Media aware FEC for Scalable Video Coding Transmission

Lukasz Kondrad[1], Imed Bouazizi[2], Moncef Gabbouj[1]
[1]*Tampere University of Technology*, [2]*Nokia Research Center*
lukasz.kondrad@tut.fi, imed.bouazizi@nokia.com, moncef.gabbouj@tut.fi

Abstract

Scalable video coding (SVC) is an extension to H.264/AVC that enables encoding a video sequence once and using it subsequently at multiple operations points by different devices and applications. Forward Error Correction techniques may then be applied to the encoded video in order to enhance the robustness against transmission errors. Media aware FEC code construction caters for scalable video streams by adjusting the protection level for each video layer.

In this paper, we discuss different approaches for media aware Forward Error Correction using Scalable Video Coded video that is transmitted over error prone channels. We describe two recent solutions for Unequal Error Protection and then evaluate those through multiple simulations.

1. Introduction

Wireless transmission channels are prone to bursty time-varying errors due to radio channel characteristics. Compressed video, an important part of multimedia data, is sensitive to those errors due to the extensive use of predictive coding in the compression process. Therefore, to enable robust transmission of multimedia data, channel coding with forward error correcting (FEC) codes are frequently used where most common FEC codes are Reed-Solomon (RS) [1] and Raptor codes [2].

Channel coding solutions are typically media unaware which means the same level of protection is applied to all data of a source block. On the other hand, compressed video data is composed of data parts that are of unequal importance to the decoding and presentation of the video. In these cases, media aware unequal error protection (UEP) has the potential to provide improved rate-distortion performance. Efficiency of any UEP method depends on accurate priority assignment to media packets. Therefore, numerous priority assignment schemes were researched and their results reported.

The new Scalable Video Coding (SVC) standard [3] was created to provide spatial, temporal, and quality adaptation. Additionally, SVC allows easy priority assignment by the exploitation of the intrinsic media data importance (e.g. based on the SVC layer those media units belong to). Consequently, SVC is suitable for unequal error protection which could provide higher error and loss resilience during transmission.

This paper is organized as follows. Section 2 gives background information on SVC and media aware FEC methods. Section 3 describes the media aware UEP algorithm proposed by the authors. In Section 4, simulation setup for comparing five different FEC methods is presented, and the results of the simulations are discussed. Finally, Section 5 contains our conclusions and future work.

2. Background

This section provides background information needed and is organized in the following way. First, basic information about Scalable Video Coding are presented. Next, two approaches for media aware FEC are described.

2.1. Scalable Video Coding

The new SVC standard is an extension to H.264/AVC standard [3] which allows for temporal, spatial and quality scalability in a video bit-stream.

The idea behind SVC is that the 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

978-1-4244-4671-1/09/$25.00 ©2009 IEEE

resolution or picture fidelity. The base layer is an H.264/AVC bit-stream which ensures backwards compatibility to existing receivers.

When temporal scalability is used frames from the higher layers 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. In the case of spatial scalability, the encoded bit-stream contains sub-streams that represent the same content at different spatial resolutions. 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 [4].

2.2. Media aware FEC

Many contributions have been made in the literature to address the issue of robustness against packet loss in multimedia data transmission. The pioneering work by Albanese et al [5] introduces media aware packet handling in the so-called Priority Encoding Transmission scheme (PET). PET was one of the first packet based data protection schemes, which operates by discarding less important media packets while retaining the more important ones when the transmission channel degrades. The PET scheme gives principles for all the rest of media aware protection methods.

Numerous modifications of PET schemes for UEP transmission were researched and their results reported in literature. However, only two approaches of UEP for SVC transmission, which later on will be evaluated in Section 4, are described here.

A novel solution for error protection for scalable media data transmission is described in [6]. Hellge et al. proposes to increase the robustness of more important layers by generating protection across layers. This approach is similar to receiver driven layered multicast (RDLM) [7]. Additionally, authors describe an extension of the Raptor FEC code which allows the implementation of a Layered FEC solution. Though, for simplicity reasons, for simulations described in Section 4 Reed-Solomon code was used. Layered FEC arrangements where FEC data from layer *n* protect the source data from layer *n* together with data from all lower layers is presented on Figure 1.

