CCTV mediated observation versus non-mediated observation: investigating perceived image quality with different test systems

Arthur BENNIS[1], Renske B. LANDMAN[2] and Dick T.M.J. LENIOR[1]

[1]HAN University of Applied Sciences, Arnhem, The Netherlands
[2]ErgoS Engineering and Ergonomics, Enschede, The Netherlands

Abstract. A basic question in practice is whether perceived image quality decreases when using a CCTV system (mediated) compared to direct observation (non-mediated). Three main characteristics were investigated: (1) test systems, (2) light conditions, and (3) image scaling. At (1) interpretable differences were found. At (2) normal light conditions (500 lux), no differences were found between mediated and non-mediated vision, while at 50 lux, mediated vision is significantly better. At (3) upscaling has a negative effect on quality of mediated observation. The results lead to clear indications for the setup of future research in laboratory as well as in practice.

Keywords. CCTV, perceived image quality, Rotakin, Vidilabs, Landolt

1. Introduction

CCTV systems are applied in more and more professional settings, for example security, lock and bridge control, traffic supervision, and remote process control. Though there are well established Human Factors (HF) guidelines for control center design in general (ISO 11064), guidelines on how to design and use CCTV systems are hardly available (Schreibers, Landman & Pikaar, 2012). Therefore, a research project was initiated to investigate the possibilities of developing guidelines for the design of CCTV systems. An overview of the entire research project can be found in Pikaar and Lenior (2014). One of the interesting issues concerns the way in which image quality is measured. This paper describes the first steps to come to a “standard” setup for the measurement of *perceived* image quality.

CCTV systems consist of cameras, interconnecting IP network hardware, recorders, support infrastructure, monitors, and human operators. The human operator sees the real scene mediated through the CCTV system. So there are many factors that influence (perceived) image quality including the hardware, scene environment, image presentation, task, and operator factors. Much literature is dealing with the technical aspects of image quality. However, technical quality of images is subordinate to what the operator perceives and what he should perceive given his tasks. We are missing a link to operator tasks.

The basic research question concerns *the differences in perceived image quality between viewing a scene using a CCTV system (mediated) or through direct observations (non-mediated)*. Concerning this basic question, the first research question is: *how can perceived quality be measured?* Literature in the CCTV field proposes the Rotakin system (Aldridge & Gilbert, 1996) and the Vidilabs Chart (Damjanovski, 2005). The Rotakin used to be the default CCTV systems test as prescribed by the UK Home Office for the UK (BS EN 50132-7:1996) and has been used by some of our sponsors (see acknowledgement). The Vidilabs test system is used in Australian CCTV approval processes and contains test images and patterns to validate the technical qualities of different parts of CCTV systems.

However, in order to get indications for the perceived image quality, we have to measure what the operator actually sees. Therefore, the optometric field provides us with several measuring techniques. The Landolt C chart is one of them, normally used for measuring the smallest critical detail a person can see (i.e. the quality of human vision in direct observation) (ISO 8596, Liu & Cho, 2002). No experimental research has been found in which these different measurement systems in relation to CCTV images are being compared. Hence, our first research question concerns *the differences between available test systems regarding perceived image quality.*

Moreover, ambient light on the scene is one of the main factors that can influence perceived image quality (APTA, 2011). Therefore, our second research question concerns *the effect of ambient lighting on perceived image quality.*

The last aspect originates from our daily practice with all kinds of control centers. We often find operators using upscaled images. When asked, they answer they see more with the upscaled images. Because the native resolution of the image does not change with upscaling, it is expected that upscaling would only be useful in case the individual pixels cannot be distinguished (given a certain pixel size and viewing distance), something hardly mentioned in literature and guidelines. So, our third research question concerns *the effect of image upscaling on perceived image quality.*

2. Methods

2.1 Experimental design

We chose for an experimental setup deploying a within-subjects design; all participants were required to participate in all conditions. Counterbalancing was applied to exclude possible order effects. Following the research questions, the independent variables are (a) type of vision: mediated or non-mediated, (b) deployed test system: Rotakin, Vidilabs, or Landolt C chart, (c) ambient lighting of the scene: 50 lux or 500 lux, and (d) ratio of upscaling: native resolution = 1:1 or 1:1.3.

