DVB-H: Technical Overview and Design Requirements for Mobile Television Broadcasting

Patrik Hummelbrunner and Werner Robitza, University of Vienna

***Abstract*—Digital Video Broadcasting - Handhelds is emerging as a digital broadcasting standard for handheld devices since 2006. Its main features are derived from DVB-T, the terrestrial version, but especially for handheld receivers new features have been added. Time slicing is introduced in order to allow the receiving device to save power due to a time division multiplexing method, MPE-FEC (Multi Protocol Encapsulation Forward Error Correction) provides a robust error correction mechanism based on Reed-Solomon Codes to error-protect the sent multi protocol encapsulated data packets in an environment where packet loss is highly common. During the last years, several studies have been made concerning the perceived quality of mobile broadcasts. Especially in a mobile environment it is necessary to find out which aspects of the presented audiovisual content are critical for a good acceptance from the viewer. Also, certain transmission parameters have an impact on visual quality. It is also important to summarize the most important criterions and design requirements for mobile television platforms and applications. Recent studies need to be analyzed and compared in order to show what has already been achieved by mobile broadcasting companies and vendors in Austria and other countries.**

I. Introduction

In the year 2004, DVB-H was specified by the ETSI, the European Telecommunication Standards Institute. Today, it is one of the most common mobile TV technologies of the whole world, with millions of receivers sold in over one hundred countries [1]. DVB-H has solved many problems related to the energy consumption issues of DVB-T and highly improved the mobile reception of audiovisual content. Behind DVB-H is a consortium of over 270 partners: from the telecommunications industry to the vendors themselves.

As mobile television deployments begin to take place all over Europe, more and more customers become test clients for a technology that has not yet achieved all goals that may have been set in advance. Primarily, the problems lie in the quality of the video content and the tune-in and channel switching times, which range from half a second to multiple seconds. Although receiving devices become smaller and more energy-efficient with each generation released, customers are faced with issues like energy saving and perception problems due to the small sized device screens and the service access times.

This paper is organized as follows: The most important technical features of DVB-H are presented in Section II, amongst them time slicing, the Multi Protocol Encapsulation (MPE) and MPEG-2 specifications for DVB-H. In Section III an overview is given on which requirements have to be met in order for a DVB-H service to be highly accepted by the viewers, which of course is an issue for mobile TV vendors. These include technical and content requirements. Also, interactivity in mobile TV is discussed. In Section IV a common approach for DVB analysis is presented.

Patrik Hummelbrunner and Werner Robitza are a students at the Faculty of Computer Science since 2007 and project researchers at the Department of Distributed and Multimedia Systems since 2009. e-Mail: patrik@hummelbrunner.net and werner.robitza@univie.ac.at

II. Technical Overview

The work on the DVB-H standard was started in the year 2002 and released by the ETSI in the year 2004 under the EN 302 304 Standard [2]. It is titled “Digital Video Broadcasting (DVB); Transmission System for Handheld Terminals (DVB-H)”.

The motivation for DVB-H was to improve the energy related problems of mobile reception that came along with the earlier introduction of DVB-T. Hence, the physical layer of DVB-H is entirely based on DVB-T [3], but it is enhanced by a new transmission mode which is explained later.

The other layers implement techniques specified in other standards: In [4] time slicing and the use of MPE are specified, both methods of DVB data broadcasting. [5] is the corresponding standard for the MPEG-2 Transport Stream, which is used at the data link layer of DVB. In [6] details are given about which MPEG-2 information tables are new to DVB-H.

A. Protocol Layers

In DVB-H, a protocol stack shall be used according to [7]. It is similar to the common ISO/OSI reference model [8], although these two models can not be matched perfectly. DVB data broadcasting specifies four methods of delivering content [4], but only one, MPE, is of interest in this paper.

Generally, in DVB-H with MPE broadcasting, any kind of application data is encapsulated in IP packets, optionally being streamed through RTP and UDP beforehand. These IP packets are then encapsulated again using the MPE technique at the data link layer. On the same layer the content is time sliced. The resulting data is multiplexed into an MPEG-2 transport stream, together with other streaming services such as DVB-T content. Therefore, DVB-H and DVB-T content can be distributed simultaneously by a single DVB-T transmitter, using the same physical layer, at which the transport stream is modulated for OFDM (Orthogonal Frequency Division Multiplexing) transmission.

B. Application Data

The DVB-H standard itself does not specify the type of application data that can be used. In fact, any kind of binary bitstream could be possibly sent to the receivers. Although so called data carousels are defined for distributing downloadable content over DVB-H [4], they are not used in real implementations. The most common application data type for DVB-H is video content, though.

As for the video, the MPEG-4 AVC / h.264 codec has to be used, since it is optimized for mobile content shown at small screen resolutions and much more bit-rate efficient than its predecessor MPEG-2 [1]. There exist technical specifications for DVB-H which define certain profiles and levels for the video streams [9]: For Standard Definition Television (SDTV) video, the main profile shall be used (level 3) - for High Definition Television (HDTV), the high profile is used instead (level 4). It is also implied for HDTV to use a 16:9 aspect ratio instead of 4:3, which is currently used for most other video content available.

