ACCEPTANCE OF MOBILE TV CHANNEL SWITCHING DELAYS

eingereicht von:

Werner Robitza

BACHELORARBEIT

Studienrichtung: Medieninformatik
Fakultät für Informatik der Universität Wien

Begutachter:
Helmut Hlavacs

Eingereicht von:
Werner Robitza
Wien, im März 2010

Statement of Authentication

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in whole or in part, for a degree at this or any other institution.

Vienna, March 2010 Werner Robitza

Abstract

User interaction has always been one of the most crucial points when evaluating the quality of experience of a service. Mobile television, especially when received over DVB-H or in other small bandwidth transmission environments such as UMTS networks, imposes some restrictions to this interaction, since switching from one channel to another requires an indefinite amount of time. In this thesis, not only these technical limitations that lead to reduced interactivity are presented, but also an experiment that has been carried out in order to determine the level of channel switching acceptability from a user's point of view. Moreover, methods are presented to increase this acceptability.

Contents

List of Figures

List of Tables

1 Introduction

Since the introduction of DVB-H in 2004 (Digital Video Broadcasting: Handhelds, [1]) it became one of the most widely used mobile television technologies aside UMTS-based streaming in Europe. All over the world, other mobile television technologies have been introduced, such as ISDB-T, T-DMB and MediaFLO.

DVB-H for instance addresses the energy consumption issues of its predecessor DVB-T by introducing features such as a burst transmission in order to compensate for the smallness and the limited battery capacity of the receivers. This burst technology – together with other factors – increases the time needed for the user to switch between several TV channels by a large amount. Also, when streaming mobile television over UMTS based networks, the limited network bandwidth might prolong the time to display a video stream correctly. Since channel switching is almost inevitable when the user does not know which program they want to see, longer channel switching times lead to a poor quality of experience.

Therefore, in this thesis the issue of too long zapping delays is assessed by describing the limitations that result from burst transmission. The acceptance of zapping delay is examined and it is shown how it can be improved by using different gap fillers.

This thesis is structured as follows: In the following section, related work in the field of waiting times is summarized. In the third chapter, an overview of channel switching in Mobile TV is given, with respect to all major technologies including UMTS. As an example of factors that lead to channel switching delay, DVB-H Time Slicing is explained in more detail in chapter four. Also, methods are presented by which channel switching delays can be estimated through basic parameters. In the fifth chapter, the design of a user survey is explained. The results of this study will be presented in chapter six. Chapter seven summarizes these results with regards to future improvements that could be achieved in increasing the acceptance of channel switching times in DVB-H and other Mobile TV technologies.

2 Related work on waiting delays

Channel switching delays leave the user waiting for a particular response to a stimulus – i.e. pressing a button and waiting for a video to start playing. Because of this, they can be looked upon as general issues of human-computer interaction (HCI) or treated in a similar way as website loading delays in internet based services. As of today, there exist many survey results on HCI delays and tolerated waiting times for internet browsing (web page loading times).

Advices to system response times generally do not seem to have changed over time. In [2] and [3], 0.1 seconds were defined as the limit for users to feel an instantaneous reaction to a response. A delay of 1 second is considered as short enough for the user's flow of thought to remain continuous. When a delay of 10 seconds or more occurs, feedback should be included so as to inform the user about ongoing progress. The use of progress bars is highly recommended for such purposes, as explained in [4]. An overview of this topic has been given in [5].

A very well known source covering the topic of web-based loading times is [6], where the so-called "8-second rule" was introduced, meaning that a waiting time of 8 seconds or more is too much. The authors also stated that a fifth of users who quit loading the website will actually never visit it again. Similar results were found in the year 2000, where the authors of [7] stated that a delay of 10 seconds would lead to bad ratings. Even more, four years later, it was estimated that the tolerated waiting time for web pages lies between 5 and 8 seconds, according to [8]. Of course it needs to be said that with the arrival of broadband internet connections such as DSL, the tolerated waiting times become shorter. This is why it is generally believed that the 8-second rule does not apply anymore.

Another report states that in broadcast television services, the zapping time should not exceed 2 seconds and a black screen should never be considered for gap intermission [9]. This is a fact that obviously has to be considered for mobile television, too, although it could be possible that users are more tolerant to increased waiting times. Lastly, in [10] a model for mapping zapping time and MOS in IPTV is presented. A similar model could be created for mobile television.

A handful of works already address the special problem of channel switching time in DVB-H. One of them is the DVB-H specification itself [11], which presents basic calculations on Time Slicing and burst transmission. The authors of [12] have also analyzed the optimal channel switching delay with respect to an optimal power saving curve, since these two factors affect each other in a nonlinear fashion, as explained later. They also analyzed the decoder and buffer refresh times with respect to a given video quality. In [13], a video splicing method for reducing the channel switching time was proposed – however, it is questionable if the method is feasible, because two separate video streams would have to be encoded at the distribution source.

