

Can Eye Movements Be Quantitatively Applied to Image Quality Studies?

Tero Vuori, Maria Olkkonen, Monika Pölönen, Ari Siren, Jukka Häkkinen

Nokia Research Center

P.O. Box 407, FIN-00045 Nokia Group, Helsinki, Finland

{tero.vuori/ ext-maria.olkkonen/ monika.polonen/ ari.siren/ jukka.hakkinen}@nokia.com

ABSTRACT

The aim of the study is to find out whether subjective image quality evaluations can be quantified by eye movement tracking. We want to map objective or physically measurable image quality to subjective evaluations and eye movement data. Results show that eye movement parameters consistently change according to the instructions given to the user, and according to physical image quality. These results indicate that eye movement tracking could be used to differentiate image quality evaluation strategies that the users have. Results also show that eye movements would help mapping between technological and subjective image quality. We also propose to extend the widely used image quality process model, the Image Quality Circle. We suggest adding the objective measurements of a viewer (e.g. eye tracking) in parallel with customer perceptions as an option to gather information of customer perceptions of image quality.

Author Keywords

Image quality, psychometric scaling, eye movements.

ACM Classification Keywords

H.5.1 Multimedia Information Systems: Evaluation/ methodology.

INTRODUCTION

Image quality is a diverse research area that combines various research disciplines and methodologies. Because the number and variety of imaging systems has recently increased quickly, there is a need to critically evaluate the applicability of the current visual quality evaluation methods. Especially new technologies and new use contexts may change the way the users evaluate the quality of an image. For example, viewing a family photograph on a mobile phone display while walking outside is fundamentally different (emotions, needs, expectations) from viewing a DVD film on a home theater.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

NordiCHI '04, October 23-27, 2004 Tampere, Finland
Copyright 2004 ACM 1-58113-857-1/04/10... \$5.00

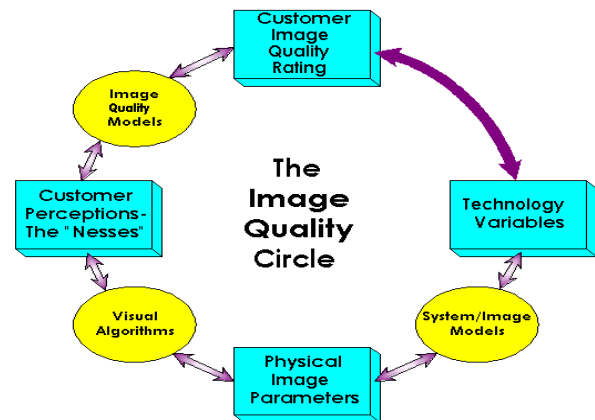


Figure 1. The Image Quality Circle.

Our approach can be conceptualized within the Image Quality Circle by Engeldrum [3]. The circle presents image quality research process from technology variables via physical image parameters and customer perceptions to customer quality preference (see Figure 1). The most important issue is the user, i.e. customer quality preferences. Knowledge of these preferences is usually derived from psychometric user studies, which have well defined methodology that quantifies the mapping from physical parameters to subjective quality impressions. We propose that this relation is not so simple, and the aim of this study is to quantify the physically measurable differences in image quality evaluation process and try to find a way to derive preferences from this data.

Human Visual System and Imaging

One way to conceptualize human visual perception is by top-down and bottom-up processes. On the one hand, there are low-level saliency mapping processes guided by the low-level features of the viewed stimuli (e.g. color, spatial frequency components). On the other hand, perception is also affected by high level cognitive factors (e.g. semantic features, motivational and emotional factors). Saliency mapping is a good determining variable with meaningless images, whereas cognitive factors are more influential with meaningful and emotional images [8]. The different levels of perceptual processes probably affect image quality perception. It is also known that people rate different image categories differently and have different quality

expectations for different image types. For example, images presenting a human face can have more blur than the images of a scene containing small details [5]. Therefore, a typical image quality test must contain various image contents. However, it would be interesting to find out what is the reason for differences between image contents and differences between observers. One possible way of measuring this is to simply ask the users to tell the reason for their choices. However, verbalizing the reasons for image quality preferences may be difficult and can affect the results especially with difficult tasks [10], so a more objective method should be used. Eye movement tracking is a good candidate for such method, because eye-movements are influenced by the characteristics of the visual scene and the nature of the visual task [1].

