

Image Quality Loss and Compensation for Visually Impaired Observers

Sophie Triantaphyllidou*, Edward Fry*, Vicent Sanchis**, Álvaro Pons**;

**University of Westminster, Computational Vision and Imaging Technology, Faculty of Science and Technology, UK;

***University of Valencia, Dept. of Optics, Optometry and Vision Science, Faculty of Physics, Spain

Abstract

The measurement and modeling of image quality are aimed to assist the design and optimization of systems, typically built for 'normal' observer vision. But in reality image viewers rarely have perfect vision. There have been few attempts and no universal framework for measuring image quality loss due to visual impairments. The paper presents initial experiments designed to measure still image quality losses, as experienced by observers with visual accommodation problems, by proposing modifications to the Quality Ruler method described in ISO 20462-3:2012. A simple method is then presented, which compensates directly on the display for some of the quality lost due to the impairment. It uses a purpose-built image equalization software. The compensated image is finally examined in terms of quality gained. The losses and gains in image quality are measured on a Standard Quality Scale (SQS), where one unit corresponds to 1 JND. Initial results show that the quality lost due to visual accommodation impairments can be accurately measured with the modified ruler method. The loss is scene-dependent. Partial or full quality compensation can be achieved for such impairments, using image contrast equalization; the level of quality gained also scene-dependent.

Introduction

The measurement and modeling of image and video quality are very vibrant areas of research that assist the design and optimization of systems. New image quality models appear frequently in the literature, all striving to best predict the observer perception of image goodness. An increasing number of image databases with Mean Opinion Scores (MOS) are made available for benchmarking such models. In all the relevant psychophysical experiments it is ensured that visually fit observers participate since, after all, their purpose is to measure the system's ability to produce high quality imagery. In the visual modeling camp, great efforts are put to model appropriately the average, normal human observer for implementation in image quality modelling. So, it is fair to say we design systems for the 'perfect observer', who has typically correct (or corrected) spatial vision and no color deficiencies.

In reality viewers rarely have perfect vision when they encounter imagery on displays or printed media. As a result of significant progresses in imaging technologies, today systems - even relatively inexpensive ones - are capable of producing images of high objective quality, yet they are not necessarily perceived as such due to a variety of observer visual deficiencies.

In the Western world more than 40% of the adult population have visual deficiencies due to refractive problems (myopia, astig-

matism, hyperopia), whilst the majority of the people above 40 years old have some degree of impairment due accommodation problems (presbyopia) [1]. Refractive and accommodation issues affect mainly spatial image quality dimensions (i.e. luminance contrast, sharpness and acuity). More than 8% of the European Caucasian population have color deficiencies (with red-green being the most prevalent) [2], affecting color perception and discrimination. There are a number of deficiencies relating to old age (e.g. cataract, macular degeneration) or diseases (e.g. glaucoma, diabetes), resulting often to severely impaired vision and thus affecting all image quality dimensions to a different degree.

Given that we rely so much on our vision and the multiplicity of devices that produce imagery, it is surprising that there is relatively little research on quantifying the losses in image quality due to visual problems, or the gains resulting from various image compensating techniques. The overarching aims of our work are to examine whether we can quantify image quality losses experienced by visually impaired observers using standardized evaluation methods and to examine to what extent we may optimize displayed images to compensate for the losses.

The focus of this paper is on spatial, still image quality as experienced by observers with accommodation impairments (presbyopia). In the following sections we present and evaluate results from initial experiments, designed to measure the quality lost due to the impairment. For the purpose, we propose modifications to a standard method for image quality quantification (i.e. ISO 20462-3) [3]. Further, we present a simple method for compensating for some losses in quality directly on image displays using purpose-built image equalization software. The compensated images are then examined using the same quantification method, enabling to calculate the quality gained from image equalization. The losses and gains in quality are measured using a Standard Quality Scale (SQS), where 1 unit corresponds to 1 JND. Finally, we present conclusions and further work.

