Efficient FEC Protection of Scalable Media Streams in DVB-H

Imed Bouazizi, Miska M. Hannuksela
Nokia Research Center.
Tampere, Finland
{imed.bouazizi,miska.hannuksela}@nokia.com

Lukasz Kondrad, Moncef Gabbouj
Tampere University of Technology
Tampere, Finland
{lukasz.kondrad,moncef.gabbouj}@tut.fi

Abstract

In this paper, we discuss the deployment of SVC in DVB-H. We focus on the system configuration aspects, in order to improve the receiver performance in terms of perceived video quality and resulting power consumption.

We evaluate a set of different multiplexing and FEC protection strategies in a simulation environment. We show that interleaved transmission combined with carefully configured unequal error protection achieves best overall results, while still being backwards-compatible to the DVB-H standard.

Keywords

DVB-H, SVC, Unequal Error Protection.

INTRODUCTION

Nowadays, broadcasting technologies are witnessing a spread of their reach to the mobile area. Recent development and deployment of mobile broadcast standards, such as IP Datacast over DVB-H [1], have given a significant boost to the ubiquity of broadcast services such as TV and download services. As a consequence of this development, receiver homogeneity, a cornerstone of the success of traditional broadcast technologies, is being challenged. Contrary to the classical and homogeneous set-top boxes, a wide range of receivers with different processing and display capabilities are being targeted by mobile broadcast services. The heterogeneity in the receiver population raises significant challenges to the efficient usage of broadcast network resources that were a by-product of homogeneous receivers.

Scalable media coding is a promising approach that may be used to alleviate the drawbacks of heterogeneity. JVT has recently finalized an extension to H.264/AVC [2], defining a scalable video coding (SVC) standard [3]. An SVC stream consists of a base layer and one or more enhancement layers. The base layer is fully backward compliant with the H.264/AVC standard. This would enable legacy receivers to consume SVC content by discarding all enhancement layer media units. Enhancement layers may improve the temporal or spatial resolution or improve the video quality. Contrary to simulcast, where two or more representations of the same content at different resolutions or qualities are broadcast to receivers independently, SVC extensively exploits the redundancies between the different representations by using inter-layer prediction. Significant savings in bandwidth can be achieved compared to simulcast. In addition to the significant improvements in coding efficiency – compared to earlier scalability tools – SVC comes with significantly reduced decoder complexity due to the single loop decoding.

In recognition of the great potential of SVC, the Digital Video Broadcasting (DVB) organization adopted SVC as one of the video codecs to be deployed by DVB broadcast applications. However, existing standards for broadcast applications may need to be tailored to the transport of scalable video streams, in order to take advantage of the full potential of SVC.

The rest of this paper is organized as follows. In the next section background information about DVB-H and SVC standards are provided. Subsequently, the different options for multiplexing and transmitting SVC coded video streams as well as the optimal Forward Error Correction protection configuration are presented and developed. Simulation results are presented and discussed in section 4. The paper is concluded in section 5.

BACKGROUND

This section provides background information needed and is organized in the following way. First, details of the DVB-H features, such as time-slicing and the link layer FEC are examined. Next, the main concept of SVC codec is provided.

Digital Video Broadcasting – Handheld

Broadcasting to handheld devices is enabled by the DVB-H standard [4] which is based on Digital Video Broadcasting – Terrestrial (DVB-T) [5] and is backward compatible with it. However, together with mobile devices new challenges, such as limited power supply and more error-prone transmission have occurred. To solve those problems DVB-H, compared to DVB-T standard, introduces new solutions. The problem of limited battery life is overcome by introduction of time-slicing concept and the new link layer FEC, multi-protocol encapsulation FEC (MPE-FEC) is added to provide more robust transmission.

Time-Slicing

Time-slicing was introduced in DVB-H standard to reduce the power consumption in handheld terminals. The concept of time-slicing focuses on sending data in burst using significantly higher bit rate in comparison to the bit rate required if the data is transmitted continuously. Between the bursts the data of the elementary stream is not transmitted, allowing other elementary streams to use the bit rate otherwise allocated. This enables the receiver to stay active only for a fraction of time, while receiving burst of a requested service. To get a reasonable power saving effect, the burst bit rate should be at least 10 times the constant bit rate of the delivery service. In Figure 1 the concept of the time-sliced transmission is depicted.

