Good News for Mobile TV

Knoche, Hendrik; McCarthy, John D.

Published in:
Proceedings of WWRF14

Publication date:
2005

WIRELESS WORLD
RESEARCH FORUM

Good News for mobile TV

Hendrik O. Knoche, John D. McCarthy

***Abstract*— Previous studies have found mobile TV usage durations to be shorter than 10 minutes [1]. News programmes fit this requirement and were identified by mobile phone users in focus groups and surveys as the most interesting content type [2]. This paper presents two studies that investigate human requirements on the delivery and presentation of TV news footage at different video and audio encoding bitrates. The effects of resolution, picture size, different mobile presentation devices, encoding schemes and audio on the news content are presented. To ensure validity both studies were carried out on mobile devices. Study 1 examined video and audio encoding bitrates on a 3G phone and an IPAQ PDA whereas study 2 looked at various resolutions and picture sizes ranging from 120x90 to 240x180 on an IPAQ.**

***Index Terms*— Mobile TV, image resolution, image size, text in videos**

INTRODUCTION

NEWS was identified by users in focus groups and surveys as the most interesting content type for mobile consumption. [2].The typical length of a News bulletin also fits well with observed use of mobile TV in bursts of less than ten minutes. However, to ensure an optimal supply of news content to a mobile audience the following questions need to be addressed:

1. Is it possible to repurpose standard TV news without reediting?
2. How does the reduced resolution of the mobile screen affect the acceptability of the service?
3. How do different encoding bitrates of both audio and video affect the perceived quality?
4. What price are people willing to pay for this news service?

The answers to these questions have ramifications for both the production and the delivery of news.

Since is not clear whether the audience of mobile TV will support the cost involved in editing content especially for mobile TV the simplest and cheapest solution would be to deliver TV material without additional editing. However, little is known about the technical requirements to deliver content at an acceptable quality to a mobile audience.

The visual resolution of the content affects the required bandwidth for transmission and the amount of visual detail that can be displayed, and consequently limits the content that can be presented. Sensible resolutions for the content depend on the resolution of the target device, the physical size of the picture and the viewing distance which – if too close – may contribute to eyestrain [3]. The portability of mobile devices is also a major concern of mobile users and violating this for the sake of bigger screens may inhibit uptake of mobile TV [2].

Lowering the video resolution might have two opposing effects. A reduction in the amount of pixels lowers the amount of detail that can be represented in the video. On the other hand a reduced resolution increases the encoding bitrate per pixel and might therefore raise the clarity of the presented video. The dynamics of this trade-off have not been studied in previous research. Whether people reduce their viewing distance to counter small screen sizes and small resolutions over a prolonged period also remains an open question.

With reduced screen sizes screen clutter could become an issue for mobile users at a much earlier point than it does at traditional TV resolutions where screen clutter has been shown to impede attention and comprehension [4]. Another factor is text. Text is an important tool in the presentation of news and text sizes that work on standard television sets might pose a problem for smaller resolutions and screen sizes.

Effects of Size and Resolution

From an economic standpoint, videos with lower resolutions require less bandwidth and hence reduce the price of delivery. It is clear that image resolution and the image size of a viewing device cannot be reduced indefinitely as important detail will be lost.

As the size of the display in the viewer's visual field depends on both the size of the screen and the distance between viewer and the screen, the viewing ratio (VR) is defined as the viewing distance divided by the picture height (H). When planning the production of their programs, television producers consider the typical resolutions of TV sets, their sizes and the seating distance at which the audience is sitting [5]. In a mobile context, some factors restrict the range of possible combinations though. Mobile devices are typically operated at 'arm's length' and continued viewing at distances closer than the resting point of vergence – approx. 89cm, with a 30º downward gaze – can contribute to eyestrain [3].

Previous research has examined the impact of increasing the image size in the viewer's visual field by means of large physical displays or projection areas. Typically these studies have compared very large size screens (e.g. 46") to standard sized TV screens (15-20") [6] [7]. The results show that larger image sizes are more arousing than smaller ones, better remembered, and better liked. Other studies also show that users generally prefer bigger image sizes – ideally depicting people and objects up to life-size [8].