Background

The quantification of image quality lost due to various visual impairments, and the quality gained when displayed images are enhanced to directly compensate for such losses, have seen relatively little investigation. Some of the most important early [4-6] and on-going [7-9] research in these areas comes from E. Peli and colleagues, who have produced a large volume of work and many relevant publication (mainly on patients with macula degenerations and cataract). Peli's work assumes that the ratio of the impaired to normal contrast sensitivity function (CSF) characterizes the spatial frequency losses due to the impairment. It has revealed that, compensation for the impairment(s) is achieved

by enhancing band-limited luminance contrast to improve visibility of information (not quality, per se). High band-pass filtering is required to enhance spatial frequencies that are not detected by the visually impaired, whereas little low frequency compensation is needed due to contrast constancy effects. Individually tuned enhancement depends on the patient's CSF, the image's contrast spectrum and image size. In more recent work Peli and colleagues reviewed their approaches to image enhancement, and the difficulties encountered in the evaluation of the benefits of enhancement [8].

Examples of latest advances in correcting image displays for the visually impaired include Hwang and Peli's implementation of an augmented vision system on Google Glass to provide contrast-improved central vision [9] and computational light field displays that correct for refractive errors [10].

Measuring Quality Loss due to Impairment

The basis of the proposed method for quantifying the quality lost due to a visual impairment is detailed in the Part 3 of the ISO 20462 - Psychophysical experimental methods for estimating image quality [3]. The series of ISO 20462 papers address the need for determining image quality in a calibrated fashion. Part 3 of the series - referred to as the *quality ruler method* - is 'a psychophysical method that involves quality or attribute assessment of a test stimulus against a series of ordered, univariate reference stimuli that differ by known numbers of JNDs'. It is particularly suitable to measuring larger differences that exceed 1 JND. The paper describes the measurement of quality differences in JNDs from the ratings given by observers in real time. Results from the method are reported using the Standard Quality Scale (SQS), a numerical scale that is anchored against physical standards, has a zero point and one unit corresponding to 1 JND in quality.

The standard provides standard reference stimuli (SRS), but also describes how experimenters can generate their own quality ruler calibrated digital images. The method was implemented successfully in our labs on a soft display (soft copy ruler implementation), in a study that examined JPEG 2000 image quality [11], and using an image database developed in [12]. Observers were asked to match the compressed image quality to that of a set of reference stimuli varying in sharpness. Sharpness variations were introduced by modifying the modulation transfer function (MTF) of the complete imaging system used to generate and display the images, in a method compatible to that described in the ISO paper [11].

Methodology

The methodology we propose in this paper for measuring image quality loss in observers with accommodation impairment applies modifications to the standard soft copy ruler paradigm, as previously implemented in [11]. It uses two identically calibrated very high quality displays positioned side by side, with a 15 degree tilt toward each other, separated by a black screen. Display 1 displays the test image in its central area; display 2 displays the ruler image in the same area, corresponding to the slider bar position situated below the image area (Figure 1). The test image is of a pre-defined high objective quality of known SRS value. In our set up the test image quality was of value 24 on an SQS2 (calibrated based on average scene relationship [3, 11]), with SQS2 values ranging between 1 and 33, corresponding to the sharpest

and the blurriest images respectively. The test image is viewed first from the set viewing distance (for which the ruler images are calibrated) by the impaired observer without corrective prescription spectacles. The observer is asked to examine the test image in detail, rotate slowly his/her head only and look at the second display while putting their spectacles on. The ruler images are examined next by the observer, using the slider bar, who is now wearing spectacles prescribed for the given viewing distance. The original slider bar position - and respective image - is randomly selected, as described in the ISO paper. The observer is asked to balance the quality loss due to his/her impairment to the quality loss due to unsharpness in the ruler images - this is achieved by flicking through the images using the slider bar. The observer is permitted to repeat this process maximum three times; second and third viewings allow the observer to refine his/her previous choices, if necessary.

