Effect of experimental methodology on the JND of the black level for natural images

Citation for published version (APA):

Qin, S. L., Ge, S., Yin, H. C., Liu, L., & Heynderickx, I. E. J. (2009). Effect of experimental methodology on the JND of the black level for natural images. *Journal of the Society for Information Display*, 17(8), 687-694. <https://doi.org/10.1889/JSID17.8.687>

DOI:

[10.1889/JSID17.8.687](https://doi.org/10.1889/JSID17.8.687)

Document status and date:

Published: 01/01/2009

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Effect of experimental methodology on the JND of the black level for natural images

ShaoLing Qin

Shu Ge

HanChun Yin

Lu Liu

Ingrid Heynderickx (SID Fellow)

Abstract — The effect of experimental methodology on the Just Noticeable Difference (JND) of the black level (BL) is assessed using a set of representative natural images. Two methodologies, known in psychophysics to determine JNDs, are mutually compared: an adapted method of adjustment with a same-difference task and a one-up-two-down staircase method with a two-alternative forced-choice task. Thus far, the literature has shown contradictory results for which of the two methods yields the lower difference thresholds. Experimental results show that the second method yields lower JNDs, but is slightly less efficient. Based on this second method, it was found that the JND of the BL for natural images ranges from 0.2 to 1.0 cd/m², varying with image content significantly.

Keywords — Adapted method of adjustment (AMOA), staircase method (SM), two alternative forced choice (2AFC), just-noticeable difference (JND), black level (BL).

DOI # 10.1889/JSID17.8.687

1 Introduction

Image quality is considered as one of the important decisive factors when buying a display.¹ The aspects dominating image quality include color rendering, brightness and contrast.¹ They are all affected by not only the black level (BL) of the display, but also characteristics, such as the white level (WL), the chromaticity of the primaries, and the gamma of the display function. In order to determine the relative importance of each of these characteristics to image quality, they all should be varied in a perceptual equivalent way. One option for doing so is using JNDs.^{2,3}

In principle, the JND of BL and WL can be deduced from the DICOM standard.⁴ This standard gives an increase in luminance for one JND as a function of luminance. This luminance increase is determined for a side-by-side comparison of two homogeneous squared patches (one test patch and one reference patch) under the condition that the human eye is fully adapted to the luminance level of the reference patch. Since these JNDs are determined with specific patterns under well-controlled viewing conditions, they are not necessarily directly applicable to natural images viewed under more practical viewing conditions. First support to this hypothesis is already given in Refs. 2 and 3, in which visibility thresholds in BL and WL are reported for critical, though realistic, natural image content, viewed under practical viewing conditions. In general, the JNDs found there are higher than the ones predicted by the DICOM standard. Moreover, they are strongly affected by image content.

Since an accurate value of the JND of the BL is prerequisite for our future research,^{2,3} we want to investigate the effect of measurement methodology. For this investigation, we limit ourselves to the determination of the JND of

the BL. In the research field of psychophysics, the three classical methodologies to measure a difference threshold are: (1) the method of constant stimuli, (2) the method of limits, and (3) the method of adjustment.^{5,6} For the method of constant stimuli, stimuli with different degrees of difference are repeatedly shown to the viewer, and the viewer is asked to respond whether he sees the difference. Although this method can achieve high accuracy, yet it is inefficient.^{7,8} Considering the variation in image content that we want to include, the method of constant stimuli becomes impractical. With the method of limits, descending and ascending series of difference between images are repeatedly shown to the viewer, and he is asked to indicate at what level he starts seeing the difference. This method is less accurate than the method of constant stimuli, but it is far less time consuming, and therefore, used much more extensively.^{5,9,10} To improve its efficiency further, a variation on the method of limits is developed; namely, the staircase or up-down method. In this method, the descending and ascending series of stimuli do not always start at extreme values, but alternate and start at the reversal point indicated by the viewer. The advantage of the staircase method is that it is still accurate, but its disadvantage is that it only yields reliable results for stabilized perceptions (*i.e.*, it is not affected by, *e.g.*, adaptation phenomena).¹¹ The third method is the method of adjustment, in which the viewer can freely adjust the difference until it is just not perceived.^{5,6,12} Compared to the staircase method, the method of adjustment is even more efficient. It gives the observer a larger amount of active participation in the experiment, which may help to prevent boredom, and therefore, to maintain high performance.^{5,12} Nevertheless, it is considered to be inaccurate because the threshold is not necessarily identical for ascending and descending changes.^{5,6,13}

Received 12/11/08, accepted 04/09/09.

