Error tolerance schemes for H.264/AVC: an evaluation

Tuan Tran Thai, Jérôme Lacan, Hugo Méric
University of Toulouse; TéSA; ISAE/DMIA; Toulouse, France

***Abstract*—Video transmission is sensitive to losses due to high compression efficiency. To tolerate the quality degradation from losses, Forward Error Correction (FEC) and error resilience schemes are commonly used. In this paper, we evaluate the performance of error tolerance schemes with the latest video coding standard, H.264/AVC. The analysis in three zones of packet loss rates (PLR) shows that no FEC scheme outperforms the others in a wide PLR range. We also compare the equal and unequal FEC schemes with the Flexible Macroblock Ordering (FMO) error resilience mechanism and find that FMO performs well in moving videos while FEC codes are better in rather static videos. Our results and analysis would give insights to design flexible applications which are able to adapt to the network dynamics.**

I. Introduction

Video traffic recently has a significant growth in the Internet. Real time video transmission in the networks becomes a challenging task since it requires effective data compression to save bandwidth and storage. The current standard video codec, H.264/AVC (Advanced Video Coding), has shown better compression performance than the previous standard codecs such as MPEG-4 Part 2, H.263 [1]. However, the high compression makes the video coding more sensitive to errors and losses in the networks. Even a small number of losses can degrade significantly the video quality.

Different approaches have been proposed to mitigate the video quality degradation due to losses. One of the approaches uses the coding tools provided by H.264/AVC standard (i.e., data partitioning, FMO) to tolerate the errors or losses in the networks [2]. However, these tools usually come with a price of lower coding efficiency. Another approach uses FEC codes by adding the redundancy to protect the video in error prone environments. A known proposal, Priority Encoding Transmission (PET) system [3], protects the video based on its priority function. In [4], the authors propose the Dependency-Aware Unequal Erasure Protection (DAUEP) scheme which protects the video based on the relationships between video frames. However, these schemes are proposed and assessed with MPEG-4 Part 2 video coding standard. Furthermore, the extension of H.264/AVC, H.264/Scalable Video Coding (H.264/SVC) [5], has not been widely used.

In this paper, we evaluate the error tolerance schemes with the latest standard video codec, H.264/AVC. First, we compare the performance of DAUEP, PET and other FEC codes with two H.264/AVC video profiles (Baseline and Extended). We find that no scheme outperforms the others in a wide PLR range by the analysis in three zones of PLR. The wide PLR range shows the different environment characteristics from less lossy networks (i.e., the Internet with losses mostly due to congestion or link outage) to error prone networks (i.e., wireless LAN with loss rate up to more than 30% in some periods [6], [7]). We also show the impact of tuning the redundancy parameters on video quality. Then, we compare the FEC schemes with the FMO in different reference videos. To the best of our knowledge, we do not find any work providing such comparison. Our analysis shows the benefits of FMO in the videos with moving background and objects, while FEC protection schemes perform well in rather static videos.

The rest of this article is organized as follows. Section II introduces the different error tolerance schemes applied to H.264/AVC standard. Section III provides the detail setup and configuration for evaluating. In Section IV, we present our simulation results and analysis. Then, Section V provides some discussion about the results interpretation. We conclude and provide our future work in Section VI.

II. Error tolerance schemes

A video is composed by a sequence of consecutive frames. Typically, video coding has 3 frame types namely I, P and B. I frames are intracoded, thus, they can be independently decoded and displayed. While P frames are coded based on previous one or more reference frames (I or P). Hence, the proper decoding of P frames requires the correct decoding of previous frames to which they refer. On the other hand, B frames are coded based on a pair of previous and following reference frames (I or P). A sequence of I frame followed by P and B frames and terminated by another I frame composes a Group of Picture (GOP). For instance, a video with a GOP size of 10 can be composed by IBPBPBPBPBIBPB... By this dependency nature, the loss of B frame does not cause the distortion propagation to other frames in a GOP while the loss of P frame causes the error propagation from this frame to the end of GOP. More importantly, the loss of I frame leads to an error propagation on that entire GOP. Therefore, I frame is the most important frame in a GOP followed by P and then B frames. The unequal importance of frame types in a GOP motivates the proposal of unequal error protection (UEP) schemes.

