Getting the Big Picture on Small Screens: Quality of Experience in Mobile TV

Hendrik Knoche
Department of Computer Science
University College London
Gower St,
London, WC1E 6BT
+ 44 20 7679 3642
+44 20 7387 1397
h.knoche@cs.ucl.ac.uk

M. Angela Sasse
Department of Computer Science
University College London
Gower St,
London, WC1E 6BT
+44 20 7679 7212
+44 20 7387 1397
a.sasse@cs.ucl.ac.uk

Getting the big picture on small screens: Quality of Experience in mobile TV

ABSTRACT
This chapter provides an overview of the key factors that influence the Quality of Experience (QoE) of mobile TV services. It compiles the current knowledge from empirical studies and recommendations on four key requirements for the uptake of mobile TV services: handset usability and its acceptance by the user, the technical performance and reliability of the service, the usability of the mobile TV service (depending on the delivery of content) and the satisfaction with the content. It illustrates a number of factors that contribute to these requirements ranging from the context of use to the size of the display and the displayed content. The chapter highlights the interdependencies between these factors during the delivery of content in mobile TV services to a heterogeneous set of low resolution devices.

KEYWORDS: MOBILE TV, QOE, SOCIO-TECHNICAL SYSTEMS, MULTIMEDIA APPLICATION, VIDEO COMPRESSION, VIDEO QUALITY

INTRODUCTION

It is the second time around for mobile TV. In the 1980s, Seiko introduced a TV wristwatch that was capable of displaying standard TV channels on an LCD wrist watch. It seemed like a great idea at the time. Many people wore watches, a growing number of people used LCD or digital watches and it was possible to display anything on an LCD display. However, the watch was not a success. One of the biggest problems was high energy consumption - the watch wearer had to separately carry the battery, which was part of a box that housed the TV receiver and connected to the watch through a cable. This setup gave the wearer around one hour of viewing time. The screen was monochrome and had low contrast. Furthermore, watching TV while wearing the watch resulted in an unnatural wrist posture. Last but not least, the TV wristwatch was expensive.

Twenty years later, mobile TV is back. Many people now carry inexpensive mobile phones with built-in LCD screens. This allows the display of moving images, which can be received in a more energy efficient way these days and mobile TV is making its second appearance. Today, mobile TV services are available in a number of countries. While Asian consumers already have access to broadcast services, Western countries have finished trials and are aiming to move from unicast, i.e. individual delivery services, to broadcast solutions. Portable play-stations and video Ipods provide alternative platforms for playing pre-stored content.

So far, the deployment of these services has been driven by technical feasibility and matching business models. The wireless domain is one of limited bandwidth resources, and service providers have to decide on broadcasting more content at lower quality or vice versa in search of optimal configurations for people's QoE that are financially viable. The content is produced by companies with a specific primary target medium i.e. cinema, TV or mobile in mind. This choice influences the selection of shot types, length and the type of programme. Cameras can be chosen from a wide selection delivering different resolutions, aspect ratios, contrast ranges and frame rates. After post-production the content is delivered to audiences through various channels. For example, TV broadcast companies adapt cinema content to the TV and mobile service operators adapt TV content for mobile TV distribution. Uptake of existing mobile TV services lags behind expectations, possibly because customers are not willing to pay high premiums for content (KPMG, 2006). To assist service providers in improving their service offerings, we need to understand how people might experience mobile

TV services in their entirety. Quality of Experience (Jain, 2004), (Aldrich, Marks, Lewis, & Seybold, 2000), (McCarthy & Wright, 2004) is a broad concept that encompasses all aspects of a service that can be experienced by the user. In the case of mobile TV, QoE includes the usability of the service, the restrictions inherent in the delivery, the audio-visual quality of the content, the usage and payment model, the social context as well as possible parallel use of standard TV. According to (Mäki, 2005) the following four requirements are the most important for adoption of mobile TV services:

1. Handset usability and acceptance
2. Technical performance and reliability
3. Usability of the mobile TV service
4. Satisfaction with the content

We will address each of these factors in turn in more detail in the following sections in order to provide a comprehensive view on the Quality of Experience of mobile TV services.

HANDSET INTEGRATION AND USABILITY

Currently, the mobile phone is the most likely platform for mobile TV but PDAs, portable game consoles and music players are attractive alternatives. In 2003 a total of 70% of the people in Europe owned or used mobile phones. The importance of mobile phones in people's lives means that most owners carry it with them wherever they go. Mobile TV consumption on mobile phones allows for privacy of consumption, because of short viewing distances and the viewing angle afforded by many mobile devices. However, people perceive the battery consumption of mobile TV as a threat to more important communication needs. The application should provide warnings when the battery is drained beyond a certain threshold (Knoche & McCarthy, 2004) and service providers should set user expectations about battery drain induced by e.g. watching live content (Serco, 2006).