In this thesis, insight is provided into the causes of channel switching time, focused on DVB-H. The effectiveness of different gap filling types are investigated as a contribution to mobile television recommendations.

3 Channel switching delay in Mobile TV

In this thesis, channel switching time is defined as the time between a user input (i.e. a key or button press) and the playback of the desired video channel. In terrestrial analogue television, there is no inherent delay when zapping through channels. Switching appears to happen instantaneously or at least within 1 second. With the advent of digital television transmitted by satellite, cable or over terrestrial channels, these delays have become noticeable, because more time is needed to decode the video from the digital input stream. Yet, the delays seem acceptable, because they don't exceed the length of a few seconds at maximum.

3.1 Broadcast-based mobile television

In broadcast-based mobile television, every device receives the same signal, and all available channels are transmitted over one carrier from which the user can choose one. The most important technologies today are DVB-H, ISDB-T, T-DMB and Qualcomm's MediaFLO. In all of these products, the channel switching delay suffers from the implementation of the following techniques:

In ISDB-T[1], used in Japan and Brasil, every channel is transmitted in its own frequency range. Then, the channel signal is split into bursts of 13 segments each, where only one segment is needed for decoding Mobile TV video, but all segments need to be received for HDTV decoding. Channel switching delay occurs because it is necessary to wait for every 13th segment to assemble the stream. However, the delays are low because of the fixed frequencies where channels are located.

[1] http://www.dibeg.org/techp/techp.htm

T-DMB, mostly used in China, incorporates Micro Time Slicing, similar to DVB-H [14], but because of the smaller units of burst transmission, channel switching delays are rather low.

In MediaFLO[2], statistical multiplexing is used. Channels are transmitted as so-called Multicast Logical Channels (MLC) and multiplexing happens at the physical layer.

All of the technologies can reach a minimum delay of 1.5 seconds – in MediaFLO, most of the delays are guaranteed to be under 2 seconds. This technology seems very promising because of the low delay, however it relies on a proprietary platform. In DVB-H, this minimum delay can not be reached and channel switching delays might be even longer due to Time Slicing.

3.2 UMTS-based streaming television

In UMTS streaming television, as opposed to broadcast-based TV, channels are streamed through unicast IP connections and therefore delivered to each receiving device separately. The video streams are mostly encapsulated in Realtime Transport Protocol (RTP) packets and connected to by Session Description Protocol (SDP). Each stream is therefore accessible by a URL that identifies the resource. These URLs could be manually entered by the user or accessed by websites that include hyperlinks to the URL. Channel switching is then performed by navigating to another URL. As there is more than one possible procedure for doing so, an example focusing on one operator and mobile phone is given.

The Austrian operator A1 provides a web portal called "Vodafone Live", where TV channels can be selected from a list featured on a website. This web page is accessible from the mobile phone's browser and a few clicks, so generally, it takes at least 30 seconds to tune into a program. When the link corresponding to a TV channel is clicked, the SDP of the selected program will be opened and the mobile phone's internal video player will start connecting to the resource. When the buffer has finished loading enough content, the video is being played back.

In order to switch to another channel, the user needs to exit the video player by pressing a specific button. Then a hyperlink needs to be followed to play the next or previous channel – this hyperlink would be displayed on the website being displayed when the player has been quit. Upon following the link, the player would open again

[2]http://www.mediaflo.com

and start to connect to the video resource. Choosing another channel (i.e. not the previous or next one) requires going back in the web browser history to visit the site that displays all available channels. However, the URLs of favorite streams could be bookmarked and loaded on demand.

As clearly visible, when watching mobile television over UMTS, changing channels requires more than just the click of one button. The need to connect to a different stream increases channel switching times by a significant amount of time, because the maximum bandwidth on UMTS networks is the largest bottleneck. Several other factors include the video bandwidth, the performance of the decoder, the time needed to start the video player, and so forth. Because of this, it is not easy to calculate channel switching times.

Of course, using a dedicated mobile television application could simplify the task of switching through channels – for example, two buttons would be available for navigating to the previous or next channel. A program list from which channels can be selected could also be included.

4 Channel switching and Time Slicing in DVB-H

4.1 DVB-H Overview

Digital Video Broadcasting for Handhelds was developed by the European Telecommunication Standards Institute (ETSI) in 2002. The standard was released in 2004 [1]. DVB-H originates from DVB-T, described in [15]. Several other standards that make up DVB-H include [16] and [11]. Also, DVB-H uses the MPEG-2 transport stream technology, which has been standardized in [17] by the ISO/IEC. Several other papers specify technical recommendations and implementation guidelines.

One of the key problems originally leading to the development of DVB-H was that efficient use of mobile television requires small sized and lightweight receivers. Smaller and lighter devices also restrict the usage of extremely high capacity batteries. DVB-H therefore introduced the Time Slicing technology as an effective mean of dealing with these power consumption issues. It is explained in more detail in this chapter.