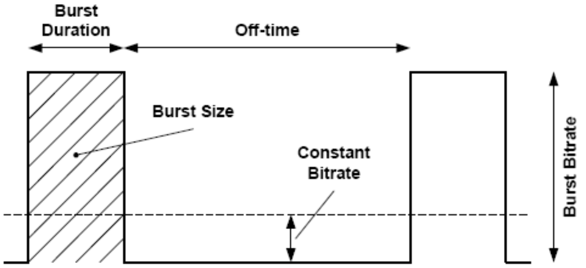


Figure 1: Time-sliced transmission

MPE-FEC

Like any typical wireless channel the DVB-H channel is also prone to transmission errors. In addition, DVB-H needs to deal with unpredictability of the propagation channel caused by mobility of the receiver. Hence, to enable better error resilient data transmission, DVB-H standard includes a link layer FEC (MPE-FEC) to counter high level of transmission errors. The parity symbols of MPE-FEC are computed over a time-sliced burst which is organized into MPE-FEC table using Reed-Solomon (RS) code [6].

An MPE-FEC table is arranged as a matrix with 255 columns and a flexible number of rows. Currently row sizes of 256, 512, 768, and 1024 bytes are defined. In Figure 2 the structure of a MPE-FEC table is shown. Each position in the matrix holds information bytes.

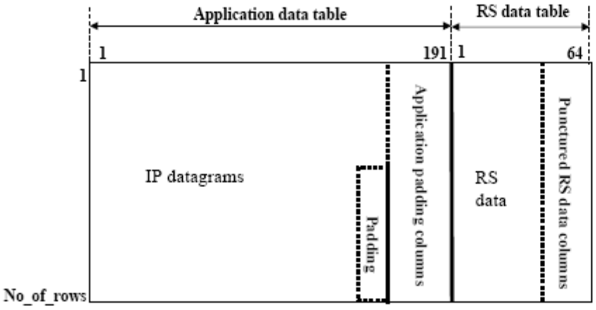


Figure 2: MPE-FEC table structure

The first 191 columns are dedicated to link layer datagrams and possible padding. This part of the MPE-FEC table is called application data table (ADT). Datagrams are placed in this part of the table column-wise and the remaining columns, when the ADT is partially filled, are padded with zero bytes thus are called padding columns. Padding is also done when there is not enough space left in the ADT to fit the next complete datagram.

The next 64 columns of the table are reserved for the RS parity information, and are called RS data table (RSDT). The RSDT is computed across each row of the ADT using RS (255,191). However, it is not necessary to compute the whole 64 columns of the RSDT and some of its right-most columns can be completely discarded. This procedure is called puncturing. The punctured and padded columns are not sent over transmission channel.

Since datagrams are placed in the table column-wise, whereas correction data is calculated row-wise, MPE-FEC table also provides virtual time interleaving.

Scalable Video Coding

SVC has been widely investigated in academia and industry for the last 20 years. Almost each of the previous video coding standards, such as H.262 [7], H.263 [8] and MPEG-4 [9], supports some degree of scalability. However, before H.264/SVC standard scalable video coding was always linked to increased complexity and drop in coding efficiency when compared to non-scalable video coding. Hence, SVC was rarely used and it was preferred to deploy simulcast, which provides similar functionalities as an SVC bit stream by transmission of two or more single layer streams at the same time. Though simulcast causes significant increases in resulting total bit rate, there is no boost in the complexity.

The new SVC standard is an extension to H.264/AVC standard which allows for temporal, spatial and quality scalability in a video bit-stream. However, in contrary to the previous implementations of scalability, SVC is characterized by a good coding efficiency and moderate complexity thanks to which it can be seen as a reasonable alternative to the simulcast.

The idea behind SVC is that encoder produces a single bit-stream containing different representations of the same content with different characteristics. An SVC decoder can then decode a subset of the bit-stream that is most suitable for the use case and decoder capabilities. A scalable bit stream consists of a base layer and one or more enhancement layers. The removal of enhancement layers leads to a decoded video sequence with reduced frame rate, picture resolution or picture fidelity. The base layer is an H.264/AVC bit-stream which ensures backwards compatibility to existing receivers. Through the use of SVC we can provide spatial resolution, bit rate, and/or even power adaptation. One main use case for SVC lies in the exploitation of the intrinsic media data importance (e.g. base on the SVC layer those media units belong to) for the purposes of achieving higher error and loss resilience through unequal error protection. Enhanced service consumers (those consuming the base and enhancement layers) may then benefit from graceful degradation in the case of packet losses or transmission errors.