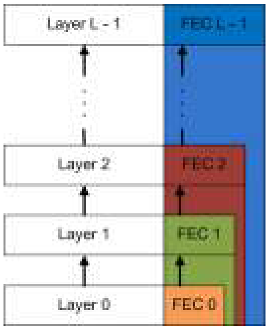

Figure 1: Layered FEC arrangement

Another approach was proposed by the authors of this publication in [8]. A novel method for constructing an UEP code targeted towards video streaming applications, where concatenation of a set of parallel outer block codes followed by a packet interleaver and an inner block code, was described. Additionally, an algorithm which calculates on the fly the optimal allocation of the code rates of the inner and outer codes was given. A description of the algorithm is presented in Section 3.

3. FEC code construction algorithm

The media data is treated as FEC source blocks. A FEC source block is a set of media packets to be delivered over a given channel with a given level of protection. The level of protection is expressed by the FEC code rate which is used to deduce the amount of repair data that is allowed for a given source block. Based on the importance of each media packet for the decoding process (and hence also for the overall quality), the source block is partitioned into m sub-blocks. To each sub-block an error protection level, depending on its priority, is then assigned. The optimal number of sub-blocks m is determined by the algorithm. The main concept of the FEC code concatenation is presented on Figure 2.

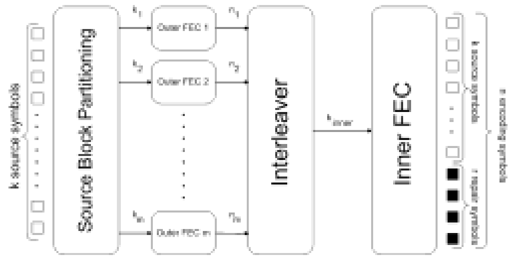

Figure 2: FEC code concatenation for UEP

A simple interleaver was deployed to re-distribute the bursty errors so they become random errors. Consequently, an assumption that packet loss over the communication channel is an i.i.d process with a packet loss probability *p* for all packets (independent of their size) can be made.

In a system where an equal error protection *(n,k)* maximum distance separable (MDS) block code such as Reed-Solomon is in use, the residual error rate is calculated as follows:

$$P_{EEP} = \sum_{l=n-k+1}^{n} \frac{l}{n} \times P(l \mid p)$$

where

$$P(l \mid p) = \binom{n}{l} (p)^l (1-p)^{n-l}$$

The result may be interpreted as the probability that a given packet is lost along with at least *n-k-1* other packets, which would mean that it is not recoverable using the EEP FEC.

The objective of the algorithm is to optimally partition the source block and allocate the share of repair data to each of the resulting sub-blocks as well as calculate optimal split of repair data between outer codes and the inner code. Let k_i be the number of packets in partition $i \mid i \in \{1..m\}$. Let **Pri***ᵢ* represent the priority of partition *i*. The resulting concatenated code minimizes the overall expected distortion D_{total} as defined by the following equation:

$$D_{total} = \sum_{i=1}^{m} \mathrm{Pri}_i \times P_{loss}(i)$$

$P_{loss}(i)$ represents the residual packet loss rate for packets that belong to partition *i*. $P_{loss}(i)$ depends on the code rate of the outer code of partition *i*, k_i / n_i, and of the code rate of the common inner code. $\frac{\sum_{i=1}^{m} n_i}{n}$