When it comes to TV images, the general message from these studies is, '*the bigger the better'*. This clearly presents a challenge to mobile TV where there is a tradeoff between the screen size and the portability of the device. These concerns have been noted in focus groups assessing the potential uptake of mobile TV services [2]. Users want a screen as large as possible for viewing, but they do not want their phones to be too big. Some of them would opt for a dedicated mobile TV device to circumvent the former dilemma. Moreover, it is not clear whether users will want higher arousal and immersion in a mobile context, because of the increased risk of errors and accidents.

In one of the few studies that specifically examined smaller screens, Reeves et al. found no difference in arousal and attention between users watching 2" and 13" screens, although arousal and attention were larger with a very large screen (56") [9].

Other studies have even shown that smaller image resolutions can improve task performance. For example, Horn showed that lie detection was better with a small (53x40) than a medium (106x80) video image resolution [10]. In another study, however, smaller video resolutions (160x120) had no effect on task performance but did decrease satisfaction when compared to 320x240 image resolutions [11]. In a study by Barber et al., a reduction in image resolution (from 256x256 to 128x128) at constant image size led to a loss in accuracy of emotion detection especially in a full body view [12].

To investigate these factors we adopted a method recently used to evaluate quality tradeoffs with mobile sports coverage [13] to investigate the effects of resolution, picture size and encoding bitrates on the perceived video quality News.

Study 1

The first study was designed to research encoding bitrates that would deliver acceptable video quality to mobile users receiving news on two devices with different resolutions. The aim of this study was to identify bandwidth requirements to deliver an acceptable service on the two different devices.

Equipment

The experiments were carried out on two devices:

1. iPAQ 2210 PDA with an additional 1GB CF storage card, and a pre-installed windows media player, a 3.5" TFT liquid crystal display, 64K colours, (240x320 pixels)
2. NEC e616 3G phone, with 19MB internal memory, capable of decoding/playing 3GPP MPEG-4 video, a 35x43mm TFT liquid crystal display, 64K colours, (240x176 pixels).

Content and Encoding

Test material used for quality evaluation is usually selected from a video or audio test set. For example, VQEG uses a test set of 20 8-second clips [14] to represent a range of difference types of motions, content and camera position. While such test sets are suitable for comparing technical performance differences between codecs, they are less useful in evaluating the perceived quality of service as the clips do not have audio and are not representative of the experience users would have with mobile TV.

Consequently this study used a 15 second clip from a CBS news bulletin, similar in length to those used in the VQEG test set. This clip had no text ticker and used big text only for headlines.
It was encoded for display on both devices. The encoding bit-rates ranged from 28Kbps to 256 Kbps depending on the device and are presented in Table 1 and Table 2.

Table 1: Parameters for clips on 3G phone at 176x144

Total bitrate	Encoding bitrate: video	fps	Encoding bitrate: audio
28 Kbps	16 Kbps	6	6.7 Kbps
64 Kbps	45.4 Kbps	10	12.2 Kbps
128 Kbps	11 Kbps	15	12.2 Kbps
256 Kbps	243 Kbps	25	12.2 Kbps

Table 2: Parameters for clips on PDA at 320x240

Total bitrate	Encoding bitrate video	fps	Encoding bitrate audio
64 Kbps	52 Kbps	10	12 Kbps
128 Kbps	112 Kbps	15	16 Kbps
256 Kbps	224 Kbps	25	32 Kbps
350 Kbps	285 Kbps	25	64 Kbps
512 Kbps	448 Kbps	25	64 Kbps

The iPAQ clips were encoded with *MS Media Encoder* in Microsoft Video format. For the 3G phone, clips were encoded using a trial version of the 3GP MPEG-4 encoder (*PV Author*) available online from Packet Video (packetvideo.com).

Participants and Procedure

Twenty paid participants were tested using the described material. There were equal numbers of men and women and the mean age of participants was 26.

On the iPAQ the video clips were arranged in a play list in the windows media player. On the phone each clip was selected separately from a folder. The participants watched all videos first on one device then again on the other. The order of devices on which the participants watched the videos first was randomized.

Each clip was shown to the participants at all of the different encoding bitrates starting with the highest encoding bitrate and descending to the lowest encoding bitrate.

The participants alternately watched a clip and then rated the acceptability of the video and the audio quality on a rating sheet. These were binary decisions as participants had to tick either yes or no as a reply. Additionally, the participants rated both qualities on a scale from 0 to 100.