4.2 Time slicing

Time Slicing is a form of time division multiplexing that allows the transmission of multiple DVB-H channels on one single MPEG-2 elementary stream (ES). As described in [11], it works as follows: Video data is streamed through RTP and IP. Already at this stage, it could be defined as a broadcasting technology, because the IP stream is a broadcast stream. Then, the datagrams are encapsulated in so-called MPE datagrams (Multi Protocol Encapsulation). MPE packets are collected to form an MPE-frame, which is forward-error corrected by using Reed Solomon codes.

A single MPE-FEC frame (maximum size 2 MBit) is then transmitted in a burst, then a pause occurs. After this gap, the next frame is transmitted. This is repeated periodically for every DVB-H channel that should be transmitted, in a way that each burst is sent within the gap of another burst. The time between the end of one burst and the beginning of the next one is often called "Time Slice Off Period"; equally, the time of the burst duration is referred to as "Time Slice On Period". All the bursts form an MPEG-2 elementary stream.

A schematic of this technology is visible in Figure 1. Because both DVB-H and DVB-T rely on MPEG-2 transport streams, they can be transmitted on the same carrier. In this example, DVB-H occupies one MPEG-2 elementary stream, the other three DVB-T channels use three ES. In the DVB-H PS, four channels are multiplexed. Their bursts are marked in different colors.

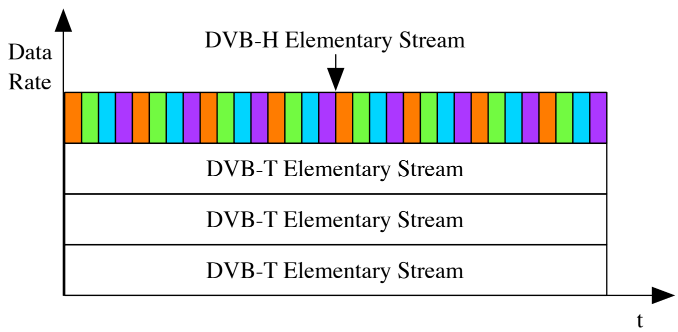

Figure 1: DVB-H and DVB-T multiplexing

When Time Slicing is used, each MPE packet in the burst carries a timestamp called "Delta T" (or ΔT). Because there is more than one MPE packed in a frame, even the loss of a single packet would still guarantee that the receiver obtains the Delta T information. Delta T equals to the absolute time difference between the start of two consecutive bursts belonging to the same logical channel. It is measured in milliseconds, therefore no clock synchronization between sender and receiver is necessary. The main advantage of this bursty transmission is the fact that the receiver knows through the parameter Delta T when to expect the next burst. In the "Off"-period, the receiving unit can be switched off automatically and therefore save energy. Depending on a set of parameters used in Time Slicing, the energy reduction can be up to 90% [18] when compared to a continuous transmission.

However, the bursty transmission leads to prolonged channel switching times. This is because of the fact that the receiver has to wait for the desired burst in order to start decoding its content.

4.3 Burst parameters

The basic burst parameters can be seen in Figure 2. The constant (or average) bitrate of the source video needs to be split into bursts. Because of the gap between the bursts (OFF-Time, or T_{OFF}), their bitrate is larger than the constant bitrate itself. The burst size is then calculated by the burst duration (also ON-Time or T_{ON}) and the burst bitrate.

Most of these parameters could be defined freely, but of course, some are implied already. For example, the constant or average data bitrate is known before burst scheduling. Then, the burst time can be calculated by the size and the bitrate of each burst:

$$T_{ON} = \frac{Size_{Burst}}{Rate_{Burst} \cdot Overhead} + T_{Sync} \tag{1}$$

The Off-Time will be calculated by the burst bitrate and the average or constant video bitrate, and subtracting the On-Time:

$$T_{OFF} = \frac{Rate_{Burst}}{Rate_{Const}} - T_{ON} \tag{2}$$

Overhead is a factor that compensates for the additional headers included in the MPEG-2 Transport Stream. It should be set to 0.96. T_{Sync} is the synchronization time needed for the receiving unit to tune into the DVB signal again. Its value can

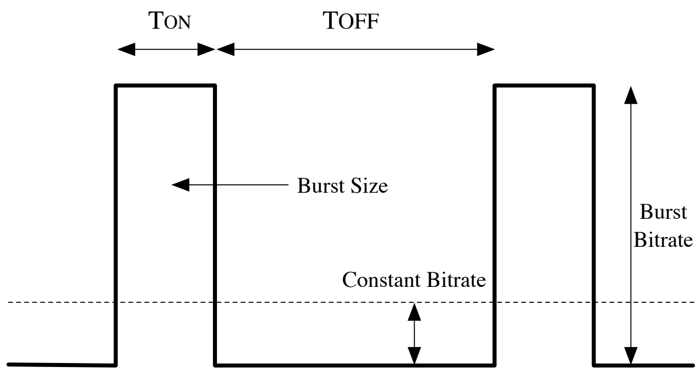

Figure 2: Common burst parameters in DVB-H

change depending on the device used and lies in the range from 80ms to 200ms (see [18] and [12], respectively).