The mobile TV application should not get in the way of communication but alert users of incoming calls, text or other messaging and provide means to deal with them in a seamless manner. On inbound communication this includes automatic pausing of the TV service if possible and offering to resume once the user has finished communicating. Likewise, important indicators e.g. for battery status or menus should not unnecessarily obstruct the TV screen but could use semi-transparent menus as suggested in (Serco, 2006).

Depending on the technical realisation, having TV reception might require a second receiver unit in the handset that would allow for parallel reception of TV content and making and receiving telephone calls at the same time. A single receiver unit, for example, would not be able to record live TV content during a phone call.

Display

The screen should have high contrast, backlight and a high viewing angle to support viewing in different circumstances and by multiple viewers. Due to size and power constraints liquid crystal displays are currently the preferred technology to present visual information on mobile devices. LCD displays come in a range of shapes, sizes and resolutions, from VGA PDAs (480x640 pixels) and high end 3G or DVB-H enabled phones (320x240) to more compact models with QCIF size (176x144) and below. Users want as large a screen as possible for viewing, but they do not want their phones to be too big (Knoche et al., 2004). Landscape oriented use of the display might be preferred (Serco, 2006) over the typical portrait mode that mobile phones are used in. In general, pictures subtending a larger visual angle in the eye of the beholder make for a better viewing experience. Results from studies on TV pictures revealed that larger image sizes are generally preferred to smaller ones (Reeves & Nass, 1998), (Lombard, Grabe, Reich, Campanella, & Ditton, 1996) and perceived to be of higher quality (Westerink & Roufs, 1989) but that there is no difference in arousal and attention between users watching content on 2" and 13" screens (Reeves, Lang, Kim, & Tartar, 1999).

Which resolutions best support the different screen sizes of mobile TV devices is subject to current research and the pros and cons of different resolutions will be discussed in the section on video quality. Besides the size of the device, the visual impact can be increased by head mounted displays and projection techniques. Whereas the former results in an experience of greater immersion and requires additional equipment to be carried around the latter reduces the anonymity of visual consumption.

Device use

A stand for continuous viewing is beneficial for mid- to long-term use. In public places, the use of headphones, which increase immersion (see below), might be required. Many people already use head phones for portable music players and standard headset jacks on mobile TV devices would make switching between devices easy. Dedicated buttons would be valuable for mobile TV access, basic playback controls and content browsing, e.g. channel switching or selecting. On touch screens many people value on-screen buttons that do not decrease the viewing area of the content instead of having to use a stylus which requires two-handed operation.

Immersion

Users are worried about becoming too absorbed in what they are watching, and thus distracted from other tasks while being on the move, e.g. missing trains or stops (Knoche et al., 2004). They require a pause/mute facility to cope with likely interruptions. In the case of broadcast content, this requirement places demands on the device's storage capacity. Volume control should possible preferably without the need to access menus. The question whether a separate means to mute the volume and would let the video play in the background will be necessary or might confuse users more in conjunction with the pause button which pauses both audio and video has to be addressed by future research. An easy way to set alarms or countdowns might help mobile users to not loose touch with the world around them.

TECHNICAL PERFORMANCE AND RELIABILITY

The way the content is delivered has a major effect on the possible uses of a mobile TV service. The perceived video and audio quality will depend on the Quality of Service (QoS) provide by the network that is delivering the packets that carry the content and might noticeably degrade the content by introducing errors, loss and varying delays to those packets. For an example of how loss influences the perceived video quality of mobile TV content see (Jumisko-Pyykkö, Vinod Kumar, Liinasuo, & Hannuksela, 2006).

Service delivery

From the user point of view, TV is commonly understood as an 'any time' service: turn it on and it will deliver content at any time of day. Mobile TV services implicitly suggest being available anywhere at any time. Mobile phone users have been wary of this promise (Knoche et al., 2004).

There are four content delivery models that significantly shape the experience of the mobile TV service: media charger, streaming (unicast), broadcast and pre-cached broadcast. The video Ipod is an example of a *media charger*. The user has no live content but does have full playback control and can watch anywhere at any time. In order to have a supply of fresh content the user has to touch base regularly.

Many of the services like MobiTV and Slingbox (Sling Media, 2006) in the US and Vodafone live! in the UK that are currently offered are unicast services, which makes them relatively expensive in terms of spectrum usage and difficult to scale. With each increase in the number of receivers in a reception cell the available bandwidth per receiver decreases. The number of users receiving a unicast mobile TV stream on demand within a wireless cell is therefore limited and the audio-visual quality degrades with the increase in receivers. However, unicast can deliver personalised content for niche interests that would not be viable through broadcasts. Broadcast approaches like DMB and DVB-H are more efficient in mass delivery as they support an arbitrary number of receivers at constant quality in the coverage area. Broadcast users have no playback control unless pausing live TV and other functions available in personal video recorders (PVR) are implemented on the user terminal. However, since being on the move results in varying levels of reception people experience varying quality and service discontinuities. This poses a problem to broadcast TV services without PVR-like functionality. People might tune into the streams at times when the programmes they want to see are not being broadcast. Similar to media chargers, pre-cached services, e.g. SDMB (Selier & Chuberre, 2005), can continuously display recently downloaded, i.e. non-live, content at higher quality through carrousel broadcasts. This is an example of TV any time, which allows users to watch broadcast content when convenient. Mobile TV services do not have to rely solely on one of these delivery mechanisms but could mix them in order to leverage their different advantages.

