The effect of image sharpness on quantitative eye movement data and on image quality evaluation while viewing natural images

Tero Vuori, and Maria Olkkonen
Nokia
P.O.Box 407, 00045 Nokia Group, Finland
Tero.Vuori@Nokia.Com

ABSTRACT

The aim of the study is to test both customer image quality rating (subjective image quality) and physical measurement of user behavior (eye movements tracking) to find customer satisfaction differences in imaging technologies. Methodological aim is to find out whether eye movements could be quantitatively used in image quality preference studies. In general, we want to map objective or physically measurable image quality to subjective evaluations and eye movement data. We conducted a series of image quality tests, in which the test subjects evaluated image quality while we recorded their eye movements. Results show that eye movement parameters consistently change according to the instructions given to the user, and according to physical image quality, e.g. saccade duration increased with increasing blur. 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. Furthermore, these results give some empirical emphasis to top-down perception processes in image quality perception and evaluation by showing differences between perceptual processes in situations when cognitive task varies.

Keywords: Subjective image quality, eye movements, eye movement tracking, image sharpness, psychometric scaling, salience mapping, top-down theory of perception, image quality evaluation

1. INTRODUCTION

The usual aim of imaging industry is to provide a customer with satisfying image quality by developing new technologies to make even better image quality possible. Candidates for these new technologies are measured various ways. Because of this, image quality measurement has an important role in research and development done in industry, e.g. camera, display, and paper technology areas. Probably, because of the important role in these businesses, image quality has various approaches. These approaches used in image quality industry are conceptualized in image quality circle model [3] that links technological parameters to customer image quality preference with several links, e.g. physical measurements. Examples of widely used approaches to image quality are technical (resolution, MTF), or customer perception, or customer preferences. For these image quality approaches, there are various methodologies to measure the specific image quality conception. Image quality can be measured technically by measuring MTF or other physical variables, and these technical measurements can already give some clue of amount of added value to the customer. However, in business where customer is always right and where the end user usually is a human being, the user related approaches and methodologies give valuable and more direct information of customer preference. Knowledge of these preferences is usually derived from psychometric user studies, which have well defined methodology that quantifies the mapping from technological and physical parameters to subjective quality impressions. However, the relation between technical or physical parameters and customer perceptions and preference is not so simple. Moreover, preference measurements, yet the best way to find out reliable user preference, have their own known methodological weaknesses, which though possible to minimize but still leave room for searching other alternatives, advanced, or additive possibilities to measure what kind of added value is possible to give to the user in image quality point of view. We would like to try one approach more to image quality processes, i.e. physical measurements of a user, i.e. eye movement tracking. Eye movements are not yet fully applied to the study area of image quality evaluation process. The aim of this study is thus to quantify the physically measurable differences in image quality evaluation process and try to find a way to derive preferences from this data. Should we apply eye movement tracking more in image quality testing studies to get information on image quality preferences, or is it just a method to use in theoretical studies when forming new theories on image quality perception processes?

1.1. Human Visual System and Imaging

There is a long and strong research tradition in the human visual system (HVS) and perception process in the ground of image quality research. We know many characteristics of human vision, thresholds for various visual capabilities, sensitivity function for various features of visible targets. Also, a tradition of subjective scaling of image quality has a long history, and strong methodology. These methods are widely used in image quality industry, but they come from scientific studies and understanding of human perception process. That is why even very fundamental characteristics of human visual perception process have very central role in image quality measurement.

One fundamental way to classify human visual perception is to divide it into the top-down and bottom-up processes. On the one hand, there are theories that emphasize low-level salience mapping processes guided by the low-level features of the viewed stimuli (e.g. color, spatial frequency components). These theories suggest that essential thing in an image is what kind of small details there are in an image. 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]. Meanings, emotions, or semantic factors are not that essential in salience mapping theories. The different levels of perceptual processes probably affect image quality perception. In this study, we want to use eye movement methodology to give some answers to this problematic question of perceptual processes. Understanding the fundamentals of perceptual processes while evaluating image quality give important feedback to experimental setup, test images, and evaluation task.

It is 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. This reason may be based on low level feature differences in images, or it may be some high level cognitive or semantic process in image quality evaluation process while viewing different types of images. 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 [12], 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]. Therefore, we apply eye movements to image quality research.

1.2. Eye Tracking and Image Quality Research

Eye movement measurement gives a good indication of the focus of attention, since accurate viewing is possible only in the central fovea area only where the density of photoreceptive cells is highly concentrated [4]. By measuring saccades (rapid eye movements) and fixations (the time when the eye is fairly motionless), we can get detailed information of attention paid to the image.