According to [18], the potential for saving energy can be calculated by the following equation, in percentage P, where $J_{\Delta T}$ denotes the jitter of Delta T:

$$P = (1 - \frac{(T_{ON} + (3/4 \cdot J_{\Delta T}) \cdot Rate_{Const} \cdot Overhead)}{Size_{Burst}}) \cdot 100\% \qquad (3)$$

The authors of [12] however propose an equation that does not include jitter:

$$P = (1 - Rate_{Const} \cdot (\frac{1}{Rate_{Burst} \cdot Overhead} + \frac{T_{Sync}}{Size_{Burst}})) \cdot 100\% \qquad (4)$$

4.4 Calculating channel switching times

In a simplified model, only On and Off times derived from the parameters stated before are considered for calculating channel switching delay. In fact, there are more factors possibly increasing the delay.

4.4.1 Simple Model

In the simplified model it can be assumed that upon initiating the channel change, the receiver has to wait for the next desired burst to arrive. The burst has to be received as a whole because the data contained has been interleaved and error-corrected, so the smallest possible delay would be equal to T_{ON}. When the burst is received completely, the video playback starts immediately. To perform calculations in this model, we assume that

- the average synchronization time T_{Sync} is close to 0 (the time needed by the receiver to power up and lock to the signal)
- each burst contains at least one picture frame that allows instantaneous decoding without buffering,
- receiving conditions are good enough to allow reconstruction of all received data in case of a transmission error, and
- there is no inherent delay in the decoding unit.

In the optimal case, the receiving unit is ready at the time the desired burst is transmitted. Then, the absolute minimum delay would be equal to T_{ON}. The worst case scenario results in the receiving unit starting immediately *after* the begin of a desired burst the waiting time is as long as $2 \cdot T_{ON} + T_{OFF}$. However, in weak reception conditions, frame losses lead to much longer delays than mentioned, with each lost frame resulting in an additional delay of $T_{ON} + T_{OFF}$. This is not considered here.

Since the user input is absolutely independent from the actual physical transmission, we can model the channel switching delay as a uniform distribution. Let us denote

$$D \in [T_{ON},\ 2T_{ON} + T_{OFF}] \tag{5}$$

as the delay. The average delay (and its variance) equals to

$$E(D) = \frac{3 \cdot T_{ON} + T_{OFF}}{2} \tag{6}$$

$$Var(D) = \frac{(T_{ON} + T_{OFF})^2}{12} \tag{7}$$

and the probability of a delay smaller than D is given by

$$P(x \leqslant D) = \frac{x - T_{ON}}{T_{ON} + T_{OFF}}. \tag{8}$$

4.4.2 Extended model

As already mentioned in Section 4.4, some assumptions have to be made for such a simple model. This has also been mentioned in [12]. In order to statistically model channel switching times more accurately, one would need to consider the following aspects: When tuning into a DVB-H stream, a synchronization time has to be added to the absolute user input time when calculating delays. This time includes delays from the operating system of the receiving device as well as the DVB-H sync time T_{Sync}. We expect this time to range between 100ms up to 500ms (depending on the receiver).

Upon receiving a whole burst it is necessary to decode the data stream, especially the forward error correction bits. In a lossy transmission environment, it is also sometimes necessary to recover lost bits by performing Reed Solomon calculations. DVB-H specific headers have to be decoded too, since they include the data necessary to access the IP layer. Then, IP datagrams with their included UDP and RDP data have to be analyzed. The video and audio stream contained have to be synchronized before displaying them.

Also, video playback can only start when the burst contains enough data to be displayed, meaning that an IDR (Instantaneous Decoder Refresh) picture should be inserted into every burst [13]. However, the stage of encoding the video happens before Time Slicing is applied, so depending on the coding parameters, a burst might not carry an IDR picture. Then, the waiting time would be increased by another burst cycle.

4.5 Measuring channel switching times

We measured DVB-H channel switching times in the city of Vienna using a Nokia N96 mobile phone. These results have been used for the user survey explained later. Samples have been taken in a building near a window.

The average delay resulted to be 7.4 seconds (median 6.8), with a minimum of 4.5 and a maximum of 15 seconds. The standard deviation was 1.8.

Other sources show similar results. In [19], for example, DVB-H channel switching delays are distributed from 3 to 39 seconds with an 80% decile of 6 seconds.

5 User study

5.1 Motivation

To our knowledge, there is no data available dealing with the acceptance of channel switching in Mobile TV. Although the common delays have been described, it is not known to which extent longer waiting times will affect the acceptance of Mobile TV services. We can however safely assume that longer waiting times result in lesser satisfaction from mobile television users, just like longer web page loading times have negative influence on the quality of experience. It is also valid to say that once a user is highly disappointed by a service, it is very likely that they will not use it again or at least less often. The interest to pay money for such a service might become lower, and currently, most of the available Mobile TV services require paying a monthly fee. Therefore, less acceptance results in fewer users and also in a monetary loss for the Mobile TV distributor and the device vendors. Especially in the early stage of Mobile TV deployment, good service quality needs to be assured for people to adopt the technology.