The content has to be delivered through one of these transmission schemes to a range of devices with different display capabilities. There are three main ways to address the problem of multiple target solutions: 1) Sending multiple resolutions, which requires more bandwidth if broadcast. 2) Broadcasting at the highest resolution and resizing at the receiver side. 3) Employing layered coding schemes that broadcast a number of resolution layers from which every receiver can assemble the parts it can display.

Resolution, image size and viewing distance

Human perception of displayed information has been studied for a long time, see (Biberman, 1973) for an overview. Spatial and temporal resolution are key factors for the perceived quality of video content. Whereas temporal resolution below 30 frames per second (fps) results in successively jerkier motion, lowering the number of pixels to encode the picture reduces the amount of visible detail. Excessive delays and loss during transmission of the content may affect both the spatial and temporal resolution resulting in visible artefacts and or skipping of frames causing the picture to freeze.

The higher the resolution in both of these dimensions, the more bandwidth is required to transmit it. Service providers only have a limited amount of bandwidth available and want to maximise the content they can offer to their customers while still delivering the quality that the customers expect. They face the trade off between visual quality and quantity of the content.

Mobile TV will be consumed at arm's length. Paper, keyboard and display objects are typically operated at distances ranging from 30cm to 70cm. Continued viewing at distances closer than the resting point of vergence – approx. 89cm, with a 30º downward gaze – can contribute to eyestrain (Owens & Wolfe-Kelly, 1987). When viewing distances come close to 15cm, people experience discomfort (Ankrum, 1996). Normal 20/20 vision is classified as the ability to resolve 1 minute of arc (1/60º) (Luther, 1996) and translates to 60 pixels per degree. The amount of pixels p that can be resolved by a human at a given distance d and a picture height h can be computed by the following equation: $p = \frac{h}{d \cdot 2\tan(1/120)}$.

In the typical TV viewing setup at a seating distance of 3m, the benefits of HDTV can only be enjoyed on relatively big screens. On handheld devices, people could easily enjoy HDTV resolutions on a screen of 8cm height. However, mobile TV does not exceed QVGA resolution at present. In addition, people are able to identify content that has been up-scaled from low resolutions to higher resolution mobile screens. So far no research has addressed the potential effects of up-scaling low broadcast resolution content to a screen with a higher resolution. Research on these topics is proprietary. Philips uses a non-linear up-scaling method called Mobile PixelPlus to fill a screen with higher resolution than the broadcast material.

Some studies have addressed the perception of low resolution content on small handheld screens (Song, Won, & Song, 2004), (Knoche, McCarthy, & Sasse, 2005b) and (Jumisko-Pyykkö & Häkkinen, 2006). Content shown on mobile devices at higher resolutions is generally more acceptable than lower resolutions at identical encoding bitrates. However, the differences are not uniform across content types (Knoche et al., 2005b). All content types received poor ratings when presented at resolutions smaller than 168x126. Other studies have shown that smaller image resolutions can improve task performance. For example, (Horn, 2002) showed that lie detection was better with a small (53x40) than a medium (106x80) video image resolution. In another study, however, smaller video resolutions (160x120) had no effect on task performance but did reduce satisfaction when compared to 320x240 image resolutions (Kies, Williges, & Rosson, 1996). 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 (Barber & Laws, 1994). The legibility of text has a major influence on the acceptability of the overall video quality (Knoche & Sasse, 2006) and should be sent separately and rendered at the receiving side.

Frame rate

Low video frame rates are common in recent mobile multimedia services especially in streamed unicast services. Frame rates as low as 5 fps and lower were avoided at all costs in a desktop computer based study by Pappas & Hinds (1995). Another study, conducted by Apteker et al. assessed the watchability of various types of video at different frame rates (30, 15, 10, 5 fps). Compared to a benchmark of 100% at 30fps, video clips high in visual importance dropped to a range of 43% to 64% watchability when displayed at 5 fps, depending upon the importance of audio for the comprehension of the content and the static/dynamic nature of the video (Apteker, Fisher, Kisimov, & Neishlos, 1994). Participants who saw football clips on mobile devices found the video quality of football content less acceptable when the frame rate dropped below 12fps (McCarthy, Sasse, & Miras, 2004). Comparable displays on desktop computers maintained high acceptability for frame rates as low as 6fps. The reason for the higher sensitivity to low 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. The proprietary Natural Motion approach by Philips supposedly reduces the jerkiness of low-frame rate content by generating intermediate frames from the broadcast set of frames at the receiver side (de Vries, 2006).