Results

The yes-no responses on the acceptability of the video quality from the participants were averaged for each encoding bit rate and are presented for both devices in Figure 1.
News was acceptable to all participants when it was encoded at 285kbps and higher and on both devices News quality was acceptable to the vast majority of participants at encoding bit rates above 100kbps.
However, we did not find a statistical difference between acceptability at any comparable bitrates: [52 vs. 45; $Z = -0.33$, ns:

112 vs. 111; Z=-1.11, ns: 224 vs. 243; Z=-0.57, ns]. Thus, at any particular bitrate there was evidence that the iPAQ has a quality advantage because of the larger screen, and there was no evidence that the phone has a quality advantage because of the higher bitrate per pixel ratio.

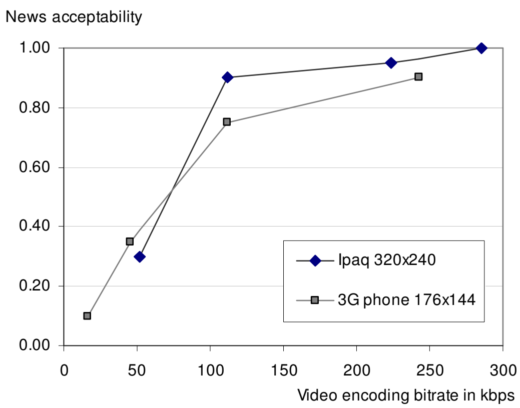

Figure 1: News video quality acceptability

Discussion

The results from Study 1 indicated that at a bitrate of at least 100kbps is required to deliver an acceptable News service and that screen resolution and size have little effect on the boundaries of on acceptable service.
However, there are a number of limitations to this study that caution against over generalisation. First, the sample is very small so it is possible that the tests are not sensitive to small differences in service acceptability. Second although we showed the videos on both devices at the same frame rates for comparable encoding bitrates, the comparison is based on different video and audio codecs and the devices have different hardware implementations such as screen brightness and contrast. Third, the News clips used was much shorter than a standard News broadcast and was also repeated many times in the test. It is not clear that this procedure is representative of the experience users would have from watching an actual News bulletin on a mobile device.

Study 2

The second study was designed to look more closely at the effects of reduced image resolution on video quality acceptability. The study addresses the limitations of study one by:

1. Testing a much larger sample
2. Controlling for hardware differences
3. Using representative News material

An additional factor examined was the effect of varying audio quality. News content relies heavily on the intelligibility of spoken words and any loss of visual detail might be compensated for by having high audio quality. The study also includes a summary on qualitative feedback, interest in different mobile TV content types and the willingness to pay as an indicator for general interest in mobile TV.

Equipment

The experiments were carried out on the same iPAQs that were used for study 1. Because a pilot of this study showed that videos played from the CF card would stall intermittently no compact flash cards were used in study 2. The content was stored on a 512MB SD card.
The iPAQ was equipped with a set of Sony MDR-Q66LW headphones to deliver the audio. A customized application was programmed in C# using the Odyssey CFCOM middleware to integrate the Windows Media Player for Windows CE. It presented the clips along with a volume control and two response buttons to indicate acceptable and unacceptable quality.

Content and Encoding

For current mobile TV services, there is usually an additional editing process to prepare the material for mobile consumption. This involves removing certain shots that would not render or compress well for a mobile device. Bespoke editing takes time (which means access to topical content such as news is delayed) and is expensive; thus, many service providers favour immediate re-use of TV material. For the purposes of this study, we investigated the acceptability of directly recorded TV material without any special editing steps. Clips of this type have been successfully used to examine quality tradeoffs for football coverage on mobile TV [13].

Another consideration was the length of the test clips. Mobile TV viewing will typically be considerably longer than 8-15 seconds, and composed of a mixture of different foreground and background motion in a scene, content and camera angles. Whereas the clips used in Study 1 were 15 seconds long to compare the different codecs we introduced a more realistic setup in Study 2 by using clips that lasted two minutes and 20 seconds.

The video clips were prepared as follows: We recorded news footage from DVB-T TV (BBC24 news). All extracted clips were chosen such that after 2:20 minutes (or shortly thereafter), a story line would end. In addition, the news material chosen for this study was recorded from TV and included a variety of text information in the form of news ticker, clock, logos, graphs and captions.