Eye Tracking and Image Quality Research

Eye movement tracking is applied in various research areas such as neuroscience, psychology, human factors, marketing, and computer science [4]. In the image quality research area, eye-tracking systems have been used e.g. when collecting information of perceptual processes while watching print advertisements to get information of how the effective advert should look like [9]. Eye movements differ when watching meaningful and meaningless images [8]. Viewer's purposes of looking at an image have an effect on eye movements. For example, when trying to evaluate the aesthetical values of an image versus finding out what is happening in an image, eye movements differ [6]. It is also known that the evaluation task (e.g. rank order, or paired comparison) has an effect on eye movements [2]. Näsänen et al. [7] have also shown that saccade amplitude increases with increasing contrast. However, mapping technical image quality and customer perceptions have not been widely studied with eye movement tracking.

In this paper, our purpose is to find out whether eye movement information can be applied to predicting subjective image quality preferences.

METHODS

Stimuli

Four representative image contents were carefully selected (see Figure 2). Each image was used as an original and as contrast and blur manipulated versions. Adobe Photoshop's blur and contrast algorithms were used to achieve five levels of blur and five levels of contrast (see Figure 3). Modular transfer functions (MTF) were calculated for the simulated images (see Figure 4) as follows:

$$MTF = M_{out} / M_{in}$$

where M_{out} is the modulation of a simulated image, and M_{in} is modulation of the original image.

There were a total of 44 images (4 contents * 2 manipulations * 5 levels + 4 originals). Image resolution was 611x469 pixels for landscape images and 606x504 for

portrait images. Images subtended 17 x 14 (portrait images) and 13 x 17 (landscape images) degrees of visual angle at a viewing distance of 83 cm.



Figure 2. Test images. Cafeteria (up-left), Girl (up-right), Fruits (down-left), and Group (down-right).



Figure 3. An example of a quality manipulated image. Group with added blur.

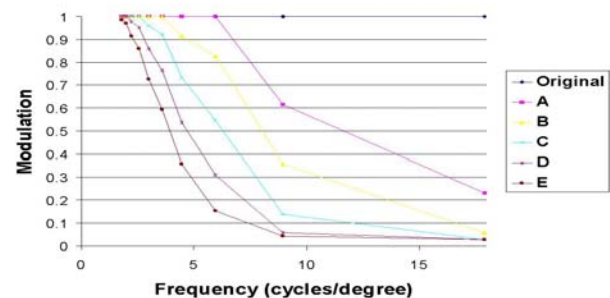


Figure 4. MTF functions used in simulated blur images.

Test Subjects

Eight subjects (six males, two females) with normal or corrected to normal vision participated in the test. Every subject participated in all test conditions.

Apparatus

The stimuli were shown on a 21-inch CRT monitor with Neurobehavioral Systems' Presentation 0.50 software. Eye

movements were recorded with the Eyelink II head-mounted gaze tracking system. The participants' eyes were monitored using miniature infrared cameras with infrared LEDs illuminating the eyes. A Dell Latitude C640 computer functioned as the host computer through which the image analysis, eye event parsing, and data recording functions were performed. Eye movements were recorded in the pupil only mode at a sampling rate of 500 Hz with an average gaze position error of smaller than 0.5° . The eye-movement system was calibrated and validated with 9-point biquadratic calibration software. Only the pupil of one eye was recorded (left with seven, right with one subject).

Testing Procedure

Head was stabilized with a chin rest to avoid disturbing head movements. The stimuli were randomized in four sequences, which contained 11 stimuli each. Order of stimuli within a sequence and the order in which sequences were presented were counterbalanced between participants. The instruction was given on the screen before a sequence. A fixation cross was displayed in the center of the screen for 2 seconds between every stimulus. A stimulus was shown for 5 seconds, after which the subjective evaluation was given with the keyboard. The program waited for a response to continue the sequence.

