The Big Picture on Small Screens: Delivering acceptable video quality in mobile TV

Hendrik O Knoche
University College London
Gower St
London, WC1E 6BT, UK
+44 207 679 3642
h.knoche@cs.ucl.ac.uk

M Angela Sasse
University College London
Gower St
London, WC1E 6BT, UK
+44 207 679 7212
a.sasse@cs.ucl.ac.uk

ABSTRACT

Mobile TV services can deliver up-to-date content on the move, but to deliver a good viewing experience on hand-held devices, service providers need to adapt content in frame rate, resolution, encoding bitrate and the viewable area. Viewers can change the viewing distance and on some devices scale the picture to their preferred viewing ratio - trading off size for angular resolution. We investigated optimal trade offs between size and resolution through two lab and one field study. In a first lab study 35 participants watched clips of different content and shot types on a 200ppi PDA display at a resolution of either 120x90 or 168x128 at six sizes. They selected their preferred size and rated the acceptability of the visual experience. Despite low resolution participants preferred image sizes that resulted in viewing ratios (9.8) similar to those in average living room TV setups and much larger size than ITU recommendations for assessing picture quality suggest. The minimal angular resolution people required and which limited the up-scaling was 14 pixels per degree. The second study had the same participants watch 168x128 video clips zooming into extreme long shots of football. Participants preferred players to be at least 0.7° in height. A third experiment examined the acceptability of video at different sizes and encoding bitrates at constant angular resolution video encoding bitrates and compared them to an earlier lab study. Participants in the field appreciated bigger pictures (≥35mm) more than the participants in the lab. Our results show that current recommendations on TV picture quality research underestimate the contribution of video size to the experience of mobile TV.

Categories and Subject Descriptors

H.5.1 Multimedia Information services – Evaluation/ methodology. H5.2 User Interfaces [user-centered design]

General Terms

Human Factors, Experimentation, Measurement, Design

Keywords

Mobile multimedia consumption, resolution, size, trade-off

MM'08, October 26–31, 2008, Vancouver, British Columbia, Canada.
Copyright 2008 ACM 978-1-60558-303-7/08/10...$5.00.

1. INTRODUCTION

Advances in the development of displays have equipped mobile devices with 200ppi displays offering VGA (640x480) resolution. This capability goes some way towards reducing worries that viewing of TV content on mobile devices may be marred by coarse displays at short viewing distance. At a constant viewing distance (D), perceived video quality is determined by the size of the depicted image and image resolution. By increasing the viewing distance, the image size decreases while the angular resolution of the picture increases. Viewing distances and video quality and have been researched for traditional TV displays, and more recently for HDTV. When people were free to choose their viewing distance for consuming TV content on fixed displays the resulting viewing ratio (VR) was not based on the best attainable subjective video quality as identified in [1], but depended on the picture size [2]. Due to living room layouts viewing distances in the home [3] could be considered fixed [4] (see Figure 1, left) and do not match the preferred viewing ratios obtained in lab settings. With mobile TV, the viewing at about arm's length distance might be fixed, too, but the provider has to choose which resolution to deliver and the size of the picture could adjustable on the device, changing the angular resolution along. Currently, it is unclear which configurations of viewing ratio and angular resolution would result in the best experience for the viewer.

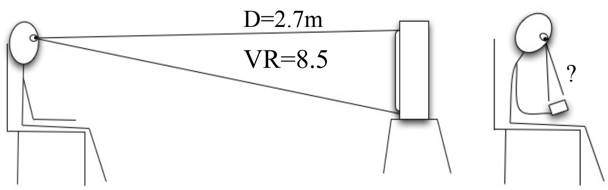

Figure 1: Standard TV (left) **and mobile device viewing** (right)

In this paper we describe three studies conducted to determine viewing preferences for mobile TV in two different settings. The first two studies investigated user preferences for trading off image size and resolution for different content and shot types in a lab-based setting. The second study looked specifically at zooming into extreme long shots (XLS) important for adaptation of sports content. The third study evaluated these parameters in a field trial, in which users viewed the content while traveling on underground trains. The results provided insights for 1) the development for displays, and 2) the optimal delivery of mobile TV content, especially how mobile TV content should be presented (scaled up) on displays that are used at close distance on small mobile devices.

In the next section we present the background literature on TV viewing distances, resolution and shot types. It motivates the parameters chosen for the three studies through our previous research into consumption of multimedia content on mobile devices. Sections 3, 4 and 5 describe the studies and are followed by an overall discussion of the results in section 6. Conclusions for researchers and practitioners are presented in Section 7.

2. BACKGROUND

Most people who have not experienced mobile TV are concerned that screen size would be obstacle to a satisfying experience [5], [6]. Talking about the size or the resolution of videos and displays only makes sense in conjunction with a viewing distance (*D*). When considering the viewing distance at which mobile devices are used, the relative size in terms of the viewing ratio (*VR*) – the quotient of *D* and screen height *H* – is not radically different from those in a home or PC TV setting. The real difference is the resolution of the content that is delivered to mobile devices, compared to standard television (SDTV). Historically, the viewing distance was the only way for people to adjust their preferred *VR* - the angular size of the display - and the angular resolution, since most devices would only depict content at a fixed resolution. With advanced coding schemes and presentation devices that can stretch video, the question of the preferred angular resolution and viewing ratio can be reconsidered, because many displays will allow for adjustment of the size of video content depending on user preferences.

2.1 Viewing Close Up

The amount of detail resolvable by the human eye is primarily limited by the density of the light-sensitive rods and cones on the eyes' retina. Normal 20/20 vision is classified as the ability to resolve 1 minute of arc (1/60º) [7] and translates to 60 pixels per degree (ppd). Viewing distance is often expressed in terms of the ratio between the distance of the observer to and the height of the visible screen. A viewing ratio of 5 describes a viewing distance five times the picture height (*H*). The visual angle (*VA*) or angular size (AS) θ expresses the viewing ratio in degrees regardless of *D* as illustrated in Figure 2.

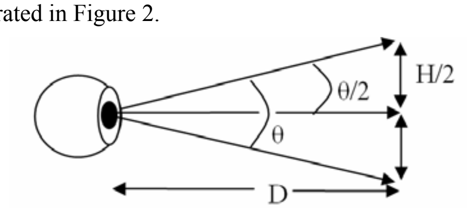

Figure 2: Viewing ratio and the visual angle

The human visual system uses two mechanisms to focus on objects: convergence and accommodation. Convergence denotes the eyes moving inward when focusing on nearby objects, and accommodation describes the focusing on objects of different distance by means of physically deforming the lens of the eye. The resting point of accommodation (RPA), i.e. the default distance at which objects appear sharp, e.g. when opening the eyes, is around 75cm for younger people and increases in distance with age [8]. The resting point of vergence (RPV) is 114cm when looking straight ahead, and drops to 89cm when looking 30 degrees down. This is a posture (cf. Figure 1, right) often seen in mobile TV consumption because people use their legs or bags as support for the hand holding the device [6]. The stress of convergence contributes more to visual discomfort than the stress of accommodation. Continued viewing at distances closer than the RPV can contribute to eyestrain [8]. When viewing distances come close to 15cm, people experience discomfort [9]. Boff & Lincoln showed that visual acuity decreases as viewing distance increases [10] so for close viewing distances people's acuity is at its maximum. Clearly, multimedia consumption on mobile devices happens at close range but its exact preferred viewing distance (PVD) has not been researched in relation to different sizes and resolutions. Kato *et al.* obtained typical viewing distances of approx. 35cm (VR=11) from both standing and sitting people using a 166ppi mobile device [11].

2.2 TV Viewing

From research on traditional TV viewing we know that a number of factors can influence people's PVD. Thompson found that people chose their viewing distance so that the TV lines were not visible anymore [12]. More recent research, however, has refuted this assumption. Due to the layout of the average living room, people typically watch TV at the so-called Lechner (US, 9ft) or Jackson (Europe, 3m) distance [13]. Unfortunately, both of these values are poorly documented, and their original sources are not readily accessible. As recently as 2004, the median viewing distance for standard definition TV of BBC employees was reported as 2.7m (VR=8.5) [14]. Nathan *et al.* showed that the viewing distance of regular TV in the home varied with the age of the viewers. The average viewing distance for 17 year olds and younger was 2.25m (VR=7.8), whereas adults watched from 3.37m (VR=11.7) [3]. They did not explain this difference, but reported that children were more mobile than adults, and much less likely to sit or lie on furniture while watching TV.

In a series of five studies, Lund [2] found that participants' preferred viewing ratio was not a constant 7. With increasing image size, and independent of resolution, the preferred viewing ratio approached 3 or 4. Based on Yuyama's [15] and his own results, Lund hypothesized that viewers might select their viewing distance not to maximize perceived visual quality, but "*to optimize a sense of presence or reality*" [2]. Ardito [16] found that when brightness was reduced, there was a trend of participants sitting closer to the screen. When watching HDTV content on a 38 inch screen (in a completely dark room), the average preferred viewing ratio was 3.8, compared to 6.3 when viewing the same footage in brighter surroundings. Ardito predicted a viewing distance (in cm) of $D = (3.55H + 90)/H$. Although he did not test small mobile screens, he interpolated from a range of HDTV screen heights from 198cm to 15cm that for screens with a screen height close to zero the viewing distance would be 90cm [16]. Ardito *et al.* [17] found the viewing distances for moving picture content to be further away than for still picture content. However, the effect of brightness and screen resolution on the PVD were smaller than the effect of the size of the screen. For HDTV content, Ardito predicted a viewing distance (in cm) of $D = (3.55H + 90)/H$. Although he did not test small mobile screens, he interpolated from a range of HDTV screen heights from 198cm to 15cm that for screens with a screen height close to zero the viewing distance would be 90cm [16].

