MOBILE TV USING SCALABLE VIDEO CODING AND LAYER-AWARE FORWARD ERROR CORRECTION

Cornelius Hellge\textsuperscript{1}, Thomas Schierl\textsuperscript{1}, and Thomas Wiegand\textsuperscript{1,2}

\{hellge|schierl|wiegand\}@hhi.fraunhofer.de

\textsuperscript{1}Fraunhofer Institute for Telecommunications Heinrich-Hertz-Institut (HHI)

Einsteinufer 37, D-10587 Berlin, Germany

\textsuperscript{2}Image Communications Group

Technical University of Berlin

Einsteinufer 17, D-10587 Berlin, Germany

\section*{ABSTRACT}

A new approach to error protection for scalable media is presented for mobile TV applications. Mobile TV is typically characterized by a number of receiver capabilities and connection qualities. A broadcast service should preferably work for multiple receiver capabilities without the need for downscalining or transcoding at the battery-powered mobile devices. Moreover, a media quality that gracefully degrades with reception quality instead of a complete signal loss is also a desirable feature. The scalable video coding (SVC) extension of H.264/AVC offers an efficient way to support the aforementioned features. In mobile broadcast channels forward error correction (FEC) is used to overcome packet losses. This work proposes a layer-aware forward error correction (L-FEC) approach in combination with SVC. L-FEC increases robustness of the more important layers by generating protection across layers. L-FEC is integrated as an extension of an Raptor FEC implementation. It is shown by experimental results that L-FEC outperforms traditional FEC and UEP protection schemes.

\textbf{Index Terms}— DVB-H, SVC, Raptor, Application Layer FEC

\section{1. INTRODUCTION}

Digital video broadcasting for handhelds (DVB-H) \cite{1} is becoming an increasingly popular solution for mobile TV. Due to the variety of different device capabilities, e.g. different display resolutions or computational power, transmitting only one video quality could be problematic due to extra computation like downscaling or transcoding at the battery-powered mobile devices. Moreover, due to the mobility of the users, reception quality varies making a media quality that gracefully degrades with reception quality instead of a complete signal loss also a desirable feature.

An approach similar to receiver driven layered multicast (RDLM) \cite{2} in combination with the recently approved SVC extension of H.264/AVC \cite{3}\cite{4} offers an efficient way to provide multiple video signals over a broadcast channel as shown in previous work \cite{5}. In such an RDLM scenario, a client joining the service only requests the scalable layers, which provide a signal that the device is capable or chooses to process. Transmission of multiple video signals using SVC is much more efficient in terms of bit rate compared to simulcast transmission \cite{6}.

A mobile broadcast channel is typically showing burst errors. With the delay constraints of a broadcasting service, reliable transmission is a big challenge. Since broadcast services only provide a unidirectional downlink channel, a possible solution to increase reliability is the use of additional forward error correction (FEC) at the link or application layer. FEC is applied in DVB-H at the link layer with the optional multi protocol encapsulation FEC (MPE-FEC) and using Raptor coding \cite{7} as application layer FEC. Although MPE-FEC is intended to be used for streaming services and application layer FEC for file download, the application layer FEC offers more flexibility for media aware protection. Possible approaches are unequal error protection (UEP) \cite{8}, priority encoding transmission (PET) \cite{9} or dependency aware FEC as proposed in \cite{10} or \cite{11}. These approaches do not take the existence of quality layers and the multiple dimensions of dependencies in a video stream into account.

SVC allows up to three different scalability dimensions within one bit-stream. Scalability in SVC can be applied to the temporal, spatial and quality dimension. The proposed layer-aware FEC (L-FEC) approach generates redundancy symbols following dependencies in the scalability dimensions, i.e., symbols of layers of lower importance can be used to correct symbols of layers of higher importance. The proposed transmission scheme uses RDLM in combination with SVC to serve different device capabilities. Furthermore, this work introduces L-FEC as extension of the Raptor code as defined in DVB-H to increase the reliability of an SVC transmission in a RDLM scenario.

The rest of the paper is organized as follows. Section 2 gives a very brief overview of the SVC standard. Section 3 introduces the Raptor L-FEC defined in DVB-H and in Section 4, we outline the proposed layer-aware FEC extension. In section 5, we apply the L-FEC to the Raptor defined in DVB-H and in 6 we show selected simulation results.