Analyzing the perceived quality of mobile television often takes place in the application layer, since the effects of transmission errors become visual (i.e. in forms of blocking or frame losses). There are many quality assessment metrics for video, PSNR (Peak Signal to Noise Ratio) being one of the most common. However, PSNR values can not perfectly describe whether the quality of a video is acceptable for a human being or not. In the second part of this paper, more on this topic is explained.

C. Streaming

In order to distribute the video content to a large number of receiving devices, it is streamed via IP broadcasting or multicasting, the latter meaning the IP addresses of the devices can be limited. As for most of streaming applications today, RTP is used at the application layer. The RTP packets are encapsulated by UDP, since in DVB-H, there is no feedback channel that would allow a two-way connection between sender and receiver (which is necessary for the TCP handshake).

MPEG-4 AVC / h.264 introduces a new technology that facilitates mapping video data to RTP packets compared to MPEG-2 video: The NAL (Network Abstraction Layer) generates NAL Units from the VCL (Video Coding Layer) that can be mapped directly to an RTP packet. One RTP packet can hold multiple NALUs, one part of a NALU, or in the best case exactly one NALU [10]. Mapping one NALU to one RTP packet improves the receiving quality since upon loss of one RTP packet, only one NALU is lost, in contrast to the other cases of mapping where parts of a NALU would have been lost, hence rendering the NALUs useless for decoding.

The RTP packets are then encapsulated in IP packets and forwarded to the MPE encapsulator.

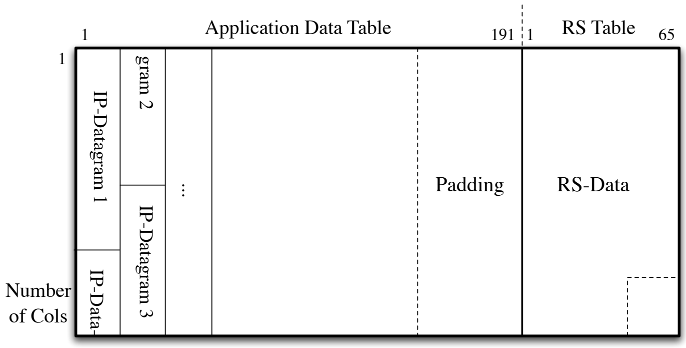

Figure 1. The MPE-FEC frame

D. Multi Protocol Encapsulation and FEC

The data link layer introduces the key technologies to DVB-H [4]: Multi Protocol Encapsulation adds another abstraction layer by simply encapsulating the IP packets into a so called MPE section. Its header though provides new fields that are currently allocated for real time transmission parameters that are needed by the receiver. Also, MPE works as a translator between addressable IP data and the underlying DVB structure [11].

At the same layer, a new forward error correction mechanism is used, called MPE-FEC. While the physical layer already implements two FEC stages, the third one is especially suited for mobile receiving conditions with high packet loss probabilities.

In MPE-FEC, frames are created for allocating the data bits (see Figure 1). These frames are divided into 255 columns and a variable number of rows limited to a maximum of 1024. Hence, the maximum FEC frame size equals 2 MBit. The frame is also divided into an application data table and a RS-data table. The application data table is filled column-wise with the IP-packets. If one datagram is larger than a single column, its remaining bits are inserted into the next column. At the end of the table, padding bytes can be inserted. This is the case when an IP packet would not fit into the frame as a whole.

Then, a Reed Solomon code (using the input parameters 255,291) is calculated row-wise and filled into the RS-data table. The redundant RS data allows the decoder to correct multiple erroneous bits in the application data table. The MPE-FEC frame is now transmitted row-wise by forming MPE sections, which results in an interleaving effect, achieving an even stronger error protection against bit errors during transmission.

The usage of MPE-FEC is entirely optional in DVB-H, although strongly recommended. Receivers not capable of decoding the FEC data can skip the RS-data part of the frame, which makes MPE-FEC a technology that is fully backwards compatible to not yet compliant receivers.

E. Time Slicing

Time slicing is a time division multiplexing method. This means that multiple programs can be transmitted through

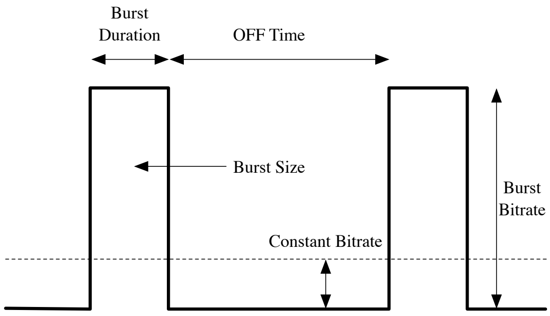

Figure 2. Burst parameters

one signal, the single components being sent at specific times or within certain time intervals - thus the term "time multiplexed".