As mentioned before, the waiting times are implied on a technical level and can not be improved easily, especially when it is mandatory to adhere to a standard like DVB-H. Although newer standards are planned, it is necessary to render channel switching times more appealing to the users than for instance just showing a black screen for an indefinite amount of time until the next program is available. There are no technical limitations to these methods. We also state the hypothesis that a more “interesting” gap type will render the subjective duration shorter.

A recent experiment addressed the same channel switching issues that occur in IP-based television [20]. The authors have also introduced similar methods to those presented herein (advertisement logo), but we believe that

- it is necessary to conduct such an experiment on a mobile phone, focused on the restrictions of mobile transmission,
- more different methods of increasing QoE need to be used, and
- more randomization in the experiment is needed.

To conclude, the primary goal of this experiment was to find out how users would rate the different gap types in comparison to each other and on a subjective scale. The individual channel switching times have not been rated. The secondary goal was to assess whether they felt that one or an other gap type was annoying. Furthermore, user experience related questions have been asked.

5.2 Experiment rules and methodology

We conceived an experiment that adopts the following rules: The test session should last up to half an hour, but not longer. All participants should be non-expert and therefore not involved with the particular technical details.

The experiment begins with an introduction on the topic and an explanation of the rules. Then, the first part of a questionnaire is filled out. Observers will see a test sequence which they can abort when they feel confident enough. After that, they choose the content type they want to see as explained in section 5.3.1. They are presented four sequences of videos, each sequence using a different gap type (see section 5.3.2). They are allowed to switch to the next "channel" whenever they want to. In between each simulated channel, a gap occurs. An example of two single experiment session can be seen in Figure 3. After all the sequences are finished, the questionnaire is completely filled out.

5.3 Experiment parameters

Three parameters have been selected, which are explained in more detail in this section.

5.3.1 Content type

The content type is the first parameter: From the experience of previous subjective tests it could be noted that personal preference of specific content types can have noticeable impact on the resulting quality scores. In order to prevent users from giving a bad score to the gap that occurs before or after a content type they would not like to see, they can choose between four different types, which reflect some of the most common contents seen on television and mobile television: News, Sports, Music Videos and Soap Operas (Sitcoms) [21,22]. Samples from real television programs and widely known music videos have been used for presentation.

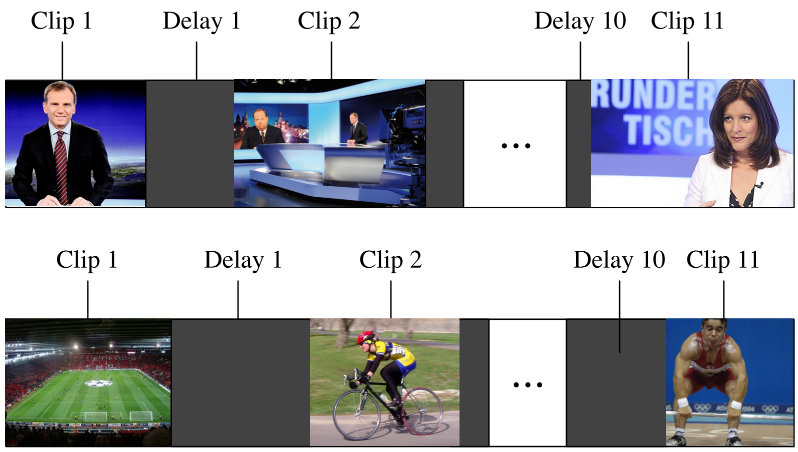

Figure 3: Example of two experiment sessions – Content: News, Sports; Gap type: Black screen

5.3.2 Gap type

During the experiment four different gap types were shown to the users:

- Black screen
- "Please wait" animation
- Static advertisement logo
- Short commercial movie

Their order of presentation is randomized for each participant. So, for example, observer 1 will see {A,B,C,D}, whereas observer 2 will see {A,B,D,C} and so on.

The most simple form of a gap intermission consists in showing a *black screen* until the next channel starts. Here, the system provides no audiovisual feedback to the user stimulus at all. We state the hypothesis that showing a black screen is possibly the least distracting gap type, however we also assume that it results in a feeling of losing control over the device and therefore will not be rated as good as other gaps.

An *animation* showing the words "Please wait..." is a more commonly used gap type which serves one purpose: it provides feedback to the user. Its message is to inform the user that something is going on and there is progress being made. Scrolling messages and progress bars are used often in these contexts [4]. However, an indeterminate progress bar has to be used in DVB-H, e.g. without a “remaining seconds” or “percent done” indicator, because it can not be known when the next desired burst will appear and if it includes all the necessary data to display the video.