For subjective video quality assessment of multimedia applications, the ITU suggests viewing ratios between one and eight in their recommendation series P.910 [18]. But for subjective video quality assessment of TV material in recommendation series BT.500 [19], the ITU specifies preferred viewing distances (PVD) depending on the screen height. The recommendation contains a graph that illustrates the relationship between screen height and preferred viewing distance for screen sizes between 18*cm* and 2*m*. A power function $f(x)=76.5\,x^{-0.41}$

describes (R^2 = 0.97) the relationship of screen height in *mm* to PVD. According to ITU, these values should be applied for both SD- and HDTV *'as very little difference was found'* between the two resolutions. Screen heights smaller than 18cm and lower resolutions are not covered by the ITU's recommendations but considering the trend their PVD should be 11H and higher.

2.3 Size and Resolution

Since the early works of Kell *et al.*, a number of studies have investigated the impact of TV resolution and size on the experience of the viewers [8]. Jesty found evidence for a preferred viewing distance [20] when participants chose their distance from which to watch projected still pictures - their preferred viewing ratio was constant for a given resolution. Ribchester argued that this could be merely attributed to conditioning to existing physical setups in the home [21]. Westerink & Roufs conducted the most comprehensive study on the effects of picture size, viewing distance and resolution on subjective image quality with still pictures. They achieved a range of angular resolutions (5, 17, 50 and 64ppd) by defocusing a projector lens and of heights 24, 48, 72 and 92cm at three viewing distances 2.9*m*, 3.9*m* and 5.4*m* – equivalent to viewing ratios between 22.5 and 3.2 - using projected square format picture slides in a dark room. At a constant viewing distance, the subjective quality was influenced independently by both the resolution and the size of the pictures [1]. Maximum subjective quality resulted when the resolution equalled 16cpd (cycles per degree), independent of the picture width (which for pixel-based displays translates to 32ppd). This indicates that the gains in perceived visual quality from achieving a higher visual resolution beyond 16cpd were not big enough to compensate for the reduction in picture angle. However, these results were based on quality ratings of still pictures and viewing distances between 2.9m and 5.4m and might not apply to consumption on rasterized mobile devices and preferred viewing conditions.

Sugama *et al.* found that - for pictures of identical angular resolution of 27ppd, on a 100ppi display, all shown at a VR of 6 - subjective video quality was higher when they were viewed at a close distance of 40cm, compared to viewing distances of 80cm and 1.6m [22]. However, the study did not control for the 'graininess' or pixilation of the display. At the closest distance, the angular resolution of the display was 27ppd, but for the largest viewing distance (1.6m) with medium (54ppd) and large (100ppd) images the pixels were close to and above the human discrimination threshold. In a study that used different viewing distances (1.5, 2 and 3*m*). Hatada *et al.* showed that the angular size of the display was not sufficient to describe the effect of display size, but that the absolute picture size or the absolute viewing distance needed to be considered [23]. Yu *et al.* found no statistical difference with assessors judging video quality impairments when the material was presented at a VR of 3 or 5 [24]. Lombard *et al.* found that bigger screens (115cm height) displaying standard definition television (SDTV) resulted in more intense experiences for the audience compared to smaller screens (30cm) but the level of enjoyment remained unchanged. Reeves *et al.* found that large HDTV presentations with high audio resulted in more positive evaluations in comparison to SDTV, e.g. in terms of excitement [25]. Apparently, larger depictions need to be coupled with higher resolutions in order to achieve higher levels of enjoyment.

In summary, almost all of the existing research on size and image resolution preferences for TV content was based on large screens, and there is a gap in the understanding of what constitutes people's preferred angular resolution and viewing ratio when watching video on mobile devices. Current models of video image quality e.g. [26] are not based on empirical results in the domain of mobile viewing devices

Some models such as Barten's [27] *square-root integral* (SQRI) describe the effect of picture resolution, contrast, display resolution, luminance, display size, viewing distance and noise on subjective image quality. Barten showed that SQRI accurately models the subjective image quality results obtained by Westerink & Roufs and predicts the preferred viewing distances of Jesty [20] the maximum of achievable subjective quality. It does, however, not explain the preferences in viewing ratios observed by Lund [2] and Ardito [16].

2.4 Shot types

In this paper we use Thompson's classification [28] for Medium Close-Ups (MCU), Medium Shots (MS), Long Shots (LS), Very Long Shots (VLS) and Extreme Long Shots (XLS) – see Figure 3 for examples. Faced with the more constrained visual real estate, content producers are considering a different mix of shot types for mobile TV. In Asia, content creators produce soap operas especially for mobile devices, which are short and rely heavily on close-up shots with little dialogue. Most emotions have to be conveyed by means of facial expressions and "*there is very little dialogue and a lot of close-ups of characters striking exaggerated poses*" [29]. ESPN minimizes the use of extreme long shots In sports coverage for mobile devices [30]; instead it uses more high-lights with close-up shots.

Hands' [31] multimedia model contains a notion of shot type. According to his research multimedia quality depends on what content is presented. For head-and-shoulder (HS) content similar to an MS both audio and video contribute equally to the multimedia quality. But for high-motion action scenes video quality becomes more important than audio quality.

In the domain of pictures XLS appear less sharp than shots that depict closer objects. Frieser & Bierman reported that portrait pictures (equivalent to an MS in Figure 3) received consistently higher quality ratings in comparison to other scenes (similar to XLS) despite the fact that they were of equal objective quality [32]. Kingslake pointed out that people can detect blurring for distant objects more readily and ascribed this to the fact that they are small [33]. Corey *et al.* reported that "*subjective print quality* [of pictures] *depends upon scene magnification (or equivalently camera-to-subject distance)"* [34]. Very high resolution does not necessarily equate to the highest perceived quality. For a portrait the highest resolution in terms of the modulation-transfer function (MTF) was judged of worse quality than a slightly lower quality in a study by Frieser *et al.* [32].

2.5 Zooming

The research community has embraced the idea of zooming into pictures to improve the viewer's experience on small screens e.g. [35], [36], [37]. But a number of concerns remain about the range of zooms that can and should be used. The resolution of standard TV footage is limited and therefore can only afford a finite amount of zoom. The coverage of many popular sports makes extensive use of XLS which cover a large amount of the pitch, and the audience can benefit from seeing potential pass receivers or other strategic information. Zooming can remove such valuable context information. Viewers have higher standards for the sharpness of smaller and far away objects in still pictures [32], [33], [34] might expose the perceptual imperfections of XLS more.

2.6 Previous Research

The studies presented in this paper are based on a line of previous research on multimedia consumption on mobile devices which have built on one another. A first study looked at the trade-off between frame rate and resolution. Five frames per second (fps) as a minimum frame rate for video-conferencing [38], observation [39] or intellectual tasks [40] but comparable knowledge lacked for passive consumption of content on mobile devices. Guidelines [41] suggested prioritizing frame rate over resolution for high motion sports content but some research had shown that this content was not very sensitive to frame rate changes [42],[43]. The study found that for watching sports content on a mobile devices resolution was more important than a smooth frame rate such that people can identify players and the ball. People rated the video quality of football content less acceptable when the frame rate dropped below 12fps [44]. Comparable displays on desktop computers achieved high acceptability for frame rates as low as 6fps. The reason for the need for higher frame rates on mobile devices is not yet fully understood, but highlights the importance to measure video quality in as realistic setups as possible to the real experience. In all subsequent studies we encoded our videos at a nominal frame 12fps and presented them mobile devices. This study introduced the binary method of *acceptability* for the measurement of video quality. It provides critical points at which quality becomes unacceptable and can easily be translated into utility curves.

In a series of focus groups in three countries [5] we identified news, sports and music as the content types people preferred to watch on mobile TV. In a large mobile TV field trial [45] children's programmes and cartoons had proved were very popular and we added animation as a fourth content type and focussed our subsequent studies on these.

In [46] we focused on the effects of video and audio bitrates and size on the four content types at constant angular resolution on a 115ppi mobile device. We found no evidence that people change their viewing distance in response to varying video sizes at – instead, they kept an average estimated viewing distance of about 27cm resulting in angular resolution for all clips of 21ppd. The perceived acceptability decreased non-linearly when the dimensions decreased from 168x128 (VR=9.6) to 120x90 (VR=13.5). Either the video size of 20mm height was to small or its resolution of 120x90 too low for an acceptable mobile TV experience. Since the study used constant angular resolution, size and resolution were correlated. Study 1 in this paper disambiguates these results. Text quality and legibility affected the acceptability of video quality especially in news content. Text quality is often omitted in objective video quality models. The study found that higher audio quality could adversely affect the acceptability of video quality. The acceptability of videos encoded between 32-224kbps and accompanied by 32kpbs audio were judged worse when they were presented with 16kbps audio. A non-additive effect of audio and video quality that runs counter to current multimedia models [31], [47].

Due to the importance of text on video quality found in [46] and the fact that downsizing popular news content can render text illegible we devised a study [48] which insured text legibility at four different sizes. The news clips in study 1 are taken from those clips. The study showed that despite legible text news content below QCIF resolution was not acceptable when displayed natively on a 115ppi mobile device.

We conducted an initial investigation on the effect of size and resolution on shot types [49] by including shot types in an additional analysis of the results in [46]. The acceptability of shot types varied with content type. The XLS of soccer content was the least acceptable shot type when presented natively on a 115ppi screen below 208x156 (height smaller than 4cm; VR≥7.8) [50]. The reasons for that were mostly ascribed to a lack of visual detail which could have been due to either the small size or the low resolution of the content. The need for including different shot types commonly used in TV material into study 1 stemmed from this research.

A follow-up study investigated preferred zoom ratios into XLS of football content at constant angular resolution of 35ppd. The content had to be presented at a viewing ratio larger than 8.5 for a large majority (more than 60% of the participants) to prefer zoomed over non-zoomed material [51]. Zoom factors larger than 1.3 were rejected by a majority of participants when watching QVGA content at a VR of 8.5. Viewing ratios of 11.3 (240x180 resolution) and 14 (176x144) made people prefer larger zooms. The condition in terms of size and detail must good enough and the information was left out by larger zooms too valuable to justify higher zooms. Since the study used content at constant angular resolution study 2 presented in this paper was devised to disambiguate the confounding effects of size and resolution.