In time slicing the data of exactly one MPE-FEC frame is transmitted as one burst (time slice ON period), then a pause occurs (time slice OFF period). This pause is referred to as delta T, or ΔT, since it is the difference between start of burst two and the start of burst one. Its value is therefore measured relatively (i.e. in milliseconds), which allows the sender and receiver to stay unsynchronized. Also, constant transmission delays have no affect on time slicing.

One or more DVB-H programs can be multiplexed by simply sending one burst after each other. A burst scheduler (a software also referred to as "time slicer") would then fill all time slice OFF periods with another burst ON period, thus resulting in a constant DVB-H stream.

One of the most important features of time slicing is the power saving effect. Since the receiver knows through the parameter delta T when to expect the beginning of the next burst, it can switch to a power saving mode during the burst OFF times.

Common time slicing parameters can be found in Figure 2. The *Burst Length* itself (T_{Burst}) could be calculated simply by dividing the *Burst Size* S_{Burst} through the *Burst Bitrate* R_{Burst}. However, a factor of 0.96 is also introduced in order to compensate for the transport packet and section headers [7]:

$$T_{Burst} = \frac{B_{Burst}}{R_{Burst} \cdot 0.96} \tag{1}$$

Furthermore, a *Synchronization Time* T_{Sync} is added to the burst duration to get the *Time Slice ON period* T_{ON}. In this synchronization time, the receiver will tune into the signal again before receiving the burst.

$$T_{ON} = T_{Burst} + T_{Sync} \tag{2}$$

By choosing a pause duration (i.e. the *Time Slice OFF Period*) T_{OFF}, the possible *Constant Bit rate* R_{Const} for the stream can be calculated as follows:

$$R_{Const} = \frac{R_{Burst}}{T_{ON} + T_{OFF}} \tag{3}$$

Vice versa, choosing a constant bit rate in advance leads to a specific OFF period time. This calculation seems to be more intuitive, since the constant (or average) bit rate of the video stream is already known from the coding stage:

$$T_{OFF} = \frac{R_{Burst}}{R_{Const}} - T_{ON} \tag{4}$$

By choosing a constant and burst bit rate, the authors in [12] suggest that the *Power Saving Percentage* P can be calculated as:

$$P = \left(1 - R_{Const} \cdot \left(\frac{1}{R_{Burst} \cdot 0.96} + \frac{T_{Sync}}{S_{Burst}}\right)\right) \cdot 100\% \tag{5}$$

Also, DVB-H implementation guidelines [7] provide another formula for calculating the power saving percentage, where $J_{\Delta T}$ refers to the jitter of the ΔT parameter:

$$P = (1 - \frac{(T_{ON} + (3/4 \cdot J_{\Delta T}) \cdot R_{Const} \cdot 0.96)}{S_{Burst}}) \cdot 100\% \tag{6}$$

It is clearly visible that the power saving effect is dependent on the chosen bit rate as well as the durations chosen for burst ON and OFF periods. These durations again have an impact on the channel switching delay, which according to [12] is calculated as:

$$D = \frac{1}{2} \cdot (T_{ON} + T_{OFF}) + T_{ON} \tag{7}$$

As an example, some values have been chosen: The burst size is the maximum MPE-FEC frame size of 2 MBit. The bit rate for the burst is 4096 kBits/s. With a synchronization time typically less than 80 ms, this results in a burst ON time of approximately 0.6 seconds. A video coded with an average bit rate of 500 kBits/s therefore requires an OFF time of 7.5 seconds. Hence, the overall estimated power saving percentage is about 90%.

The channel changing delay can be estimated to about two seconds using the formula above. Of course, in reality the delay can take much higher values depending on various factors not mentioned here. Since it is clearly visible that a higher power saving percentage results in higher channel switching times and vice versa, these parameters affect the quality of experience for the end user. Channel switching delay will become another aspect of the second part of this paper, the design criterions.

F. Transport Stream

The ISO specified MPEG-2 transport stream [5], which was introduced years before DVB-H, provides methods for multiplexing the time sliced and DVB-H programs together with other services. Although the DVB-H content could be sent without this adaptation, MPEG-2 TS allows addition of program specific information and service information (later referred to as PSI/SI) to the stream, which is necessary for each receiver to locate streams in a network.