The residual packet loss rate Ploss (i) for partition class *i* is calculated according to the following equations:

$$P_{loss}(i) = \sum_{l=n_i-k_i+1}^{n_i} \frac{l}{n} \times P(l \mid P_{EEPinner})$$

where

$$P_{EEPinner} = \sum_{l=n-k_{inner}+1}^{n} \frac{l}{n} \times P(l \mid p)$$

Based on the previous residual packet loss calculations, we propose the following low-complexity heuristical algorithm for exploring the optimal code construction.

```
1. set $m = 1$ and $P_{total_{\min}} = \infty$
2. set $k_i = \frac{k}{m} \forall i \in 1..m$
3. set $\mathrm{Pri}_i$ as the average priority of all packets of partition $i$
4. set $k_{inner} = k + 1$
5. allocate $n_i$ for each partition $i$ so that $\sum_{i=1}^{m} n_i = k_{inner}$ and
   proportional to $\mathrm{Pri}_i$
6. calculate $P_{loss}(i) \forall i \in 1..m$
7. calcualte $D_{total} = \sum_{i=1}^{m} \mathrm{Pri}_i \times P_{loss}(i)$
8. store $D_{total}$ if smaller than $D_{total_{\min}}$
9. set $k_{inner} \leftarrow k_{inner} + 1$
10. if $k_{inner} = n$ set $m \leftarrow m + 1$ and goto step 2
11. else goto step 5
```

The algorithm checks all practical constructions from a single partition to $m_{\max}$ partitions. For each construction, the amount of repair data assigned to each partition is proportional to the average priority of that partition. Finally, the estimated overall distortion for the construction is calculated and the construction is marked as interesting if the resulting distortion is below all prior constructions. As a result, the construction with the minimal distortion expectation is determined. This algorithm is of low complexity and may run in real-time scenarios for each source block without significant additional delay.

4. Evaluation

In this section, we present and discuss the results of a comparison of two different FEC methods, which were introduced in Section 2, together with traditional EEP approach. Their impact on the residual frame loss and consequently perceived quality is shown.

A simulator for generic wireless transmission channel with Gilbert-Elliot error model was used throughout the simulations. The City sequence was encoded using SVC encoder with a H.264/AVC base layer at QVGA resolution and a spatial, quality, and temporal enhancement layers. The code-rate and the operation points of the layered SVC stream are depicted in Table 1. The sequence contained 300 frames and was played with 30 frames per second, the maximum slice size was set to 1200 bytes and Intra Decoder Refresh frames were inserted every 30 pictures.

Layer	Resolution	Frame rate	Bitrate [kbps]	PSNR [dB]
L0	320x240	15	100	31,597
L1	320x240	15	510	34,350
L2	640x480	15	775	31,509
L3	640x480	30	830	31,441
L4	640x480	15	2100	34,240
L5	640x480	30	2200	33,934

Table 1: Operation point of the SVC stream

During simulations five different configurations were analyzed. In each of the configurations the same amount of repair data was used. The following arrangements of FEC repair data were examined:

1. equal error protection (eep) with code rate 5/6
2. unequal error protection (uep) where repair data is divided between layers by the algorithm described in Section 3. Packet loss probability *p* in the algorithm is fixed to 0.2.
3. layered FEC (lfec) where repair data is assigned to each layer based on original code rate 5/6 and amount of source data in a layer (data from depended layers is not taken into account during division of the repair data).
4. layered FEC (lfec2) where repair data is divided equally between all layers independently of amount of source data in each layer.
5. layered FEC (lfec3) where repair data is split between all layers respectively: L0-33%, L1-25%, L2-16%, L3-12%, L4-8%, L5-6%.

The resulting frame loss rate after FEC decoding for each layer is depicted in Figures 3 to 7 respectively. Frame has been marked as lost if any of packets belonging to the frame or any depended frame from lower layers has been marked as lost. The results are mean value from 200 tests runs.