The ITU guidelines for video quality assessment suggest control of e.g. lighting conditions [19], which is not feasible in the field. Evaluation of subjective video quality in the field is time-consuming and difficult because interruptions, movement, lighting and sound conditions are hard to control [52]. However, results that are obtained in more realistic settings have greater predictive validity [53]. The only previous study that compared video quality assessments on mobile devices in the field and the lab is a study by Jumisko-Pykköö & Hannuksela, who evaluated the effect of packet loss on audio-visual quality [54]. In general, participants rated the acceptability and of audio-visual quality higher in the field than in the laboratory for all four tested error ratios (from 1.7% to 20.6%). The same was true for the satisfaction ratings except for the 1.7% error ratio condition. Satisfaction ratings in the field (6.5) were lower than in the lab (7.5) while the participants' acceptability ratings for the same clips were higher for the field (89%) than in the lab (82%). The reason for this difference is not yet understood but study 3 replicates the study [46] on public transport.

3. STUDY 1: SIZE AND RESOLUTION

3.1 Material

From previous studies on multimedia consumption on mobile devices we obtained news [48], animation, music and sports video clips [46]. These base clips were based on footage from DVDs and terrestrial digital video broadcasts (DVB-T). For each of the four content types we used four shots of each of the shot types MS/MCU, LS, VLS and XLS (see Figure 3 for examples). Our material did not provide consecutive shots of all types lasting for more than ten seconds, so in order to control for effects due to shot types each shot lasted for only 8-10 seconds. Due to differences in content, the most detailed shot types were not identical. For example, the football shot with the most detail was closer to a mid-shot than a medium close-up, and the most detailed shot in animation was closer to a close-up. Another difference worth pointing out is that the XLS of football and news depicted people far away whereas the XLS of music (moving camera) and animation (static shot) depicted a landscape.

Using previously tested material allowed for a comparison with earlier results [48] and [46], which partly addressed size and resolution concerns. The clips were originally encoded at 192kbps WMV V9 at a nominal 12.5 fps at two respective resolutions 120x90 and 168x126 with Audio V8 at 32kbps. The encoding bitrate of 192kbps was chosen based on previous results [46] that showed that for 168x126 clips this resulted either in high acceptability ratings (over 70%) or that the ratings reached a plateau at this value. The two sets of clips were then encoded at the six dimensions of 480x360, 400x300, 320x240, 240x180, 168x126 and 120x90 at a higher encoding bitrate in order to ensure that the resulting clips had the same visual quality. They appeared bigger on the screen but at the same resolution. The original pixels stretched over more pixels on the display. The text contained in the news clips was legible at all above sizes.

Figure 3: Shot type examples of animation, music football and news from left to right: MCU/MS, LS, XLS, VLS

The dimensions and values for angular size, angular resolution, and viewing ratio are summarized in Table 1. All values are based on 32cm viewing distance, which was the average observed in Study 1. AR and VR are rounded values.

Table 1: Experimental setup values (at D=32cm)

Video dim. in mm	Pixels used on display	*VR*	Ang Size	AR ppd 120x90	AR ppd 168x126
60 x 45	480 x 360	7	7.4°	11	16
50 x 37.5	400 x 300	8.5	6.1°	13	19
40 x 30	320 x 240	11	4.9°	17	23
30 x 22.5	240 x 180	14	3.7°	22	31
21 x 16	168 x 126	20	2.6°	31	45
15 x 11.25	120 x 90	28	1.8°	45	45‡

‡ The resolution of this footage was only 120x90 limited by the resolution of the display.

3.2 Apparatus

The clips were presented on an iPaq hx4700 with a 200 pixel per inch (ppi) 640x480 (VGA) resolution transflective TFT display with 64k colours. For content played at its native resolution and a viewing distance of 32cm this resulted in an angular resolution of 45ppd. The sound was delivered through a set of Sony MDR-Q66LW headphones.

We checked that all clips played in the application *The Core Pocket Media Player* (TCPMP version 0.71) at their nominal frame rate using the included benchmarking tool. Each set of the six different size clips was arranged in a play list. Benchmarking videos encoded at 640x480 pixels showed that videos did not play at their nominal frame rate of 12.5 fps. The highest resolution that played at the nominal frame rate was 480x360, which was then chosen as the maximum for this study.

3.3 Procedure

The participants watched 16 clips (the order was randomized) on a couch in a lab with ambient light of 345 lux. The instructions stated that the participants could assume any position sitting on a couch and that they deemed appropriate for following mobile TV.

Each block of four clips had each content and shot type appear at least once. The presentation assured that each content type and shot type combination was used at least once as the first clip. After the first 16 clips we showed the four XLS clips in randomized order which assured that the same base clips were not played twice in a row. Each of the 20 clips (play lists) started playing at the smallest size. The participants were told to find their *favourite size* and point out which sizes they deemed *acceptable* and *unacceptable* in terms of the visual experience. They could use buttons to increase or decrease the size. On each button press, the video started over from the beginning. We encouraged and prompted the participants to explain why they found certain sizes unacceptable. Finding one's preferred size was similar to the *method of adjustment* which was successfully adopted in previous video quality research by, e.g. Richardson *et al.* [55].

We tested participants for visual acuity with a Snellen chart, and for colour-blindness with an Ishihara test. To capture participants' comments and measure viewing distances, participants were audio and video recorded. Viewing distance measures were also taken by means of a measuring stick that was occasionally held at the side of the participants, which did not seem to interfere with the participants' task.

3.4 Participants

A total of 35 paid participants (18f, 17m) with an average age of 25 took part in this study. Thirty participant had a visual acuity was 100% or better, 95% (1), 85% (1), 80% (1). Two male participants were colour-blind.

3.5 Results

For each video clip we obtained three measures - the *favourite size* at which participants preferred to watch, the *minimal size* and the *minimal angular resolution* (derived from the largest acceptable size) at which watching was still acceptable. We ran three mixed factor ANOVAs on *favourite size, minimal acceptable size* and *minimal angular resolution* as the dependent variables each with *content type* and *shot type* as *within*- and *resolution* as a *between-subjects* factor. The results are based on a total of 4200 acceptability and 700 favourite size ratings. The qualitative results are based on the 1030 comments we received. Angular sizes are reported in degrees and angular resolutions in pixels per degree (ppd).

3.5.1 Viewing distance

Only one participant systematically varied the viewing distance with the six different size videos – pulling it closer for the smaller images. All other participants generally assumed the same posture when flicking through the different sizes. When they were unsure about the acceptability of a small size clip they occasionally pulled it closer for inspection but then usually changed back into their preferred position. We averaged the viewing distances of each participant during the trial. Both the average and the median of those average viewing distances were 32cm with a standard deviation σ of 6.8cm. Although the average viewing distance in the 168x126 resolution group was slightly higher (32.7cm; σ=6cm) than in the 120x90 group (31.8cm, σ=7.6cm) a t-test showed that this difference was not significant: t(33)=-0.372, n.s.

3.5.2 Acceptability of video quality

We averaged the acceptability scores of all participants for the six different sizes in both resolution groups (depicted in Figure 4). The acceptability of the video quality varied tremendously with the size of the video. The averaged acceptability values of the video quality for both resolutions increased greatly for the larger sizes in comparison to the smallest size (picture height 11.25mm). However, the acceptability then reached a local maximum – 80% at 30mm picture height for the 120x90 resolution and 90% at 37.5mm picture height for the 168x126 resolution – after which the acceptability dropped off.

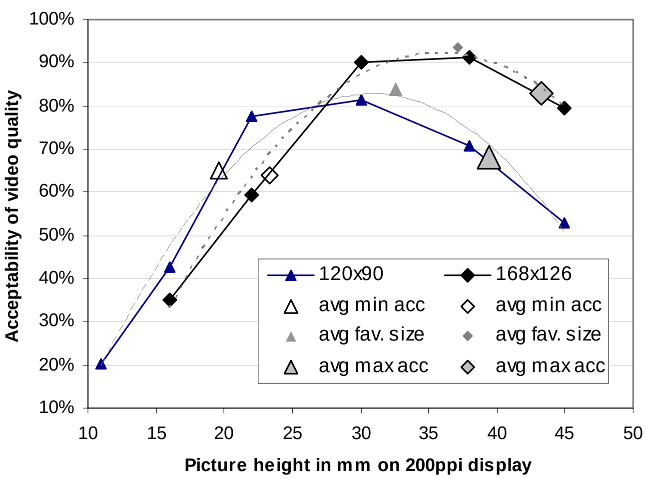

Figure 4: Video quality acceptability dependent picture height by resolution along with avg. min., favourite and max. size

The second order polynomial trend lines of the averaged acceptability scores were:

120x90: $y = -0.0016x^2 + 0.0988x - 0.6948$; ($R^2 = 0.985$),
168x126: $y = -0.0015x^2 + 0.1075x - 1.0051$; ($R^2 = 0.973$).

They result in local maxima of acceptability of video quality at a picture height of 31mm for 120x90 (VR=10.3, 16ppd) and 35.5mm for 168x126 (VR=9, 20ppd). In Figure 5 the acceptability ratings are plotted dependent on the angular resolution. For angular resolutions higher than 20ppd curves seem to differ only by a constant offset larger picture sizes resulting in higher acceptability. For viewing ratios 14 and 20 we can see that for a constant size decreasing the angular resolution resulted in higher acceptability.