Each sequence was shown twice with two different instructions. The participant was either instructed to evaluate the general quality of the image, or to evaluate the quality of the colors of the image, both in a 7-step scale (very bad - very good). Each of the original images was also shown with a third content specific instruction (see below). In sum there were three instructions:

- 1) Evaluate the overall image quality.
- 2) Evaluate the color quality
- 3) Task:

- How many buildings are there in the image? (Cafeteria)
- Evaluate the emotional state of the person. (Girl)
- How many apples are there in the basket? (Fruits)
- How many thumbs can you find? (Group)

Eye movements were recorded while the image was on the screen.

Data Analysis

Paired samples t-test and repeated measures analysis of variance (ANOVA) were used for statistical analyses.

RESULTS

We chose the saccade duration as the test parameter, because it represents eye movements well and corresponds to all other saccade and fixation parameters. Differences between instructions were significant (Figure 5). Instruction 3 (task) caused highly significantly shorter saccade duration than instructions 1 (overall image quality) ($t(7)=8.87$,

$p<0.001$) and 2 (color quality) ($t(7)=8.89$, $p<0.001$). Equally importantly, the difference between instructions 1 and 2 was significant ($t(7)=3.29$, $p<0.05$).

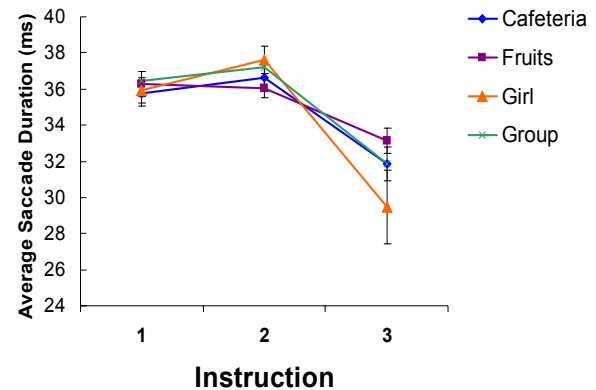


Figure 5. Saccade duration as a function of instruction for all image contents. Instructions: 1="Evaluate image quality." 2="Evaluate color quality." 3=task. Error bars are standard errors of the mean.

Also, image quality had highly significant effects on eye movements (Figure 6). Images with added blur caused longer saccade duration ($F(5,2)=7.95$, $p<0.001$). Also, contrast had a significant effect on saccade duration ($F(5,2)=2.39$, $p<0.05$). Saccade duration was longer when there was low contrast or too much contrast added.

Image quality affected subjective evaluations significantly. Subjective quality deteriorated with increased blur ($F(5,2)=13.23$, $p<0.001$). The effect was slightly stronger when evaluating general image quality than when evaluating color quality. Contrast had a similar effect on subjective image quality ($F(5,2)=7.55$, $p<0.001$); evaluations were best when a little contrast had been added, and poorest when there was too little or too much contrast added. There was no clear difference between evaluations of image quality and color quality.

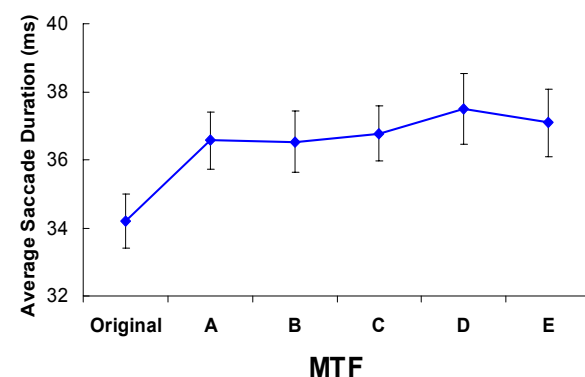


Figure 6. The effect of blur on saccade duration. X-axis is the MTF scale and blur is increasing to the right. Error bars are the standard errors of mean.