\section{2. SCALABLE VIDEO CODING - SVC}

The SVC design, which is an extension of the H.264/AVC video coding standard, can be classified as a layered video codec. An SVC bit-stream can be structured so that devices with different capabilities can decode parts of it that have a quality very similar to the case when the bit-stream for each device would be a single-layer H.264/AVC bit-stream. In SVC, the hybrid video coding approach of motion-compensated transform coding of H.264/AVC is extended in a way that a wide range of spatio-temporal and quality scalability is achieved. The base layer is an H.264/AVC compliant bit-stream that ensures backward-compatibility for existing receivers. The temporal scaling functionality of SVC for high delay configurations is typically based on a temporal decomposition using hierarchical bi-predictive pictures. The spatial scalability is achieved by different encoder loops with an over-sampled pyramid for each resolution. For details of SVC, see \cite{3}\cite{4}.
3. RAPTOR CODE IN DVB-H

The approach proposed in this work extends the Raptor code as specified in DVB-H [7]. Raptor codes belong to the category of Rateless or Fountain codes. Such a type of an FEC code can produce a theoretically infinite number of encoding symbols (ESs) from a limited number of source symbols (SSs) with linear complexity. The receiver can recover the original data by an inverse encoding process after receiving an amount of ESs only slightly larger than the number of SSs.

\[ \begin{bmatrix} 1 & 0 & 0 & 1 & \ldots & 0 \\ 0 & 0 & G_p & 1 & \ldots & I \\ 0 & 1 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 1 & 1 & \ldots & 1 \\ 0 & 0 & 1 & 1 & \ldots & 1 \\ 0 & 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 1 & 1 & \ldots & 0 \\ 1 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 1 \\ 0 & 0 & 0 & 1 & \ldots & 1 \end{bmatrix} \]

**Figure 1: The Raptor constraint process**

The Raptor code specified by DVB-H is a systematic code based on a concatenation of a Luby-Transform (LT) Code, and an additional pre-code, producing intermediate pre-coded symbols (PSs) used as input symbols of the LT-Code. Below is a brief description of the Raptor encoding process. We mainly focus on the LT- and pre-coding, since these parts have to be modified for the approach proposed in this work to keep the code systematic. More details about the Raptor design can be found in [12].

\[ \begin{bmatrix} 1 & 0 & 0 & 1 & \ldots & 1 \\ 0 & 0 & 1 & 1 & \ldots & 1 \\ 0 & 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 1 & 1 & \ldots & 0 \\ 1 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 1 \\ 0 & 0 & 0 & 1 & \ldots & 1 \end{bmatrix} \]

**Figure 2: LT encoding matrix \( G_{LT} \)**

For Raptor encoding, first the input data is divided in \( k \) source symbols \( SS_{0:k-1} \) each of a size \( t \). Solving the constraint process in equation (2) generates the pre-code symbols \( PS_{0:p-1} \) which are used as input for the LT encoding process (Figure 2). The PSs are generated by an \( s \times k \) generator matrix \( G_p \), where \( s \) denotes the number of parity symbols. In a non-systematic code, the parity symbols would be used together with the SSs as input for LT encoding. To compensate the XOR’ing of the PSs in the later LT-coding process, i.e. the SSs are no longer present in the ESs, the generator matrix \( G_{LT} \) is also introduced in the pre-code generation. Therefore an additional constraint is introduced with the \( G_{LT} \) matrix causing the LT-encoding of the first \( k \) pre-coding symbols \( PS_{0:k-1} \). Results in the original source symbols \( SS_{0:k-1} \) and therefore in a systematic code. If a sub matrix \( I \) depicts an \( s \times s \) identity matrix, the constraint matrix \( G_{pSYS} \) is defined in (1).

\[ G_{pSYS} = \begin{bmatrix} G_p & I \\ G_{LT} \end{bmatrix} \]  

The pre-code symbols \( PS_{0:k-1} \) can be computed solving the system of equations built by the constraint process in equation 2.

\[ G_{pSYS} \cdot PS = \begin{bmatrix} 0 \\ SS \end{bmatrix} \]  

Different XOR combinations of the PSs compute each ES. These combinations build the LT encoding matrix \( G_{LT} \) as depicted in Figure 2, whereas all PSs with a Boolean 1 in the matrix \( G_{LT} \) are XOR’ed. The complete encoding operation can be expressed with

\[ ES[0:n-1] = G_{LT}(0,1, \ldots, n-1) \times PS[0:p-1] \]  

The Raptor code can produce a theoretically unlimited number \( n \) of the ESs whereas the code rate \( c \) of the resulting bit-stream is defined by \( c = \frac{k}{n} \).