It is important to note that, there is a time interval in-between the viewing of the two stimuli, due to i) the time taken for the observer to rotate his/her their head and wear/remove their spectacles, and ii) the time taken to accommodate comfortably once the spectacles are worn/removed. This arrangement relies on short-term observer memory - the memory of the impressed quality of the images. One of the goals is to examine whether this arrangement can produce reliable results.

The quality difference between the test and the selected ruler image indicates the quality lost due to the impairment for the given viewing distance, measured in JNDs on the SQS2. Differenced in SQS2 are still separated by 1 JND in image quality, but are not anchored against a physical standard and have no zero point. They can be converted in SQS, if necessary, following the procedures indicated in the ISO paper [3].

Experimental Set-up

Two identical, very high quality wide gamut 24-inch EIZO LCDs with built in hardware calibration displayed the image stimuli; they were driven by a 1 GB NVIDIA Quadro 2000 graphics card, set to display 10 bits per pixel resolution. The displays were calibrated to sRGB colorimetry [13], but with a white point of 100 cd/m². Their colorimetry and transfer characteristics were examined after calibration using a Konica Minolta CS200 luminance and color meter. Colorimetric and luminance discrepancies between them were negligible and below the levels of perceptibility in the dark viewing environment where the tests were conducted.

Experiments were repeated for 8 images, selected from the database developed in [12]. They were selected to vary significantly in content (people, nature, objects) and had important variations in spatial frequency and color information. Image dimensions were 588 by 882 pixels, shown in landscape or portrait orientations (subtended angles of 23° horizontally - 16° vertically, or vice versa). They occupied approximately half of the display area, with the remaining area displaying a neutral gray of 50% luminance. They were presented in a random order in each experimental run.

Observers sat in front of the two displays, facing directly the black screen separating them. They had to rotate their heads by approximately 30 degrees to face the center of each display. The observer-to-centre-of-the-display viewing distance was set to 60 cm - the distance for which the ruler images were calibrated [11]. A chin rest was not used, instead observers were asked to main-

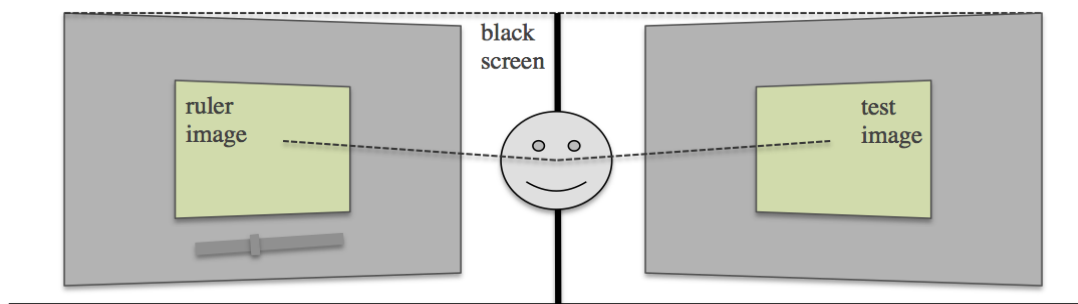


Figure 1. Experimental set-up.

tain their viewing distance by rotating only their head, whilst sitting with their upper back in contact with the rear of the seat. If they felt that their eyes were strained, they could rest their eyes upon a neutral gray screen. Each round of observations took approximately 15 minutes, yielding results for one observer, for all 8 images. Five rounds of experiments maximum were allowed per day, with breaks between them, to avoid observer fatigue. The experimental set up is illustrated in Figure 1.

Results

The results are preliminary, their purpose is to test the feasibility of the experimental paradigm. They were obtained by conducting tests with two observers with different degrees of accommodation impairment: a 52 years old female (observer #1), an expert observer, with previous experience in image quality experiments, including the soft copy ruler; and a 48 years old male (observer #2), a non-expert observer, with no previous experience using the ruler. Information on their eye prescriptions are presented in Table 1.