As an example, a DVB-H stream can be multiplexed with several other DVB-T streams, sharing the available bandwidth

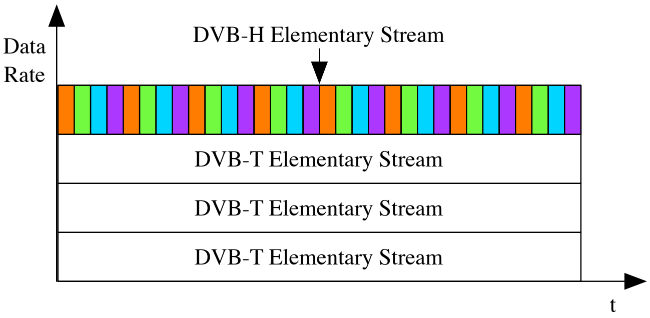

Figure 3. Multiplexing in DVB-H

of the whole TS. This is shown in Figure 3. Each separate stream in the MPEG-2 TS is called elementary stream (ES). The overall available bandwidth between ES can be distributed freely, although the maximum recommended bit rate per ES is limited to about 10 Mbps [7].

Multiplexing works as follows: Data is encapsulated into transport stream packets, of which every one has the size of 188 Bytes and is assigned a PID (packet ID) in the header. Packets belonging to the same ES have the same PID, thus the majority of possible PID values is freely assignable for the broadcaster. However, certain PID values can not be assigned. They refer to another payload type of the packet: PSI/SI tables. These tables are sent repeatedly throughout the whole transmission. Their intervals range from seconds to a few minutes, depending on the content.

Carrying the PID 0x0000, the Program Association Table (PAT) for example gives a list of all the possible services in the multiplex by listing all the Program Map Tables (PMT). A service again is defined by this PMT, which lists the PID values that together form the ES. By looking up the PID value(s) for a specific service in the PMT, the demultiplexer can filter for this single PID and combine the packets in order to decode the ES.

Another important PSI/SI table is the Network Information Table (NIT), which describes the whole DVB-H network, including its cells, cell dimensions, and so forth. Each PSI/SI table can carry one or more descriptors, which add more details to the provided information. The PSI/SI tables are important to DVB-H, since they allow not only proper tuning and reception for the receiver, but they also specify the content (i.e. like a program guide): For each event, there shall be at least one descriptor for the content type and its duration. Also, these tables can be used to notify the user of certain alerts or program changes.

As mentioned in the introduction to this section, [6] specifies all the PSI/SI tables that can or must be included in a DVB-H transmission.

G. Physical Layer Adaptation

The physical layer of DVB-H is based on DVB-T, as specified in [3], and independent from the content. The input for a DVB-T modulator consists of transport stream packets that are exactly 188 Bytes long - the output is an OFDM modulated bitstream.

The methods applied at the physical layer are as follows:

1) randomizing the TS packets (“scrambling”)
2) adding an outer Reed-Solomon code to the TS packets
3) interleaving the packets
4) adding an inner punctured convolutional code
5) interleaving the packet data

Randomizing will scramble the whole packets, leaving their Sync Bytes (the first bytes in the header) at the place they have been in before. Every eight packets, the scrambler is re-initialized.

Then, two forward error correction coding stages are applied. The (outer) Reed-Solomon code allows for up to eight bytes to be recovered upon data loss. It enlarges the packets from 188 to 204 bytes. The inner convolutional code adds redundancy to the signal by using a special algorithm that splits the bitstream into two streams, depending on a chosen code rate, and then puncturing (i.e. leaving out) certain bits from this stream.

Interleaving refers to transmitting data in a way that the contiguous bits are sent separately, so that the impact of a burst-error (which affects a few bits in a row) is reduced.

H. OFDM Transmission

As the bitstream resulting from the adaptation is DVB-T compliant, the existing DVB-T infrastructure can be used for DVB-H transmission. The DVB-H standard introduces a new transmission mode, the 4k mode [2], which can be used for OFDM transmission. DVB-T already makes use of the 2k and 8k mode, which all define parameters for the physical layer adaptation. The name of these modes refers to the number or carriers used in the OFDM transmission. The main difference between 2k and 8k mode is the maximum transmitter distance and maximum receiver movement speed for a single frequency network. 2k mode allows 17 km distance with fast moving receivers, whereas 8k mode allows up to 67 km for slow moving receivers. The 4k mode instead provides a tradeoff between those two by allowing a 33 km distance [13], aiming for more flexibility in the transmission of DVB content.

Also, DVB content can be multiplexed into one high quality (also high priority, HP) and one low quality (LP) stream upon physical layer adaptation. This allows for a better performance degradation, because high quality streams need not to be received, but given enough signal strength, can increase the quality of service.

In conclusion, DVB-H services provide much more sustainable quality than older video broadcasting standards, because of a high data rate that can be achieved during broadcast, a very well developed error correction mechanism as well as a well scalable infrastructure.

III. Design Requirements

There are two different aspects with importance for the QoE (quality of experience) of mobile TV. One is about the technical requirements (the technical specifications have been explained in the first part of this paper), the other one is about the requirements for the content. But before explaining the details of the requirements, it is necessary to know how and where people watch mobile TV.