Two additional gap types will be introduced: The first is a *static logo* of a random brand, e.g. a mobile phone vendor or a fast food chain. These logos will not move in any way, they will be presented until the start of the next channel without any sound or moving elements.

The second type is a *short commercial movie* that is played until the end before the next channel starts. For example, advertisements could be preloaded and buffered by UMTS connection or through a DVB-H data carousel (for users without paid UMTS data traffic). Upon switching a channel, they would be played back instantaneously from the mobile phone memory. Channel switching ads have already been implemented by the online TV vendor “Zattoo”[3], but there is no data available for experimental analysis. As for advertisements, we assume that they might distract the user, nevertheless they could render the prolonged channel switching times more acceptable because there is an audiovisual stimulus trying to attract their attention.

5.3.3 Gap length

As presented in Sections 3 and 4.5, there is no fixed channel switching time in DVB-H. Instead, the delay varies between certain boundaries, although in theory, there is no upper boundary, given enough packet loss. We chose to show a set of channel switching times based upon our measurements in Section 4.5, but with a lower bound as described in Section 3. These are the selected delays (in seconds): $\{1.5, 2.5, 3.5, 5, 7.5, 10, 15\}$. The subset of $\{3.5, 5, 7.5\}$ is shown twice per session. This results in 10 gaps per session. There are two randomized sets for each combination of content and gap type from which each user randomly saw one.

5.4 Technical details

For presenting the videos, a Nokia N96 mobile phone was used. It features a 320x240 pixel backlit color screen and also allows for receiving DVB-H streams through built-in DVB-H hard- and software. In order to simulate a Mobile TV application, videos and gaps were randomly ordered in a playlist and played through the commercial

[3]http://www.zatoo.com

software *Core Player*[4], which is an audio and video player that supports the playback of playlists. The participants would watch for a few seconds, then press the "next"-button to jump to the next video, see a simulated pause, and then wait for the next video to play.

6 Results

In the questionnaire handed out to the participants, their personal data was captured (sex, age). They were also asked about their general feeling at the day of the experiment. To this date, ten valid subjects have taken part in the experiment, aged between 21 and 39, with a mean of 26 years. There have been two female participants.

6.1 Experience and wishes for Mobile TV

The first topic-related question was whether users had any experience with mobile television yet. Also, they were asked which content type they would like to see if they owned device capable of receiving Mobile TV. Then, the observers could choose the specific content type of the experiment.

To a majority, participants have not had any experience with mobile television yet, although most of them would describe themselves as at least a bit "technophile" (i.e. always open to new technologies). 3 of 10 persons had used mobile television before (in any context). This seems not much when considering the fact that DVB-H has been largely promoted in Austria during the European Soccer Championship 2008. Also, UMTS-based TV streaming services are available from every major network operator and sometimes even included in monthly fees. Most of the common mobile phones sold support at least one form of Mobile TV.

After the test, the participants were asked how they rated the overall quality of the simulated Mobile TV service, especially in terms of video quality. The specific question was "Would you use it in real life?". The Mobile TV service simulated in the experiment was received reasonably well. It has to be noted that most of the critique addressed the "cheap feeling" of the receiving device (especially the plastic buttons). The results can be seen in Table 1.

[4]http://coreplayer.com

	Answer count
Very likely	3
Likely	2
Maybe	3
Probably not	2
Definitely not	0

Table 1: Service quality ("Would you use the service again?")

When asked which content type they would like to see, the following categories were mentioned (with their appearance count). The results show that the initial content choice seemed appropriate and are shown in Table 2.

	Answer count
News	4
Sports	3
Series, Soaps	2
Music videos	2
Documentations	1

Table 2: Desired content types ("Which content type would you like to see?")

6.2 Subjective sense of time

The experiment subjects were asked how long they thought the gaps were (minimum, maximum and average length). Also, it was examined how long they thought they were willing to wait for the next channel.

	Guessed	Actual
Minimum	1.3	1.5
Maximum	7.6	15
Average	3.65	6.4

Table 3: Guessed vs. actual gap lengths (in seconds)

Interestingly enough, the minimum gap length was estimated to be 1.3 seconds by average (+/− 0.39 at 95% confidence interval), whereas the actual minimum length was at least 1.5 seconds, including delay from the operating system and the video player itself.

The maximum gap length however was thought to be 7.6 seconds (+/− 1.21). This is almost exactly half of the real length (15 seconds + delay). Also, the guessed average length was 3.65 seconds (+/− 0.55) compared to the actual 6.4 seconds. In average, people said that they would wait for 3.15 seconds (+/− 0.97).

This shows us that the observers' subjective feeling of time does not correlate with the real passed time. This effect may actually lead to a higher acceptability of waiting times because users might not be aware of their length. Gaps even as long as 15 seconds may seem shorter than they are.

6.3 Gap types

After the experiment, people were asked to order the different gap fillings presented from best to worst. Moreover, they were told to score the fillings in detail on a MOS based scale as proposed in [23], with a set of {Excellent, Good, Fair, Poor, Bad}. Scores were also taken after presentation of each gap type for comparison purposes.