Table 1: Eye prescriptions of observers.

Observer	Eye	SPH	CYL	Axis	ADD
#1	Left	+0.75	-0.25	155	+2.50
#1	Right	+0.75	-0.25	10	+2.50
#2	Left	-1.00	-0.25	5	+1.50
#2	Right	-1.75	-0.25	165	+1.50

Both observers carried out the experiment with spectacles that corrected their presbyopia only, which were optimized for the display viewing distance. Observer #1 has mild hyperopia (i.e. + SPH), which adds on to the effect of presbyopia (i.e. + ADD), making her near vision impairment effectively equal to 3.25 diopters. Observer #2 has mild myopia (i.e. - SPH) and mild presbyopia and completed the experiment wearing far distance vision contact lenses that corrected fully for his myopia. It is common for people with myopia, when they are wearing contact lenses to correct for it, to wear also spectacles for correcting presbyopia when they operate their near vision (e.g. for reading, or computer work). Both observers have a negligible degree of astigmatism (i.e. - CYL), which was not corrected during the experiments.

Figure 2 illustrates mean JND differences between the test image quality and the quality of the selected ruler image, for all image stimuli for observer #1. JND differences indicate the loss in image quality due to (mainly) the degree of presbyopia; they were obtained by averaging results from 10 experimental rounds. Error bars indicate ± 0.5 standard deviation from the mean. The results demonstrate a small degree of scene dependency, with average JND differences for each stimulus ranging from -16.2 to -18.2. Also, the magnitude of the error bars varies slightly, implying that the relative quality of some stimuli was possibly relatively easier to quantify than of others. But errors are overall very small. This suggests that the experimental task should be straightforward for observers experienced with the quality ruler (i.e. those who are more familiar with the concept of equating image quality of a test stimulus to a blurred version of the same stimulus).

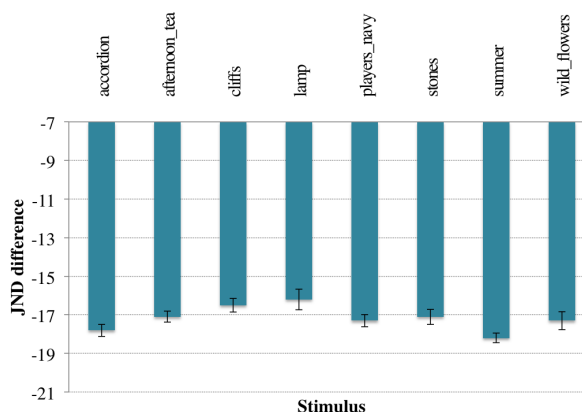


Figure 2. JND differences between test and ruler image for observer #1.

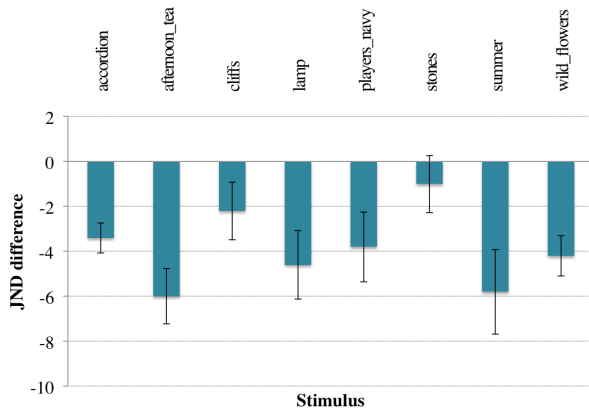
The mean and median JND difference from all images, along with the mean standard deviation after 3, 6 and 10 experimental rounds for observer #1 are listed in Table 2. The results indicate again great stability for this observer.