We used Virtualdub to segment these source clips into seven 20 second long clips at the different resolutions at 12.5fps. These segments were encoded with Windows Media Encoder (WME) using the Microsoft Windows Media Video V8 codec with the different bitrates for the different segments as shown in Table 3. Each group of seven WMV segment files were then converted and concatenated to one AVI file with TMPGEnc Express. Finally, these files were encoded with WME again to alter the audio encoding to either 32 or 16kpbs using the Windows Media Audio V9 codec. The video was encoded at a higher bitrate than the maximum of the first WME encoding in order to prevent significant alterations in the video quality in any of the segments.

It should be noted that the setting of the frames per seconds as shown in Table 3 only represent the target for the encoder. The actual frame rates of the final videos with the lower encoding bit rates deviated from this value.

Table 3: Video segments encoding bitrates

Seg	Time (secs)	Encoding bitrate video	fps	Encoding bitrate audio
1	1-20	224 kbps	12.5	16 / 32 kbps
2	21-40	192 kbps	12.5	16 / 32 kbps
3	41-60	160 kbps	12.5	16 / 32 kbps
4	61-80	128 kbps	12.5	16 / 32 kbps
5	81-100	96 kbps	12.5	16 / 32 kbps
6	101-120	64 kbps	12.5	16 / 32 kbps
7	121-140+	32 kbps	12.5	16 / 32 kbps

Participants and Procedure

Most of the 128 paid participants (83 women and 45 men) were university students. The age of the participants ranged from 18 to 67 with an average of 24 years. They came from a total of 26 different countries. English was the first language for 72 of the participants.

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 quality for watching different types of content. The instructions stated

"If you are watching the coverage and you find that the quality becomes 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 then please click the button labelled 'Acc'.

Once it was clear that they understood the instructions, participants were provided with headphones and an iPAQ and given a minute to practice pressing the buttons on the display. When they were ready the experiment began and the participants watched 16 clips in succession. Sixty-four participants saw the clips increasing in size and the other 64 saw the clips decreasing in size.

During the session we recorded the participants' interactions with the devices on video. The video was later used to measure viewing distance at the different image resolutions. The participants' ratings, i.e. the taps on the 'Unacc.' and 'Acc.' buttons were recorded on the device.

At the end of the video rating session, we interviewed the participants to find out what aspects of the video quality they found unacceptable. We administered a survey on their activities in dead time situations, interest in a mobile TV service in general and the content types in particular, current mobile phone expenses and the willingness to pay for a mobile TV service.

Results

Before analyzing the results, we conservatively coded each 20 second interval of a clip as *unacceptable* if they had given a rating of unacceptable at any point during that period. Non parametric tests were performed to test for differences between the image resolutions and the video and audio encoding bitrates. A post-hoc analysis examined differences between those who were interested vs. not interested in mobile TV services.

There were significant differences in service acceptability between the four resolutions. [$\chi^2(3)=59.04$, $P<0.001$]. However, the largest size was not the most acceptable. Instead a small reduction in size from 240x180 to 208x156 gave significant gains in ratings of service acceptability at bitrates greater than 64kbps [$Z = 2.11$, $P<0.05$] (see Figure 2).

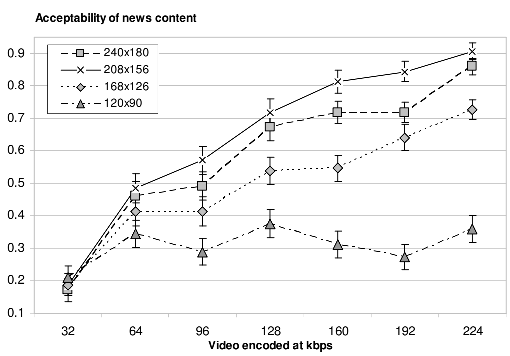

Figure 2: Acceptability of News content

As expected there were also significant differences between the seven video bitrates tested [$\chi^2(6)=163.84$, $P<0.001$] with higher video bitrates leading to higher service acceptability. However, the same pattern was not observed with audio bitrates. Here the trend was that participants were less likely to rate services as unacceptable when they had a lower audio bitrate. The difference however was not significant [$Z = 0.43$, ns].

Finally, the post-hoc analysis revealed no differences in ratings between those who were interested vs. not interested in mobile TV services [$Z = 0.58$, ns].