Mobile TV users mostly do not watch more than ten minutes at a time according to studies carried out in [14]. This is important to know for the content that should be provided in DVB-H. Also interesting to know is what is watched. According to [15], the user's most favored content types are:

- cartoons
- music video
- news
- sports

Also, there are many places where people watch mobile television, but the most important places are:

- on bus or train station
- in the bus or the train
- in bars or restaurants while waiting or eating alone

It is also worth mentioning that people least enjoy mobile TV at their homes. This is most probably because of other means of watching television at home, like a normal TV receiver or on a computer.

A. Technical Requirements

1) Transmission and network issues: The most important technical problem we face with mobile transmission is that the receiver is being moved (buses and trains are moving at high speeds), so the technology should be able to handle this without burst errors and lags. DVB-H is able to provide rather good service quality - especially in closed buildings and rooms - due to the newly designed 4k-mode, nevertheless most of the reception problems faced in reality exist due to multi path fading, which is caused by large buildings in cities. Especially in urban environments, DVB-H network planning results in the usage of a high number of antennas/cells. In rural areas, though, covering a wide area is a hard task because it takes too many antennas to ensure a high network coverage. Because of the small device size, antennas are restricted in their operation and need higher gain levels to compensate for this effect. Using the device in close range to the human body can also affect reception quality because of the body acting as a shield, just like large buildings do.

The high number of cells requires switching from one to another more often than usual. The receiver performs offline scanning of other cells in between the burst intervals. If there is a stronger signal compared to the currently tuned in one, the receiver may change to the better signal. This is called handover. Time slicing assists this seamless handover, because during OFF-periods no relevant packets are sent to the receiver. Because of this, it is also possible to receive DVB-H while moving, given the precondition that the DVB-H network covers the whole journey route with high enough signal strength.

As shown in Figure 4 we can see that currently mobile TV reception with UMTS/HSDPA is better than with DVB-H. DVB-H coverage is very good in cities but in rural areas will not be very economical [16]. In [15] the authors calculated which error rate would be acceptable for a certain content. The following rates are based on an overview of all content types.

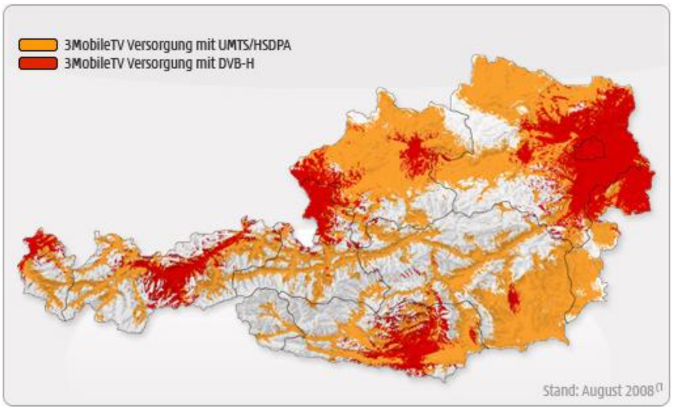

Figure 4. Network coverage in Austria

For a given error ratio (PER) of 1,7% the acceptance is about 89%, for 6,9% PER it is 81%, for 13,8% PER the acceptance drops to 56%, and for 20,8% it is 54%. Most people find an error ratio of 20,8% still acceptable. It is also interesting that there is a big gap between 6,9% and 13,8% and only a small between 13,8% and 20,8%, thus giving a hint on a certain minimum PER that the broadcasting companies should try to guarantee.

2) Energy saving: Another technical problem comes from the fact that a mobile device gets its energy from a (mostly rechargeable) battery. Because most of the mobile television receivers are also used as means of communication, people will not trade in battery life for mobile TV usage, since there is not always the possibility of charging the battery. The technology to solve this problem is energy saving at the receiving stage. As said before, time slicing is one of the biggest advantages of DVB-H over DVB-T: It is possible to save about 90% of the energy.

3) Video resolution: The next technical requirement is a high enough video resolution. In [17] resolutions were tested up to 240x180 pixels, the results of this paper will be discussed later. One can presume that the QoE normally rises with the resolution. The standard resolution for DVB-H is 320x240 pixels [9], [17], which seems to be enough for most of the current receivers. For example, the Nokia N96 provides a resolution of 320x240 pixels on a 2.8 inch screen. A higher pixel (and therefore screen) size would lead to larger devices which are not likely to be considered mobile anymore.

B. Content Requirements

1) Content length: As already mentioned above, normal mobile TV consumption lasts for ten minutes. It means that long movies are not preferred by users, also series that run for 30 minutes might be too long for the average user. Normal television programs, though, focus on exactly these content types. Short mobile TV content therefore needs to be created separately. Because of their nature, this does not apply to music videos, news or most of the sports programs, which is why the are rated so high in comparison to other content types.