Figure 3 illustrates mean JND differences between the test and ruler image quality for all image stimuli for observer #2. They were obtained after 6 experimental rounds. Table 3 presents statistics from this observer. It is evident that, the quality lost due to presbyopia for observer #2 is considerably smaller than for observer #1. The degree of visual impairment for observer #2 is less

Table 2: Statistics from 8 stimuli - observer #1.

Statistic	10 rounds	6 rounds	3 rounds
Mean JND difference	-17.2	-17.3	-17.4
Median JND difference	-17.0	-17.0	-17.0
Standard deviation	0.93	0.90	0.92

than half of that of observer #1 (in total number of diopters), but the mean quality loss is more than four times smaller. Error bars in Figure 3 and mean standard deviation from observer #2 experiments show a larger variability when judging image quality, which may be due the observer's lack of experience with the ruler experiment. Further, the measurement variability, along with the greater scene dependency seen in the results of observer #2 may indicate that, for a very mild amount of impairment it is more difficult for an observer to consistently identify the quality loss that arises from it; and that the scene content of the stimulus plays a large role when evaluating the perceived loss of quality.

**Figure 3.** JND differences between test and ruler image for observer #2.**Table 3: Statistics from 8 stimuli - observer #2.**

Statistic	6 rounds
Mean JND difference	-3.9
Median JND difference	-4.0
Standard deviation	2.49

Quality Compensation and Gain

Spatial band contrast equalization was carried out to compensate for (some of) image quality lost due to the observers' accommodation impairment. The basis of the enhancement method is found in references [5, 14]. A purpose-built image equalization Graphics User Interface (GUI) was developed for optimizing the visual quality of the test image (i.e. the image that the observer views without spectacles in the modified soft copy ruler implementation). The equalization GUI operates as follows:

Once a still image is loaded, spatial frequency decomposition is carried out to separate the image's frequency content into eight spatial frequency bands. This was achieved using Peli's log

cosine filters [15], for the range of visual octaves of 0.125, 0.25, 0.5, 1, 2, 4, 8 and 16 cy/deg with respect to the observer's viewing distance (60 cm). After decomposition, high and low frequency residuals are added to the highest and the lowest frequency bands respectively. Individual band-limited image adjustment (equalization) is achieved by multiplying the individual band RMS contrast of the luminance component of the YCbCr image by a chosen weighing factor. Conversion from sRGB (images' color space) to YCbCr is achieved in MATLAB using the build-in routine. Eight sliders, with values ranging from 1.0 (no adjustment) to 33.0, allow the desired weighting factor to be selected for each band.

A final slider adjusts the global chroma of the uploaded image. Chroma adjustments are carried out in the CIELAB space, where the $|a^*|$ and $|b^*|$ image channels are multiplied by a chosen weighing factor between 1.0 (no adjustment) and 2.0 (double the original chroma). The inbuilt MATLAB routine was used for conversion from sRGB to CIELAB. Pilot tests with the equalizer showed that minor chroma adjustments improve the contrast enhanced image, which otherwise loses some of its original perceived colourfulness due to band contrast enhancement. Further, the enhanced color contrast also means enhanced visibility for the impaired observer.

The user interface of the equalizer includes: a window for displaying the image/stimulus; a slider per spatial frequency band for choosing weighting factor for band contrast modifications; a slider for choosing the weighting factor for chroma modifications; and a 'show image' button that displays the manipulated image as per the current settings of the sliders. Other buttons show the previous image (i.e. the version prior to the currently displayed manipulation), and if needed, the two previous versions too. The user/experimenter can save the equalized image and the chosen slider's settings.

Methodology

The equalization GUI was used to display each stimulus individually to the observer, who sat at the set viewing distance from the display without spectacles. The image loaded in the equalizer was decomposed into individual band-limited images and reconstructed by adding them all back. It was displayed in the first instance without contrast or chroma modifications. Then, for each frequency band a simplified staircase method was implemented to adjust the RMS contrast of the band-limited image (i.e by multiplying it by a weighting factor). The weighting factor was first set to maximum (x33.0). The weighting factor (step) was then reduced to half of the difference between no contrast adjustment (x1.0) and the previous adjustment, and the observer was asked whether the adjustment was still too strong, or should be further decreased. Each time, the observer compared the previous adjustment to the current adjustment by swapping between the two images on the GUI, until he/she decided that the band contrast would be increased or decreased. The decision for ending the adjustments was based on 3 consequent down-up (i.e decrease-increase) observer responses.