For both increasing and decreasing image resolution groups, there was no significant difference in the distance at which the iPAQ was held at the start or end of the study.

Qualitative Results

When asked why they rated the video quality of news as unacceptable, participants mentioned a number of factors. Across all 128 participants, a total of 290 comments described the reasons for the unacceptability of news coverage. Of these comments, 34% related to *text detail*, i.e., the legibility of the news ticker, the headline text, the clock, the logo, or the captions for the people being interviewed by the newscaster. Other problems people reported were *facial details* and expressions in the switch from anchor person to field reporter. A summary of these problems and the frequency with which they were mentioned is presented in Figure 3.

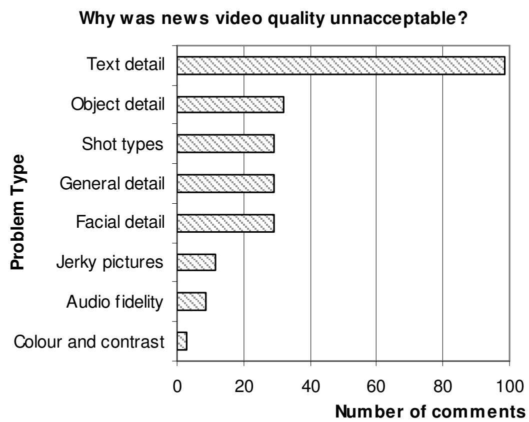

Figure 3: Reasons for unacceptable news quality

Survey Results

Ninety-eight participants responded to the survey after the experiment. Asked about their usual activities during dead time situations participants mentioned most often reading. Listening to music was the second most popular activity. These results are presented in Figure 4. Preparing text messages, maintaining the address books, and weeding out text messages are summarized within communication in Figure 4. It should be noted that making phone calls is currently not possible on the London underground system.

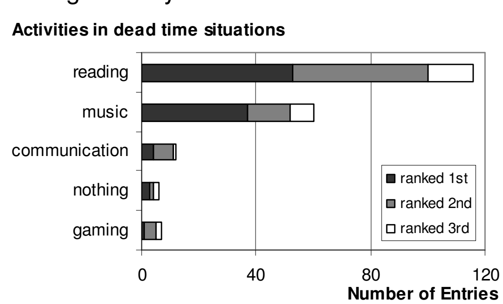

Figure 4: Activities in dead time

Books were the most popular reading material, which was closely followed by newspapers. The number of participants that buy their newspapers is about equal to the number of participants that read newspapers distributed free of charge (e.g. METRO).

Of the 98 participants that replied to this question 55% said that they were interested in mobile TV, 18% maybe, 13% not really and 12% were not interested.

Asked which of the presented or other content types would be interesting to them the participants chose news as the most interesting content type. The replies are summarised in Figure 5 and match results of previous studies [2].

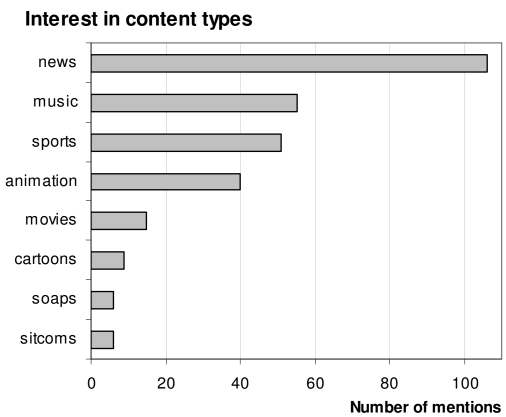

Figure 5: Interest in content types

We asked the participants how much they were willing to pay for unlimited mobile TV viewing per day and per month. The medians of what the participants were willing to pay were £2.50 and £10 respectively. These numbers should not be taken as a direct market measure but as a measure of the participants' attitude towards the service. On average they were willing to spend six times the daily amount for a month's access. There was no correlation between the participants' current monthly expenses and the prices they were willing to pay for mobile TV.

Discussion

Both quantitative and qualitative results indicate that the primary effect of reducing image resolution was a loss of visual detail. The effect of reducing image resolution was more pronounced when bandwidth was abundant. When bandwidth was scarce, there is little or no effect of reducing the image resolution, as visual detail is already low. At 128kbps and above, there was a sharp reduction in acceptability when image resolution was dropped from 168x126 to 120x90.