In Figure 3 an example of the temporal scalability is presented. The frames T_0 represent the base layer while frames T_1, T_2, and T_3 correspond to enhancement layers. The frames from the higher layers (T_3, T_2, T_1) can be discarded, which results in a lower value of frames per second but does not introduce any distortion during play out of the video. This results from the fact that the hierarchical bi-predictive frames are used.

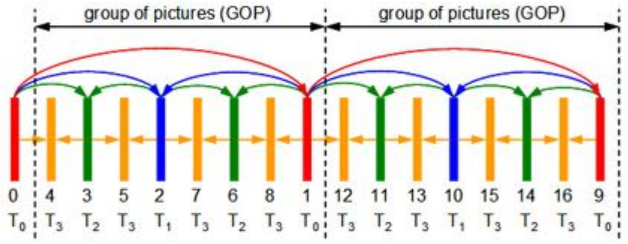


Figure 3: Temporal scalability

Other modes of scalability that SVC supports are the spatial scalability and quality scalability. In the case of spatial scalability, the encoded bit-stream contains sub-streams that represent the same content at different spatial resolutions. Spatial resolution is a major motivation behind the introduction of SVC to mobile TV services. It addresses a heterogeneous receiver population, where terminals have different display capabilities (e.g. QVGA and VGA displays). Coding efficiency in spatial scalability is achieved by exploiting inter-layer dependencies while maintaining low complexity through a single loop decoder requirement. Quality scalability enables the achievement of different operation points each yielding a different video quality. Coarse Granular Scalability (CGS) is a form of quality scalability that uses the same tools as the spatial scalability. Medium Granular Scalability (MGS) achieves different quality encoding by operating on the transform coefficients.

For detailed information about architecture, system and transport interface for SVC, the reader is referred to the Special Issue on Scalable Video Coding in IEEE Transactions on Circuits and Systems for Video Technology [10].

SVC OVER DVB-H

In this section, different options for the deployment of SVC over DVB-H are investigated. We proceed by presenting approaches as discussed in the literature and conclude by introducing a new proposal for improved SVC transmission over DVB-H.

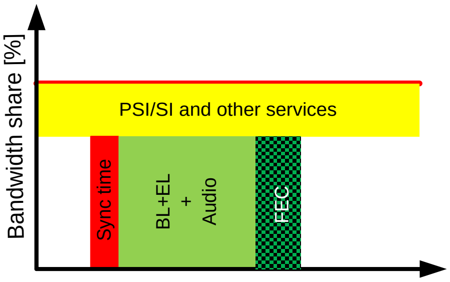
In [11] Schierl et al. describe gain achievable by adopting the SVC in DVB-H. First, authors point out use cases which can bring benefits. Those cases include: serving terminals with different capabilities at a time, conditional access to particular video qualities, graceful degradation using unequal error protection, and backwards-compatible introduction of new services. Next, by comparing transmission of two single layer streams with no FEC protection, two single layer streams with FEC protection, and a SVC stream with four layer unequal error protection ,a noticeable gain in using unequal error protection on SVC over an equal protection for the single layer streams is shown.

A novel solution for error protection for scalable media data for mobile TV application is described in [12]. Authors propose to increase the robustness of more important layers by generating protection across layers. Additionally, authors describe an extension of the Raptor FEC code which allows the implementation of a Layered FEC solution. Furthermore, results confirming that L-FEC with multiple layers performs better than the standard FEC scheme where each layer is protected separately are provided.

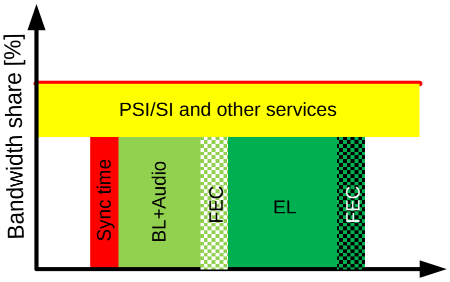
We propose a new and standard-compliant approach for multiplexing the SVC video streams of a mobile TV service. For each configured operation point of in the SVC video bit-stream, the corresponding set of additional layers is transmitted in a separate elementary stream, which is mapped to its own data burst in DVB-H. By consequence, a receiver that desires to operate at a certain operation point may have to receive more than one data burst simultaneously. The impact on buffering delay and power consumption at the receiver depends on the data burst arrangement and on the selected operation point. Receivers wishing to consume a service at a higher operation point, e.g. a higher spatial resolution, have to take into account the resulting penalty. Figure 4 depicts different possible burst arrangements for a mobile TV service offering that provides spatial scalability with two SVC layers.