In the family of block-based erasure codes, Maximum-Distance Separable (MDS) codes show their advantages in erasure coding capability per block [8]. They can recover k information packets when receiving any k among n packets where $n-k$ are repair packets. In the context of video transmission, MDS codes equally protect all packets in a GOP where $n-k$ repair packets are built from the linear combination of all k packets of I, P and B frames. Thus, MDS codes are considered as Equal Error Protection (EEP) schemes. MDS time-sharing (MDS-TS) codes [9] can be used for unequal error protection to deal with the unequal importance of frame types. With given redundancy ratios for I, P and B frames, MDS-TS builds 3 sub-matrices where each sub-matrix equally protects all packets of each frame type.

Let us show an example where a GOP pattern and the number of packets required for each frame are shown in Table I. Let us assume that $n-k=5$ repair packets protect $k=10$ source packets in a GOP. We assign two repair packets protecting I-frame, two others protecting P-frames and last repair packet protecting all frames. The generator matrix of MDS-TS is shown in Figure 1 where I_{10} is the identity matrix of size 10 and $\alpha_{i,j}$ ($i \in \{1...k\}, j \in \{1...n-k\}$) are coefficients in finite field $\mathbb{F}_q$.

TABLE I
NUMBER OF PACKETS REQUIRED FOR EACH FRAME

Frame	I	B	P	B	P	B
Number of packets	3	1	2	1	2	1

The idea of DAUEP is to build the repair packets based on the dependency relationships between different frame types in a GOP. DAUEP constructs the generator matrix based on the priority of each frame type in a GOP. In this article, we use the principle of DAUEP to build the generator matrix which distributes repair packets protecting P frames in a GOP with decreasing level. This implies that all packets of P frames are protected and the first P frames are more protected than the later ones. The generator matrix of DAUEP is shown in Figure 1. It should be noted that we limit the coding block length to the GOP size in order to avoid the delay induced by long blocks. For this reason, we do not evaluate other block codes like LDPC [10] or Raptor codes [11] which have lower correction capability for blocks with small size that we consider in this article.

The well-known proposal for unequal protection is Priority Encoding Transmission (PET) [3]. The PET system encodes the video data into different classes of given importance according to the priority function. For instance, with a priority fraction of 60, 75, 90 for I-frame, P-frames and B-frames in a GOP, respectively, PET system can recover all packets in I-frame class when receiving 60% of total packets and so on. A nice property of PET is that the decoding is not required if the data arrived at the receiver are not disturbed. The detail packetization of PET is explained in [12]. However, according to the published articles, PET system uses "all or nothing" strategy [12] which means all frames of a certain class are recovered or not at all depending on the priority function and the fraction of packets received. This leads to more severe quality degradation when the losses exceed the priority function of a class in comparison with other unequal protection schemes.

$$G_{10,15}^{MDS-TS} = \left[I_{10} \;\; \begin{matrix} \alpha_{1,1} & \alpha_{1,2} & 0 & 0 & \alpha_{1,5} \\ \alpha_{2,1} & \alpha_{2,2} & 0 & 0 & \alpha_{2,5} \\ \alpha_{3,1} & \alpha_{3,2} & 0 & 0 & \alpha_{3,5} \\ 0 & 0 & 0 & 0 & \alpha_{4,5} \\ 0 & 0 & \alpha_{5,3} & \alpha_{5,4} & \alpha_{5,5} \\ 0 & 0 & \alpha_{6,3} & \alpha_{6,4} & \alpha_{6,5} \\ 0 & 0 & 0 & 0 & \alpha_{7,5} \\ 0 & 0 & \alpha_{8,3} & \alpha_{8,4} & \alpha_{8,5} \\ 0 & 0 & \alpha_{9,3} & \alpha_{9,4} & \alpha_{9,5} \\ 0 & 0 & 0 & 0 & \alpha_{10,5} \end{matrix} \right]$$