3.5.3 Favourite size

The participants in the higher *resolution* group had larger *favourite sizes* ($F(1,33)=5.47$, $p<0.05$). The average favourite sizes of all participants of the two resolution groups were 32.6mm (VR=9.8, 15ppd) and 37.2mm (VR=8.6, 19ppd) – slightly larger than the computed maxima of the polynomial trend lines in Figure 4 based on the averaged acceptability results. There was a significant main effect for *content type* $F(3,99)=5.5$, $p<0.01$. The Bonferroni adjusted pair comparisons showed that this effect was due to the news content with an average *favourite size* of 33mm (VR=9.7) - significantly smaller than football and music with an average *favourite size* of 35mm (VR=9.1). No significant effect was found for *shot type*. The interaction between *content type* and *shot type* was significant ($F(9,297)=3.35$, $p<.01$) but only due to the football XLS. Participants preferred to these at 39mm (VR=8.2, 18ppd) a significantly larger size compared to the XLS of animation and news at 35mm (VR=9.1, 20ppd).

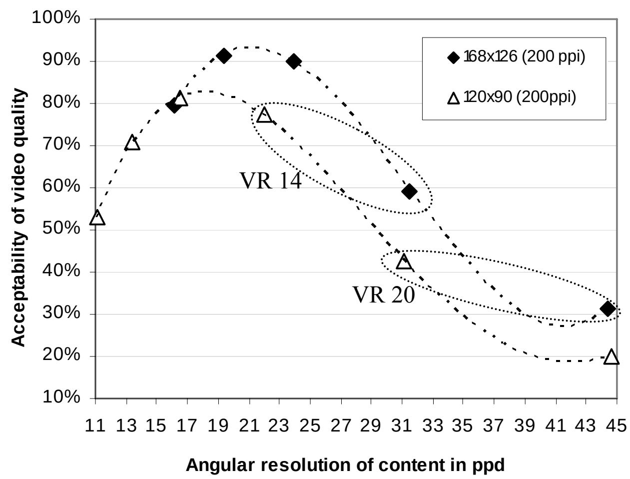

Figure 5: Acceptability of video dependent on angular resolution

3.5.4 Minimal size

Higher *resolution* content had to be presented at a larger size than lower *resolution* content in order to be acceptable (*cf.* Figure 4). For the high *resolution* video clips at 168x126 the *minimal acceptable size* was 23.4mm (VR=13.9) – significantly larger than the 19.6mm (VR=16.3) for the low *resolution* clips ($F(1,32)=7.32$ $p<0.05$). We found a significant main effect for *shot type* ($F(1,32)=40.71$, $p<0.001$). The average minimal acceptable size of the two more detailed shots was 19.5mm (VR=16.4) LS and 21mm (VR=15.2) for the MCU/MS significantly smaller than for XLS and VLS (both around 23mm, VR=13.9). An interaction effect between *shot type* and *resolution* ($F(1,288)=10.78$, $p<0.001$) (illustrated in Figure 6) showed that for the low *resolution* clips the differences between shot types as described in the main effect for *shot type* were smaller. The only difference that remained significant was the required minimal size for XLS (20.8mm) in comparison to the MCU/MS (18.2mm) for low resolution.

There was a significant effect for *content type* ($F(1,32)=7.32$ $p<0.05$) on *minimal acceptable size*. The average *minimum acceptable size* for football content was 23mm (VR=13.9) but all other *content types* were still acceptable at 21mm (VR=15.2). This was due to the football's XLS which required larger sizes than the

XLS of the other content types. Similarly, an interaction effect between *shot type* and *content type* was based on individual clip differences - the animation's VLS, a relatively dark shot, the news' LS with the presenter being occasionally occluded and the football's XLS. They all required larger sizes to be acceptable. The animation's static LS was acceptable at smaller sizes than the other LS shots.

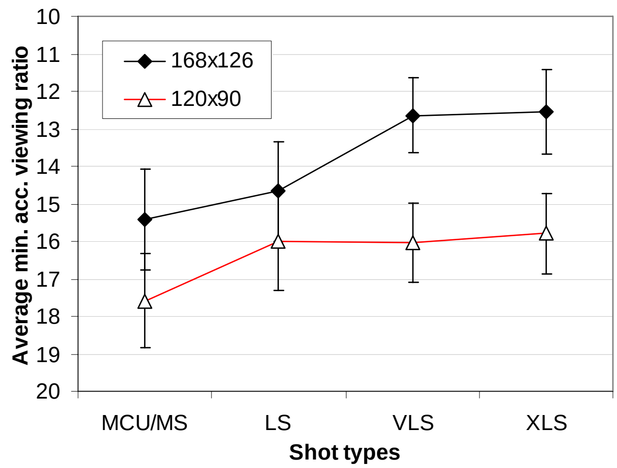

Figure 6: Interaction effect of *shot type* and *resolution* on *minimal acceptable size* of the picture

3.5.5 *Minimal angular resolution*

Resolution was the only factor that had a significant effect on the acceptable minimal angular resolution ($F(1,33)=7.05, p<.05$). The average lower bound was higher for the 168x126 group (17ppd) than for the 120x90 group (13.5ppd). We discuss the possibility of this being due to a ceiling effect in Sec 3.6.4.

3.5.6 *Qualitative results*

In terms of qualitative feedback people deemed the smaller sizes *"too small"*, *"couldn't figure out what's going on"*, "*hard to identify people*" and *"hard to look at"*. The number of these complaints (depicted in , left) dropped off once the size reached 30mm in height (VR=11). Some participants commented that although the definition seemed high - the image size was not big enough to appreciate it. With the larger image sizes, the experience was rated unacceptable because of the lack of definition or resolution. For both groups, complaints about definition started once the viewing ratio was 14 (equating to angular resolutions lower than 31ppd for the 168x126 and lower than 24ppd for the 120x90). Once the angular resolution fell below 20ppd (see , right), the number of complaints increased significantly. Lack of definition was a common complaint about text albeit to a lesser degree. With small image sizes (<22mm), participants complained about the effort required to read the text: with larger sizes and lower angular resolution (<17ppd), the quality of the text became too '*blurred*', '*pixelated*' or '*fuzzy*'. Other problems mentioned in connection with smaller images were dark scenes, insufficient contrast, and movement (either of the camera, or in the scene). For all angular resolutions lower than 24ppd (*cf.* , right) the higher resolution group (which saw a larger picture than the lower resolution group but at the same angular resolution) made more complaints about insufficient definition than the low resolution group.

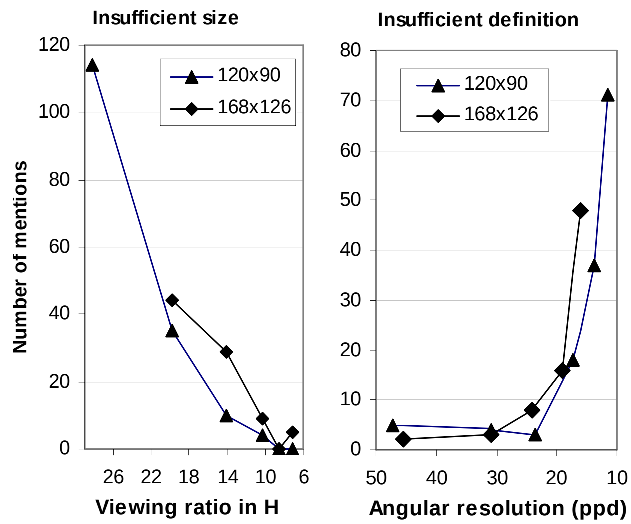

Figure 7: Participants' complaints about insufficient size (left) and insufficient definition (right)

3.6 Discussion

3.6.1 *Viewing distance*

The viewing distances observed in this study are in line with earlier research by Kato *et al.* [11]. The fact that the viewing distance observed in this trial was larger than in the previous study [46] that used longer clips of the same content could be attributed to several factors.

1) The previously reported measures were obtained by estimating viewing distances based on observational video recordings. The measurements obtained in this study are more accurate.
2) The higher resolution of the 200ppi display in comparison to the 115ppi display in the previous study. In the previous study this resulted in an angular resolution of 21ppd for all viewing ratios, which ranged from 6.8 to 13.5.
3) In the previous study the participants saw on average smaller picture sizes, had no control over the size of the clips, had to sit through the whole video and had to use a stylus to contribute their feedback. In this study they could quickly flick through and discard sizes that they did not find acceptable.
4) In this study the participants were told that they should assume a comfortable posture that they would assume if they were watching mobile TV. They were seated on a sofa not a chair with an armrest, which might have affected their posture differently.

3.6.2 *Acceptability of video quality*

When *trading off* size and definition, the acceptability of the video clips increased until the VR reached 10.6 for 120x90 at 16.5ppd and a VR of 8.7 (19.4ppd) for the 168x126 video clips. Acceptability declined, and complaints about definition increased, as angular size increased and angular resolution declined further. Our participants commented on the '*high definition*' at small image sizes, but did not try to achieve Westerink & Roufs' optimal viewing ratio with 32ppd. Although angular resolution of 32ppd was possible to attain in both groups, the resulting sizes were deemed too small. Trading off resolution below 32ppd for a gain in size increased the video acceptability. Apparently, size must play a different role for acceptability ratings, because the computed acceptability maxima were close to the favourite sizes chosen by the participants. In line with Westerink and Roufs results, complaints about resolution in the 168x126 group started after the angular resolution dropped below 31ppd (*cf.* , right). But at that point, image size was a bigger concern, and increasing it at the cost of a reduced angular resolution resulted in higher overall acceptability. The acceptability results showed that at small image sizes, lower resolution was more acceptable than higher resolution content. In practice, this suggests that - if screens are not big enough - it is counterproductive to deliver high resolution content; using lower resolution would result in higher acceptability. This has tremendous implications for service providers, who could save on bandwidth and deliver a better experience at the same time.