This methodology was applied to adjust the contrast of all frequency bands, starting with the highest band and ending with the lowest band adjusted. Note that, pilot studies with the equalizer showed that the chosen settings of the sliders of the equalizer were similar (but not identical) if these bands were adjusted in the reverse order. The final stage in the stimulus enhancement

involved global chroma adjustments, where the same staircase method was used to choose an adequate chroma weighting factor.

It is important to bear in mind that the aim of this experiment was not to identify the best possible adjustments for compensating for the observer impairment directly on the image display, but to test the adequacy of the equalizer for the purpose and produce an enhanced image with clearly better visual quality than the original, when viewed without spectacles. This aim was achieved; during a last comparison, both observers confirmed that all equalized (compensated) images were of superior image quality than their respective uncompensated test images. Note that, neither observer had previous experience in image contrast equalization.

Results

Figures 4 and 5 present the amount of contrast enhancement per spatial frequency band that observers #1 and #2 respectively chose during the process of equalization. These are plots indicating the RMS contrast weighting factor chosen for each band-limited image, and for eight out of the fifteen original stimuli that we employed in this experiment. We observe that, enhancement trends are similar for both observers, but that observer 1 required more contrast boosting, as expected, seeking especially to compensate for the high frequency information lost due to their visual deficiency. The results exhibit some level of scene dependency, which is again observer dependent, with observer #1 showing a more consistent equalization profile, which may be due to both previous experience in judging image quality and the higher level of impairment.

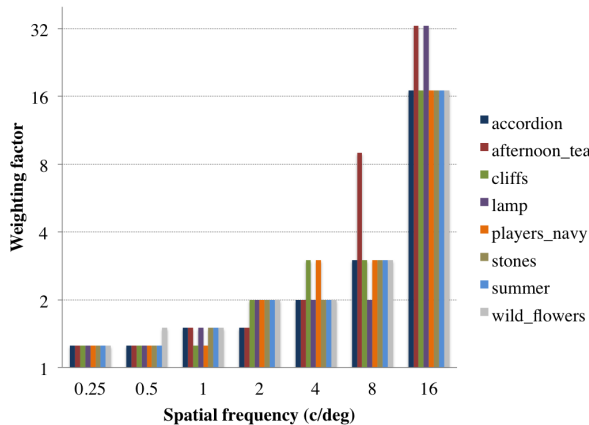


Figure 4. Image enhancement: band RMS contrast weighting factors chosen by observer #1, for eight stimuli.

Quality Gained by Compensation

The equalized images were finally used as test stimuli in a second round of experiments with the modified soft copy ruler set-up. Experiments were undertaken by both observers, under the same experimental conditions as those mentioned earlier; only the test image in this round was replaced by the previously compensated stimulus, as equalized by each observer. This way it was possible to estimate the quality gained by the equalization. It is readily obtained from the difference between the ruler image

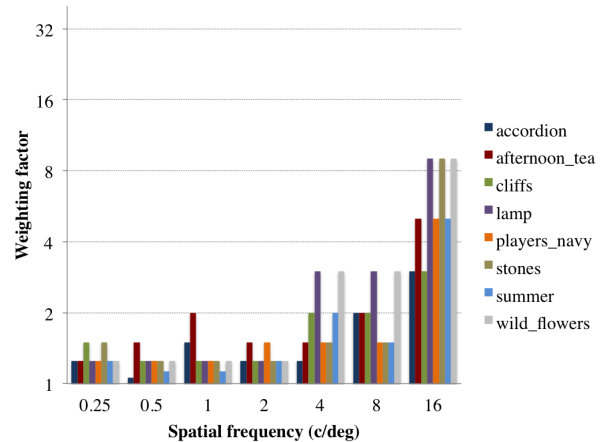


Figure 5. Image enhancement: band RMS contrast weighting factors chosen by observer #2, for eight stimuli.