Some eye movement concepts are presented in Table 1. Eye-trackers can record them all but in the analysis phase it is common to concentrate on some, because they are highly correlated. In this case, we concentrated in saccade duration.

Table 1. Definitions of basic eye movement concepts.

Fixation	Relatively stable stage when information is obtained for about 200 – 300 ms
Saccade	Rapid (about 30 – 50 ms) movement when eyes are moved to a new location, no information is obtained
Number of fixations	Amount of fixations in a certain unit of time or a certain task
Fixation duration	Average fixation duration in a task (milliseconds)
Saccade duration	The time taken to complete the saccade (milliseconds)
Saccade amplitude	The size of saccade (deg of visual angle)
Saccade peak velocity	The highest velocity reached during a saccade (deg of visual angle/ms)

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 perceptional processes while watching print advertisements to get information of how the effective advert should look like [11]. Eye movements differ when watching meaningful and meaningless images [10]. 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 [7]. 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. [9] 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. There may be several reasons behind that fact. Eye movement trackers have been very expensive, they need a lot of space, their usability has been very poor from both scientist's and a test subject's point of view. However, there is always a need to find suitable methods to link subjective and objective image quality, and eye movement measurement devices are developing all the time in terms of accuracy and usability. There may be a suitable use of eye movements in image quality research area. In this study, we try pure quantitative eye movement analysis for we wanted to find an easy applicable procedure for applying eye movement information in image quality studies. Quantitative approach can easily be automated so that results would be ready soon to be used when evaluating perceived image quality of a system in image quality industry.

In this paper, our purpose is to find out whether eye movement information can be applied to predicting subjective image quality preferences. Our main research question is how would it be possible to link physical image quality, and subjective evaluation data to quantitative eye movement data. We are interested in finding some practical future eye movement application to image quality area. Behind of these research questions is a broad interest in human perception processes, and processes happening while image quality evaluation. We like to find out whether eye movement data could tell us fundamental characteristics of perception and evaluation process. This means that we want to test the salience mapping theories and see what kind of impact the high level cognitive factors (like image content, or given evaluation task) may have on perception and evaluation process.

2. METHOD

2.1. Stimuli

Four representative image contents were carefully selected from ISO standard 12640. They represented different contents and types of images. Two of them were portrait, and two of them landscape images. They contained widely noticed important test image features, like memory colors, skin color, sky tone, and various compositions like landscapes, and human face. Image contents were called "girl, "cafeteria", "fruits", and "group". (see Figure 1).

Figure 1. Test images from the left: Girl, Cafeteria, Fruits, and Group.

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

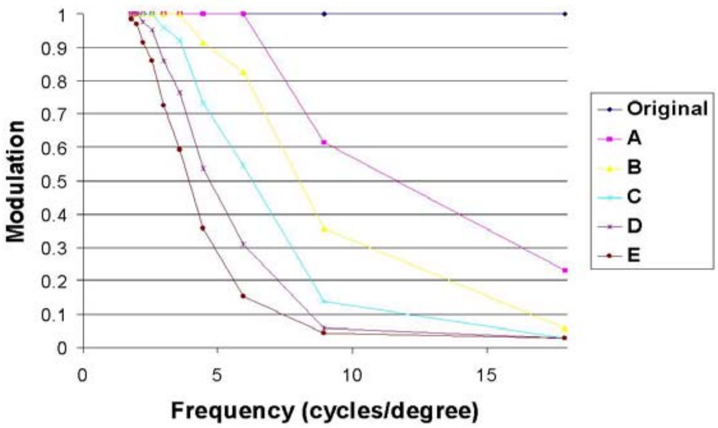

Figure 2. MTF functions used in simulated blur images.

There were a total of 24 images (4 contents * 6 levels, including 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 3. An example of a quality manipulated image. Group with added blur on MTF level E (Figure 2).

2.2. Observers

Eight observers (six males, two females) with an average age of 28 (range 24-37) participated in the experiment. All subjects had normal or corrected to normal (glasses) vision. Only color normal observers participated. Color vision was tested using the Ishihara color plates [6]. Every subject participated in all test conditions.

2.3. Room setup

The experimental room windows were covered with dark grey curtains, and it was illuminated by fluorescent lights (3500K). The illumination level was about 300 lux. Colors of the interior were neutral grey.

2.4. 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 (Figure 4). 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).