ShaoLing Qin, Shu Ge, HanChun Yin, and Lu Liu are with Dong Fei Display R&D Center, School of Electronic Science and Engineering, Southeast University, #2 Sipailou, Nanjing, Jiangsu 210096, P.R. China; telephone +86-25-83792449, fax +86-25-83363222, e-mail: shaolingqin@gmail.com. I. Heynderickx is with Philips Research Laboratories, Eindhoven, The Netherlands, and Technical University Delft, Delft, The Netherlands.

© Copyright 2009 Society for Information Display 1071-0922/09/1708-0687\$1.00

When choosing a psychophysical methodology, the sensitivity of the test, as well as the number of trials and the total duration of the procedure needs consideration. Hence, it is important to understand the differences among methodologies with respect to these aspects. Nonetheless, direct comparisons between the classical methodologies are limited in the literature. To determine the most comfortable loudness level (MCL) for speech, Loftiss¹⁴ and Kopra *et al.*¹⁵ found that the results with a method of limits were not different from or only insignificantly higher than those obtained with the method of adjustment. To diagnose glaucoma in the field of ophthalmology, T. Maddess *et al.*¹⁶ found that the thresholds obtained with the staircase method and the method of adjustment were markedly different; the thresholds with the method of adjustment were higher than those with the staircase method. Nonetheless, both methods were effective in diagnosing glaucoma.¹⁶ Lakshminarayanan *et al.*¹⁷ evaluated the effect of methodology on speed discrimination using the staircase method, the method of adjustment, and a modified method of limits. Their results showed that the staircase method yielded significantly lower thresholds than the other two psychophysical methods. In summary, literature, discussing comparisons between different methodologies for measuring a difference threshold, show conflicting results or suggests a context dependency. Therefore, we found it appropriate to compare the effect of methodology for the determination of the JND of the BL.

Regarding the task, both the same-different task and the two-alternative forced-choice (2AFC) are basic procedures. In the same-different task, each trial consists of the presentation of two stimuli, which are either the same or different, and the observer has to report whether in his opinion the two stimuli are indeed the same or different. In the 2AFC, the observer must choose which of the two stimuli satisfies a given question.^{18,19} The 2AFC is always preferred over the same-different task for determining thresholds

because the 2AFC procedure discourages response biases and produces an especially high level of performance.^{5,6,9}

Based on the knowledge available in the literature, we chose two methods to determine the JND of the BL for natural images: the method of adjustment and the staircase method. The first method (AMOA) is an adapted version of the method of adjustment with a same-different task, in which the subject is forced to go up and down for several times. The subject is asked to adjust the test stimulus so that it is perceived the same as the reference. Subsequently, the adjustment direction is reversed, and the subject is asked to adjust the test stimulus until it is just distinguishably different from the reference. This procedure is repeated a couple of times, depending on a predefined number of reversals. The threshold is calculated as the averaged value over the reversal points. The second method (2AFCSM) is an extension of the staircase method; namely, the one-up-two-down staircase method.^{11,18–21} In this case, subjects are not asked whether they see the difference, but they are asked to indicate the image with, *e.g.*, the higher BL. Since they have to choose one out of the two images shown, it is based on a 2AFC task. The difference in BL is only decreased when the subject gives two consecutive correct responses (*i.e.*, twice in a row the image with the higher BL is selected). The difference in BL is increased after each incorrect response (*i.e.*, as soon as the image with the lower BL is selected). Hence, it is named as the one-up-two-down staircase method. This process is also repeated for a predefined number of reversals in adjusting the difference in BL, and the threshold is determined as the average over these reversals.

2 Perception experiment

We performed two experiments with the aim to evaluate the effect of experimental methodology on the JND of the BL for natural images. In the first experiment, an adapted

FIGURE 1 — The nine original images used in the subjective experiments.

method of adjustment with a same-different task was used, while in the second experiment, a one-up-two-down staircase method with a 2AFC task was employed.