Some programs have sign language interpreters signing to make the programme understandable for deaf people. This is one of the few applications that require high frame rates for comprehension of the visual content. Spelling sign language requires 25 frames to be able to capture all letters in at least one frame (Hellström, 1997).

Temporal vs. spatial resolution

Whereas earlier guidelines suggested the use of higher frame rates for fast moving content, e.g. sports, (IBM, 2002) recent findings show that users prefer higher spatial resolution over higher frame rates in order to be able to identify objects and actors in mobile TV content (McCarthy et al., 2004). Wang et al. reported on a study in which they manipulated both frame rate and quantization with an American football clip. They concluded that *"quantization distortion is generally more objectionable than motion judder"* and that large quantization parameters should be avoided whenever possible (Wang, Speranza, Vincent, Martin, & Blanchfield, 2003).

Audio-Visual Quality

A number of studies have found that the combined quality of audio-visual displays is not simply based on the sum of its parts e.g. (Hands, 2004), (Jumisko-Pyykkö et al., 2006). In a study on audio-visual interactions, Winkler & Faller found that selecting mono audio for a given bitrate gives better quality ratings and that more bitrate should be allocated to the audio for more complex scenes (Winkler & Faller, 2005). As a byproduct in a study on TV viewing experience, Neuman et al. discovered that the perceived video quality was improved by better audio (Neumann, Crigler, & Bove, 1991). However, it was only the case for one of the three used content types. Similarly, a study by Beerends et al., using a 29cm monitor, found that the rating of video quality was slightly higher when accompanied by CD quality audio than when accompanied by no audio (Beerends & de Caluwe, 1999). The effect, however, was small and has not been replicated with small screens. However, in the same study participants judged the two lower video quality levels (where the video bandwidth was limited to 0.15 MHz and 0.025 MHz) worse when they were presented with audio, than without audio. Similarly, in a study by Knoche et al., the visual quality of video clips displayed on mobile devices was more acceptable to participants across all video encoding bitrates when it was supported by lower (16kbps) than with higher audio quality (32kbps) (Knoche et al., 2005b).

Synchronous playback of sound and video affects the overall AV-quality (Knoche, de Meer, & Kirsh, 2005a). For 30 frames per second video the window of synchronization is ±80ms (Steinmetz, 1996). The temporal window of synchronisation depends on the video frame rate (Knoche et al., 2005a), (Vatakis & Spence, 2006). At lower frame rates audio-visual speech perception is more sensitive to audio coming before video and the presentation of the audio relative to the video should be delayed (Knoche et al., 2005a).

In general, human perception is more sensitive to degradations and fluctuations in audio information than in video information. In situations with imperfect reception service providers should use coding schemes that can deliver uninterrupted audio and prioritise its delivery over the video.

USABILITY OF THE MOBILE TV SERVICE

In order to understand what makes for a usable mobile TV service, we need to know about the context of the user including the motivation for use and the location. Many of the guidelines that apply to mobile application design equally apply to mobile TV see Serco (2006) for an overview.

Motivation of use

Whereas the drivers behind standard TV consumption are fairly well understood, we lack comparable knowledge in mobile TV. Peoples' watching of standard TV is driven by ritualistic (Taylor & Harper, 2002) and instrumental motives (Rubin, 1981) as in 'electronic wallpaper' (Gauntlett & Hill, 1999), mood management (Zillman, 1988), escapism, information, entertainment, social grease, social activity, and social learning (Lee and Lee 1995). For many of these drivers watching TV constitutes a group activity. Mobile TV is, due to its nature and limitations, more likely to be an individual consumption activity. The restricted viewing angle of the screens, the (for some people uncomfortable) proximity with others to share it and the fact that the mobile phone is a rather personal device might curb group usage.

Location

According to (Mäki, 2005) the most common places for mobile TV use are (in descending order):

1. in public transport
2. at home
3. at work

This is supported by other studies, in which many participants of mobile TV trials used the device as an additional TV set at home (Södergård, 2003). While at home users' perception of the mobile TV service might depend on the comparison with standard TV in terms of delay (mobile broadcasts might incur additional delay due to processing or delivery, e.g. through satellite), program availability, audio-visual quality, responsiveness, ease of use, interoperability with other media solutions including recording devices - such as personal video recorders (PVR) - that allow for easy recording of television shows, content sharing and user-controlled storage.

People are able to compare the different experiences of consuming TV content at home. Some might object to the inherent delay (approximately 1 minute) between the live broadcast TV signal and the mobile TV signal as currently seen in MobiTV (Lemay-Yates Associates Inc., 2005). What is more important, perhaps, is that the delay disadvantages the mobile audience in interactive game shows or betting services.

Usage patterns

Previous research has shown that peoples' average usage of mobile TV is less than ten minutes long (Södergård, 2003). This window of consumption places demands both on the length of consumable content and the time that users might be willing to spend to access and navigate through it. Data from SDMB trials in Korea for example show that people use mobile TV throughout the day with peaks in the morning, at lunch time, in the early evening and very late in the evening.