quality that matched the quality of the test image in the first round (the image prior to equalization) and the ruler image quality that matched the quality of the equalized image. This final experiment demonstrates that, the quality of an image that has been modified to compensate for a certain visual impairment can be compared with the uncompensated image quality using the modified soft copy ruler paradigm.

Figures 6 and 7 show JND differences between test and ruler images for observers #1 and #2 respectively, before equalization and after equalization. The mean quality gained from all stimuli from the process of equalization was 4.5 and 3.6 JNDs for observers #1 and #2 respectively. Image equalization partially compensated for the quality loss due to the impairment of observer #1 (with relatively higher level of impairment). For observer #2 (with milder level of visual impairment) the quality of some equalized stimuli was judged as nearly, or fully compensated for, whilst overall, the results show an important degree of scene dependency. Furthermore, for at least two stimuli, equalization resulted in superior quality with respect to the original image viewed with spectacles. The outcomes suggest that, for observers with minor visual impairments, image equalization may be used, not only to fully compensate for the loss of image quality due to the impairment, but also produce a user-optimized image quality. The power of image equalization would be, of course, dependent on the objective performance limits of the display device.

Conclusions and Further Work

Results from initial validation of the modified image quality ruler implementation indicate that the proposed method can be used to accurately quantify the image quality lost due to visual accommodation deficiencies. The degree of intra-observer variability was very low for the expert observer, who had a higher level of impairment. Measurement variability was larger for the non-expert observer, who's impairment was very mild in comparison; results were also more scene-dependent. This suggests that observers with milder deficiencies may find it more difficult to consistently evaluate the quality lost due to their impairment, and that scene content of the stimuli may play a larger role in such evaluations.

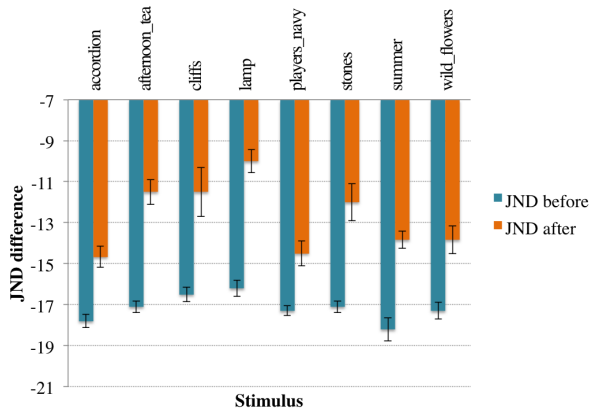


Figure 6. JND differences between test and ruler image for observer #1, before and after equalization.

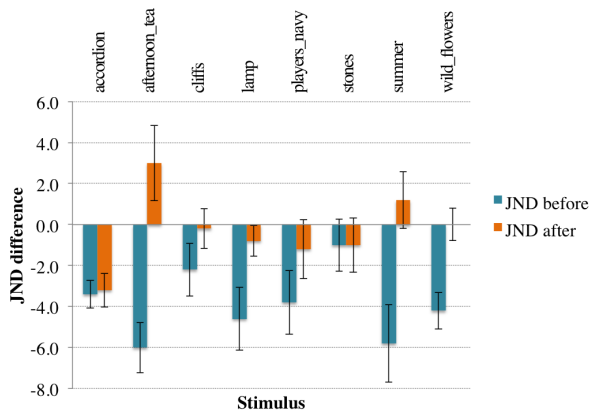


Figure 7. JND differences between test and ruler image for observer #2, before and after equalization.