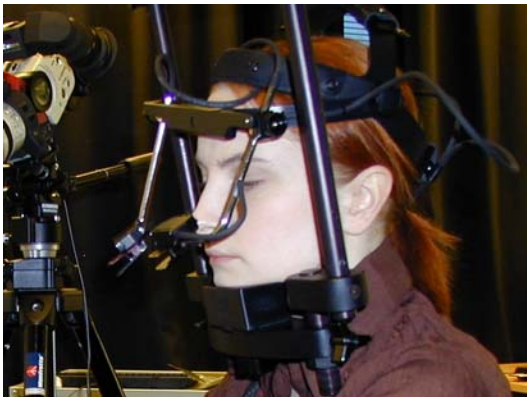

Figure 4. A test subject with head-mounted gaze tracking system.

2.5. Testing Procedure

Head was stabilized with a chin rest to avoid disturbing head movements. The stimuli were randomized in four sequences. 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.

2.6. Data Analysis

As the aim has been set in the title, we analyzed eye movement data only quantitatively in this study. It means that we did not pay attention to the locations of fixations this time but to the quantitative metrics calculable from the eye movement data. We did this decision because of applicability of quantitative analysis in research and development in industry. Quantitative analysis can easily be automated so that testing could be taken as a part of image quality testing process. Paired samples t-test and repeated measures analysis of variance (ANOVA) were used for statistical analyses.

3. RESULTS

We chose the saccade duration as the test parameter, because all eye movement parameters correlate highly and because saccade duration represents eye movements well and corresponds to all other saccade and fixation parameters.

Differences between instructions (tasks) 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 ($t(7)= 3.29$, $p<0.05$). Differences in saccade durations between image contents were not significant when task remained the same.

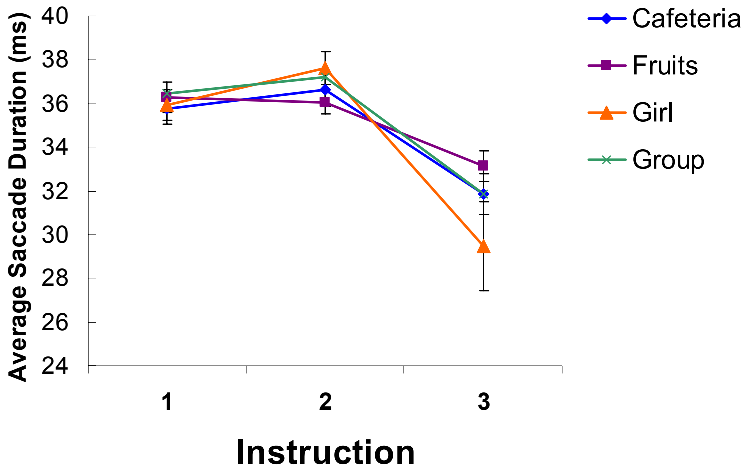

Figure 5. Saccade duration as a function of instruction for all image contents. Instructions: 1="Evaluate image quality." 2="Evaluate color quality." 3=task. Error bars are standard errors of the mean.

Image quality had highly significant effects on eye movements (Figure 6). Images with added blur caused longer saccade duration ($F(5,2)=7.95$, $p<0.001$). However, differences in saccade durations between blurred images were not significant.

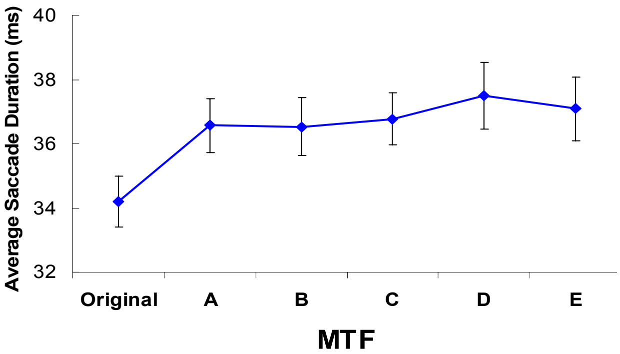

Figure 6. The effect of blur on saccade duration. X-axis is the MTF scale and blur is increasing to the right. Error bars are the standard errors of mean.

Physical image quality (MTF) affected subjective evaluations significantly (Figure 7). Subjective quality deteriorated with increased blur ($F(5,2)=13.23$, $p<0.001$). Image contents alone did not differ significantly in subjective image quality evaluations ($F(5,3)=2.64$, ns). However, there were significant interaction between subjective evaluations and image contents and MTF levels ($F(5,1)=459.06$, $p<0.001$), which means that image contents were evaluated differently at different MTF levels.