In 4a, a single burst is used for the transmission of base and enhancement layers. Due to the limited maximum size of a data burst, this arrangement will result in reduced off-time and higher power consumption for the basic service consumers. Enhanced service receivers will benefit from better synchronization and a slight reduction in power consumption attributable to the omitted burst synchronization time. Both base layer and enhancement layer data will benefit from the same FEC protection, which results in equal error protection.

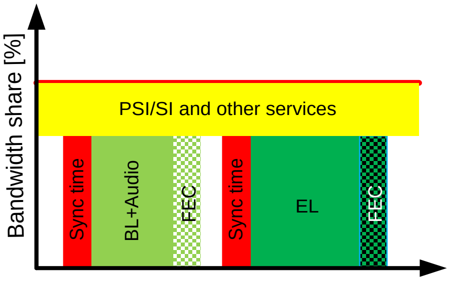
4b and 4c represent two similar approaches for carrying basic service and enhanced service in different data bursts. In 2b, the bursts are arranged to be adjacent, so that to omit the synchronization time needed for receiving the data burst of the enhancement layer. In 4c, the data bursts are transmitted in a time-shifted manner.



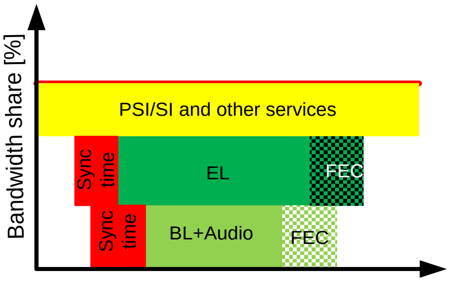
a) single data bursts



b) adjacent data bursts



c) time-shifted data bursts



d) interleaved data bursts

Figure 4 Different options for burst arrangement

A novel multiplexing approach is presented in 4d. The multiplexing occurs at the MPEG-2 transport packet level, where TS packets of the basic service and of the enhanced service are transmitted in an interleaved manner. This approach achieves better resilience against burst errors due to the increased interleaving. It also balances the power consumption penalty between consumers of the basic service and those of the enhanced service.

4b,c, and d enable the achievement of unequal error protection in a backward compatible manner, due to the separate protection of the data bursts. For each data burst, FEC code puncturing and shortening may be used to adjust the amount of source and repair data to the desired FEC code rate. This paper strives at determining the most efficient repartition of repair data among the basic and enhanced service. Due to the inherent dependencies between the SVC enhancement layer and the base layer, a differentiated protection against transmission errors promises to achieve an improved overall quality for a mixed receiver population by supporting graceful degradation for enhanced service consumers and a robust basic service.

RESULTS

In this section, we present and discuss the results of a comparison of different burst arrangements and FEC resource repartition strategies on the perceived quality.

A simulator for DVB-H multiplexing and transmission has been used throughout the simulations. It enables the adjustment of the FEC code rates for each data burst. For simulating channel conditions, real TS packet error trace files are fed to the simulator. The results shown in this section are measured based on a relatively challenging channel conditions, where the receiver is moving at a speed of 100km/h and the measured SNR is 20db.

The data bursts were configured in way to allow for sufficient accuracy in the synchronization at the receiver by maintaining equal amount of media duration at the base layer and the enhancement layer. Consequently, the data burst sizes are set to 1 Mbit (512 rows) for the basic service and 2 Mbits (1024 rows) for the enhanced service. MPE-FEC (based on Reed-Solomon) was used in order to provide backwards compatible service configurations. A virtual application data table was used in the case of L-FEC at the enhanced service data bursts. The base layer data was appended to the enhancement layer data in the virtual application data table and then FEC protected. Only enhancement layer data and FEC sections were ultimately transmitted in the enhanced service data burst. L-FEC decoding was performed iteratively to exhaust the full benefit of double protection of the base layer. The procedure starts with FEC decoding at the data burst of the basic service and then proceeds to FEC decoding at the enhanced service data burst and then iterates as long as partially successful corrections were performed.