2.2 Apparatus and layout

To check for color vision deficiency (or color blindness) the Ishihara test was used. Rotakin, Vidilabs, and Landolt C were used to test visual acuity, being one of the main component of perceived image quality. To determine visual acuity, the wedges on the Rotakin were used (see figure 1). On the Vidilabs, we used the 5-line wedges (similar to Rotakin) and from the Landolt C chart the displayed optotypes were used. Next to these systems, the Freiburg Vision Test (FrACT) (Bach, 1996), a digital Landolt C chart, was deployed to check this test for future research applicability in comparable settings.

Figure 2: Wedges on the Rotakin

For the mediated conditions, the test charts were recorded with a Sony DSR-PD170P camera with a F= 6-72mm 1:1.6 lens, a 1,3” optical sensor (ratio 4:3), and a resolution of 786 x 576 pixels. The camera was connected with a firewire connection to a computer (Windows XP) which presented the camera image by Media Encoder 9 on a 24” Philips Brilliance TFT monitor with a resolution of 1920 x 1200 pixels (pixel size 0.27 mm). The size of the presented image was 786 x 576 pixels (native resolution). The camera was set on default settings; all image enhancements disabled.

The experiment was conducted in a HCI laboratory including two rooms. In one room, video images were shown to participants on a computer monitor for the mediated conditions. In the other room, the charts were set up for the non-mediated conditions. The viewing distances were adjusted to obtain the same viewing (eye) angle for both mediated and non-mediated condition for each chart.

The ratio of upscaling was achieved by either using a software upscaler (mediated condition) or by placing the participant closer to the chart (non-mediated condition). A photometer was used to ensure a 50 lux and 500 lux ambient light level on the charts. The light level at the location of the participant was 500 lux at eye height.

2.3 Adjustments

A pilot experiment showed promising results for the Landolt C chart. However, results from Vidilabs were found hard to interpret and needed extra research. The Rotakin did not show any added value. The wedges on the test systems allowed the participants to easily develop tactics in order to obtain a higher score on visual acuity. In follow-up experiment the Rotakin was excluded, and on the Vidilabs subjects were asked to identify the point where the converging lines were not individually distinguishable (figure 2 shows the used elements). Furthermore, the turnaround time of the experiment was optimized from 40 minutes to 20 minutes to prevent effects from fatigue. Finally, the lighting of the charts was improved by using LED studio lights.

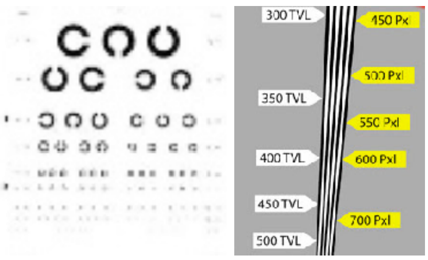

Figure 3: Landolt C optotypes and converging lines on Vidilabs

2.4 Participants

The experiment consisted of 14 participants (10 male and 4 female, 21-61 years; mean age: 36), half sampled from the student population and the other half among employees of the university.

3. Results

The design of the experiment is within-subjects. Because of the ratio level of the scores, the relatively low n, and the expectation that the data will not be distributed normally, a non-parametric Wilcoxon signed ranks test is used to analyze all results.

Results show the number of negative differences (-), positive differences (+) and the number of participants (n). The first p-value is asymptotic, the second p-value is exact (generalized p-value). Next to that, mean scores of all subjects are noted for a rough indication. For the Landolt scores the mean of the visual acuity is shown, for Vidilabs the mean of the merging point of the converging lines. The last column shows whether there is a significant difference. All results are reported with two-tailed p-values and an α of 0.05. M stands for *mediated* results and NM for *non-mediated* results.