Interactivity of the mobile TV service

The interface needs to provide the user with controls to use the different kinds of interactivities offered in mobile TV. Users expect the entry points to the mobile TV to be available from prominent places in the mobile phone user interface (Stockbridge, 2006).

Participation interactivity and payments

One of the potentially biggest advantages of mobile TV over regular TV is the existence of a return channel with built-in billing possibilities for premium and subscription services, as well as transactions involved in interactive services such as voting and betting.

Participants in mobile TV trials favoured the flat-rate payment model, i.e. a single payment for unlimited mobile TV use during a billing period (Mäki, 2005). Flat rates do not place additional barriers between the users and the content. In South Korea, early payment models greatly influenced the use of mobile TV. When mobile TV usage was billed in the amount of kilobytes received, each one-minute part of a programme made especially for mobile TV had to be confirmed for delivery, which resulted in a discontinuous viewing experience (Knoche, 2005).

Distribution interactivity and content navigation

Taylor et al. (2002) argued that channel surfing is inherently associated with the act of watching TV. The methods to select a program used in traditional TV viewing depend on the time of day. But the method used generally escalates – if nothing of interest is found – to strategies that require more effort on behalf of the user. The order of strategies is:

1. Channel surfing
2. Wait or search for a TV program announcement
3. Knowledge of weekly schedules or upcoming programmes
4. Paper-based or onscreen guides

Since mobile TV usage spurts are rather short, waiting for and searching for announcements or upcoming programs might not be feasible. Information on what is currently playing and what will come up next might be valuable and should be easy to access.

Ideally dedicated buttons or soft keys will allow users to switch channels. Long waiting times after a requested channel switch will result in lower user satisfaction. Tolerable switching delays between mobile TV channels have not been thoroughly researched but should be as short as possible since users are accustomed to almost instantaneous switches on standard TV. First results for digital TV indicate that 0.43 seconds might be the limit beyond which users will be increasingly dissatisfied (Ahmed, Kooij, & Brunnström, 2006). In digital TV, the switching delays depend to a large part on the video codec, e.g. in MPEG encoded content on the occurrence of so-called key frames. Fewer key frames in a video broadcast result in smaller amounts of bandwidth required to transmit the content but the receiver has to wait for the arrival of the next key frame in order to be able to display a newly selected channel. Service providers could exploit the fact that the human visual system is inert. An average recovery time of 780msec between scene changes was acceptable to even the most critical observers, when visual detail was reduced to fraction of the regular stream (Seyler & Budrikis, 1964). Further research would be needed to see if this period applies equally to channel switching on mobile devices and how which codecs could make use of this period. Displaying the logo of the upcoming channel or other tricks might perceptually shorten the wait time for users. Long wait times, for example, for downloading or on demand streaming

content should be accompanied with progress bars to help users assess the remaining time (Serco, 2006).
Because of the strong brand recognition of current TV broadcasters (e.g. CNN, BBC), it is likely that channel-centric content organisation under those brand names will prevail in mobile TV. But they could be replaced by virtual channels (Chorianopoulos, 2004) or category-centric content organisation which would group similar content from various sources under one category (e.g. news, music, movies etc.). Because of the limited space and need for fast access users will be interested in arranging content and channels according to their preferences. An electronic programme guide (EPG) which shows what programmes are currently available for viewing and what will come up might become a more important content navigation tool in mobile TV than in digital TV settings as reported in (Eronen & Vuorimaa, 2000).
Different video skipping approaches (Drucker, Glatzer, De Mar, & Wong, 2004), skimming video (Chistel, Smith, Taylor, & Winkler, 2004) and overall gist determination and information seeking (Tse, Vegh, Marchionini, & Shneiderman, 1999) have been studied in digital and standard TV settings but not in the mobile domain. When selecting from a range of programs represented by video clips playing in parallel on mosaic pages of a digital TV study found that viewers preferred interfaces that gave fewer choices and bigger pictures (Kunert & Krömker, 2006). This would have to be traded off with the necessary navigation required between pages or scrolling in order to display all possible channels of a big bouquet. Mobile TV services, which provide a mixture of live, pre-cached and downloadable content need to communicate these differences through the user interface.

Information Interactivity
Accessing additional information on mobile TV programmes is a challenge to design because of the limited screen estate. While watching regular TV some people are already making use of their mobile phones by sending SMS messages to friends to comment on what they are watching on TV. This kind of distributed co-viewing experience would be feasible on mobile TVs with large enough screens to show both the content and the textual conversation.

Digital rights management (DRM)
People have a strong sense of ownership about the content that resides on their mobile devices. Many expect to be able to capture and transfer the content to and from computers for back-up purposes or for sharing with friends (Knoche et al., 2004). Restrictive DRM approaches that run against perceived user needs will affect the experience of mobile TV.