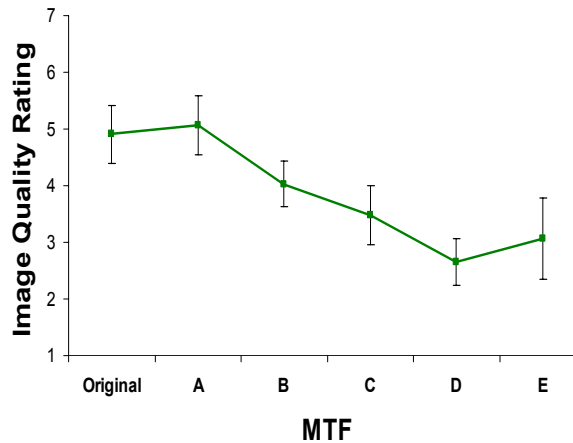


Figure 7. Subjective evaluations of image quality with levels of blur. Error bars are the 95% confidence intervals.

CONCLUSION

Image quality and instruction have a significant effect on eye movements and subjective evaluations of image and color quality. In our image quality study, saccade duration increased with decreasing subjective and objective image quality.

We conclude that eye movement measurements provide an interesting addition to image quality research. Subjective evaluations can be contaminated by emotional and motivational factors, such as social desirability. Eye movements are largely controlled by unconscious processes, so they can give more direct information of the perceptual and cognitive processes lying behind image quality evaluations.

In our image quality study, saccade duration increased with increasing blur. This means that saccade durations were longer with poorer subjective and objective image quality. Based on our results, it can be hypothesized that shorter saccade duration reflects perception of better subjective image quality. This hypothesis should be studied further and could be applied in future image quality studies.

We also have a suggestion for developing the Image Quality Circle. Knowing that effective and consistent results could be acquired by using objective measurements of the user, these methods should be added to the model as a separate item. At the moment they are included in the customer perceptions ("the Nesses"), which also include subjective image quality methods (questionnaires etc.). Objective measurements of a viewer could be added in the model in parallel with Customer Perceptions. It is obvious that physical image parameters have an effect on viewer measurements and now it has been shown that viewer measurements correlate with customer perceptions and image quality ratings.

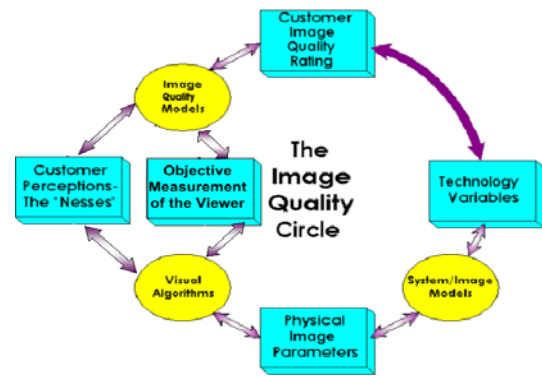


Figure 8. Our suggestion to add objective measurements to the Image Quality Circle.

REFERENCES

1. Andrews, T.J. & Coppola, D.M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research* 39, 2947 – 2953.
2. Babcock, J.S., Pelz, J.B. & Fairchild, M.D. (2003). Eye Tracking Observers During Rank Order, Paired Comparison, and Graphical Rating Tasks. *PICS 2003*, Rochester, NY, May 13, 2003 , Vol. 6.
3. Engeldrum, Peter G. (2000). *Psychometric Scaling: A Toolkit for Imaging Systems Development*. Winchester, USA: Imcotek Press.
4. Duchowski, A. (2003). *Eye Tracking Methodology: Theory and Practice*. London: Springer.
5. Freiser, H. & Biederman (1963). Experiments on image quality in relation to the modulation transfer function and graininess of photographs. *Photographic Science and Engineering*, 7, 28-34.
6. Molnar, F. (1981). About the role of visual exploration in aesthetics. *Advances in Intrinsic Motivation and Aesthetics*. New York: Plenum Press.
7. Näsänen, R., Ojanpää, H. & Kojo, I. (2001). Effect of stimulus contrast on performance and eye movements in visual search. *Vision Research* 41, 1817-1824.
8. Parkhurst, D. Law, K. & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42, 107-123.
9. Radach, R., Lemmer, S., Vorstius, C., Heller, D. & Radach, K. (2003). Eye Movements in the Processing of Print Advertisements. *The Minds Eye: Cognitive and Applied Aspects of Eye Movement Research*. Amsterdam: Elsevier Science BV.
10. Wilson, T.D., Schooler, J.W. (1991) Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology* 60, 181-192.