2.1 Methodology

For the first experiment, a reference image, a test image, and a tuning button were shown to the subject. The subject was asked to indicate whether the two images – side-by-side shown on the panel – were the same or different. For the first pair of images, the test image was obviously different from the reference image, and a “different” response was expected. Then, the subject was instructed to decrease the difference in the BL by means of the tuning button until the response would be “same.” Subsequently, this point was recorded as a reversal and the adjustment direction was reversed. Then, the subject was requested to increase the difference with the tuning button until the response would be “different” again. Also, this reversal point was recorded, and the adjustment direction was reversed again, so that the difference was decreased again. This procedure was repeated using gradually smaller step sizes, and finally stopped after a sufficient number of reversals was reached (see below for more details).

In the second experiment, a test image and a reference image were displayed side-by-side. The subject was asked to indicate which of the images had the higher BL. Even when they did not see a difference, they were forced to make a choice. For the first pair of images in the experiment, the test image was picked out easily since its difference with the reference image was obvious. When a subject selected the right image out of the pair (*i.e.*, the one with the higher BL), the two images were shown again with their position at the left or right side of the panel randomized. Once a subject gave two consecutive correct responses, the difference in BL was decreased for the next pair of images. However, as soon as an incorrect response was given, the difference in BL was increased again. Each change from a decrease in difference to an increase in difference or vice versa was recorded as a reversal point. The whole process continued until a given number of reversals was reached (for details see below).

For both experiments, the step size of increasing or decreasing the difference in BL changed at different reversals. The starting step size was eight intervals (of the maximum of 49 intervals; for an explanation see Sec. 2.2). After each reversal, the step size was halved, until a minimum step size of one interval was used. The procedure of increasing or decreasing the difference in the BL was stopped after

TABLE 1 — Summary for JND of BL per image as determined with an AMOA with a same-different task and a one-up-two-down staircase method with a 2AFC task.

Image	Mean chromaticity	Mean luminance	Method	Mean	Num. of subjects	Std. Deviation
Balloon	42.3	38.79	AMOA	.80	18	.70
			2AFCSM	.58	15	.35
Carriage	26.7	8.43	AMOA	.30	18	.12
			2AFCSM	.24	17	.15
Children	28.6	56.18	AMOA	1.02	20	.76
			2AFCSM	.61	18	.51
Flower	45.9	46.14	AMOA	.65	20	.43
			2AFCSM	.61	20	.52
Globe	44.0	29.10	AMOA	.62	20	.36
			2AFCSM	.48	20	.36
House	39.0	33.70	AMOA	.58	20	.51
			2AFCSM	.46	17	.29
Parrot	30.8	54.18	AMOA	.41	18	.33
			2AFCSM	.24	16	.12
Rabbit	17.9	64.26	AMOA	1.04	18	.68
			2AFCSM	.75	16	.61
Toy	60.0	40.81	AMOA	.77	19	.54
			2AFCSM	.32	14	.12

FIGURE 2 — (a) Illustration of the process used to change the BL of a natural image and (b) illustration of the linear scaling of Y in the xyY color space.

nine reversals. The threshold was calculated as the average over the last six reversals.

2.2 Stimuli

To evaluate a possible image-content dependency, nine images with a spatial resolution of 620×700 pixels were chosen as a representative mix, covering a broad distribution of averaged brightness, colorfulness, and level of detail. The stimuli are shown in Fig. 1, while the mean brightness and mean chromaticity of these stimuli are shown in Table 1.

Starting from the original images, a series of stimuli were generated adapting the level of the BL. To do so, 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 Fig. 2(a). First, the RGB value of each pixel in the original image was scaled from the range $[0, 255]$ (for 8-bit images) to the range $[0, 1]$. Additionally, they were transferred to the linear light domain by applying a gamma (*i.e.*, power) of 2.2, as specified in the Rec. 709 standard.²² The RGB values were then transformed to XYZ values using the primaries and white point specified in the Rec. 709 standard. These XYZ values were translated into xyY values, the Y component of which was used to change the BL of the image. The way the BL was changed is shown in Fig. 2(b). First, the minimum luminance was increased to a certain value indicated as $1 - \alpha_b$, and then the Y component per pixel was scaled to the new range $[1 - \alpha_b, 1]$. This can be formulated as $Y_o = \alpha_b Y_i + 1 - \alpha_b$, where the scale factor α_b was varied in the range $[0.981, 1]$, and Y_i and Y_o were 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 the RGB values using the inverse characterization model of the display panel of the experiment.