Table 1: Mediated versus non-mediated conditions

Condition	-	+	n	p	p_{exact}	mean NM	mean M	significance
Landolt 50 lux	0	11	14	.00	.001	1.3214	1.7143	M > NM
Landolt 500 lux	2	1	14	.79	1.000	1.8750	1.8570	None
Landolt 500 lux 1:1.3	8	0	14	.01	.008	2.5710	2.2321	NM > M
Vidilabs 50 lux	8	5	14	.29	.312	640.36	619.29	None
Vidilabs 500 lux	11	1	14	.01	.006	694.29	620.71	NM > M

The results in table 1 show that at 50 lux, the Landolt M scores significantly higher than NM. The light emitting nature of the monitor is expected to cause this finding. The Vidilabs was expected to show the same result because it also measures visual acuity. However, the results from the Vidilabs do not support this. On the other hand, at 500 lux, the Landolt does not show a significant difference between M and NM conditions but the Vidilabs does. At ratio 1:1.3, the NM condition scores significantly higher than M. The system cannot match the advantage a participant has when moving closer towards the Landolt chart in the NM condition.

Table 2: Low light versus standard light conditions

Condition	-	+	n	p	p_{exact}	mean 500	mean 50	significance
Landolt NM	14	0	14	.00	0.000	1.8750	1.3214	500 lux > 50 lux
Landolt M	6	1	14	.11	.141	1.8570	1.7143	None
Vidilabs NM	10	2	14	.00	.002	694.29	640.36	500 lux > 50 lux
Vidilabs M	6	7	14	.94	.973	620.71	619.29	None

Results of both Vidilabs and Landolt show significantly lower scores in the NM condition at 50 lux than at 500 lux ambient lighting (see table 2). This was expected because it is harder to see the test systems in lower light conditions. The mediated scores however do not show a significant difference; the light emitting characteristics of a monitor might explain this effect.

Table 3: Actual increase versus expected increase at ration 1:1.3

Condition	-	+	n	p	significance
Landolt 500 lux M	11	3	14	.02	Actual increase < expected increase
Landolt 500 lux NM	6	8	14	.24	None

Upscaling of the Landolt was done by decreasing viewing distance by 30% in the NM condition and upscaling the image with 30% in the M condition. Therefore, the participants are expected to score 30% higher in acuity. When comparing the expected increase (30%) with the actual increase, no significant difference was found in the NM condition (see table 3). However, for the M condition the actual increase was significantly lower than expected. Hence, improving the image by 30% (in size) did not lead to an improved perception of 30%. This is supported by the third result from table 1 (Landolt M versus NM at 500 lux and ratio 1:1.3).

This result shows that when the ratio was increased to 1:1.3, the M condition significantly underperformed compared to the NM condition.

In addition to the experimental conditions, some tests with FrACT were conducted to investigate the possible alignment of results of the FrACT with results of the Landolt NM, see discussion.

4. Discussion and Conclusion

In practice, we found that work on testing image quality of CCTV systems is focused on the technical features of the images. The visual qualities of the observer are hardly considered. Therefore, a set of experiments was carried out in which CCTV mediated vision was compared with direct (non-mediated) vision especially focusing on the *perceived* image quality.

Concerning our first research question about the differences between Rotakin, Vidilabs, and Landolt C, the pilot study showed that the wedges on the Rotakin were prone to tactics developed by participants to obtain higher scores. Rotakin was therefore excluded in the follow-up. The results from the Vidilabs were difficult to interpret; our future research will focus on whether Vidilabs can be used for those factors (contrast, color etc.). The Landolt C provided clear outcomes. We found that the Landolt testing system is not only a method for measuring critical detail but it can also be used as an instrument to test the visual effort of subjects in our future research. Then we will investigate other aspects of *perceived* image quality while using the Landolt C as measurement task for *perceptual* effort, a step towards a "standard" experimental setup.

About the second research question, while ambient light did have an effect on the non-mediated results, we found that at low lighting level (50 lux) mediated vision was significantly better than non-mediated vision. The light emitting characteristics of a monitor compensate for the lack of light at the test charts. Brightness and contrast are clearly factors that can influence perceiving quality of images. In future research, ambient lighting (at the recording end) will be maintained at 500 lux and variations in brightness and contrast conditions will be investigated further.