The video sequences have been encoded in SVC format enabling a base layer in QVGA@15Hz at 300kbps and an enhancement layer in VGA@30Hz at 700kbps.

The following configurations have been analyzed in the simulations:

1. Single data burst for base and enhancement layer as depicted in 4a
2. Separate data bursts for base layer and enhancement layer, which are transmitted in an adjacent manner as depicted in 4b. Only base layer

is FEC protected with the same amount of FEC data as configuration 1.

3. Separate data bursts for base and enhancement layers with interleaved transmission as in 4d. Similarly to configuration 2, only the base layer is protected with the same amount FEC repair data.
4. Data burst arrangement similar to configuration 3. Amount of FEC data is split unequally, 70% of the FEC data is used to protect the base layer.
5. Data burst arrangement is similar to configuration 3. Equal FEC code rate for base layer and enhancement layer with total amount of FEC data equal to that of all the previous configurations.
6. Separate data bursts transmitted in time shifted manner as depicted by 4c. Layered FEC is used to protect base layer together with the enhancement layer at the second burst. Total amount of FEC data is unchanged compared to all previous configurations. 25% of the FEC data is assigned to for base layer protection in the base layer data burst. The rest of the FEC data is used for protection of the base layer and enhancement layer at the enhancement layer data burst.
7. Similar to experiment 6 but with different FEC data partitioning rates. 33% of the FEC data is assigned to base layer data burst.
8. Similar to experiment 6 but with different FEC data partitioning rates. The FEC data is split equally between the data bursts of the base layer and of the enhancement layer.

The simulations were performed for a total FEC code rate (applying to the sum of source data from base layer and enhancement layer) of 3/4.

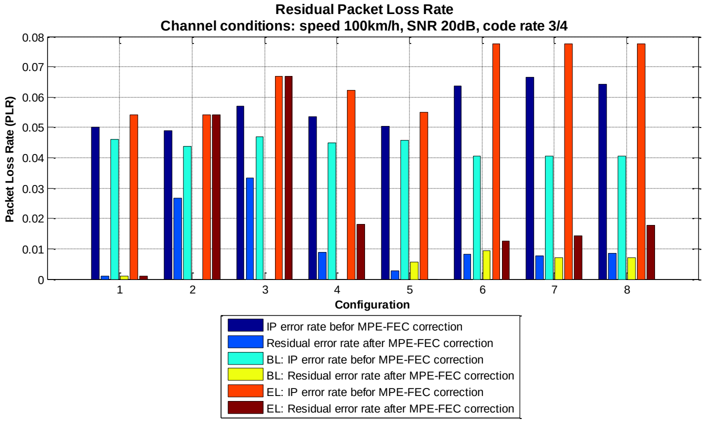


Figure 5 Residual packet loss rate after FEC correction for the different configurations

The resulting packet loss rate before and after FEC decoding at the receiver are depicted for the 6 different configurations in Figure 5. For each configuration, the plots show the total packet loss rates as well as the packet loss rate for base and enhancement layers separately. Although the same error patterns have been used in the simulations, differences in the resulting transmission errors are visible. Those differences are due to the different data burst arrangements. The residual error rate after FEC correction is also different for the different configurations. In the case of configuration 1, the single data burst results in equal packet loss rate and residual packet loss rate for both base and enhancement layer. In configuration 2, all the protection is dedicated to the base layer, so that minimal residual packet loss rate for the base layer at the cost of no correction at all for the enhancement layer is achieved. Layered FEC (configurations 6 through 8) perform worst, as the residual error rate remains high for base and enhancement layer compared to the other configurations.

Figure 6 shows an evaluation of the achieved video quality after transmission and FEC correction for a different set of simulations (reflecting lower packet loss rate averages). The video quality is measured for each configuration for base layer and base and enhancement layer. Configuration 1 shows an equivalent drop in video quality for both the basic service and the enhanced service. Configurations 2, 3, and 4 show the best quality for the basic service. Given the importance of the basic service quality, configuration 4 achieves the best total quality (ranked first based on basic service quality and then on enhanced service quality). This is mainly due to the interleaving of the data bursts and the unequal error protection. Figure 7 shows the measured video quality for a set of video frames and the different configurations. The PSNR curves are indicative of the locations of packet losses that could not be correctly recovered at base and enhancement layers.