3.6.3 Minimal size
Shot types depicting objects from closer up could be watched at smaller image sizes. Similar to the results on *favourite size* (see next section), higher resolution required larger sizes to be acceptable. More research is required to explore the full extent of the interaction between resolution and shot types with small images. We can explain the effect of *content type* on *minimal size* by the football's XLS, which was different from all other XLS. It depicted small actors on a field that people want to be able to see. The music XLS had no actors and the actor in the animation XLS did not move and was hard to see. In the XLS of the news content the people were quite large compared to the football players. In study 2 we will look in more detail into what sizes of actors are favoured in XLS. Considering that for both resolution groups, acceptability for all participants at the averaged minimal size was around 66%, service providers would loose a large share of their potential viewers when designing close to these minimal sizes with viewing ratios of 14 and higher.

3.6.4 Minimal angular resolution
The *minimal angular resolution* depended neither on *content* nor on *shot types*. The lowest acceptable angular resolution was around 14ppd - the same for all content and shot types. For up-scaling we do not have to consider these.

The effect of *resolution* – the minimal angular resolution for 168x126 was 17ppd and for 120x90 was 13.5ppd - could be due to a ceiling effect. The 168x126 group could not select larger sizes than were available and thereby reduce the angular resolution more. The theoretical minimum at the largest size for 168x126 was 16ppd (11ppd for 120x90). The ceiling effect is also supported by the fact that the acceptability obtained from the polynomial trend line of the average maximum acceptable size (84%) (depicted in Figure 4) is much larger than the values of the three other bounds – the two on minimal size and the minimal angular resolution of 120x90 (between 63% and 71%). If we assumed that the acceptability of the maximum size should be in the same range, and used the polynomial trend line as an approximation, we would reach 66% at 49mm picture height. At this point, the angular resolution would be 10ppd. This is close to the 11ppd derived from Lund's results. He had obtained minimal viewing distances at which participants were willing to watch projected video content in a dark room [2]. Taken together this would suggest a border for angular resolution irrespective of screen size exists beyond which people's willingness to watch video content declines abruptly (*cf.* Figure 8). From our qualitative feedback we can deduct that angular resolution starts to affect video acceptability once it goes below 25ppd. In this study further reductions below 25ppd were up to a point compensated by the larger sizes.

3.6.5 Favourite size
The *favourite size* depended on the *resolution* of the content. Higher resolutions were preferably watched at larger sizes than lower resolutions. The average *favourite size* of *news* content (33mm) was smaller than of other content types. This could be rooted in perceived quality of text. People made the fewest complaints about text either being '*illegible'* or '*too hard to read'* at 30mm picture height. In a previous study smaller depictions of news had received higher acceptability scores than larger depictions despite constant angular resolution [46]. Only football's XLS was preferred at significantly larger sizes than the XLS of other content types. The XLS depicted a far away pitch in which actors were only 12 pixels in height in the original footage. At the *preferred size* the actors were about 0.7º tall. The fact that we found significant interactions between content type and shot type could stem from potential confounding factors. The qualitative feedback suggested an influence of dark scenes, text, camera movement and the presence or absence of actors.

In Figure 8, we have collated the preferred (PVD) and minimal viewing distances (VD) from the studies by Lund, Ardito *et al.* and Nathan *et al.* and plotted them in terms of the resulting VR and angular resolution. Results obtained in dark rooms are marked with shadows. The assumed lower limit of angular resolution is marked with a dotted black line. Our results were based on preferred viewing sizes (PVS) all others on PVD (except for Lund-2 which was based on a minimal acceptable viewing distance). People are willing to watch video content across a large range of both sizes and resolutions. Before this study people had only chosen viewing configurations with angular resolutions lower than 22ppd (the minimum observed by Lund on a 7" screen presenting Q-NTSC resolution video) in darkened rooms with large picture sizes (VR≤ 4).

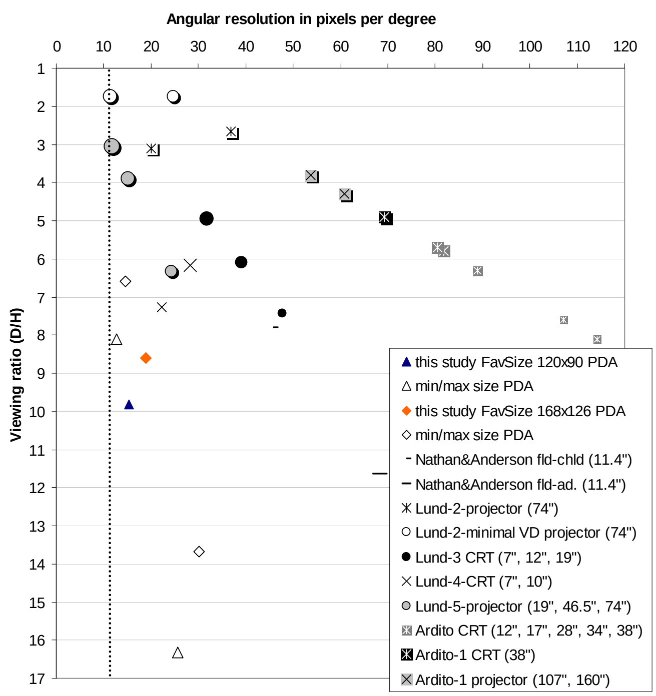

Figure 8: Comparison of results obtained by Lund, Ardito *et al.* and Nathan *et al.* on PVD with our results on preferred size

4. STUDY 2: ZOOMING INTO XLS

Study 2 was devised to better understand the contribution of actor size and overall size on the acceptability of XLS on mobile devices. We used four video clips, which had been produced in the context of a previous study that had looked at preferred zooms in content adaptation to mobile devices [51].

4.1 Material

Two base XLS clips depicted football at two distances (the size of the depicted actors was different) and the other two clips were 1.6 times zoomed in versions of the base clips. The zoomed clips did not show all that was visible in the base clips but showed the content of a moving zoom window at a larger size (see Figure 9 for an example). For further details on the preparation of the zoomed material consult [51].

Figure 9: Zoomed (left) and non-zoomed material (right) with a zoom of 1.6 at 176x144

This provided us with four different sizes of actors in the footage: 11, 15, 18 and 24 pixels in height in the original resolution 168x126. This would allow us to find out whether participants' preferences in terms of preferred size are due to the absolute size of the clips or depicted objects within the video clips. We prepared these four clips at the same six dimensions (see Table 1) but only one resolution 168x126 and at its original encoding bitrate - 350kbps WMV V9 at 12.5 fps and WMA V8 at 32kbps.

4.2 Participants & Procedure

The same participants as in study 1 watched these clips as a second session. The procedure was identical.

4.3 Results

We followed the same approach in the analysis as study 1. We averaged the acceptability scores of all participants at all picture heights for the four clips to obtain the curves presented in Figure 10. XLS clips depicting actors that were larger in size – either through zoom or the fact that the original scene was closer to the players - were generally more acceptable at all sizes smaller than 37.5*mm* (VR>8.5). Once the viewing ratio reached around 8 the benefits of the zooms diminished – the four clips' acceptability scores are at similar levels. At viewing ratios larger than 14 even the clip with the largest depictions of actors achieved only an acceptability of 60%. The acceptability dropped off for viewing ratios larger than 11.

The acceptability of all four clips reached its maximum at the two largest sizes. This means that the measures *favourite size* and *minimal angular resolution* might be subject to ceiling effects.

We ran *repeated measures one factor* ANOVAs on *favourite size*, *minimal* size and *minimal angular resolution* with *actor size* as the sole factor. As expected, clips with smaller *actor sizes* required higher *minimal acceptable sizes* than those with larger depictions – a significant linear effect ($F(3,102)=13.58$, $p<.001$). The angular size of the depicted actors for this lower bound ranged from 0.5º to 0.8º. Analogously, the clips with smaller depictions of actors also yielded a higher *favourite size* than the clips with larger actor sizes ($F(3,102)=8.54$, $p<.001$). The participants favoured to watch the clip with the smallest actors (12pixel) at a picture height 42.5mm. The clips in which the actors were twice as large (24pixel) had preferred height of 38.5mm. The favourite angular size of an actor in XLS varied between 0.7º and 1.3º. There was no significant effect of *actor size* on the *minimal angular resolution*.

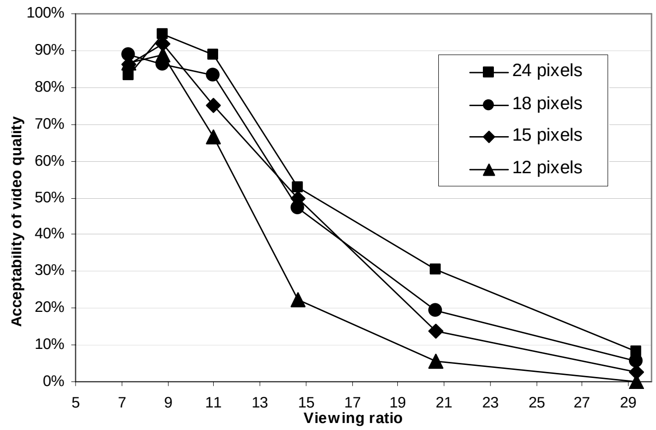

Figure 10: The acceptability of sports XLS clips at 350*kbps* by *picture height* for different *actor sizes*

Most qualitative complaints were about size and insufficient definition. As in study 1 once the viewing ratio was 8.5 there were no complaints about insufficient size. Although complaints about definition started once the angular resolution had dropped below 31ppd there were much fewer complaints than in study 1. For the clip with the smallest actor size the participants started complaining about a lack of clarity only once the angular resolution was as low as 16ppd. We received a number of comments from the participants who remarked that the quality of these football clips were higher than the ones they had seen in the previous 16 clips. The zoomed clips were originally prepared at 350kbps and the clips that were shown in study 1 at 192kbps only.