4. LAYER-AWARE FORWARD ERROR CORRECTION

Using layered multicast, typically the redundancy symbols are generated separately for each layer. The idea of L-FEC is, to follow media coding dependencies in the media stream, to generate redundancy across layers. Using the proposed approach, ESs of less important layers can be jointly used with ESs of more important layers for recovering the SSs of all participating layers. Due to the dependencies within the SVC bit-stream, lower priority layers and the associated redundancy symbols cannot be used without successfully decoded higher priority layers.

A dependency path (DP) contains all referenced layers for decoding a particular frame in the order of importance. Using layered FEC, all redundancy symbols in the same DP can be jointly used for error correction. For the sake of convenience, Figure 3 depicts the layered FEC generation for only one scalable dimension. Note that the presented approach uses multiple dimensions of dependencies as present in SVC.

**Figure 3: Simplified L-FEC for only one dimension**

After partitioning of the media bit stream into \( L \) dependency layers, the redundancy symbols FEC 0 of the highest priority layer \( l=0 \) are typically generated given by the FEC coding technique \( T \). Considering the dependency structure, redundancy symbols of the enhancement layer \( l=x \) are calculated incorporating \( SSs \) of all layers \( l \leq x \). FEC 1 symbols are generated over \( SSs \) of layer \( l=0 \) and layer \( l=1 \). Furthermore, FEC 2 symbols are generated over \( SSs \) of layer \( l=0 \), layer \( l=1 \) and layer \( l=2 \) and so on up to FEC \( L-1 \), which is generated over \( SSs \) of all lower layers \( l\leq0 \) to \( l=L-1 \). Using the \( L \)-FEC approach, the number of redundancy symbols remains constant but the redundancy symbols of different layers, but same DP, can be jointly used for error correction.
5. SYSTEMATIC RAPTOR CODES FOR L-FEC

To apply the idea of L-FEC to the presented systematic Raptor FEC, the encoding and pre-coding matrices, \( G_{LT} \) and \( G_{PS} \), have to be modified for dependency layers \( l > 0 \) following the layered FEC procedure shown in section 4. The standard LT encoding matrix \( G_{LT} \) has the dimensions \( pxn \). To extend the XOR’ing to lower layers and to keep the code rate constant, the modified generator matrix \( G_{LTm} \) of layer \( m \) has to be extended with the number of the \( PSs \) of all lower layers \( l < m \) to a \( \binom{G_{LT}}{G_{P0}} \times n \) matrix. To extend the encoding matrix \( G_{LTm} \) of layer \( m \), \( G_{LTm} \) can be concatenated with the encoding matrices of the lower layers \( G_{LTi} \) building the layered encoding matrix \( G_{LayeredLT}(m) \):

\[
G_{LayeredLT}(m) = [G_{LT0}; G_{LT1} \ldots ; G_{LTm}]
\]

(4)

This procedure allows the use of the LT-coding algorithms defined in DVB-H. Figure 4 depicts the LT encoding matrices \( G_{LT} \) and the extended \( G_{LayeredLT}(1) \) for two layers. Note, that the proposed extension can be applied to multiple dependency layers over multiple dimensions.

![Extended LT encoding matrix](image)

Figure 4: Extended LT encoding matrix

\( ESs \) of layer \( l=0 \) are generated in typical way, as shown in section 3, and the associated encoding matrix \( G_{LT0} \) corresponds to the standard matrix in Figure 2. \( ESs \) of layer \( l=1 \) are generated by a concatenated matrix \( G_{LayeredLT}(1) \) following equation (5).

\[
ESs_{1[0:n-1]} = G_{LayeredLT}(1)[0,1,\ldots,n1-1] \cdot [PSs_{0[0:p0-1]}PSs_{1[0:p1-1]}]
\]

(5)

For a systematic code, the pre-code must guarantee, that the \( ESs \) generated by the LT-encoding contain the original \( SSs \). Following equation (1), the modified pre-coding matrix \( G_{PS_{Layered}}(1) \) of layer 1 is defined as described in equation (6) and depicted in Figure 5.

\[
G_{PS_{Layered}}(1) = \begin{bmatrix}
G_{P0} & I \\
G_{LT0} & 0 \\
0 & G_{P1} \\
G_{LayeredLT}(1)
\end{bmatrix}
\]

(6)

Adding \( G_{LayeredLT}(1) \) instead of \( G_{LT}(1) \) to the pre-coding matrix compensates the matrix extension of the LT encoding matrix in Figure 4. Figure 5 shows a suitable extended pre-coding matrix for two layers. \( PSs \) are calculated in typical way, but \( PSs \) of layer 1 are calculated by solving the layered constraint process in equation (7).

