ACCEPTANCE OF MOBILE TV CHANNEL SWITCHING DELAYS

W. Robitza, S. Buchinger, H. Hlavacs

University of Vienna
Faculty of Computer Science, Department for Distributed and Multimedia Systems
Lenaugasse 2/8, 1080 Vienna, Austria

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. It is therefore necessary to describe the technical foundations leading to these delays. Moreover, methods need to be developed to render channel switching times more appealing.

Index Terms— Mobile TV, Channel Switching, Channel Delay, User Interaction, Student Experiments

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. DVB-H addresses the energy consumption issues of DVB-T by introducing features such as 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 DVB-H 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.

In this paper, we want to address the issue of too long zapping delays by describing the limitations that result from burst transmission. Then, we examine if the acceptance of zapping delay and show how it can be improved by using different gap fillers.

2. RELATED WORK
Channel switching delays can be treated in a similar way as delays in internet based services. As of today, there exist many survey results on tolerated waiting times for internet browsing (web page loading times). For instance, the authors of [2] stated that a delay of 10 seconds could lead to bad ratings. Even more, it was estimated that the tolerated waiting time for web pages lies between 5 and 8 seconds, according to [3]. A very well known source is [4], where the “8-second rule” was introduced. The authors stated that a fifth of users who quit loading the website will actually never visit it again. 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 [5]. 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 [6] a model for mapping zapping time and MOS in IPTV is presented.

A handful of works already address the problem of channel switching time in DVB-H. One of them is the DVB-H specification itself [7], which presents basic calculations on Time Slicing and burst transmission. The authors of [8] 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. They also analyzed the decoder and buffer refresh times with respect to a given video quality. In [9] a video splicing method for reducing the channel switching time was proposed.

We provide insight into the causes of channel switching time of DVB-H and investigate the effectiveness of different gap filling types.

3. CHANNEL SWITCHING DELAY IN MOBILE TV
In this paper, channel switching time is defined as the time between a user input (i.e. a key 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. With digital television, these delays have become noticeable, because more time is needed to decode the video. Yet, the delays seem acceptable, because they don’t exceed the length of a few seconds at maximum.

In mobile television, channel switching delay is increased due to several aspects. The most important technologies today are DVB-H, ISDB-T, T-DMB and Qualcomm’s MediaFLO. Moreover, UMTS-based streaming TV is also very common.
All of these products need to address the fact that the receiving devices are small and not capable of decoding video for a long time because of their limited battery capacity:

In ISDB-T\(^1\), every channel is transmitted in its own frequency range. Then, the channel is split into 13 bursts, where only one burst is needed for decoding mobile TV video. T-DMB, mostly used in China, incorporates Micro Time Slicing, similar to DVB-H \(^2\). In MediaFLO\(^2\), statistical multiplexing is used. Channels are transmitted as so-called Multicast Logical Channels and multiplexing happens at the physical layer.

All of these technologies can reach a minimum delay of 1.5 seconds – in MediaFLO, most of the delays are 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.

4. CHANNEL SWITCHING IN DVB-H

4.1. Time slicing

Digital Video Broadcasting for Handhelds was developed by the European Telecommunication Standards Institute (ETSI) in 2002 and released in 2004 \(^1\) as a successor of DVB-T. DVB-H introduced the Time Slicing technology as an effective mean of dealing with power consumption issues on mobile devices.

Time Slicing is a form of time division multiplexing that allows the transmission of multiple DVB-H channels on one single MPEG-2 program stream. As described in \(^7\), it works as follows: A single forward-error corrected data frame (maximum size: 2 MBit) is 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”.

When Time Slicing is used, the burst carries multiple timestamps called “Delta T” (or $\Delta T$). This information equals to the absolute time difference between the start of two consecutive bursts belonging to the same logical channel. It is measured in milliseconds. 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% \(^11\).

\(^1\)http://www.dibeg.org/techp/techp.htm  
\(^2\)http://www.mediaflo.com

4.2. Burst parameters

The basic burst parameters can be seen in Figure 1. The constant bitrate of the source video needs to be transmitted in 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 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. The needed Off and On period time could then be calculated by the data bitrate and the burst bitrate.

4.3. Calculating channel switching times

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

4.3.1. 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, 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 (i) the average synchronization time $T_{Sync}$ is close to 0 (the time needed by the receiver to power up and lock to the signal), (ii) each burst contains at least one picture frame that allows instantaneous decoding, (iii) receiving conditions are good enough to allow reconstruction
of all received data in case of a transmission error, and (iv) 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 $2 \cdot T_{ON} + T_{OFF}$.

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

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

(1)

as the delay. The average delay equals to

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

(2)

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

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

(3)

4.3.2. Extended model

As already mentioned in Section 4.3, some assumptions have to be made for such a simple model. This has also been mentioned by [8]. 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 $T_{sync}$.

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.