CONTENT
The content distributed to mobile devices ranges from interactive content, specifically created for the mobile, to material that is produced for standard TV or cinema consumption. A number of studies have identified news as the most interesting content for mobile consumption (Mäki, 2005), (Knoche et al., 2004). Considering the fact that many users watch mobile TV at home there is not much reason why programmes on regular TV would not be popular on mobile devices unless they prove impractical to watch on small screens. Whereas news is of interest throughout the day participants want to watch sports, series and general entertainment, music and films on specific occasions (Mäki, 2005). Many people expect that their standard TV channels will be available on mobile TV (Serco, 2006). Time will tell whether relaying standard TV channels will be good enough for a mobile audience that is constrained when to watch, have short viewing periods and small display sizes.

Made for mobile content
Currently, content made especially for mobile use is expensive as the audience compared to broadcast television is relatively small. However, content producers adapt their content with respect to low resolutions and the typical use time, e.g. short versions of the popular TV series 24. In sports coverage for mobile devices ESPN minimises the use of long shots in their coverage (Gwinn & Hughlett, 2005) and instead uses more highlights with close-up shots.

Others produce soap operas for mobile devices that rely heavily on close-up shots with little dialogue (Guardian, 2005). However, the gain of these changes is not fully researched or understood. Research has shown that differences in the perceived quality of shot types depend on the displayed content (Knoche, McCarthy, & Sasse, 2006). Further research is required to evaluate the potential benefits of cropping for mobile TV resolutions.

Recoded content

Relatively cheap in comparison to the made-for-mobile content is the pre-encoding of cinema or TV content both in length and in size. Automatic highlight extraction from TV content (Voldhaug, Johansen, & Perkis, 2005) is a promising technique that needs to be evaluated with end-users on mobile devices.

Content-based pre-encoding can improve on the visual information and detail by: (1) cropping off the surrounding area of the footage that is outside the final safe area for action and titles and does not include essential information. (Thompson, 1998). (2) zooming in on the area that displays the most important aspects (Dal Lago, 2006), (Holmstrom, 2003). (3) visually enhancing content, e.g. by sharpening the colour of the ball in football content (Nemethova, Zahumensky, & Rupp, 2004). Research is required to rule out possible negative side-effects caused by these automated approaches.

FUTURE TRENDS

Video encoders will further reduce the amount of encoding bitrates required and will result in better perceived quality. Memory will continue to drop in price and make full PVR functionality with ample amounts of storage capacity available on mobile TV devices. Designing a mobile TV service on the edges of the coverage area might be another challenge. When viewers move in and out of the coverage area or the kind of delivery service that is provided the application will have to feature a way to gracefully switch between these different service concepts, e.g. DVB-H live streams and pre-cached content in SDMB. Intelligent cropping algorithms that enlarge parts of the content might become a solution if the content depicted on mobile TV screens is too small for the viewer. Mobile phones with video camera capabilities might make for a very different mobile TV experience if peers or groups of people start providing each other with video clips on the go.

CONCLUSION

Mobile TV is a very promising service for both customers and service providers. In order to provide the former with a satisfying Quality of Experience during potentially short interaction periods service provider will have to take into consideration a range of aspects in the creation, preparation, delivery and consumption of content on a variety of mobile platforms. It will require cooperation between all involved parties to make mobile TV as appealing as the standard TV that constitutes a necessity in many households. This chapter has presented the key factors that determine QoE for mobile TV along with previous research results which can help improve the uptake of mobile TV 2.0 in a mobile and diversified market place.

REFERENCES

Ahmed, K., Kooij, R., & Brunnström, K. (2006). Perceived quality of channel zapping. In *ITU-T Workshop on QoE/QoS 2006*.

Aldrich, S. E., Marks, R. T., Lewis, J. M., & Seybold, P. B. (2000). *What kind of the Total Customer Experience Does Your E-Business Deliver?* Patricia Seybold Group.

Ankrum, D. R. (1996). Viewing Distance at Computer Workstations. *Work Place Ergonomics,* 10-12.

Apteker, R. T., Fisher, A. A., Kisimov, V. S., & Neishlos, H. (1994). Distributed multimedia: user perception and dynamic QoS. In *Proceedings of SPIE* (pp. 226-234).

Barber, P. J. & Laws, J. V. (1994). Image Quality and Video Communication. In R. Damper, W. Hall, & J. Richards (Eds.), *Proceedings of IEEE International Symposium on Multimedia Technologies & their Future Applications* (pp. 163-178). London, UK: Pentech Press.

Beerends, J. G. & de Caluwe, F. E. (1999). The influence of video quality on perceived audio quality and vice versa. *Journal of the Audio Engineering Society, 47,* 355-362.

Biberman, L. M. (1973). *Perception of displayed information*. Plenum Press.

Chistel, M., Smith, M., Taylor, C., & Winkler, D. (2004). Evolving video skims into useful multimedia abstractions. In *Proceedings of CHI '98* ACM Press.