4.4 Discussion

The value of zooming diminished once the viewing ratio of the whole picture reached 8. For viewing ratios of 14 and larger there was still a large benefit for zooming but the overall acceptability (60%) was low. Viewing ratios between and a maximum of 11 should result in the best experience of XLS. The participants' favourite sizes resulted in depiction of actors between 0.7º and 1.3º much larger than in a previous study [51] in which the preferred size of actors was between 0.5º (176x144) and 0.7º (320x240). However, in that study increased actor sizes had to be traded off for a reduction in visual context by cropping the picture while the angular resolution of the picture was constant (35ppd). In the study at hand increasing the size of players did not reduce context - only the angular resolution of the depicted video. In this study the participants chose *angular sizes* of actors in XLS between 0.5º and 0.7º as the *minimal acceptable size*. But their favourite size yielded *angular sizes* of actors between 0.7º and 1.3º. To achieve these sizes they were willing to watch the footage at an angular resolution between 19ppd and 17ppd. For other shot types in study 1 this would have already resulted in a reduced acceptability due to the insufficient definition of the picture.

5. STUDY 3: PERCEIVED QUALITY ON THE MOVE

In this study we evaluate the ecological validity of results obtained from two previous lab-based studies [46] and [48] by having participants watch and rate the same clips while travelling on public transport (London Underground).

5.1 Design

The experimental design followed the one used in [46] and [48]. We ran two groups: each group of 16 participants viewed 16 clips in groups of four, at each of the four sizes. The groups differed in whether they experienced *increasing* or *decreasing* image sizes.

Table 2: Experimental design

Group	Size Order	Image size	Content Clip			
A (16)	Decreasing	240x180	N1	S1	M1	A1
		208x156	N2	S2	M2	A2
		168x126	N3	S3	M3	A3
		120x90	N4	S4	M4	A4
B (16)	Increasing	120x90	N1	S1	M1	A1
		168x126	N2	S2	M2	A2
		208x156	N3	S3	M3	A3
		240x180	N4	S4	M4	A4

Within each group, we ran eight variations to control for content using a Latin squares design. This ensured that the different content clips were tested at each of the image sizes across participants.

5.2 Material

The video clips were encoded at four resolutions (240x180, 208x156, 168x126 and 120x90). Within each clip, the bitrate allocated to video was degraded from a maximum of 224kbps down to 32kbps, every 20 seconds, in steps of 32 kbps,. The boundaries of these intervals were not pointed out to the participants - they were told that the quality would vary over time. Participants watched with 16video clips, each of which gradually decreased in quality as descrived above, and audio was encoded at 32kbps in stereo (WMV V9). (A more detailed description of process of producing the video clips can be found in [46], and for the news clips in [48].) The text included in the news ticker and inserts was legible at all for sizes.

Table 3: Image sizes used on PDA

Video dim. in mm	Pixels used on display	VR	Ang. Size	Ang. res. (ppd)
53 x 40	240 x 180	10	5.7 º	31
46 x 34.5	208 x 156	12	4.9 º	31
37 x 28	168 x 126	14	4 º	31
26.5 x 20	120 x 90	20	2.9 º	31

5.3 Equipment

The test material was presented on an iPAQ 2210 with a 400Mhz X-scale processor, 64MB of RAM and a 512MB SD card. The screen was a 115ppi transflective TFT display with 64k colours and a resolution of 240x320. The iPAQ was equipped with a set of Sony MDR-Q66LW headphones to deliver the audio. We used the same interface as in the previous studies [46] and [48] - a customized application in C# using the Odyssey CFCOM software [56] to embed the Windows Media Player. It presented the clips along with a volume control and two response buttons to indicate acceptable and unacceptable quality. The participants could switch back and forth between these two states with little effort. The program recorded at what time in which clip a participant clicked acceptable or unacceptable.

5.4 Participants

Most of the 32 paid participants (11 women and 21 men, aged 20 to 65 with a median of 28 years) were university students. The majority came from the UK (20). English was the first language for 28 of the participants. Visual acuity was 100% or higher for 24, 95% for six, 90% for one and 85% for one of the participants.

5.5 Procedure

Before boarding the London Underground trains, participants were instructed by the experimenter, who accompanied them throughout. The participants were told that a technology consortium was investigating ways to deliver TV content to mobile devices, and that they wanted to find out the minimum acceptable video quality for watching news. The instructions stated: *"If you are watching a sequence and you find that the video quality unacceptable at any time, please click the button labelled 'Unacc'. When you continue watching the clips and you find that the quality has become acceptable again, please click the button labelled 'Acc'… you can hold the PDA at any distance that is comfortable for you."* Each clip started with the interface in the 'Acc.' state.

The participants watched eight clips on the outbound journey, and another eight clips on the return train. The train journeys included both underground and over-ground segments. We video-recorded all participants while they were watching the clips. As in the lab experiment the session concluded with a debrief interview about what aspects of the video quality they had found unacceptable. Additionally, we asked whether they had had any specific problems watching while riding on a train.

5.6 Results

We combined the data obtained in this experiment with data from two previous lab studies – from 64 participants from [46] and 32 participants [48]. The results were analyzed based on each 20-second bitrate segment. If a participant judged the quality unacceptable at any time during a segment, it was classified as unacceptable. We used a binary logistic regression to test for main effects and interactions between the independent variables of the previous studies – *Image Size*, *Video Bitrate* and *Content Type* and *Context*. Context denoted whether the data was obtained in the lab or the train. Control variables *Gender, isNativeSpeaker* and *Size Order* were included in the analysis.

As in the two previous studies significant predictors of *acceptability* were *Image Size* [$\chi^2(1)=221.1, p<0.001$], *Video Bitrate* [$\chi^2(1)=16.7$, $p<0.001$] and *Content Type* [$\chi^2(3)=1027.9.$, $p<.001$]. Larger *Image Sizes* and higher *Video Bitrates* resulted in higher *acceptability*. But at the lowest *Video Bitrate,* the benefits of larger *Image Sizes* diminished. t *Context* was a significant predictor of *acceptability* [$\chi^2(1)=20.6$, $p<.001$] – the participants viewing the clips on the trains rated them more acceptable than those viewing in the lab, but there is an interaction of *Context* with *Image Size* [$\chi^2(1)=16.4$, $p<0.001$]. For the smaller *Image Sizes,* there was no significant difference between the lab and the train ratings, but a non-parametric Kruskal-Wallis test [$\chi_2(1)=24.56$, $p<.001$] showed that the participants on the train found the larger two sizes more acceptable than the participants in the lab. This finding is depicted in Figure 11.

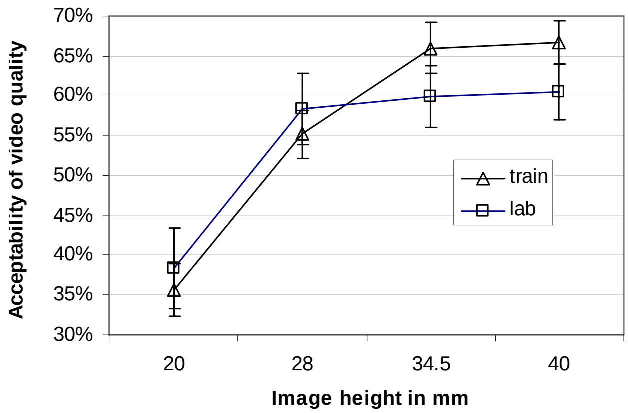

Figure 11: The interaction of *image size* and *context* at a constant angular resolution of approx. 31ppd (192kbps video)

The interaction of *Context* with *Video Bitrate* was another significant predictor [$\chi_2(1)=20.2$, $p<.001$] of *acceptability.* At high *Video Bitrates,* there was no difference between lab and train but for low *Video Bitrates* (<96kbps) the participants on the train found the video quality more acceptable than the participants in the lab. Figure 12 illustrates this interaction. The regression revealed significant effects on all of the control variables:

1. Female participants generally rated the video quality more acceptable than men [$\chi_2(1)=17.1$, $P<.001$],
2. non-native speakers more [$\chi_2(1)=8.6$, $P<0.001$] than native speakers and
3. the people whose clips increased in size more than those whose clips decreased in size [$\chi_2(1)=119.9$, $p<.001$].

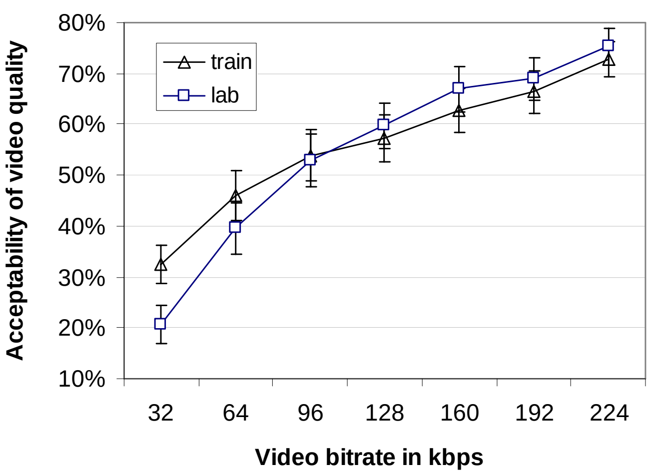

Figure 12: The interaction of *video bitrate* and *context*

5.7 Discussion

The acceptability ratings for video quality were generally lower than those obtained on the train. This is in line with results of Jumisko-Pyykkö *et al.*, whose participants rated the audio-visual quality of clips impaired by packet loss consistently higher in the three contexts in the field (bus, train station, cafe) compared to the lab. The difference was most pronounced at the lowest quality – the highest loss ratio [54]. We found the same to be the case for low encoding bitrates. For service providers delivering content in medium to high video quality, this means that previous lab results provide them with conservative estimates of what the levels of quality their customers find acceptable when viewing on the move.

In terms of the size requirements the story was different. Our results showed that on the train, the larger sizes (208x156 and 240x180) yielded a higher acceptability than in the lab. The acceptability of depictions smaller than 34.5mm resulted in equally reduced experiences both in the lab and the train. Further research is required to find the reason behind this.

6. OVERALL DISCUSSION

In order to compare the results from Study 1 with the train results from Study 2 and the lab results from [46] and [48], we weighted the acceptability scores of the shot types of Study 1 and 2 according to their relative occurrence in the footage in [46] and [48]. Figure 13 collates these results by picture height.