2.3 Experimental conditions

In both experiments, a 19-in. LCD monitor (Philips190P5EG/93) with a spatial resolution of 1280×1024 was used. Its maximum and minimum luminance was 238.0 and 0.33 cd/m^2 , respectively. The white point of the display was adjusted to D65. Since the size of the stimuli was 620×700 pixels, two images could be displayed side-by-side without scaling or cropping. The asymmetry in luminance in the horizontal direction over the display screen was compensated as good as possible by software; *i.e.*, by linearly scaling the maximal light output to an approximately constant value in the horizontal direction.

The user interface used in each experiment is shown in Figs. 3(a) and 3(b), respectively. They were the same, except for the two adjustment buttons added to the interface for the first experiment. In both experiments, the test and reference image were displayed side-by-side with their position random between the left and right side for each presentation. The right image was flipped horizontally with

FIGURE 3 — User interface with the test and reference image displayed simultaneously on the display screen against a gray background; (a) for the first experiment using the AMOA and (b) for the second experiment using the 2AFCSM. The only difference is the presence of the adjustment buttons in the interface for the first experiment.

respect to the left image to circumvent the influence of a remaining slight non-uniformity in the display.

The viewing distance for the participants was four times the screen height, which resulted in a viewing distance of 1.2 m from the panel. The ambient illumination was 20 lx measured on the screen in the direction to the viewer, and about 18 lx measured on the background.

2.4 Protocol

Before the experiment, each subject was asked to carefully read the written instructions about the purpose of the experiment and the tasks to finish. Each experiment consisted of two tasks: (1) subjects were asked to judge the difference between the two images shown with a method depending on the experiment (see Sec. 2.1), and (2) subjects were requested to mark the area in the image that they used for their judgment. It should be noted that the latter results are not further discussed in this paper. They are not used to evaluate the effect of measurement methodology on the JND of the BL, but will be used later to model the content dependency of the JND of the BL.

The subjects were informed that the difference between the two images shown was a change in the BL, but they were not informed on what the reference image was. Before the actual start of the experiment, we conducted a training allowing the participants to get used to the experimental set-up as well as to the task. During this training, each subject was asked to assess the JND of the BL with one of the tasks for one image, which was selected randomly out of the nine images. When the subject was indicated to fully understand the task, the experiment was started. In each experiment, the order in which the images were shown was randomized over the subjects in order to avoid time effects. Since the second experiment was performed about one month after the first experiment, no influence of one experiment on the other was expected.

2.5 Participants

Twenty subjects with their ages ranging from 22 to 27 participated in the two experiments. This was substantially more than the minimum requirement of ten observers and three scenes as prescribed in the ISO 20462 standard.²³ Twelve males and eight females performed the first experiment, while 11 males and nine females were involved in the second experiment. Six of the females and five of the males performed both experiments. All volunteers had normal or corrected to normal vision (≥ 1.0 on a Landolt chart) and no color deficiencies (tested with the Ishihara color vision test).

2.6 Statistical analysis

As mentioned before, the JND values were calculated from the raw data by averaging over the last six reversal points.

These JND values were then imported in SPSS version 13 for further statistical analysis. First, the data were explored on their reliability computing various descriptive statistics and by means of graphs. The boxplot for the JNDs of the BL over subjects obtained in both experiments were plotted per image, and outliers were excluded. Then the boxplot was plotted again with the filtered data to exclude the remaining outliers. This procedure was repeated until there were no longer any outliers. Note that this is not a standard procedure; 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 1 for each experiment; it is still substantially more than the minimum requirement of ten observers as prescribed in the ISO 20462 standard.²³ It should be remarked that we found more outliers for the 2AFCSM than for the AMOA. A possible reason may be that for the 2AFCSM the last six reversals were more stable, resulting in a smaller within-subject standard deviation over reversals per image. As a consequence, deviations between subjects in the mean over the reversals were more easily significant. To get an impression of how this data exclusion procedure affected the JND values in both experiments, the boxplot for the JND of the BL before and after excluding outliers, is plotted in Figs. 4(a) and 4(b). It shows that the data exclusion procedure mainly shrinks the body of the

FIGURE 4 — Boxplot of the JNDs of the BL assessed with two methods. The blue boxes represent the results for the AMOA, while the green boxes represent the results by the 2AFCSM. (a) The data before exclusion; (b) the data after excluding the outliers.