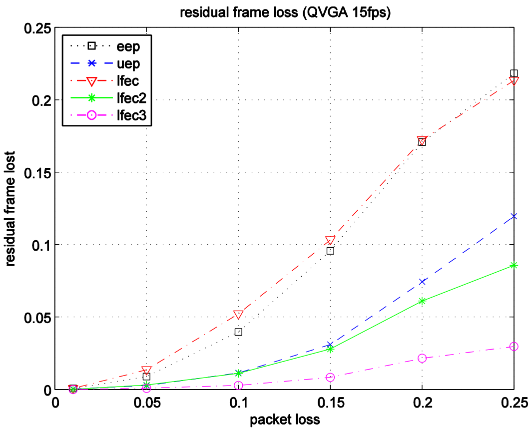

Figure 3: Residual frame loss for layer L0

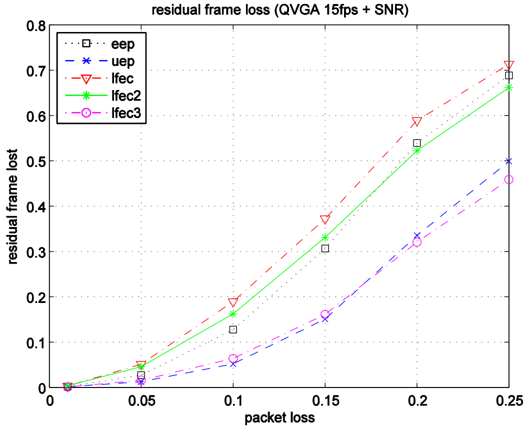

Figure 4: Residual frame loss for layer L1

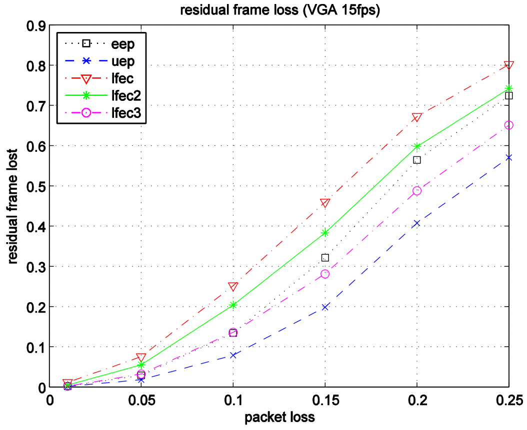

Figure 5: Residual frame loss for layer L2

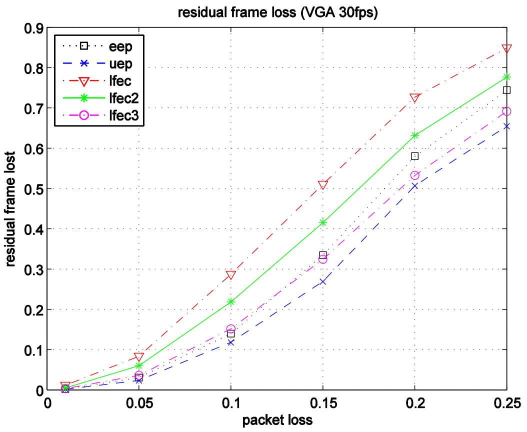

Figure 6: Residual frame loss for layer L3

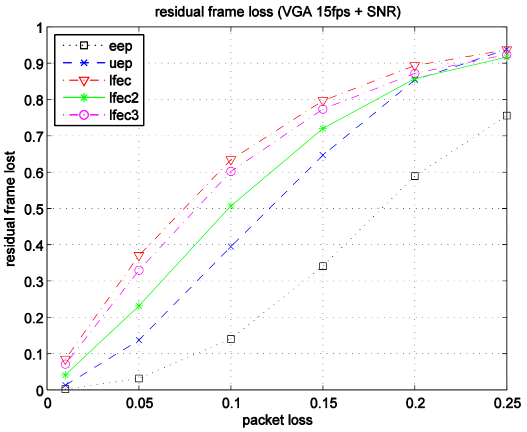

Figure 7: Residual frame loss for layer L4

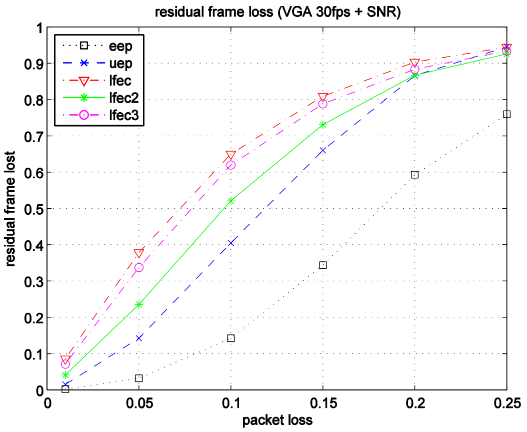

Figure 8 Residual frame loss for layer L5

Although the same amount of repair data was used in all configurations, differences in the residual frame loss are visible. Those differences are due to the different repair data assignment approaches.