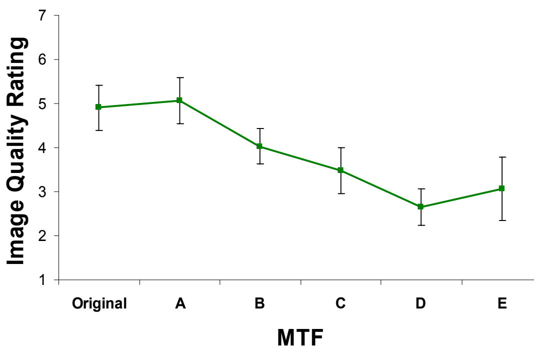

Figure 7. Subjective evaluations of image quality with levels of blur. Error bars are the 95% confidence intervals.

4. CONCLUSION

Objective and subjective image quality, and instruction have all a significant effect on quantitative eye movements. Already this sets some high expectations to quantitative eye movement methodology in image quality studies.

Saccade duration increased with decreasing subjective and objective image quality. Saccade duration increased with increasing blur. This means that saccade durations were longer with poorer subjective and objective image quality. However, small quality differences or artifacts did not differentiate by image quality data, but original quality caused significantly shorter saccade durations than manipulated images. Based on our results, it can be hypothesized that shorter saccade duration reflects perception of better subjective image quality. Is there some kind of fatal difference between original image and manipulated image? According to subjective evaluation data, the original and the less manipulated image got about the same subjective score in scale. So, observers did not consciously differentiate those images, although their eye movements were significantly different with the original and manipulated images. Could eye movements differentiate some good-enough quality from bad quality? Longer saccade durations with lower image quality could probably tell something about task load or stress in perceptual processes between high and low image quality viewing. In other words, it could mean image quality as an ergonomic sense. Higher image quality probably causes less effort for an observer.

We conclude that eye movement measurements provide an interesting addition to image quality research. It could be added to a methodological toolbox in some timescale. In one point of view, eye movement tracking gives a different point of view to the image quality information and added value compared to scaling methodology. Subjective scaling, and 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. So, eye movements are promising, and this research and methodological development should be continued.

Despite of the promising results from study fields, eye movement tracking is not a realistic method to be widely used in image quality industry at the moment. Not many of the companies working in image quality industry are willing to pay for eye movement trackers now that devices are expensive and hard to use and usually still uncomfortable for users. However, these practical problems are getting solved gradually. Measurement devices are getting better all the time, and measurements will be more accurate and ecologically valid. Not all the devices need a helmet-mounted system to measure accurate eye movements. For the future methodology, one should keep an eye on eye movement tracking as a method for studying image quality related issues, at least as an option in one’s image quality toolbox. The quantitative approach used in this study enables researcher to automate the process quite easily to get valuable information or clue of perceived image quality of a new system.

Image contents did not significantly differ in eye movement characteristics when the task of the user remained the same. This still does not mean that content does not matter. It is known that image content has an effect on subjective evaluation, so image content is an important factor from semantic and cognitive perspectives. Moreover, in this study, image contents were evaluated differently at different MTF levels. It means that the best possible MTF is not required by the observer (also in our other study [8]) when the content is certain that the best possible sharpness is not needed, in this case for example girl image. However, it is interesting that eye movements did not differ between contents. Probably just the upper level perceptional processes are just more important in evaluation process, because contents still differed when the task was different.

From the result that image contents did not differ in eye movement metrics when the task was the same, we can derive some contribution to the importance of the top-down perceptual theories in image quality research. Our result means that low level image characteristics did not differentiate eye movement when evaluating image quality, so salience mapping and bottom up processing theories did not get support from this study. Conversely, eye movements were different at different tasks when viewing images. This gives some more empirical emphasis to top-down perceptual theories as a theoretical basis when forming image quality theory. The contribution of this finding is that we should pay even more attention to the task, instruction, and other cognitive aspects when conducting image quality studies. The study conditions should simulate the normal user case as good as possible, and the test subject should have a common understanding with the researcher about what the test subject is doing when evaluating image quality, and what for the tested images are going to be used.

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. Ishihara, S. (1977). *Tests for Colour-Blindness*. Kanehara Shuppen Company Ltd. Tokyo, Japan.
7. Molnar, F. (1981). About the role of visual exploration in aesthetics. *Advances in Instrinsic Motivation and Aesthetics*. New York: Plenum Press.
8. Nyman, G., Radun, J., Leisti, T., and Vuori, T. (2005). From image fidelity to subjective quality: a hybrid qualitative/quantitative methodology for measuring subjective image quality for different image contents. *IDW/AD'05*, Dec 6 – 9, Takamatsu, Japan.
9. Näsänen, R., Ojanpää, H. & Kojo, I. (2001). Effect of stimulus contrast on performance and eye movements in visual search. *Vision Research 41*, 1817-1824.
10. Parkhurst, D. Law, K. & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research, 42,* 107-123.
11. 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.
12. 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.