Regarding our third research question, the results show a clear negative effect of scaling in mediated observation. An enlargement of 30% did not meet expected 30% higher scores. In future research we will rule out problems regarding scaling by maintaining the native resolution. (The use of recorded video seems more practical).

We tested with FrACT (a digital Landolt C test, displaying single optotypes directly on screen) in addition to the Landolt C chart. Our preliminary test showed consistent and significant higher scores (0.5 points higher) on the mediated Landolt C at 500 lux than on FrACT. Both systems are discrete, however FrACT uses smaller steps. This might explain the difference. The Landolt C chart might be prone to learning effects or derivable optotype directions which also can explain the difference. By presenting single optotypes, as presented by the FrACT, this effect might be countered. The FrACT will be deployed in future research to determine visual acuity.

In future research, several questions of our engineering practice will be tackled. For instance: How many camera images can an operator monitor, without getting mental overload? This depends on the task to be fulfilled and the properties of the surrounding images. Three 'sorts of properties' of these images will be considered.

On the one hand the *content*: to what extent is there a clear relation between the surrounding image and the image where the operator is focusing on at a certain moment. This will be investigated in practice (being executed in another part of the project) and will give a good understanding of *scene perception:* which images form a logical unity for the operator?

The second 'sort of properties' concerns the question to what extent the surrounding images cause distraction from the main task. This question will be investigated in an experiment. The main task performance will be scores on electronic Landolt C optotypes. The parameters to be varied in the surrounding images are distance, brightness/contrast and motion.

The third 'sort of properties' is what we have called the link between the tasks the operator has to fulfill and the features of his viewing task. What is the relation between task distinction and perceived image quality as defined above? For the latter a standard test setup seems to emerge.

Acknowledgement

This research would not have been possible without the funding by H.I.T.T. Traffic, IHC Dredgers/ IHC Beaver Dredgers, Nedap Security Management, Rijkswaterstaat, Rijksbelastingdienst, Dutch Railways, ProRail ICT Services, Royal Haskoning DHV, Total E&P, Vopak Management Netherlands B.V., Waterschap Hollandse Delta, DG-Organisatie Bedrijfsvoering Rijk and Province of North-Holland.

References

Aldridge, J. & Gilbert, C. (1996). Performance testing of CCTV perimeter surveillance systems: Using the Rotakin Standard Test Target (Publication 14/95). Sandridge, UK: Police Scientific Development Branch. Home Office - Police policy directorate.

APTA (2011). Selection of cameras, digital recording systems, digital high-speed networks and train lines for use in transit-related CCTV systems (Recommended practice: APTAIT-CCTV-RP-001-11). Washington.

Bach, M. (1996). The Freiburg Visual Acuity Test-automatic measurement of visual acuity. Optometry & Vision Science, 73(1), 49-53.

BSI, (1996). BS EN 50132-7:1996 Alarm systems. CCTV surveillance systems for use in security applications Application guidelines. London, United Kingdom: BS

Damjanovski, V. (2005). CCTV Networking and digital technology. Burlington: Elsevier Butterworth – Heineman.

ISO, (2009). ISO 8596:2009 Ophthalmic optics -- Visual acuity testing -- Standard optotype and its presentation. Geneva, Switzerland: ISO.

ISO 11064. Ergonomic Design of Control centres – Multi part standard.

Liu, L., & Cho, J. (2002). The Effect of Pixel Nonlinearity on Cathode-Ray Tube–Based Visual Acuity Tests. *Optometry & Vision Science*, 79(11), 724-734.

Pikaar, R. & Lenior, D. (2014). Human Factors Guidelines for CCTV control centre design. In O. Broberg et al. (Eds.), Proceedings of the XI Conference on Human Factors in Organizational Design and Management (pp. this issue). Copenhagen.

Schreibers, K.B.J., Landman R.B., & Pikaar, R.N. (2012). Human Factors of CCTV: Part 1. Technology and Literature review. Ergos, HAN, Intergo, & vhp-hp. Enschede: ErgoS Engineering & Ergonomics