$$G_{10,15}^{DAUEP} = \left[I_{10} \;\; \begin{matrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} & \alpha_{1,4} & \alpha_{1,5} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} & \alpha_{2,4} & \alpha_{2,5} \\ \alpha_{3,1} & \alpha_{3,2} & \alpha_{3,3} & \alpha_{3,4} & \alpha_{3,5} \\ 0 & 0 & 0 & 0 & \alpha_{4,5} \\ 0 & 0 & \alpha_{5,3} & \alpha_{5,4} & \alpha_{5,5} \\ 0 & 0 & \alpha_{6,3} & \alpha_{6,4} & \alpha_{6,5} \\ 0 & 0 & 0 & 0 & \alpha_{7,5} \\ 0 & 0 & 0 & \alpha_{8,4} & \alpha_{8,5} \\ 0 & 0 & 0 & \alpha_{9,4} & \alpha_{9,5} \\ 0 & 0 & 0 & 0 & \alpha_{10,5} \end{matrix} \right]$$

Fig. 1. Generator matrices of MDS-TS and DAUEP

Another way to tolerate the packet losses is to consider the protection at video coding level. The H.264/AVC encoder can use the error resilience schemes such as the use of slices, data partitioning, FMO to tolerate the losses. In [13], S. Wenger showed that FMO outperforms the other schemes. Thus, in this paper we only compare the FEC codes with FMO. Indeed, H.264/AVC standard allows encoder to intelligently assign macroblocks (MBs) of size 16x16 pixels into a slice group with a different order than the scan order. For instance, a picture with two slice groups in checker board mode is shown in Figure 2. Obviously, if one of two slice groups is lost, the remaining slice group can be used to conceal the lost information.

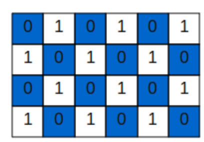


Fig. 2. FMO checker board with 2 slice groups

III. EXPERIMENT SETUP

We use the H.264/AVC Joint Model (JM) reference software [14] to encode and decode in CIF format the reference sequences (Akiyo, Foreman, etc.) which provide various video characteristics. All videos are coded with a GOP size of 30 frames and up to 3 reference frames can be used for inter-prediction. The frame rate is set to 30 frames per second (fps). Frame copy mechanism is used by the JM decoder to conceal the losses. The Quantization Parameter (QP) is set to 28 for all frames and there is no rate-distortion optimization (RDO). The encoded video is packetized in Real-Time Transport Protocol (RTP) format [15]. We evaluate the video quality using Peak Signal to Noise Ratio (PSNR) in dB through a wide PLR range from 1% to 40% and show appropriately the Mean Opinion Score (MOS) scale. We perform 50 runs for each experiment in order to achieve representative results.

The video sequences are encoded with Baseline and Extended profiles to evaluate different applications [16]. The Baseline profile (only I and P frames) is designed for low delay, low complexity applications (i.e., conversational communications) while the Extended profile is suited for streaming applications. In Baseline profile, the videos are encoded with the basic coding where one frame is coded by one RTP packet. If the size of RTP packet is greater than the Maximum Transmission Unit (MTU) which is set to 1500 bytes, it will be fragmented into several segments at the network layer. The RTP packet is considered as lost if any of segments can not be recovered. The FMO mechanism which is also encoded with Baseline profile has 2 slice groups in checker board mode as in Figure 2. The Extended profile is encoded with one B frame between I and P frames or two consecutive P frames. The PET scheme is simulated with the priority function 60, 75, 90 for I-frames, P-frames and B-frames classes, respectively. The priority function leads to different overall redundancy ratio depending on video sequences. The redundancy ratio is specified in each simulation.