In Table 4 and Figure 4 average opinion scores are presented. The first row refers to the scores taken immediately after a video session, the second row shows the scores taken after the experiment. As clearly visible, there is no significant difference between both measurements. Also, it has to be noted that the gap types lie in an ordinal scale, where (from left to right) the amount of information provided to the user is increased.

	Black	**Please Wait**	**Logo**	**Commercial**
After viewing	1.7	2.9	3	3.8
After experiment	1.6	2.8	3	3.6

Table 4: MOS for gap types

Based on a contingency table (independent multinomial samples) a Fisher exact test on both available datasets shows significant results: The null hypothesis stating that the gap type does not have an effect on the opinion scores has to be rejected for both measurements with $p = 0.011$ for measurement 1 (immediately after presentation) and $p = 0.006$ for measurement 2 (after the experiment). The effect of changing the gap type can also be seen in Figure 4.

It is very obvious that the commercial scores best for all gap types. On the rank scale, it was also voted best six out of ten times – in contrast to the black screen, which

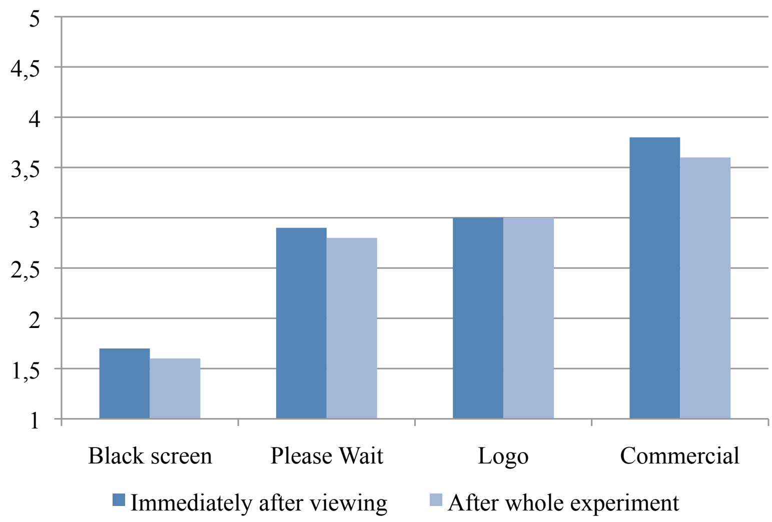

Figure 4: MOS for gap types

scored fourth place eight times. A reasonable alternative to the commercial is the logo, which scored second place six times.

6.4 Annoyance level

The final question was: "Have you found any of the gap fillings annoying or unpleasant?". For each gap type, a set of the following answers was available: {very annoying, little annoying, not annoying, okay}. Also, they were given the choice to say "I don't know.".

The results are given in Table 5 and Figure 5 and clearly show that the black screen was found to be very or at least little annoying. Also, the "please wait" animation was judged to be a little annoying by the subjects. The same applies to the static logo. Only the commercial was not considered annoying by the majority of observers – the two occurrences of "very annoying" are from those observers who generally disapprove of advertisement because they find it annoying in every life situation. This leads to the conclusion that the tolerance for the need to watch commercials in between channel switching is rather high. Only one person admitted that seeing commercials in every break would be stressful after some time.

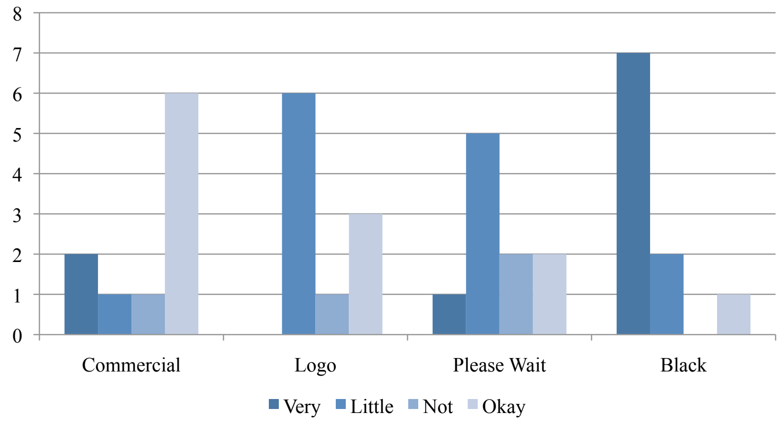

Figure 5: Annoyance level for each gap type

	Very	Little	Not	Okay
Commercial	2	1	1	6
Logo	0	6	1	3
Please Wait	1	5	2	2
Black Screen	7	2	0	1

Table 5: Annoyance for each gap type (occurences)

7 Conclusion and future work

From the currently available experiment data the following conclusions can be derived:

Firstly, it is undeniable that changing the gap intermission when switching channels has a result on user experience. A static black screen appears to be the most dissatisfying gap type because it provides no feedback to the user stimulus and there is no sign of progress. It should be avoided altogether. However, a simple animation showing a message like “please wait” increases the acceptability, because there is a feeling of progress. Also, a static logo will be interesting and therefore shorten the felt delay.