FIGURE 5 — Mean and 95% confidence interval for JNDs of the BL assessed by two methods. The blue triangles correspond to the AMOA and the green circles to 2AFCSM.

boxplot by removing the less reliable results, but only slightly changes the median and hardly affects the relative difference in the JND between the two methodologies.

3 Experimental results

The JND of the BL is shown in Fig. 5 for the two methods with the mean and 95% confidence interval (CI) per image. The mean JND varies from 0.2 to 1.0 cd/m^2 over all images for both methods. The 95% confidence interval is around 0.5 cd/m^2 for most images in both experiments.

The mean and standard deviation of the JND of BL per image and methodology is also included in Table 1. Generally, the JND values assessed by the AMOA are larger than those assessed by the 2AFCSM. This can also clearly be observed from Fig. 6, in which the relationship between the JNDs evaluated by the two methods is plotted. Also, the standard deviation of the JNDs is, in most cases, comparable or larger for the AMOA than for the 2AFCSM.

Figure 7 gives the number of sample presentations required for assessing the JND per image with either method, and averaged over all subjects. The averaged

FIGURE 6 — Relationship between the JNDs of the BL evaluated with the AMOA and the 2AFCSM.

FIGURE 7 — Number of sample presentations averaged over the subjects, including the 95% confidence interval per image in both experiments. The bars without patterns are for the AMOA and the bars with patterns for 2AFCSM.

number of presentations needed is between 30 and 45. For most images, on average the 2AFCSM needs a comparable or larger number of sample presentations than the AMOA. The images “children” and “toy” form the exception.

To investigate the effect of methodology (two levels) and image content (nine levels) on the JND further, a full ANOVA is performed taking the JND of the BL as a dependent factor, methodology and image content as independent factors, and subjects as a random factor. The two-way interaction between methodology and image content is also included. The JND of BL is significantly affected by image content ($F = 7.21$, $df = 8$, $p < 0.001$) and methodology ($F = 1.72$, $df = 1$, $p < 0.001$). The interaction between image content and methodology is not significant ($F = 0.77$, $df = 8$, $p = 0.63$). The significant effect of methodology results from the fact that the JNDs assessed by the 2AFCSM (mean = 0.48 cd/m^2) on average are lower than those assessed by the AMOA (mean = 0.69 cd/m^2). The effect of image content is further analyzed with a Tukey post-hoc test, and reveals:

carriage<parrot<house<globe<toy<flower<balloon<children<rabbit

The images underlined together are not significantly different from each other.

FIGURE 8 — JNDs of the BL assessed with the 2AFCSM as a function of display luminance averaged over the image content. Each dot represents one of the original images.

The increase in JND of the BL with image content seems to be related to the averaged luminance in the original image. To evaluate this hypothesis, the averaged luminance in the image is plotted in Fig. 8 as a function of the JND of the BL, measured with the 2AFCSM. The JND of the BL highly linearly correlates with the luminance for all images except for the image “parrot.”

4 Conclusions and discussions

The performance of two methods, being the AMOA with a same-different task and the one-up-two-down staircase method with a 2AFC task, has been studied in this paper by assessing the JND of BL for a set of representative natural images.

Before discussing the difference in methodology, we evaluate the possible impact of having partly different participants in both experiments. Since only half of the participants (six female, five male) have performed both experiments, the remaining half of the participants is different for both experiments. To check whether this has an effect on our results, we repeat the ANOVA analysis with only those participants who have performed both experiments. Still, the effect of image content ($F = 3.82$, $df = 8$, $p < 0.001$) and methodology ($F = 7.89$, $df = 1$, $p = 0.006$) are both significant, and, so, these conclusions are not affected by the particular group of subjects participating in the experiments.