Chorianopoulos, K. (2004). *Virtual Television Channels Conceptual Model, User Interface Design and Affective Usability Evaluation.* PhD Unpublished thesis: Athens University of Economics and Business.

Dal Lago, G. (2006). Microdisplay Emotions. Internet, from: http://www.srlabs.it/articoli_uk/ics.htm

de Vries, E. (2006). Renowned Philips picture enhancement techniques will enable mobile devices to display high-quality TV images. http://www.research.philips.com/technologies/display/picenhance/index.html

Drucker, P., Glatzer, A., De Mar, S., & Wong, C. (2004). SmartSkip: Consumer level browsing and skipping of digital video content. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves* (pp. 219-226). New York, NY, USA: ACM Press.

Eronen, L. & Vuorimaa, P. (2000). User Interfaces for Digital Television: a Navigator Case Study. In *Proceedings of the Working Conference on Advanced Visual Interfaces AVI 2000* (pp. 276-279). New York, NY, USA: ACM Press.

Gauntlett, D. & Hill, A. (1999). *TV Living: Television, Culture and Everyday Life.* Routledge.

Guardian (2005). Romantic drama in China soap opera only for mobile phones. Guardian Newspapers Limited, from: http://www.buzzle.com/editorials/6-28-2005-72274.asp

Gwinn, E. & Hughlett, M. (2005). Mobile TV for your cell phone. Chicago Tribune, from: http://home.hamptonroads.com/stories/story.cfm?story=93423&ran=38197

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

Hellström, G. (1997). Quality measurement on Video Communication for Sign Language. In *Proc.Of 16th International Symposium on Human Factor inTelecommunications* (pp. 217-224).

Holmstrom, D. (2003). *Content based pre-encoding video filter for mobile TV.* Unpublished thesis: Umea University, http://exjob.interaktion.nu/files/id_examensarbete_5.pdf.

Horn, D. B. (2002). The effects of spatial and temporal video distortion on lie detection performance. In *Proceedings of CHI '02*.

IBM (2002). Functions of mobile multimedia QOS control. from: http://www.trl.ibm.com/projects/mmqos/system_e.htm

Jain, R. (2004). Quality of Experience. *IEEE Multimedia, 11,* 95-96.

Jumisko-Pyykkö, S. & Häkkinen, J. (2006). "I would like see the face and at least hear the voice": Effects of Screen Size and Audio-video Bitrate Ratio on Perception of Quality in Mobile Television. In G. Doukidis, K. Chorianopoulos, & G. Lekakos (Eds.), *Proceedings of EuroITV '06* (pp. 339-348). Athens University of Economics and Business.

Jumisko-Pyykkö, S., Vinod Kumar, M. V., Liinasuo, M., & Hannuksela, M. (2006). Acceptance of Audiovisual Quality in Erroneous Television Sequences over a DVB-H Channel. In *Proceedings of the Second International Workshop in Video Processing and Quality Metrics for Consumer Electronics*.

Kies, J. K., Williges, R. C., & Rosson, M. B. (1996). *Controlled Laboratory Experimentation and Field Study Evaluation of Video Conference for Distance Learning Applications* (Rep. No. HCIL 96-02). Virginia Tech.

Knoche, H. (2005). *FirstYear Report.* Unpublished thesis: University College London.

Knoche, H., de Meer, H., & Kirsh, D. (2005a). Compensating for Low Frame Rates. In *CHI '05 extended abstracts on Human factors in computing systems* (pp. 1553-1556).

Knoche, H. & McCarthy, J. (2004). Mobile Users' Needs and Expectations of Future Multimedia Services. In *Proceedings of the WWRF12*.

Knoche, H., McCarthy, J., & Sasse, M. A. (2005b). Can Small Be Beautiful? Assessing Image Resolution Requirements for Mobile TV. In *ACM Multimedia* ACM.

Knoche, H., McCarthy, J., & Sasse, M. A. (2006). A close-up on Mobile TV: The effect of low resolutions on shot types. In G. Doukidis, K. Chorianopoulos, & G. Lekakos (Eds.), *Proceedings of EuroITV '06* (pp. 359-367). Athens, Greece: Athens University of Economics and Business.

Knoche, H. & Sasse, M. A. (2006). Breaking the News on Mobile TV: User requirements of a popular mobile content. In *Proceedings of IS&T/SPIE Symposium on Electronic Imaging*.

KPMG (2006). *Consumers and Convergence*
Challenges and opportunities in meeting next generation customer needs.

Kunert, T. & Krömker, H. (2006). Proven Interaction Design Solutions for Accessing adn Viewing Interactive TV Content Items. In G. Doukidis, K. Chorianopoulos, & G. Lekakos (Eds.), *Proceedings of EuroITV 2006* (pp. 242-250). Athens University of Economics and Business.

Lemay-Yates Associates Inc. (2005). *Mobile TV Technology Discussion* Lemay-Yates Associates Inc.