IV. Evaluation and analysis

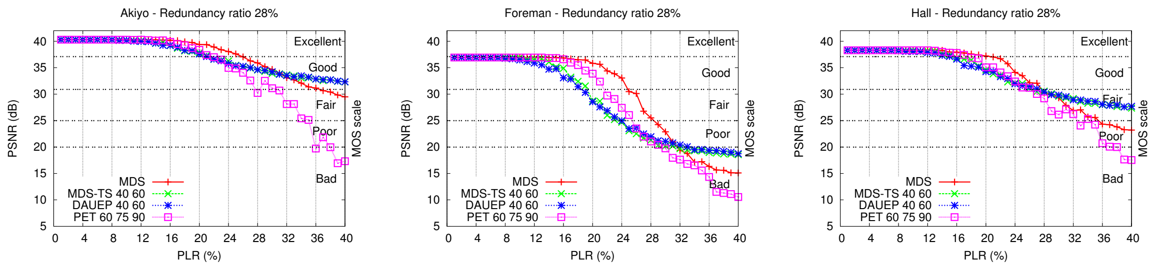
A. Equal versus unequal error protection

We first show the performance of MDS, MDS-TS, DAUEP and PET schemes in Baseline profile. The PET priority function makes the overall redundancy at 28% which is used by MDS, MDS-TS and DAUEP. For MDS-TS and DAUEP, the redundancy allocation for I and P frames is set to 40% and 60%, respectively. The result of Akiyo sequence is shown in Figure 3(a). At PLR less than 14%, all schemes achieve the maximum PSNR at 40.3 dB (in Excellent MOS scale) since they are able to recover the losses with the total redundancy of 28%. When PLR increases, the PSNR of DAUEP and MDS-TS starts decreasing with a small slope since some losses can not be recovered with low probability. MDS keeps highest PSNR at PLR from 14% to 28% since it provides high decoding capability. When the PLR exceeds the redundancy ratio, PSNR of PET decreases drastically due to its "all or nothing" strategy. Indeed, PET receives less than 75% of packets in a GOP with high probability at PLR of 30%. This implies that all P frames can not be recovered, the JM decoder makes the frame copy to conceal the losses of all P frames from I frame of GOP i^{th} to I frame of GOP $i+1^{th}$ since the GOP pattern is IPPP...IPP.... This leads to severe PSNR degradation and explains why even with an average of 50 runs, the PSNR degradation of PET is not monotonous. On the other hand, DAUEP and MDS-TS achieve higher PSNR than MDS since it is unlikely for MDS to recover the losses at PLR higher than redundancy ratio. The results for Foreman (Figure 3(b)) and Hall (Figure 3(c)) are similar to Akiyo. It is noted here the decreasing slope is higher than the case of Akiyo. This is because the background and persons in Akiyo video are not moving much, thus, the losses are well concealed by the JM decoder (still in Good MOS scale). In Baseline profile, we do not see the difference between MDS-TS and DAUEP. The higher performance of DAUEP will be shown in Extended profile.

These simulations show that there is no winner in a wide PLR range. We observe that this range forms three zones. In the first zone where the PLR is much less than the redundancy ratio, PET, MDS, MDS-TS and DAUEP perform well since the repair packets can recover most of losses. The second zone where the PLR is close to the redundancy ratio in both left and right sides is rather unstable due to probabilistic behavior. In fact, in some experiments where the losses are less than the redundancy, PET and MDS are able to recover all information packets and the video quality is good. However, in some other experiments where the losses exceed the redundancy, PET and MDS are not able to recover the information packets. PET exhibits more severe quality degradation due to its "all or nothing" strategy. In the third zone where the PLR exceeds significantly the redundancy ratio, all schemes exhibit low PSNR. It is noted here DAUEP achieves higher PSNR than all others since the first P frames are more protected than the later ones in a GOP. Obviously, no scheme outperforms the others in a wide PLR range. In principle, one repair packet can recover at most one lost packet. Furthermore, the repair packets which form the size of generator matrix are limited by the redundancy ratio. Depending on the way the protection solutions build their generator matrix, they protect well at most two zones in this wide range. These insights would help the applications to adapt to the network dynamics in presence of feedback channel.