The acceptability results of 120x90 and 168x126 clips in the previous lab studies on a 115ppi device were lower than those obtained in this study (200ppi) but follow the same trend. The discrepancy could stem from the difference in display resolution (115ppi *vs.* 200ppi), luminance and contrast, the experimental procedures (stylus *vs.* buttons) and that viewing distance was not accurately controlled for in earlier studies. The acceptability ratings of the clips at 208x156 and 240x180 from the previous lab study are below those of the 168x126 clips of this study in line with what the results in Sec. 3.5.2 would make us expect. If these higher resolution clips were increased in size they should surpass the acceptability ratings of 168x126 assuming that the bitrate of 192kbps can sufficiently encode the spatial information. This is a general limitation to our results on angular resolution requirements. We cannot know the exact resolution of the content due to the spatio-temporal compression of 192kbps.

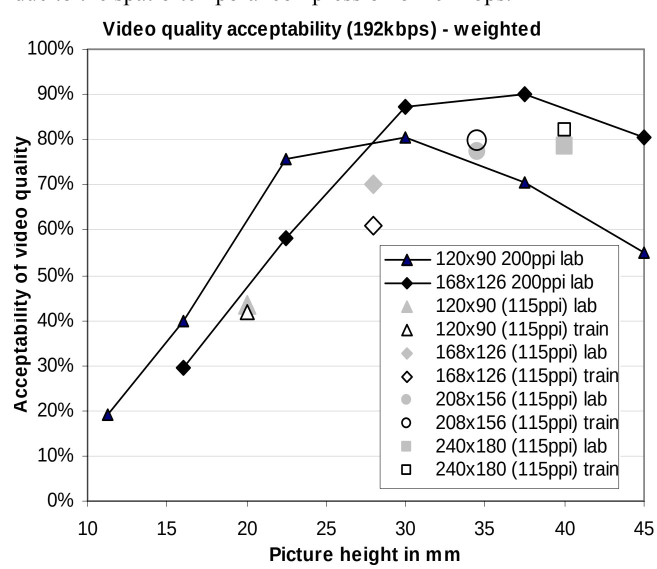

Figure 13: Video acceptability of Study 1, 2 and the combined lab results of [46] and [48] (in grey) in relation to viewing ratio

Since the 120x90 content only reached a maximum of 80% acceptability at the favourite size we can assume that it is too low to satisfy the entire market. Conservative service providers should deliver content at QCIF resolution as a minimum and match the resolution with screen heights of 4cm and larger. Observations from industry confirm this finding on size. According to Strategy Analytics [57], Samsung stated that displays of their first mobile TV phones (33mm in height; a VR of 10.6 at 35cm) were probably too small, and Nokia and Telia Sonera found that usage rates almost doubled with a screen diagonal larger than 7.6cm (a VR of 7.6 at 35cm).

Our participants preferred to watch low-resolution content at viewing ratios that were much larger in picture size then the ITU recommendations for evaluating video quality in these settings. We have plotted their recommended values in Figure 14 along with the proposed preferred viewing ratios based on our results on preferred viewing size and the results of Lund and Ardito of preferred viewing distance.

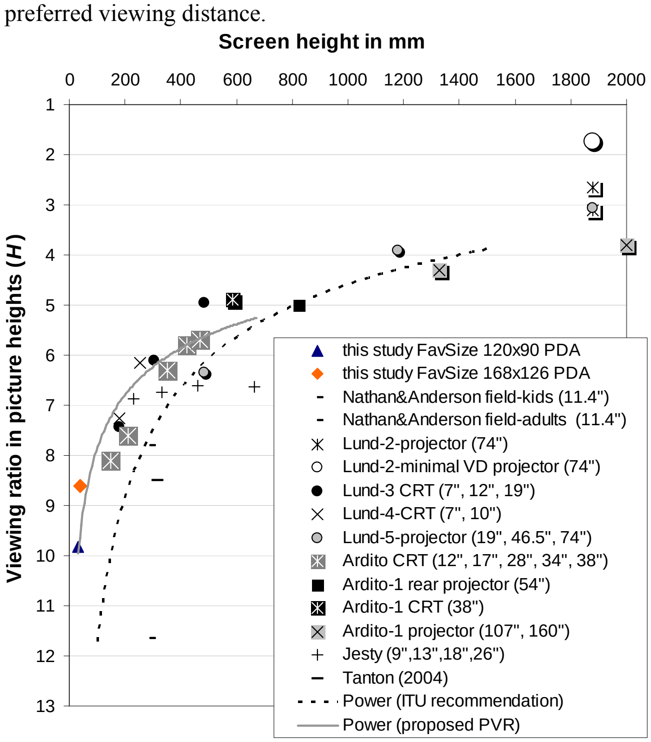

Figure 14: Viewing ratios in relation to screen size. All results based on PVD apart from our results (PVS)

Research on video quality often assumes that – people prefer to view video at the highest quality available. Westerink & Roufs suggested that people would chose their viewing distances irrespective of picture width, but to attain the best subjective quality an angular resolution of 16 cycles per degree (32ppd). This approach was based on people providing ratings on pictures of different sizes and resolution at different viewing distances. People were not asked to *choose* their preferred viewing distance.

The results presented in this paper show that participants' preferences for watching low-resolution content depends mainly on size – depending on the content's resolution they preferred viewing ratios between 8.5 and 10 which resulted in an angular resolution between 19ppd and 15ppd. From the complaints about insufficient definition in study 1 we learned that angular resolution became a concern once the viewing ratio was at least 14 or smaller. The acceptability of QCIF content will drop off for angular resolution below 20ppd. Between this threshold and 32ppd (Westerink & Roufs optimum) the acceptability of QCIF content presented on mobile devices can still be improved by increasing its size.

7. CONCLUSIONS

We wanted to find out preferred viewing ratios on mobile devices and how people trade off size, angular resolution and viewing distance as a result. Our results bear the following recommendations for service providers. Mobile TV services should be designed for close *viewing distances* - between 25cm to 50cm (a distance of 32cm was the average chosen in our study). Like [46] we found no adjustment of viewing distance depending on the resolution or the size of the footage. As a rule of thumb, service providers should target a minimal resolution about QCIF (176x144). Lower resolutions might not result in a wholly acceptable TV viewing experience on mobile devices. Most importantly the video needs to be displayed at an adequate size. People prefer to achieve a living room TV viewing ratio of 8 for QCIF content. A picture height of 4cm should result in the most acceptable experience of 4:3 QCIF content encoded at a comparable bitrate as in study 1. The angular resolution would be around 20ppd.

Apart from XLS, *shot types* were only a concern at the lower limits of acceptable size. MCU and MS could still be presented at smaller sizes than other shot types but their *favourite sizes* did not differ from other shot types. To rely on them in production would only make sense for content that would be shown on displays smaller than 22mm in height (VR>14) – the train results from Study 3, however, indicated that this would be too small. The acceptability gains for XLS by zooming were substantial for viewing ratios larger than 8.5. Content adaptation should focus on improving XLS shown at viewing ratios between 8 and 11 and target angular actor sizes of 1º. The zoom factors that need to be achieved should be between 1.1 and 1.5. Up-scaling the picture on the mobile device can be used to help achieve these sizes – if possible on the device - down to an angular resolution of 17ppd for QCIF content. Both size and the available resolution of the content have to be taken into account for the most preferred presentation of mobile TV material. Size is more important especially until a viewing ratio of at least 14 or smaller can be provided. If content of relatively high resolution is not depicted at a sufficient height this will result in lower acceptability. A general limit for up-scaling video clips regardless of content and shot types was a resulting angular resolution of about 14ppd close to the 11ppd we derived from Lund's [2] results on minimum viewing distances of large projections of TV content in a dark room.

Here are the conclusions for researchers of objective and subjective video quality and multimedia models. Current ITU recommendations on video quality assessment are suggesting viewing ratios for small screens that are much smaller and result in a poorer overall experience. We believe that services and video quality should be evaluated under conditions that resemble people's preferences. Sizes that yielded an angular resolution of 32ppd – identified as optimal picture quality in [1] - did not coincide with the participants' favourite sizes but were criticized for being too small. Acceptability ratings, however, were a good predictor of participants' favourite viewing conditions. Measures of video quality might be misguiding as indicators for people's preference and the quality of experience of a given service. Objective quality measurements and multimedia models for video content on mobile devices that do not consider the viewing ratio on a target device will make predictions that will not match people's preferences.

Lab trials should be a good approximation of home viewing conditions – between 30%-50% of field trial participants in [58], [59] used their devices at home as a "personal TV" [6]. Although our research showed that lab experiments may be a conservative estimate of acceptability of video quality consumed by people on the move, this was not true for all observed factors. It is important to test preferences in different contexts of use, especially for effects that are not fully understood yet – as in our case image size. Conducting tests of acceptability and user preferences to validate and qualify the results from laboratory results is an essential part of building our understanding of Quality of Experience of multimedia consumption. Our results were obtained on trains that induced motion, varying ambient lighting. There are many other conditions that can occur in the field which might bear different results – last but not least people only used the devices by themselves although sharing mobile devices for multimedia consumption is an option.

ACKNOWLEDGMENTS

This work was supported by the IST-2005-27034 UNIversal satellite home Connection (UNIC) project.

REFERENCES

[1] Westerink, J. H. & Roufs, J. A. *Subjective Image Quality as a Function of Viewing Distance, Resolution, and Picture Size* SMPTE Journal, 98, p.113-119, Feb1989.

[2] Lund, A. M. *The Influence of Video Image Size and Resolution on Viewing-Distance Preference* SMPTE, 102, p.406-415, May 1993.

[3] Nathan, J. G., Anderson, D. R., Field, D. E., Collins, P. *Television viewing at home: Distances and visual angles of children and adults* Human Factors, 27, p.467-476, 1985.