Figure 5. Comparison of MS (upper left corner) and VLS

2) Shot types: Mobile TV content is presented not only in a shorter, but in a smaller way than normal TV content, so there are different needs for shots on the subject of interest. In [17] the authors present different shot types, of which two can be seen in Figure 5:

- Extreme long shot (XLS): XLS means that the subject is not in focus of interest, the environment seems more important.
- Very long shot (VLS): In VLS one can see more details of the subject but the environment needs more space in the frame.
- Long shot (LS): In a LS the subject goes from top to bottom of the frame and can be seen clearly.
- Medium shot (MS): In a MS the subject is bigger than the frame an the eyes can be seen clearly.
- Medium close-up (MCU): MCU means that the focus gets more on the face of the subject.
- Close-up (CU): CU means that the whole attention of the viewer is on the face of the subject.

Each content has different needs and so it is assumable that one shot type fits better to music videos whereas another one might be used in news. The authors tested different shot types for different content with different resolutions. As variation of video resolution did not change the results in a dramatic way, they are considered not worth mentioning in this paper. So only an overview of the most important shot types for the content types is given:

In football, one of the most watched content types, the most important shot types are VLS, LS and MS. For news the most important shot types are MS, MCU and VLS. MCU and VLS result in nearly the same acceptability rates, but MS scores better. In music content LS, VLS and XLS are considered important. In animated content these are: MS, LS and VLS. Taking in account these conditions it should be possible to produce content specially suited for mobile TV, although not practiced as of the current date.

3) Zoomed content: It has already been explained that mobile content is different from normal TV content, especially in presentation length and size. Naturally, it is very expensive to produce content only for mobile TV, not only since many programs have already been recorded before the existence of mobile TV. Therefore, it is a good idea to improve the normal TV content by applying effects to the video and recoding it, as explained in [18]. One of the effects suggested is zooming. The authors tested if zoomed content would be more acceptable to users than normal content. Two displays were used, one showing zoomed content, one the normal video.

The authors tested the method only for football, but the test's outcome shows that zooming could be considered useful. Very interesting is the fact that the acceptability of zooming goes down with a higher video resolution. Also, the tests were ran with manually zoomed content, so the acceptability could become worse if the zooming regions and factors are computed automatically, which would be necessary for a real time implementation.

4) Foreground extraction: Another automatic method of video enhancement could be to separate the foreground and the background of the frame, as proposed in [19], by providing only the fore- or background of the video as an input to an effects processor. Tests showed that blurring the background and decreasing its saturation while increasing the saturation of the foreground has the best effect on most of the tested content types. Generally speaking, the visual parameters of the foreground will be accentuated and the important regions of the scene will stand out against the background. This works best with movie scenes where one or more actors are in focus while moving and the background already has been blurred by the depth of field effect. But as with the earlier proposed zooming method, it is hard to compute the background portions of the frame, whereas the human visual system could easily achieve this separation.

Separating the layers can not be performed in real time because it relies on the motion vectors. MPEG-4 AVC has to be used, since the encoder outputs information logs which allow to calculate motion vectors for each macro block in a frame. With a computer algebra system (CAS) it is possible to distinguish 'active' and 'inactive' blocks, extract the corresponding blocks and merge them into a new frame.

A general problem of image enhancement is due to timing issues. Live content needs to be delivered within seconds - one could for example imagine sports events where DVB-H users could miss a football goal due to delays, while they hear normal TV users scream. This is, generally speaking, no pleasurable experience.

C. Interactivity

Interactivity is, from the user's point of view, a very interesting component of mobile TV. Interactivity could be defined as the user's possibility to communicate and take an active role, and not just receiving information. For this type of communication, a back channel like UMTS, W-LAN or HSUPA is available on most mobile phones.

In normal TV, phones have been used for interactivity. Austrian television shows like "Dancing Stars" and "Starmania" gave users the possibility to vote for contestants by using their (mobile) phone for voting. The producers of "Talk of Town" use phone calls to receive opinions of their viewers during the show. Although present in many households, the internet is not too popular as a back channel, but still, there are some shows that rely on user's opinions from the internet, for example "derStand.punkt". German TV channel "Giga" has been using the internet more than most of the other channels. There are several shows where viewers can take part of the program, by commenting or chatting in a forum. This model could also be taken in account for mobile television.

In all of the above examples interactivity was created by using another device than the receiver, like a mobile phone or a computer. But with a mobile phone that is already connected to the internet by UMTS or other services, the interactivity potential is big, but not yet tested. The contents of [20] were chosen as an example for interactivity. The authors used only open source tools to build up their test environment. They implemented a player called *Mobile interactive Video Browser Extended Software* (MiViBES) which provides the following interactive functions:

- Tight video/browser integration (e.g. navigation items or chat as an overlay)
- Integrated push techniques for chat, event notifications, updates of additional content
- Joint-Zapping function via push channel

With these functions it is possible to assume the path the authors took with their interactivity implementation: They oriented on the successful web 2.0 concept, by building up a community. This means that users can chat with each other in a private or public chat, which are bound to the channel they are watching. The authors used DVB-H as their primary transmission method, but UMTS or W-LAN as a back channel for the chat responses. This has one big advantage: There is no additional traffic on UMTS or W-LAN. The only large problem could be the time gap due to transmission delays, limiting the interactive experience.