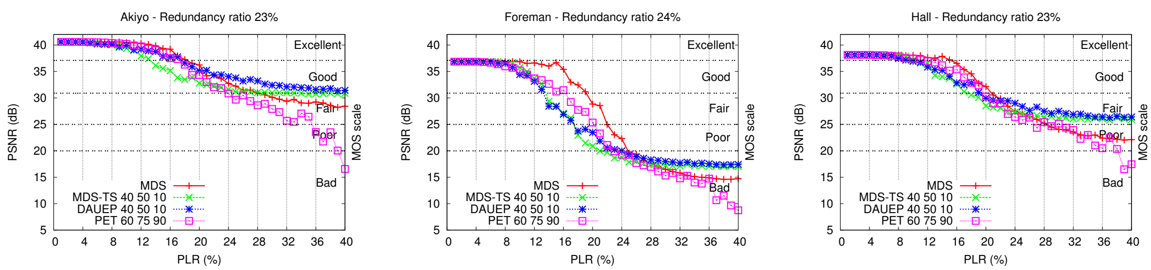
In case of Extended profile, the redundancy ratio for I, P and B frames is set to 40%, 50% and 10%, respectively, to be fair with PET priority function. The PET priority function in Extended profile generates overall redundancy between 23-24%.The results of Extended profile (Figure 4) is similar to the case of Baseline with some remarks. First, the PSNR of DAUEP is slightly higher than the one of MDS-TS since the first P frames are more protected than the latters in a GOP. Indeed, first P frames are generally more important than the latters because of their dependencies. Second, the PSNR of MDS-TS, DAUEP and PET in Extended profile decreases sooner than the one of Baseline since the redundancy for B frames (10%) is not able to recover all losses with high probability at PLR greater than 10%. Finally, the Extended profile has slightly lower PSNR than the Baseline profile at high PLR since the redundancy ratio in Extended profile is less than the one in Baseline profile.

We also evaluate the different settings by varying the redundancy allocation for I and P frames. In Figure 5, DAUEP and MDS-TS with the redundancy allocation 20% for I frame and 80% for P frames (called 20-80 allocation) achieve slightly higher PSNR than the 40-60 allocation at PLR less than redundancy ratio (28%). Indeed, it is more likely for the 20-80 allocation to recover more losses than the 40-60 one since more repair packets are used to protect P frames. However, the PSNR of 20-80 allocation is lower when the PLR exceeds the redundancy ratio. In fact, the JM decoder can not recover an I frame with high probability due to less redundancy for I frame and the overall redundancy is not able to cover the packet losses.



(a) Akiyo - Encoding rate 122 kb/s (b) Foreman - Encoding rate 531 kb/s (c) Hall - Encoding rate 306 kb/s

Fig. 3. Comparison between EEP and UEP (Baseline profile)



(a) Akiyo - Encoding rate 110 kb/s (b) Foreman - Encoding rate 447 kb/s (c) Hall - Encoding rate 251 kb/s

Fig. 4. Comparison between EEP and UEP (Extended profile)

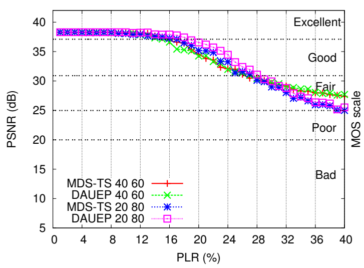


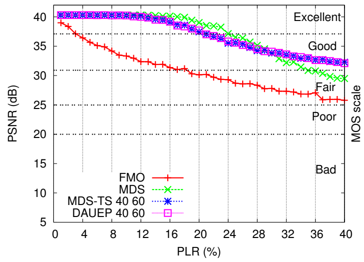
Fig. 5. Hall - Different redundancy allocations