The 2AFCSM achieves a smaller standard deviation for the JNDs and is thus more accurate than the AMOA. This confirms previous findings for these methodologies.⁵ According to informal comments on the subjects, the former method is preferred over the latter. They mentioned to feel less certain about their response with the AMOA, which might have increased the variance. When a subject did not see a difference between the test and reference image, he was forced to make a choice with the 2AFCSM. It might well be that when being forced to choose, he consistently performed above chance in selecting the test stimulus. This would result in a smaller JND for the 2AFCSM than for the AMOA. This is consistent with our results and with findings reported in Refs. 5, 12, 16, and 17, but contradictory to findings reported in Refs. 14 and 15, in which the method of limits and the method of adjustment are not significantly different. Overall, we tend to conclude that the staircase method, *i.e.*, a variation of the method of limits, yields lower difference thresholds, at least in these studies that have sufficient power for showing statistical differences.

The 2AFCSM is slightly less efficient than the AMOA. For the 2AFCSM, two consecutive correct responses have to be given before the image pair is changed and this makes the experiment lengthier. With the AMOA, the subject has full control over the decrease or increase of the difference between an image pair. This assumes that the series of sample presentations is fully adapted to his personal perceptual abilities and hence efficient. This trend, however, is only confirmed clearly for four out of the nine images used in our

experiment. For the other images, the number of sample presentations is comparable or smaller for the 2AFCSM than for the AMOA. This observation seems to be in line with the remark made in Ref. 16, that both the method of adjustment and the staircase method are as effective in diagnosing glaucoma.

The JND of the BL for natural images, determined in this paper, mainly changes in the range from 0.2 to 1.0 cd/m^2 , while the display luminance averaged over the image content varies between 8 and 65 cd/m^2 . The JND corresponding to these luminance levels as given in the DICOM standard⁴ would range between 0.1 and 0.5 cd/m^2 . As expected, the JND of the 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.

The effect of image content on the JND of the BL is significant. In this first analysis, the JND of the BL is linearly correlated to the luminance averaged over the image content, which confirms our former preliminary results.³ The image “parrots” forms the exception. According to the subjects’ comments, this image is the most difficult, because it hardly has any detail in the dark area. These comments are consistent with the findings of Fig. 7, where the number of sample presentations required to assess the JND of the BL for the image “parrot” is comparatively large for both methods. Whether the linear relationship of the JND of the BL with the averaged luminance over the image content is sufficiently stable over a larger set of natural images still needs to be further evaluated.

Acknowledgment

This project is supported by Philips Research Laboratories in Eindhoven and the 111 Project (B07027) in China. The authors are grateful to Hui Tang, Yawen Chen for execution of the subjective experiments, and to all the volunteers for participating into the subjective experiments.

References

1. Heynderickx and E. Langendijk, “Image-quality comparison of PDPs, LCDs, CRTs, and LCoS projection displays,” *SID Symposium Digest* **36**, 1502–1505 (2005).
2. J. Xia *et al.*, “The just noticeable difference in chromaticity, black level, white level, and contour rendering in natural images,” *Proc. IDW/AD ’05*, 1821–1824 (2005).
3. S. Qin *et al.*, “Just noticeable difference of image attributes for natural images,” *SID Symposium Digest* **38**, 326–329 (2007).
4. Digital Imaging and Communications in Medicine (DICOM), Standard-PS 3.14-2003, Part 14: Grayscale Standard Display Function, (2003), pp. 21–27.
5. G. A. Gescheider, *Psychophysics: The Fundamentals*, 3rd edn. (Lawrence Erlbaum Associates, London, 1997).
6. N. A. Macmillan and C. D. Creelman, *Detection Theory: A User’s Guide* (Cambridge University Press, Cambridge, U.K., 1991).
7. A. B. Watson and A. Fitzhugh, “The method of constant stimuli is inefficient,” *Perception & Psychophysics* **47**(1), 87–91 (1990).
8. P. M. Wise *et al.*, “Two rapid odor threshold methods compared to a modified method of constant stimuli,” *Chemosensory Perception* **1**(1), 16–23 (2008).

