$ \hspace{1em} \text{default}
   $K_H = 1$ \hspace{1em} \text{default}

   $K_{1} = \begin{cases} 0.045 & \text{graphic arts} \\ 0.048 & \text{textiles} \end{cases}$
   $K_{2} = \begin{cases} 0.015 & \text{graphic arts} \\ 0.014 & \text{textiles} \end{cases}$
Table 1: Experimental settings in China and in the Netherlands

<table>
<thead>
<tr>
<th>Items</th>
<th>China</th>
<th>The Netherlands</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display panel</td>
<td>19” LCD monitor</td>
<td>40” LCD TV</td>
</tr>
<tr>
<td></td>
<td>Samsung 943N</td>
<td>Sony KDX40Q005</td>
</tr>
<tr>
<td>Spatial resolution</td>
<td>1280*1024</td>
<td>1368*768</td>
</tr>
<tr>
<td>Minimum luminance</td>
<td>0.30cd/m²</td>
<td>0.40cd/m²</td>
</tr>
<tr>
<td>Maximum luminance</td>
<td>292.4 cd/m²</td>
<td>296.0 cd/m²</td>
</tr>
<tr>
<td>White point</td>
<td>D65</td>
<td>Simulated D65</td>
</tr>
<tr>
<td>Viewing distance</td>
<td>about 1.2 m</td>
<td>about 2.0 m</td>
</tr>
<tr>
<td></td>
<td>4* screen height</td>
<td></td>
</tr>
<tr>
<td>Ambient luminance</td>
<td>20 lx measured on the screen in the direction to the viewer, and about 18 lx measured on the background</td>
<td></td>
</tr>
<tr>
<td>Image materials</td>
<td>The same nine original images for WL and BL, and the same seven original images for CS</td>
<td></td>
</tr>
</tbody>
</table>
### Table 2: The breakdown of volunteer population in China and in the Netherlands

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Age</th>
<th>Male</th>
<th>Female</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL</td>
<td>11</td>
<td>9</td>
<td>22-27</td>
<td>14</td>
<td>7</td>
<td>22-60</td>
</tr>
<tr>
<td>WL</td>
<td>11</td>
<td>9</td>
<td></td>
<td>14</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>11</td>
<td>9</td>
<td></td>
<td>14</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Summary of data exclusion for BL, WL and CS in China and in the Netherlands

|     | China          |                  | Netherlands     |                  |
|-----|----------------|------------------|------------------|
|     | original       | outliers         | remaining        | original        |
| BL  | 180            | 27               | 153              | 189             |
|     |                |                  |                  | 39              |
|     |                |                  |                  | 150             |
| WL  | 180            | 2                | 178              | 180             |
|     |                |                  |                  | 11              |
|     |                |                  |                  | 169             |
| CS  | 140            | 12               | 128              | 154             |
|     |                |                  |                  | 1               |
|     |                |                  |                  | 153             |
Table 4: Post hoc analysis for JND of BL, WL and CS on image content dependency. The images underlined together indicate there is no significant difference in between.

<table>
<thead>
<tr>
<th>JND of</th>
<th>Netherlands</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL</td>
<td>globe&lt;house&lt;rabbit&lt;carriage&lt;children&lt;flower&lt;balloon&lt;parrot&lt;toy</td>
<td>parrot&lt;carriage&lt;toy&lt;house&lt;globe&lt;balloon&lt;flower&lt;children&lt;rabbit</td>
</tr>
<tr>
<td></td>
<td>____________</td>
<td>____________</td>
</tr>
<tr>
<td>WL</td>
<td>house&lt;balloon&lt;toy&lt;rabbit&lt;parrot&lt;flower&lt;children&lt;globe&lt;carriage</td>
<td>toy&lt;balloon&lt;rabbit&lt;parrot&lt;flower&lt;house&lt;globe&lt;children&lt;carriage</td>
</tr>
<tr>
<td></td>
<td>____________</td>
<td>____________</td>
</tr>
<tr>
<td>CS</td>
<td>parrot&lt;flower&lt;children&lt;globe&lt;rabbit&lt;balloon&lt;carriage</td>
<td>rabbit&lt;parrot&lt;balloon&lt;children&lt;globe&lt;carriage</td>
</tr>
<tr>
<td></td>
<td>____________</td>
<td>____________</td>
</tr>
</tbody>
</table>
Figure 1: The nine original images used in the experiments; all images were used for the JND of WL and BL, while only the first seven images were used for the JND of CS.
Figure 2: a) Illustration of the process used to change the image BL, WL and CS and b) illustration of the linear scaling of Y in the xyY color space for BL.
Figure 3: User Interface with the distorted and original image displayed simultaneously on the display screen against a grey background. The right image was horizontally flipped with respect to the left image to avoid a strong influence of luminance inhomogeneity in the monitor.
Figure 4: Distribution of the JNDs of BL in China and the Netherlands: a) raw data and b) after excluding outliers. The blue boxes with striped pattern represent the data of the Netherlands and the green homogeneous boxes represent the data of China.
Figure 5: Illustration of the saliency map for the assessment of the JND of BL in China for the image “balloon”: a) Combination of marked fixations (with Gaussian spread) for all subjects, b) Corresponding contour plot, and c) image content with brighter areas corresponding to a higher degree of attention.
Figure 6: Mean and 95% confidence interval for a) JND of BL, b) JND of WL and c) JND of CS in China and in the Netherlands. The blue triangles with whiskers are for the Netherlands and the green circles with whiskers are for China.
Figure 7: The histograms of JND values over all subjects and the corresponding saliency maps for some of the images assessed in China (CN) and in the Netherlands (NL): a) examples for which the distribution of JND values is similar for both countries, and b) examples for which the distribution of JND values is different between both countries.