B. Error protection versus error resilience schemes

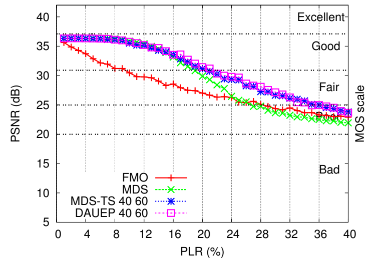
We perform an evaluation between the error resilience tools provided by H.264/AVC standard and the FEC protection schemes. The extra bit rate caused by FMO is used to calculate the redundancy ratio for FEC codes. We choose 3 pairs of reference sequences where the FMO generates high, medium and low redundancy. It is observed that the redundancy generated by FMO depends on the background, the object movement in the video, etc. Akiyo and Mother&Daughter sequences have a static background, thus, the FMO generates much higher redundancy than the basic coding. While in Container and Hall, the background is static but the objects and persons move around the background. The FMO of these sequences generate less redundancy. FMO generates low redundancy for video sequences with moving background and objects (i.e., Bus, Coastguard). In case of Akiyo and Mother&Daughter (Figure 6(a), 6(d)), all FEC codes outperform the FMO since they protect well the video at high redundancy. At medium redundancy, the PSNR of FMO is worse than all FEC codes except MDS at high PLR (Figure 6(b), 6(e)) since MDS shows its low PSNR at high PLR (Figure 3). On the other hand, FMO outperforms the FEC codes at low redundancy (Figure 6(c), 6(f)) and high PLR since the FEC schemes achieve high PSNR in the first zone. Their PSNR decreases drastically when the PLR exceeds the redundancy ratio. These simulations show the benefits of FMO in dynamic video sequences with moving background and objects while it is better to use the FEC schemes for rather static videos. This brings a special remark for the case of video conferencing which is usually experienced in a meeting room or a hall. The background and objects are not moving much, thus, it would be better to use the protection soulutions rather than error resilience schemes.

V. Discussion

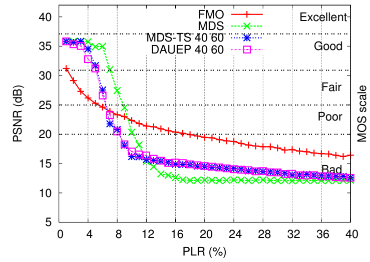
The results show that DAUEP achieves higher PSNR than other schemes at high PLR while MDS achieves the best performance at low PLR. Since the generator matrix is similarly constructed in MDS, MDS-TS and DAUEP, it is easy to switch between these schemes by changing the configuration parameters. Thus, it is interesting to design adaptive protection schemes which intelligently use the appropriate mode to provide better video quality depending on network conditions. For instance, in presence of feedback channel, an adaptive protection solution should be in MDS mode at low PLR and should switch to DAUEP at PLR close to redundancy ratio. Furthermore, DAUEP has an another advantage since the lost packets are more likely recovered sooner than both MDS



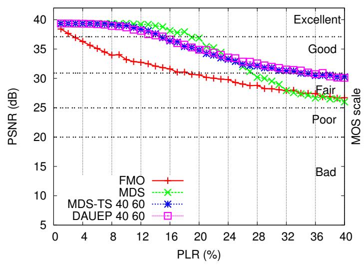
(a) Akiyo - Redundancy 26%



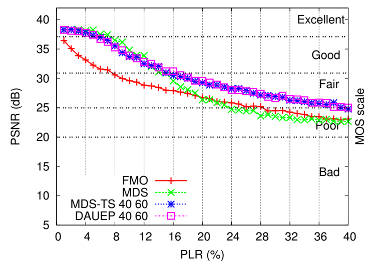
(b) Container - Redundancy 14%



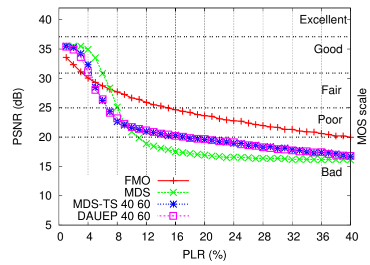
(c) Bus - Redundancy 9%



(d) Mother&Daughter - Redundancy 23%



(e) Hall - Redundancy 12%



(f) Coastguard - Redundancy 7%

Fig. 6. Comparison between error protection and error resilience schemes

and MDS-TS. In fact, the repair packet in DAUEP can be immediately sent as soon as the data packets protected by that repair packet are emitted. On the other hand, the repair packets in MDS can be sent only when all data packets of a GOP are emitted. It is noted that the adaptive scheme can also lower the encoding rate based on RDO to have more redundancy. The interesting opened questions are when and in which type of video (i.e., static or dynamic) the application should use RDO or protection schemes to provide better video quality.