- 9 P. Marvit and M. Florentine, "A comparison of psychophysical procedures for level-discrimination thresholds," *J. Acoust. Soc. Am.* **113**(6), 3348–3361 (2003).
- 10 S. Curran and J. P. Wattis, "Critical Flicker Fusion threshold: a useful research tool in patients with Alzheimer's disease," *Human Psychopharmacology* **13**(5), 337–355 (1998).
- 11 H. Levitt, "Adapted up-down methods in psychoacoustics," *J. Acoustical Soc. Amer.* **49**, 467–477 (1970).
- 12 B. L. Cardozo, "Adjusting the method of adjustment: SD vs. DL," *J. Acoustical Soc. Amer.* **37**(5), 786–792 (1965).
- 13 A. P. Ginsburg and M. W. Cannon, "Comparison of three methods for rapid determination of threshold contrast sensitivity," *Investigative Ophthalmology & Visual Science* **24**, 1035–1041 (1983).
- 14 E. W. Loftiss, "An evaluation of the effects of selected psychophysical methods and judgemental procedures upon comfortable loudness judgments of connected speech," Ph.D. Dissertation. University of Illinois (1964).
- 15 L. L. Kopra and D. Blosser, "Effects of method of measurement on most comfortable loudness level for speech," *J. Speech Hearing Res.* **11**, 497–508 (1968).
- 16 T. Maddess *et al.*, "Comparison of three tests using the frequency doubling illusion to diagnose glaucoma," *Clinical and Experimental Ophthalmology* **29**, 359–367 (2001).
- 17 V. Lakshminarayanan *et al.*, "Psychophysical estimation of speed discrimination I. Methodology," *J. Opt. Soc. Am. A* **22**(10), 2262–2267 (2005).
- 18 V. Klymenko *et al.*, "Visual psychophysics and medical imaging: non-parametric adaptive method for rapid threshold estimation in sensitivity experiments," *IEEE Trans. Medical Imaging* **9**(4), 353–365 (1990).
- 19 S. Amitay *et al.*, "A comparison of adaptive procedures for rapid and reliable threshold assessment and training in naïve listeners," *J. Acoustical Soc. Amer.* **119**(3), 1616–1625 (2006).
- 20 M. R. Linschoten *et al.*, "Fast and accurate measurement of taste and smell thresholds using a maximum-likelihood adaptive staircase procedure," *Perception & Psychophysics* **63**, 1330–1347 (2001).
- 21 B. Kollmeier, "Adaptive staircase techniques in psychoacoustics: A comparison of human data and a mathematical model," *J. Acoustical Soc. Amer.* **83**(5), 1852–1862 (1988).
- 22 ITU-R BT.709-5: *Parameter values for the HDTV standards for production and international programme exchange*.
- 23 B. W. Keelan and H. Urabe, ISO 20462, "A psychophysical image quality measurement standard," *Proc. SPIE-IS&T Electronic Imaging* **5294**, 181–189 (2004).

Shaoling Qin obtained her B.S. and M.S. degrees in electronic engineering in 2003 and 2005 from Southeast University, P.R. China. She was at the Delft Technical University, The Netherlands, with Prof. I. Heynderickx as a joint Ph.D. candidate from 2007 to 2008. She is now pursuing her Ph.D. degree at the Southeast University. Her current research area is human visual perception.

Shu Ge received her M.S. degree in physical electronics at Southeast University in April 2008. The same year, she joined GoerTek Technology Co. Ltd., in Qingdao, China, as a program manager assistant on Bluetooth headset R&D.

Hanchun Yin received her undergraduate degrees in physics from Fudan University, M.S. degree in engineering from Southeast University, and as a visiting scholar with Prof. T. Mulver at ASTON University in Birmingham, U.K. She now works at the School of Electronic Science and Engineering of Southeast University as a Professor and Advisor of Ph.D. Her specialty is physical electronics, computation of electromagnetic fields and visual perception study.

Lu Liu graduated from Automatic Control Department of Southeast University in 1985 and received her B.S. degree in computer science and technology at the same university in 2004. She is now working at the School of Electronic Science and Engineering of Southeast University. Her current research area is visual perception study.

Ingrid Heynderickx received her Ph.D. degree in physics from the University of Antwerp in December 1986. In 1987, she joined the Philips Research Laboratories in Eindhoven, and worked in different areas of research. Since 2006, she is a Research Fellow in the group Visual Experiences, where she is heading the project on Visual Perception of Display Systems. She is a fellow of the Society for Information Displays (SID), and for the SID, she was chair of the Applied Vision subcommittee. In 2005, she was appointed Visiting Professor at the Southeast University of Nanjing, and part-time Full Professor at Delft Technical University.