Lombard, M., Grabe, M. E., Reich, R. D., Campanella, C., & Ditton, T. B. (1996). Screen Size and viewer responses to television: A review of research. In *Annual Conference of the Association for Education in Journalism and Mass Communication*.

Luther, A. C. (1996). *Principles of Digital Audio and Video*. Boston, London: Artech House Publishers.

Mäki, J. (2005). *Finnish Mobile TV pilot* Research International Finland.

McCarthy, J., Sasse, M. A., & Miras, D. (2004). Sharp or smooth? Comparing the effects of quantization vs. frame rate for streamed video. In *Proceedings of CHI* (pp. 535-542).

McCarthy, J. & Wright, P. (2004). *Technology as experience*. Cambridge, Massachusetts: MIT Press.

Nemethova, O., Zahumensky, M., & Rupp, M. (2004). Preprocessing of Ball Game Video-Sequences for Robust Transmission over Mobile Networks. In *Proceedings of the CIC 2004 The 9th CDMA International Conference*.

Neumann, W. R., Crigler, A. N., & Bove, V. M. (1991). Television Sound and Viewer Perceptions. In *Proc.Joint IEEE/Audio Eng.Soc.Meetings* (pp. 101-104).

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

Pappas, T. & Hinds, R. (1995). On Video and Audio Integration for Conferencing. In *Proceedings of SPIE - The International Society for Optical Engineering*.

Reeves, B., Lang, A., Kim, E., & Tartar, D. (1999). The effects of screen size and message content on attention and arousal. *Media Psychology, 1,* 49-68.

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

Rubin, A. M. (1981). An examination of television viewing motivations. *Communication Research, 9,* 141-165.

Selier, C. & Chuberre, N. (2005). Satellite Digital Multimedia Broadcasting (SDMB) System Presentation. In *Proceedings of 14th IST Mobile & Wireless Communications Summit*.

Serco (2006). Usability guidelines for Mobile TV design. Internet, from: http://www.serco.com/Images/Mobile%20TV%20guidelines_tcm3-13804.pdf

Seyler, A. J. & Budrikis, Z. L. (1964). Detail Perception after Scene Changes in Television Image Presentations. *IEEE Transactions on Information Theory, 11,* 31-42.

Sling Media (2006). SlingPlayer Mobile. http://www.slingmedia.com

Södergård, C. (2003). *Mobile television - technology and user experiences Report on the Mobile-TV project* (Rep. No. P506). VTT Information Technology.

Song, S., Won, Y., & Song, I. (2004). Empirical Study of User Perception Behavior for Mobile Streaming. In *Proceedings of the tenth ACM international conference on Multimedia* (pp. 327-330). New York, NY, USA: ACM Press.

Steinmetz, R. (1996). Human perception of jitter and media synchronization. *IEEE Journal on Selected Areas in Communications, 14,* 61-72.

Stockbridge, L. (2006). Mobile TV: Experience of the UK Vodafone and Sky service. Internet, from: http://www.serco.com/Images/EuroITV%20mobile%20TV%20presentation_tcm3-13849.pdf

Taylor, A. & Harper, R. (2002). Switching on to switch off: An analysis of routine TV watching habits and their implications for electronic programme guide design. *usableiTV, 1,* 7-13.

Thompson, R. (1998). *Grammar of the shot*. Elsevier Focal Press.

Tse, T., Vegh, S., Marchionini, G., & Shneiderman, B. (1999). An Exploratory Study of Video Browsing User Interface Designs and Research Methodologies: Effectiveness in Information Seeking Tasks. In *Proceedings of the 62nd ASIS Annual Meeting.* (pp. 681-692).

Vatakis, A. & Spence, C. (2006). Evaluating the influence of frame rate on the temporal aspects of audiovisual speech perception. *Neuroscience Letters (submitted).*

Voldhaug, J. E., Johansen, S., & Perkis, A. (2005). Automatic Football Video Highlights Extraction. In *Proc of NORSIG-05*.

Wang, D., Speranza, F., Vincent, A., Martin, T., & Blanchfield, P. (2003). Towards Optimal Rate Control: A Study of the Impact of Spatial Resolution, Frame Rate and Quantization on Subjective Quality and Bitrate. In T. Ebrahimi & T. Sikora (Eds.), *Visual Communications and Image Processing* (pp. 198-209).

Westerink, J. H. & Roufs, J. A. (1989). Subjective Image Quality as a Function of Viewing Distance, Resolution, and Picture Size. *SMPTE Journal*.

Winkler, S. & Faller, C. (2005). Maximizing audiovisual quality at low bitrates. In *Proc.of Workshop on Video Processing and Quality Metrics*.

Zillman, D. (1988). Mood Management: Using Entertainment to Full Advantage. In L.Donohew, H. E. Sypher, & E. T. Higgins (Eds.), *Communication, Social Cognition, and Affect* (pp. 147-172). Hillsdale: Erlbaum.