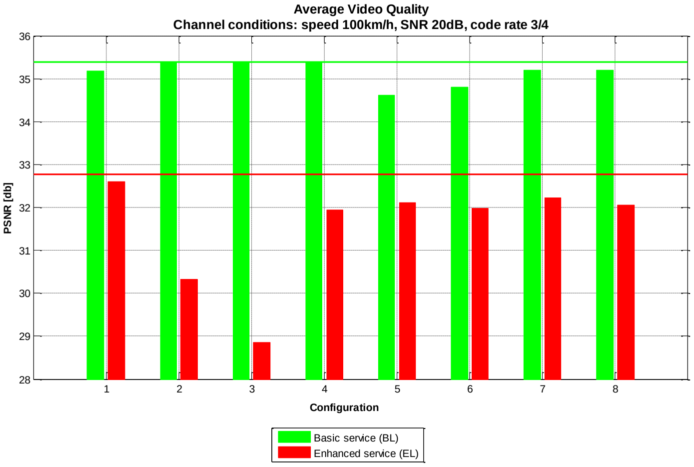


Figure 6 Video quality for the different configurations after transmission errors and FEC correction

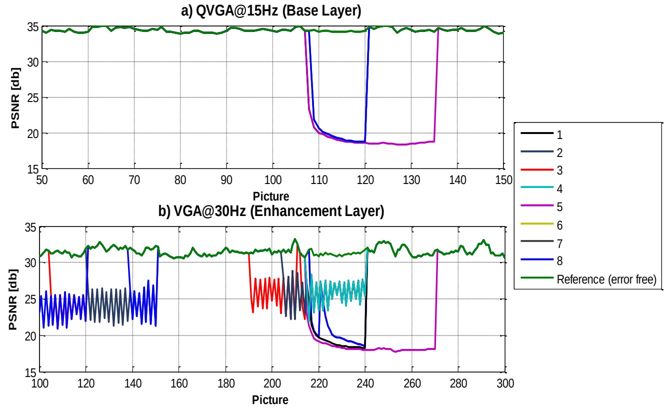


Figure 7 PSNR plot for a set of video frames and the different configurations

CONCLUSIONS

In this paper, different approaches for the multiplexing and FEC protection of SVC based mobile TV services over DVB-H. It has been shown that in a typical scenario, unequal error protection through appropriate setting of the data burst code rates as well as interleaved transmission of the data bursts yields the best compromise between basic and enhanced service. Furthermore, interleaved transmission outperforms time shifted and single burst configurations in terms of overall power consumption, taking into account consumers of the basic service and those of the enhanced service.

REFERENCES

[1] ETSI TS 102 472, *IP Datacast over DVB-H: Content Delivery Protocols*, December, 2006

[2] ITU-T Recommendation, *H.264: Advance Video Coding for generic audiovisual services*, November 2007.

[3] T. Wiegand, G. Sullivan, J. Reichel, H. Schwarz, and M. Wien*, Joint draft 11 of SVC amendment, JVT-X201*, July 2007

[4] ETSI EN 302 304, *Digital video broadcasting (DVB); transmission system for handheld terminals (DVB-H)*, November 2004

[5] ETSI EN 300 744, *Digital Video Broadcasting (DVB): Framing structure, channel coding and modulation for digital terrestrial television*, June 2004

[6] I. S. Reed and G. Solomon, *Polynomial Codes over Certain Finite Fields*, SIAM Journal of Applied Mathematics, vol. 8, pp. 300-304, 1960.

[7] ISO/IEC IS 13818-2, *Information Technology - Generic Coding of Moving Pictures and Associated Audio Information - Part 2: Video*, 1994.

[8] ITU-T Recommendation, *H.263: Video Coding for Low Bit Rate Communication*, 1998

[9] ISO/IEC IS 14496-2, *Information Technology - Coding of Audio-Visual Objects - Part 2: Visual*, 1998

[10] IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, issue 9, September 2007

[11] T. Schierl, C. Hellge, S. Mirta, K. Grüneberg, and T. Wiegand, Using H.264/AVC-based Scalable Video Coding (SVC) for Real Time Streaming in Wireless IP Networks, IEEE International Symposium on Circuits and Systems, pages 3455-3458, May 2007

[12] C. Hellge, T. Schierl, and T. Wiegand, “Mobile TV using scalable video coding and layer-aware forward error correction,” IEEE ICME, June 2008