![Extended pre-coding matrix](image)

Figure 5: Extended pre-coding matrix

Finally, \( ESs \) of the systematic Raptor for L-FEC are generated following equation (5). Resulting \( ESs \) of layer 1 are XOR combinations of layer 0 and 1, whereas the first \( k_1 \) symbols correspond to the original \( SSs \) of layer 1.

6. SIMULATION RESULTS

In this section, we present selected results for transmission of QVGA and VGA resolution using SVC and RDLM over a DVB-H channel. A Gilbert-Elliot (GE) model is used as statistical model for simulation of burst losses on the DVB-H channel as similarly used in [13]. The transmission blocks (TBs) of the wireless channel are of size 186 bytes. The mean error burst length of the GE model is about 100 TBs. The CIRCLE sequence is encoded at 25 frames per second with a length of 1297 frames using the SVC reference software JSMV8.8 with a H.264/AVC base layer at QVGA resolution and a spatial enhancement layer at VGA, a group-of-picture (GOP) size of 16 and random access point at each second GOP. In the case a VGA receiver does not receive the spatial enhancement layer, we calculated the PSNR value of an up-scaled QVGA resolution. Freeze frame error concealment is used in case of a lost base layer. The base layer is protected with a standard Raptor code, whereas the spatial enhancement layer protection uses the L-FEC approach. \( ESs \) are computed over the amount of data between the random access points. The code-rate and the operation points of the layered SVC stream are depicted in Table 1.

<table>
<thead>
<tr>
<th>Source Bit Rate</th>
<th>PSNR QVGA</th>
<th>PSNR VGA</th>
<th>Code-rate UEP/FEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>H.264/AVC Base layer (QVGA)</td>
<td>222 kbps</td>
<td>36.10 dB</td>
<td>33.26 dB</td>
</tr>
<tr>
<td>Spatial Enh. Layer (VGA)</td>
<td>275 kbps</td>
<td>-</td>
<td>36.84 dB</td>
</tr>
</tbody>
</table>

Table 1: Operation points of the SVC stream

The results for a QVGA and a VGA receiver of 200 test runs are shown in Figure 6 and Figure 7, where the Y-axis shows the mean received video quality in terms of PSNR over different TB
loss rates. We compared two settings using different code rate distributions \textit{UEP} and \textit{FEC}, as shown in Table 1. Both settings provide the same transmission bit rate of 840 kbps and are compared with standard FEC encoding for both layers (\textit{Normal}) and with the L-FEC in the spatial enhancement layer (\textit{Extended}). If the QVGA device receives insufficient symbols to decode the VGA layer, it can continue receiving the VGA layer. Due to the L-FEC, the additional symbols can be used for combined error correction (\textit{Extended+VGA}) to finally decode the QVGA layer.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{CIRCLE Sequence at a VGA receiver}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure7.png}
\caption{CIRCLE Sequence at a QVGA receiver}
\end{figure}

For the QVGA receiver in Figure 6, the \textit{UEP} scheme for \textit{Normal} and \textit{Extended} outperforms \textit{FEC} due to the higher code-rate in the base-layer. Whereas the L-FEC in the \textit{Extended} scheme does not show any gain, since the QVGA base layer is protected by standard Raptor encoding. However, the \textit{Extended+VGA} scheme shows a significant gain in terms of PSNR due to the increased protection capability by the additional reception of the VGA layer and the use of the L-FEC for combined error correction. At the VGA receiver, the L-FEC approach shows a gain in PSNR for both settings. Using the L-FEC the \textit{UEP} and \textit{FEC} show a similar performance at lower loss rates. Whereas the \textit{UEP} scheme shows a gain due to the higher protection of the base layer, the \textit{FEC} scheme shows a relative higher gain due to the higher protection in the enhancement layer, which now also protects the more important layer. In this scenario the L-FEC approach cannot show a weaker performance than normal FEC, i.e. the across layer FEC generation follows the existing dependencies.

\section{7. CONCLUSION AND SUMMARY}

In this work, we propose a layer-aware forward error correction (L-FEC) approach. L-FEC generates redundancy symbols incorporating layered structures in layered media codecs like SVC. SVC is used to transmit two different resolutions at the same time. We applied the L-FEC approach to a Raptor code. The L-FEC approach enhances the protection capability of the base layer without increasing the bit rate. Simulation results in an RDLM-like scenario show that the proposed approach outperforms a standard layered UEP scheme. The use of the L-FEC with multiple layers shows a strong performance compared to the standard FEC scheme, where each layer is protected separately.

\section{8. REFERENCES}