IV. Showcase

In this showcase it is presented how to receipt and analyze a DVB-H stream with a DVB-T USB receiver. DVB-H content is encrypted by the content provider or distributor, so it is not possible to watch a program without a service agreement (i.e. a contract with a provider) and special receiver for a PC. In Austria the market for DVB receivers is relatively small. The ZTE MF 635 data and DVB-H modem is only available at Hutchison 3G, yet this model is not compatible with dvbSAM, a software package that allows one to analyze DVB streams deeply [21]. So it is necessary to look for other possibilities to receive a DVB-H stream. Considering the fact that a DVB-H stream is mapped onto a DVB-T transport stream, while completely using the physical layer of DVB-T, reception of DVB-H is possible also with a DVB-T USB receiver. The receiver used in the experiments was manufactured by ASUS (My Cinema – U3000 Hybrid).

Figure 6. Example of a time sliced program

A. Analyzing with dvbSAM

The already mentioned software is dvbSAM (Version 3.6.22.910) from decontis. This software package is made analyzing DVB streams. It is able to decode DVB-T, DVB-S and DVB-H streams, yet decryption is not possible without a key (which can not be obtained). Because of this, it is not possible to watch the video, but information about the streams is visible. As for content details, all PSI/SI tables from the transport stream can be analyzed. For our time slicing and energy saving performance measurements, all important transmission and coding parameters are visible: the error correction used, stream data rates during the burst as well as average data rates, the average of the whole stream and many more functions.

In Figure 6 a screenshot from the dvbSAM software is shown, analyzing one single (time sliced) program stream in an Austrian DVB transmission. It shows that the values chosen in the example in II-E also occur in reality. Interesting is the fact that the figure shows 88% energy saving potential, although the burst size OFF period is relatively small compared to the chosen values in the calculations of II-E, where 7.5 seconds led to 90% energy saving. This means that the energy saving curve becomes flat for higher burst intervals, so in order to reach 90%, there is a huge trade-off against bandwidth capabilities. It should be also mentioned that calculating an optimal stream distribution for DVB-H is never trivial in real implementations, where most of the parameters may change from one second to another and therefore require a redistribution in real time.

Figure 7 shows a DVB-H multiplex, sharing more than 20 channels. The small portions of the diagram represent radio streams, which obviously do not take as much bandwidth as a TV stream.

B. Other available DVB software

In order to test DVB-H content adaptation and distribution, there also exist various tools for Linux and Unix operating systems, including IP-encapsulators and MPEG-2 multiplexers (with the ability to create PSI/SI data), as well as whole DVB suites (JustDVB-IT by CINECA [22]). FATCAPS , for example, extends the JustDVB-IT open source project by DVB-H-specific components. They can be set up to simulate a

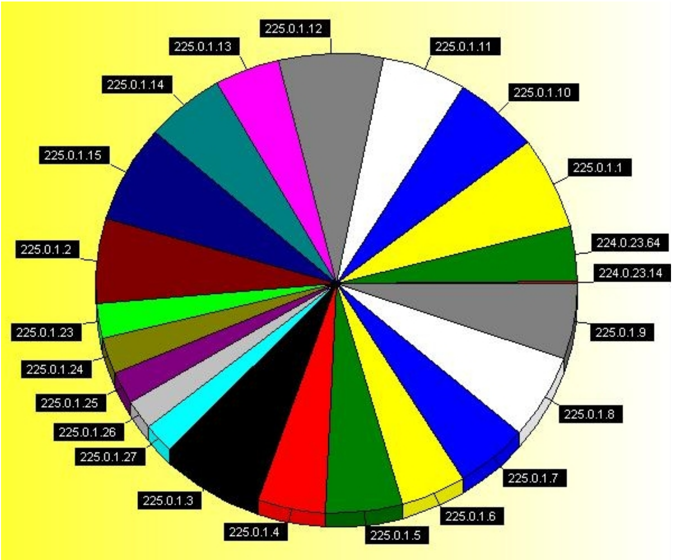

Figure 7. Channel distribution in DVB-H

DVB-H transmission over LAN or W-LAN. Also, real DVB-H streams can be used for a streaming relay by redistributing them.

The command line tool dvbsnoop [23] can also perform a deep analysis of a transport stream.

V. Future work

A packet loss model for DVB-H will be calculated throughout a project at the University of Vienna. By comparing the results with loss rate acceptance models, it is possible to create a map, for example, for the city of Vienna by measuring and analyzing a DVB-H multiplex. This map would then show the perceived quality of the transmission, rather than a packet error ratio.

Also, measurements will be taken on the channel switching time. Since it is an important factor in quality of experience for the user, switching delays need to range between certain intervals of, for example, two to ten seconds. Currently, there are no studies comparing different channel switching times. A survey will be concluded by presenting different delays to the users and then calculating the optimal delay as a general guideline to DVB-H providers and manufacturers.