[4] Poynton, C. *Digital Video and HDTV Algorithms and Interfaces* San Francisco, CA, USA: Morgan Kaufmann, 2003

[5] Knoche, H. & McCarthy, J. *Mobile Users' Needs and Expectations of Future Multimedia Services* in Proceedings of the WWRF12, 2004

[6] Chipchase, J., Yanqing, C., Jung, Y. *Personal Television: A Qualitative Study of Mobile TV Users in South Korea*, 2007.

[7] Luther, A. C. *Principles of Digital Audio and Video* Boston, London: Artech House Publishers, 1996

[8] Owens, D. A. & Wolfe-Kelly, K. *Near Work, Visual Fatigue, and Variations of Oculomotor Tonus* Investigative Ophthalmology and Visual Science, 28, p.743-749, 1987.

[9] Ankrum, D. R. *Viewing Distance at Computer Workstations* Work Place Ergonomics, p.10-12, Sep/Oct 1996.

[10] Boff, K. R. & Lincoln, J. E. *Visual Acuity: Effect of Viewing Distance and Luminance Level*. in Engineering Data Compendium, Human Perception and Performance (Ohio: AAMRL: Wright-Patterson AFB 1988

[11] Kato, S., Boon, C. S., Fujibayashi, A., Hangai, S., Hamamoto, T. *Perceptual Quality of Motion of Video Sequences on Mobile Terminals* in Proceedings of the IASTED International Conference, p.442-447, 2005

[12] Thompson, F. T. *Television line structure suppression* SMPTE, 66, p.603-606, 1957.

[13] Diamant, L. *The Broadcast Communications Dictionary* (3rd ed.) Greenwood Press, 1989

[14] Tanton, N. E. *Results of a survey on television viewing distance* (WHP 090.) Britisch Broadcasting Corporation, 2004.

[15] Yuyama, I. *Fundamental Requirements for High-Definition Television*. in High-Definition Television NHK Technical Monograph (NHK Advanced broadcasting systems research division 2008

[16] Ardito, M. *Studies on the Influence of Display Size and Picture Brightness on the Preferred Viewing Distance for HDTV* SMPTE, 103 (8) 1994.

[17] Ardito, M., Gunetti, M., Visca, M. *Influence of display parameters on perceived HDTV quality* IEEE Transactions on Consumer Electronics, 42 (1), p.145-155, Feb.1996.

[18] ITU-T *Subjective audiovisual quality assessment methods for multimedia applications* (ITU-T Recommendation P.911) , 1999.

[19] ITU-R *Methodology for the subjective assessment of the quality of television pictures.* (BT.500-11) , 2004.

[20] Jesty, L. C. *The relation between picture size, viewing distance and picture quality* in Proc.Inst.Electr.Eng., p.425-439, 1958

[21] Ribchester, E. *Discussion Before th Radio and Telecommunication Section, 19th February 1958* Proc.Institution of Electrical Engineering, 105B, p.437, 1958.

[22] Sugama, Y., Yoshida, T., Hamamoto, T., Hangai, S., Seng, B. C., Kato, S. *A Comparison of Subjective Picture Quality with Objective Measure Using Subjective Spatial Frequency* in Proc.of ICME, p.1262-1265, 2005

[23] Hatada, T., Sakata, H., Kusaka, H. *Psychophysical analysis of the 'sensation of reality' induced by a visual wide-field display* SMPTE, 89, p.560-569, 1980.

[24] Yu, Z., Wu, H. R., Ferguson, T. *The Influence of Viewing Distance on Subjective Impairment Assessment* IEEE Transactions on Broadcasting, 48 (4), p.331-336, December2002.

[25] Reeves, B., Detenber, B. H., Steuer, J. *New Televisions: The Effects of Big Pictures and Big Sound On Viewer Responses to the Screen* in 1993

[26] Barten, P. G. J. *Effect of Picture Size and Definition on Perceived Image Quality* IEEE Trans.on Electron Devices, 36 (9), p.1865-1869, Sept.1989.

[27] Barten, P. G. J. *Evaluation of subjective image quality with the square-root integral method* Journal of Optical Society of America, 7 (10), p.2024-2031, October1990.

[28] Thompson, R. *Grammar of the shot* Elsevier Focal Press, 1998

[29] Guardian *Romantic drama in China soap opera only for mobile phones*. Available at: http://www.buzzle.com/editorials/6-28-2005-72274.asp, 2005.

[30] Gwinn, E. & Hughlett, M. *Mobile TV for your cell phone*. Available at: http://home.hamptonroads.com/stories/story.cfm?story=93423&ran=38197, 2005.

[31] Hands, D. S. *A Basic Multimedia Quality Model* IEEE Transactions on Multimedia, 6 (6), p.806-816, Dec. 2004.

[32] Frieser, H. & Biedermann, K. *Experiments on image quality in relation to the modulation transfer function and graininess of photographs* Photographic Science andEngineering, 7 (28) 1963.

[33] Kingslake, R. *Lenses on Photography* New York, NY: Barnes & Co, 1963

[34] Corey, G. P., Clayton, M. J., Cupery, K. N. *Scene Dependence of image quality* Society of Photographic Scientists and Engineers, 27 (1), p.9-13, Jan-Feb1983.

[35] Kopf, S., Lampi, F., King, T., Effelsberg, W. *Automatic Scaling and Cropping of Videos for Devices with Limited Screen Resolution* in Proceedings of the 14th annual ACM international conference on Multimedia, p.957-958, 2006

[36] Sinha, A. & Agarwal, G. *A method of dynamic cropping and resizing of video frames in DVB-H to Mobile* in GPSx 2005, 2005

[37] Seo, K., Ko, J., n, I., im, C. *An Intelligent Display Scheme of Soccer Video On Mobile Devices* IEEE transactions on circuits and systems for video technology (CSVT), 17 (1), p.1395-1401, October 20072007.

[38] Tang, J. C. & Isaacs, E. A. *Why Do Users like Video? Studies of Multimedia-Supported Collaboration* Computer Supported Cooperative Work: An International Journal, 1 (3), p.163-196, 1993.

[39] Chuang, S. L. & Haines, R. F. *A Study of Video Frame Rate on the Perception of Compressed Dynamic Imagery* in Society for Information Display, 1993 International Symposium, 1993

[40] Masoodian, M., Apperley, M., Frederikson, L. *Video support for shared workspace interaction: an empirical study* Interacting with Computers, 7 ((3)), p.237-253, 1995.

[41] Assfalg, J., Bertini, M., Colombo, C., Del Bimbo, A. *Automatic extraction and annotation of soccer video highlights* in Image Processing, 2003.Proceedings, 2003

[42] Apteker, R. T., Fisher, A. A., Kisimov, V. S., Neishlos, H. *Distributed multimedia: user perception and dynamic QoS* in Proceedings of SPIE, p.226-234, 1994

[43] Ghinea, G. & Thomas, J. P. *QoS Impact on User Perception and Understanding of Multimedia Video Clips* in Proceedings of ACM Multimedia '98, 1998

[44] McCarthy, J., Sasse, M. A., Miras, D. *Sharp or smooth? Comparing the effects of quantization vs. frame rate for streamed video* in Proc.CHI, p.535-542, 2004

[45] Södergård, C. *Mobile television - technology and user experiences Report on the Mobile-TV project* (P506) VTT Information Technology, 2003.

[46] Knoche, H., McCarthy, J., Sasse, M. A. *Can Small Be Beautiful? Assessing Image Resolution Requirements for Mobile TV* in Proceedings of ACM Multimedia 2005, p.829-838, ACM, 2005

[47] Winkler, S. & Faller, C. *Maximizing audiovisual quality at low bitrates* in Proc.of Workshop on Video Processing and Quality Metrics, 2005

[48] Knoche, H., McCarthy, J., Sasse, M. A. *Reading the Fine Print: The Effect of Text Legibility on Perceived Video Quality in Mobile TV.* in Proceedings of ACM Multimedia, p.727-730, 2006

[49] Knoche, H., McCarthy, J., Sasse, M. A. *A close-up on Mobile TV: The effect of low resolutions on shot types* in Proc.of EuroITV '06 - Beyond Usability, Broadcast and TV, p.359-367, 2006

[50] Knoche, H., McCarthy, J., Sasse, M. A. *A close-up on Mobile TV: The effect of low resolutions on shot types* in Proc.of EuroITV '06, 2006

[51] Knoche, H., Papaleo, M., Sasse, M. A., Vanelli-Coralli, A. *The Kindest Cut: Enhancing the User Experience of Mobile TV through Adequate Zooming* in Proceedings of ACM Multimedia 2007, p.87-96, ACM, 2007

[52] Tamminen, S., Oulasvirta, A., Toiskallio, K., Kankainen, A. *Understanding mobile contexts.* Special Issue of Journal of Personal and Ubiquitous Computing, 8 (135), p.143, 2004.

[53] Sasse, M. A. & Knoche, H. *Quality in Context - an ecological approach to assessing QoS for mobile TV* in Proceedings of 2nd ISCA/DEGA Tutorial & Research Workshop on Perceptual Quality of Systems, 2006

[54] Jumisko-Pyykkö, S. & Hannuksela, M. *Does Context Matter in Quality Evaluation of Mobile Television* in Proc.of mobile HCI, 2008

[55] Richardson, I. E. G. & Kannangara, C. S. *Fast subjective video quality measurement with user feedback* Electronics Letters, 40 (13), p.799-801, 2004.

[56] *Odyssey software inc. CFCOM*, 2003.

[57] Strategy Analytics *TV Phones: Integration and Power Improvements Needed to Reach 100 Million Sales*, 2006.

[58] Mason, S. *Mobile TV - results from the BT Movio DAB-IP trial in Oxford* EBU Technical Review, 2006.

[59] Lloyd, E., Maclean, R., Stirling, A. *Mobile TV - results from the BT Movio DAB-IP pilot in London* (Technical Review) EBU, 2006.