The qualitative comments help to identify the source of the problems. Of the eight most frequently cited problems, four relate to identifying or distinguishing detail – such as text and facial features. We also identified particular shot types that caused difficulty. The relatively few comments about frame rate referred to picture quality as *'jerky'*. Unexpectedly, no comments were made about the audio quality.

The primary detail on which news video quality was judged was the ability to distinguish textual information – whether the news ticker, the clock, headline text or person names. It seems that the slight increase in perceived quality with a reduction in image resolution to 208x156 was caused by a perceived increase in the quality of the text. If text were coded and transmitted separately from the video we would expect clips encoded at an image resolution of 240x180 to be more acceptable than 208x156.

The dominance of text legibility would also explain the differences in acceptability between study 1 and study 2. The clip used in study 1 did not use small font text and at a video encoding bitrate of 128kbps reached an acceptability of 75% on the NEC phone at 176x144. The clips in study 2 used a lot of small font text and the acceptability of the video at 128kbps was only 53% on the iPAQ at a comparable resolution of 168x126.

CONCLUSIONS

The results of study 2 indicate that the dominant effect of reducing the image resolution was a loss of *visual detail*. This conclusion is reinforced by the qualitative comments on the problems participants experienced. This effect, however, was not universal. We found that the increased *bandwidth per pixel ratio* with a slightly smaller image gave improved acceptability ratings. However, as the image resolution was reduced further the loss of detail again dominated.

Legibility of text was an important issue and may be improved by reducing the image resolution of the content prior to encoding. The more general recommendation, however, is to stream the text information separately to the device.

No comments made by participants related to *general arousal.* Previous studies indicate that arousal is related to the visual angle subtended by the image, thus if arousal were of primary interest to participants we would expect them to adjust for the smaller image size by moving the device closer. In our laboratory setup we found no evidence of such an adjustment as image size and resolution were reduced.

The biggest limitation of the two studies at hand is that both assume perfect delivery of the data. The impact of data loss on the perceived video quality of news content at different encoding bitrates was not addressed in either study.

References

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

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

[3] Owens, D. A. & Wolfe-Kelly, K. "Near Work, Visual Fatigue, and Variations of Oculomotor Tonus" *Investigative Ophthalmology and Visual Science,* 28, p.743-9, 1987.

[4] Edwardson, M., Kent, K., McConnell, M. "Television news information gain: Videotext versus a talking head" *Journal of Broadcasting & Electronic Media,* 29, p.367-78, 1985.

[5] Weiner, R. *Webster's New World Dictionary of Media and Communications* (rev. and updated ed.) New York, NY: Macmillan, 1996

[6] Reeves, B. & Nass, C. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places* University of Chicago Press, 1998

[7] Lombard, M., Grabe, M. E., Reich, R. D., Campanella, C., Ditton, T. B. "Screen Size and viewer responses to television: A review of research" in *Annual Conf.of the Assoc.for Education in Journalism and Mass Communication*, 1996

[8] Okada, K.-I., Maeda, F., Ichikawaa, Y., Matsushita, Y. "Multiparty videoconferencing at virtual social distance: MAJIC design" in *Proc.ACM conf.on Computer supported cooperative work*, 1994, p.385-93.

[9] Reeves, B., Lang, A., Kim, E., Tartar, D. "The effects of screen size and message content on attention and arousal" *Media Psychology,* 1, p.49-68, 1999.

[10] Horn, D. B. "The effects of spatial and temporal video distortion on lie detection performance" in *Proceedings of CHI '02*, 2002

[11] Kies, J. K., Williges, R. C., Rosson, M. B. *Controlled Laboratory Experimentation and Field Study Evaluation of Video Conference for Distance Learning Applications* (HCIL 96-02) Virginia Tech, 1996.

[12] Barber, P. J. & Laws, J. V. "Image Quality and Video Communication" in R. Damper, W. Hall, & J. Richards (Eds.) *Proceedings of IEEE International Symposium on Multimedia Technologies & their Future Applications*, London, UK, 1994, p.163-78.

[13] McCarthy, J., Sasse, M. A., Miras, D. "Sharp or smooth? Comparing the effects of quantization vs. frame rate for streamed video" in *Proc.CHI*, 2004, p.535-42.

[14] VQEG *Final Report from the video quality experts group on the validation of objective models of video quality assessment* (http://www.vqeg.org), 2000.