VI. Conclusions

In this paper, an overview of the technical specifications for the DVB-H standard was given by analyzing each layer. The most important parameters that can be applied at the various protocol layers have been identified in regard to the design requirements. Details were given on how MPE-FEC is implemented in order to error correct a bursty transmission, which occurs in DVB-H, but not in DVB-T. Also, the method of time slicing was explained and simple calculation methods were presented, which would allow to calculate the burst sizes, the video bit rates and the corresponding energy saving amount. The channel switching delay, which is one of the most influencing aspects of user interaction, was analyzed.

In the second part of the paper the needs and design requirements that are to be met in a mobile TV application were summarized. They can be divided into technical requirements and content requirements, but also social and human computer interaction (HCI) aspects can be described with DVB-H: It was concluded that the average mobile TV user would preferably watch short contents not for longer than ten minutes while waiting for their bus or train to arrive. Technical requirements, though, focus on network planning and transmission problems, but also include minimum video resolution rates and energy saving. Content requirements focus on the shot types with regards to the content being presented. Methods have been proposed on how to optimize content for mobile TV. Also, the interactivity aspect was discussed briefly. A showcase was presented showing how to perform analyzing of DVB streams through combination of hardware and software.

References

[1] H. Bürklin, R. Schäfer, and D. Westerkamp, "DVB: from broadcasting to IP delivery," *SIGCOMM Comput. Commun. Rev.*, vol. 37, no. 1, pp. 65–67, 2007.

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

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

[4] *ETSI EN 301 192 V1.4.2, Digital Video Broadcasting (DVB); DVB specification for data broadcasting*. European Telecommunication Standards Institute, 2008.

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

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

[7] *ETSI TR 102 377 V1.2.1, Digital Video Broadcasting (DVB); DVB-H Implementation Guidelines*. European Telecommunication Standards Institute, 2005.

[8] *Information Technology - Open Systems Interconnection - Reference Basic Model: The Basic Model*, 2nd ed. ISO/IEC, 1994.

[9] *ETSI TS 101 154 V1.8.1, Digital Video Broadcasting (DVB); Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream*. European Telecommunication Standards Institute, 2007.

[10] S. Wenger, "H.264/AVC over IP," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, 2003.

[11] V. K. M. Vadakital, M. M. Hannuksela, M. Rezaei, and M. Gabbouj, "Optimal IP Packet Size for Efficient Data Transmission in DVB-H," *IEEE Nordic Signal Processing Symposium (NORSIG2006)*, 2006.

[12] M. Rezaei, I. Bouazizi, V. K. M. Vadakital, and M. Gabbouj, "Optimal Channel Changing Delay for Mobile TV over DVB-H," *IEEE International Conference on Portable Information Devices (PORTABLE07)*, 2007.

[13] M. Kornfeld and G. May, "DVB-H and IP Datacast–Broadcast to Handheld Devices," *IEEE Transactions on Broadcasting*, vol. 53, no. 1, pp. 161–170, March 2007.

[14] H. Knoche and J. D. McCarthy, "Mobile users' needs and expectations of future multimedia services," in *Proceedings of the WWRF12*, 2004.

[15] S. Jumisko-Pyykkö and M. M. Hannuksela, "Does context matter in quality evaluation of mobile television?" in *MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*. New York, NY, USA: ACM, 2008, pp. 63–72.

[16] H. Knoche and J. D. McCarthy, "Design requirements for mobile tv," in *MobileHCI '05: Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*. New York, NY, USA: ACM, 2005, pp. 69–76.

[17] H. Knoche, J. D. McCarthy, and M. A. Sasse, "How low can you go? the effect of low resolutions on shot types in mobile tv," *Multimedia Tools Appl.*, vol. 36, no. 1-2, pp. 145–166, 2008.

[18] H. Knoche, M. Papaleo, M. A. Sasse, and A. Vanelli-Coralli, “The kindest cut: enhancing the user experience of mobile tv through adequate zooming,” in *MULTIMEDIA ’07: Proceedings of the 15th international conference on Multimedia*. New York, NY, USA: ACM, 2007, pp. 87–96.

[19] S. Buchinger, M. Nezveda, W. Robitza, P. Hummelbrunner, and H. Hlavacs, “Mobile tv coding,” 2009.

[20] R. Schatz, W. Wiedermann, R. Ochenbauer, and I. Miladinovic, “Interactive TV on the mobile: the AMUSE 2.0 Platform for DVB-H-based next-generation mobile TV-services,” *e i Elektrotechnik und Informationstechnik*, vol. 124, no. 10, pp. 328–336, 2007.

[21] “Decontis,” cited May 2009. [Online]. Available: http://www.decontis.com/

[22] “JustDVBit,” cited May 2009. [Online]. Available: http://www.cineca.tv/labs/mhplab/JustDVb-It%202.0.html

[23] “DVBsnoop,” cited May 2009. [Online]. Available: http://dvbsnoop.sourceforge.net/