Of all the available options, a short commercial movie seems to be the most acceptable intermission. The only drawback is the fact that it will never be known in advance how long the gap will take. This has also been mentioned in Section 4.4. For instance, multiple short commercials could be shown in one longer gap, or advertisements could include hyperlinks to online content. Judging from Table 4, the logo seems to be the best compromise when commercials are not an option. Another idea would be to insert parts of an electronic program guide (EPG) into the intermission so as to inform the users about the ongoing program. Users could then switch to another channel before even seeing a video.

In general, a financial model in which Mobile TV distributors generate revenue through channel switching advertisement could be feasible. Still there are some people who might disapprove of commercials — they should have the option to turn off advertisement for a small monthly fee.

Secondly, it seems as if people misjudge the length of a pause. These errors are very significant, but they also show that people are more likely to accept longer pauses than they personally think they are.

In future work, continuous data will have to be analyzed in order to assess a time period which would constitute the maximum gap length that is still acceptable for most users. Other than that, more extensive surveys should be carried out, measuring the subjective sense of time and taking into account more types of content and gap intermissions.

References

[1] *EN 302 304 V1.1.1, Digital Video Broadcasting (DVB); Transmission System for Handheld Terminals (DVB-H).* ETSI, 2004.

[2] R. Miller, "Response time in man-computer conversational transactions," *AFIPS Fall Joint Computer Conference*, vol. 33, pp. 267–277, 1968.

[3] S. Card, G. Robertson, and J. Mackinlay, "The information visualizer: An information workspace," *CHI Proceedings*, pp. 181–188, 1991.

[4] B. Myers, "The importance of percent-done progress indicators for computer-human interfaces," *CHI Proceedings*, pp. 11–17, 1985.

[5] J. Nielsen, *Usability Engineering.* Morgan Kaufman, 1994.

[6] *The Economic Impacts of Unacceptable Web-Site Download Speeds.* Zona Research Inc., 1999.

[7] A. Bouch, A. Kuchinsky, and N. Bhatti, "Quality is in the Eye of the Beholder: Meeting Users' Requirements for Internet Quality of Service," *CHI Proceedings*, pp. 297–304, 2000.

[8] F. Nah, "A study on tolerable waiting time: how long are Web users willing to wait?" *Behaviour and Information Technology*, vol. 23, no. 4, pp. 153–163, 2004.

[9] "Triple-play Services Quality of Experience (QoE) Requirements," DSL Forum, Tech. Rep., 2006.

[10] R. Kooij, K. Ahmed, and K. Brunnström, "Percieved Quality of Channel Zapping," *IASTED Proceedings*, 2006.

[11] *EN 301 192 V1.4.2, Digital Video Broadcasting (DVB); DVB specification for data broadcasting.* ETSI, 2008.

[12] M. Rezaei, I. Bouazizi, V. Vadakital, and M. Gabbouj, "Optimal Channel Changing Delay for Mobile TV Over DVB-H," May 2007, pp. 1–5.

[13] M. M. Hannuksela, M. Rezaei, and M. Gabbouj, "Video Splicing for Tune-in Time Reduction in IP Datacasting over DVB-H," in *IEEE BMSB*, 2006.

[14] *ETSI TS 102 428 V1.1.1, Digital Audio Broadcasting (DAB); DMB Video Service, User Application Specification.* ETSI, 2005.

[15] *EN 300 744 V1.6.1, Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television.* ETSI, 2009.

[16] *EN 300 468 V1.9.1, Digital Video Broadcasting (DVB); Specification for Service Information (SI) in DVB systems.* ETSI, 2008.

[17] *ISO/IEC 13818-1, Information technology – Generic coding of moving pictures and associated audio information: Systems.* ISO/IEC, 2007.

[18] *ETSI TR 102 377 V1.2.1, Digital Video Broadcasting (DVB); DVB-H Implementation Guidelines.* ETSI, 2005.

[19] Directique, "Digital Multimedia Uses Barometer - Mobile TV - DVB-H / 3G Comparative Study," September 2009.

[20] B. Godana, R. Kooij, and O. Ahmed, "Impact of Advertisements during Channel Zapping on Quality of Experience," *ICNS*, 2009.

[21] S. Jumisko-Pyykkö and M. M. Hannuksela, "Does context matter in quality evaluation of mobile television?" in *MobileHCI.* ACM, 2008, pp. 63–72.

[22] K. O'Hara, A. S. Mitchell, and A. Vorbau, "Consuming video on mobile devices," in *CHI Proceedings.* ACM, 2007, pp. 857–866.

[23] *ITU-T Recommendation P.800: Methods for subjective determination of transmission quality.* ITU, 1996.