Enhancement of image band-limited contrast using the equalization GUI, following a simplified staircase paradigm, provided a partial compensation of image quality for the observer with the higher level of impairment. The same method provided almost full, or full compensation for the observer with milder level of impairment. The results exhibit some degree of scene-dependency; they suggest once again, that for observers with a very mild level of accommodation impairment, image content is important when judging the level of image quality compensation.

In its current form, the proposed measuring technique relies on the use of spectacles that correct for the observer's impairment. We are currently repeating observer measurements across a larger image set to further evaluate measurement consistency and examine scene-dependency. We are also exploring the applicability of the methods presented for measuring losses in image quality (and gains from equalization) for different visual impairments, and degree of impairment. The measurement of image quality lost due to visual impairments, and image enhancement methods designed to compensate for such losses, are essential research areas with outcomes applicable to all display industries.

Acknowledgments

The authors would like to thank colleagues from the University of Westminster, Dr Elizabeth Allen for making available the original soft copy ruler implementation and calibrated image stimuli, and Jan Smejkal for technical help.

References

- [1] D. Pascolini and S. P. Mariotti, Global estimates of visual impairment: 2010, Br. J. Ophthalmol., 96, pg. 614. (2012)
- [2] J. Birch, Worldwide prevalence of red-green color deficiency, JOSA A, 29, pg. 313 (2012)
- [3] ISO 20462-3:2012, Photography - Psychophysical experimental methods for estimating image quality - Part 3: Quality ruler method (2012)
- [4] E. Peli and T. Peli, Image enhancement for the visually impaired, Opt. Eng., 23, pg.47 (1984)
- [5] E. Peli, R. B. Goldstein, G M Young, C. L. Trempe and S. M. Buzney, Image enhancement for the visually impaired. Simulations and experimental results., Invest. Ophthalmol. Vis. Sci., 32, pg. 2337 (1991)
- [6] E. Peli, E. Lee, C. L. Trempe and S. Buzney, Image enhancement for the visually impaired: the effects of enhancement on face recognition, JOSA A, 11, pg. 1926 1994
- [7] E. Peli, Recognition performance and perceived quality of video enhanced for the visually impaired, Ophthalmic Physiol Opt., 25, pg. 543 (2005)
- [8] E. Peli and R. L. Woods, Image enhancement for impaired vision: the challenge of evaluation, Int. J. Artif. Intell. Tools, 18, pg. 415 (2009)
- [9] A. D. Hwang and E. Peli, An augmented-reality edge enhancement application for Google Glass, Optom Vis Sci., 91, pg. 1021 (2014)
- [10] F-C Huang, G. Wetzstein, B. Barsky and R. Raskar, Eyeglasses-free display: Towards correcting visual aberrations with computational light field displays, Proc. SIGGRAPH, ACM Transactions on Graphics, 33 (2014)
- [11] E. Allen, Image quality evaluation in lossy compressed images, PhD Thesis, University of Westminster, UK (2016)
- [12] E. Allen, S. Triantaphillidou and R. E. Jacobson, Perceptibility and acceptability of JPEG2000 compressed images of various scene types, Proc. SPIE 9016, pg. 90160W (2014)
- [13] IEC 61966-2-1:1999, Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space - sRGB (1999)
- [14] E. Fry, S. Triantaphillidou, J. Jarvis and G. Gupta, Image quality optimization, via application of contextual contrast sensitivity and discrimination functions, Proc. SPIE. 9396, pg. 93960K. (2015)
- [15] E. Peli, Contrast in complex images, JOSA A, 7, pg. 2032 (1990)

Author Biography

Dr Triantaphillidou is an Associate Professor in Imaging Science at the University of Westminster, UK and the Director of the Computational Vision and Imaging Technology research group in the Department of Computer Science. She has degrees in Computer Science and Imaging Science. Her research is interdisciplinary, exploring interrelationships between imaging system performance, image content and image perception. She has published more than 50 peer-reviewed papers, and numerous research and textbook chapters relating to digital imaging.