Configuration 1 (the EEP method) shows better performance compared to the UEP methods only for layer L4 and layer L5, Figure 7 and 8 respectively. However, the UEP methods better protect frames with lower quality and in consequence the missing frames on layer L4 or layer L5 may be replaced by frames from lower layers. The frame copy solution, however, could not bring benefits while EEP method is used due to bigger frame losses in lower layers.

Configuration 2, thanks to the optimal repair data allocation, shows the best performance for layers L1, L2, L3 and it results in lower residual losses compared to Layered FEC methods for layer L0.

Configuration 3 of Layered FEC does not bring any gains compared to EEP method. Configuration 4 shows benefits only for layer L0 and performs equally or worse than EEP method for other layers. This is due to the fact that the amount of data assigned to each layer is the same but the base layer has lower amount of source data in comparison to upper layers, so that the code rate for the base layer is lowest, which results in higher protection of the base layer.

Configuration 5 shows better results compared to EEP for the first four layers. It gives the best results for layer L0 at the cost of degraded performance for the other layers. Nevertheless, for layers L1, L2 and L3 it is only outperformed by Configuration 2.

To gain a deeper insight on how picture degradation (or frame loss) impacts the quality of the received video, a comparison of PSNR values for layer L2 over different packet loss rates was calculated. During decoding no frame copy from lower layers was used. The results are depicted in Figure 9. The UEP method maintains an acceptable video quality at L2 even at relatively high packet loss rate conditions.

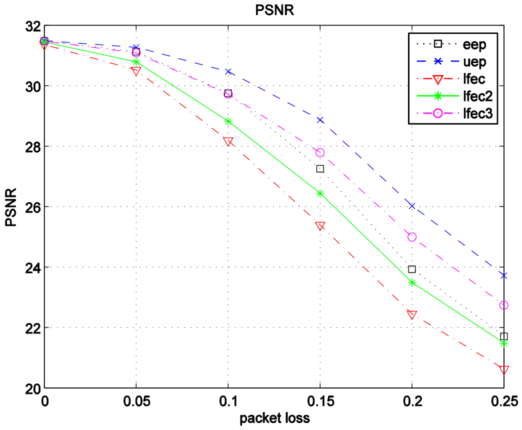

Figure 9: PSNR values for layer L2

It is also worthwhile noting that there is a trade-off in terms of which layers are perceived as important. The inter-layer dependency contributes to the complexity of this trade-off. However, the UEP strikes a good balance between the achieved video quality results at the different operation points (layers).

5. Conclusion

In this paper, we provided a description and evaluation of two methods for robust transmission of SVC stream. Moreover, we compared them against the traditional EEP solution. The results show that both of the above described UEP approaches outperformed non media aware FEC. In spite of the fact that the paper does not conclude in addressing the question of which UEP method is more efficient, the results for Layered FEC demonstrate an big impact of the proper split of repair data among layers on the performance. On the other hand, the results from Configuration 2 demonstrate the benefits from optimized repair data partitioning and concatenated codes.

6. References

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

[2] A. Shokrollahi, *Raptor codes*, IEEE Transactions on Information Theory 52, 2551–2567 June 2006)

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

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

[5] A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan, *Priority encoding transmission*, in IEEE Symposium on Foundations of Computer Science], 604–612 November 1994

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

[7] S. McCanne, V. Jacobson, and M. Vetterli, *Receiver-driven layered multicast*, ACM SIGCOMM, August 1996.

[8] L.Kondrad, I. Bouazizi, V. K. M. Vadakital, and M. Gabbouj, *Optimal FEC code concatenation for unequal error protection in video streaming applications*, In proceedings of 16’th Multimedia Computing and Networking, January 2009