All of these factors may not add much to the delay, but given the fact that not each burst might carry all necessary data to display a continuous video stream, buffering is enabled. Therefore the distribution of instantaneous decoder refresh frames in the bursts can affect the whole channel switching delay and extend the time to more than 10 seconds.

4.4. 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 [12], for example, DVB-H channel switching delays are distributed from 3 to 39 seconds with an 80% decile of 6 seconds.

5. USER STUDY

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. It is also valid to say 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. Therefore, less acceptance results in fewer users and also in a monetary loss for the mobile TV distributor and the device vendors.

As mentioned before, these waiting times are implied on a technical level and can not be improved easily. It is therefore 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. 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 television [13]. The authors have also introduced similar methods to those presented herein (advertisement logo), but we believe that (i) it is necessary to conduct such an experiment on a mobile phone, focused on the restrictions of mobile transmission, (ii) more different gap fillings need to be used, and (iii) more randomization in the experiment is needed.

The primary goal of this experiment was to find out how users would rate the different gap types in comparison to each other. The secondary goal was to assess whether they felt that one or an other gap type was annoying.

5.1. Experiment rules

We conceived an experiment that adopts the following rules:

- The test session should last up to half an hour. All participants should be non-expert.
- 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 and are presented four sequences of videos, each sequence using a different gap type. They are allowed to switch to the next channel whenever they want to. An example of two single experiment session is seen in Figure 2. After all the
sequences are finished, the questionnaire is completely filled out.

5.2. Experiment parameters

5.2.1. Content type

Of the three parameters we selected, the content type is the first one: From our experience of previous subjective tests we 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) [14, 15].

5.2.2. Gap type

During the experiment four different gap types were shown to the users. 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 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. However, an indeterminate progress bar has to be used in DVB-H (e.g. without a “remaining seconds’ count), because it can not be known when the next burst will appear.

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. 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. 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.

5.2.3. Gap length

As presented in Section 4.4, there is no fixed channel switching time in DVB-H, they vary between certain boundaries, although in theory, there is no upper boundary. We chose to show a set of channel switching times based upon our measurements in Section 4.4, but with a lower bound as described in Section 2. These are the selected delays (in seconds): \{1.5, 2.5, 3.5, 5, 7.5, 10\}. 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.

5.3. Technical details

For presenting the videos, a mobile phone was used (Nokia N96). It features a 320x240 pixel screen and also allows for receiving DVB-H streams. In order to simulate a mobile TV application, videos and gaps were ordered in a playlist and played through the commercial software **Core Player** ⁴, which is an audio and video player. The participants would then press a button to jump to the next video, which would be a gap, 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.

³http://www.zatoo.com
⁴http://coreplayer.com
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 a 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 was largely promoted in Austria during the European Soccer Championship 2008 and UMTS-based TV streaming services are also available from every major network operator.

After the test, the participants were asked how they rated the overall service quality, 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. Three users said that they would “Very likely” use the service again. Two answered “Likely”. “Maybe” was chosen three times. “Probably not” was chosen twice. “Definitely not” has never been answered.

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: (i) News (4), (ii) Sports (3), (iii) Series, Soaps (2), (iv) Music videos (2), (v) Documentations (1).

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.

Interestingly enough, the minimum gap length was estimated to be 1.3 seconds by average (+/- 0.39), 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.

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 [16], with a set of {Excellent, Good, Fair, Poor, Bad}. Scores were also taken after presentation of each gap type for comparison purposes.

In Table 1 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.

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).

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 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 2 and clearly show that the black screen was found to be very or little annoying. Also, the “please wait” animation was judged to be at least 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.

<table>
<thead>
<tr>
<th>gap type</th>
<th>after viewing</th>
<th>after exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>black</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td>please wait</td>
<td>2.9</td>
<td>2.8</td>
</tr>
<tr>
<td>logo</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>comm.</td>
<td>3.8</td>
<td>3.6</td>
</tr>
</tbody>
</table>

Table 1. MOS for gap types
Table 2. Annoyance for each gap type (occurrences)

<table>
<thead>
<tr>
<th>Type</th>
<th>very</th>
<th>little</th>
<th>not</th>
<th>okay</th>
</tr>
</thead>
<tbody>
<tr>
<td>commercial</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>logo</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>please wait</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>black screen</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

7. CONCLUSION AND FUTURE WORK

From the currently available experiment data we can derive the following conclusions:

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 because it provides no feedback. 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, so maybe multiple short commercials could be shown. This has also been mentioned in Section 4.3. Judging from Table 1, the logo seems to be the best compromise. 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.

So, 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 it shows 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.

8. ACKNOWLEDGEMENTS

We would like to thank Vittorio Baroncini (FUB, Rome) for his useful comments and ideas.

9. REFERENCES

[1] EN 302 304 V1.1.1, Digital Video Broadcasting (DVB);