VI. Conclusions and Future Work

In this paper, we have evaluated error tolerance schemes applied to H.264/AVC. First, we evaluated the MDS, MDS-TS, PET and DAUEP and found that no scheme outperforms the others in a wide PLR range. The analysis in three zones shows how a protection solutions behave in a wide PLR range. MDS shows its good performance in the first and second zones while DAUEP and MDS-TS performs well in the first and third zones. DAUEP is slightly better than MDS-TS since the first P frames are more protected than the latters in a GOP. We also showed the impact of redundancy allocation for different frame types on the video quality. Then, we compared the error protection schemes with FMO error resilience mechanism. The FMO shows its benefits in videos with moving background and objects while FEC protection schemes perform well for rather static videos. Our future work is to answer the opened questions.

VII. Acknowledgements

This work was supported by the French ANR grant ANR-VERS-019-02 (ARSSO project), CNES and Thales Alenia Space.

References

[1] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Techn.*, vol. 13, no. 7, pp. 560–576, 2003.

[2] S. Kumar, L. Xu, M. K. Mandal, and S. Panchanathan, "Error resiliency schemes in H.264/AVC standard," *J. Visual Communication and Image Representation*, vol. 17, no. 2, pp. 425–450, 2006.

[3] A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoding transmission," *IEEE Transactions on Information Theory*, vol. 42, no. 6, pp. 1737 –1744, Nov 1996.

[4] A. Bouabdallah and J. Lacan, "Dependency-aware unequal erasure protection codes," *Journal of Zhejiang University Science*, vol. 7, pp. 27–33, 2006.

[5] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Techn.*, vol. 17, no. 9, pp. 1103 –1120, sept. 2007.

[6] P. A. K. Acharya *et al.*, "Congestion-Aware Rate Adaptation in Wireless Networks: A Measurement-Driven Approach," in *In Proc. IEEE SECON Conf.*, San Francisco, CA, Jun. 2008.

[7] J. Lacan and T. Perennou, "Evaluation of error control mechanisms for 802.11b multicast transmissions," in *In Proc. of WiNMee*, Boston, Massachusetts, Apr. 2006.

[8] J. Lacan, V. Roca, J. Peltotalo, and S. Peltotalo, "Reed-Solomon Forward Error Correction (FEC) Schemes," RFC 5510, Apr 2009.

[9] S. Boucheron and M. Salamatian, "About priority encoding transmission," *IEEE Transactions on Information Theory*, vol. 46, no. 2, p. 699, Mar 2000.

[10] V. Roca, C. Neumann, and D. Furodet, "Low Density Parity Check (LDPC) Staircase and Triangle Forward Error Correction (FEC) Schemes," RFC 5170, Jun 2008.

[11] M. Luby *et al.*, "Raptor Forward Error Correction Scheme for Object Delivery," RFC 5053, Oct 2007.

[12] R. Storn, "Modeling and Optimization of PET-Redundancy Assignment for MPEG Sequences," International Computer Science Institute, Berkeley, Tech. Rep., 1995.

[13] S. Wenger, "H.264/AVC over IP," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, pp. 645–656, 2003.

[14] "H.264/AVC reference software http://iphome.hhi.de/suehring/tml/."

[15] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," RFC 3550, Jul 2003.

[16] I. E. Richardson, *The H.264 Advanced Video Compression Standard; electronic version*. Chichester: John Wiley & Sons Ltd, 2010.