A Cross Layer Scheme for H.264/AVC Video Transmission over Wireless Network

Rima Boudjadja*, Mohamed Azni** and Mohamed Nadjib Zennir***

*Computer Science Department, University of Bejaia, Algeria, rima.boudjadja@gmail.com
**Electronics Department, University of Bejaia, Algeria, m_azni@hotmail.com
***Computer Science Department, University of Bejaia, Algeria, zennir.med@gmail.com

Abstract

In this paper, we propose a cross-layer design which combines Packet Forward Error Correction (P-FEC) at the application layer with a Symbol unequal FEC in the MAC layer (S-FEC). The aim is to minimize the Packet Loss Rate (PLR) and maximize the visual quality of video transmitted over a wireless network (WN). The proposed scheme is based on a FEC compatible with the semantics of the H.264/AVC video encoding. This mechanism relies on a rate distortion algorithm controlling the channel rates under global rate constraints given by the WN. Based on a data partitioning tool (DP), the proposed optimization mechanism takes into consideration both the packet type and packet length which leads to an unequal error protection (UEP). A Reed-Solomon S-FEC is adapted to unequal protected data. Indeed the unrecoverable S-FEC blocks can be balanced by a P-FEC in the application layer. Simulation experiments over random generated noise show that UEP scheme has a Bit Error Rate (BER) greater than or equal to BER at the output of the Equal Error Protection scheme (EEP) and demonstrate various Packet Loss Rate (PLRs) in terms of packet type. Contrary, the UEP scheme has a greater computational complexity than EEP. In particular, our proposed cross-layer UEP improves the average video peak signal-to-noise ratio (PSNR) by up to 4dB, compared to an EEP scheme.

Keywords: Forward Error Correction; Unequal Error Protection; H.264/AVC Video; Data partitioning Video Transmission; Wireless video communication.

1- Introduction

The broadband digital transmission of video and their applications present an important issue nowadays. These communications systems, and the rapid development of computing capacity of computers and technical data, are enhancing significantly the number of users. Furthermore, different from traditional data communications, video applications usually impose strict end-to-end delay constraints. In addition, video packets are generally of different importance. These special features typically require an application source encoder. In order to satisfy the requirements of such applications, many CoDec (CODEC) have been developed to make the video stream more robust against transmission errors, more compressed but with a reduced computational complexity. The authors of [8] propose an intra mode decision algorithm to reduce the computation complexity of intraframe H.264/AVC encoders. The proposed algorithm achieves 18% to 70% reduction in the computational complexity, compared to many conventional methods.

In this work, we have used a H.264/AVC [14] video CoDec which aims to improve the previous standards by offering significant enhanced compression and by providing many new features able
to improve the robustness to data errors at the source coding level and recover the corrupted blocks at the decoder level. Unfortunately, these tools could not correct transmission errors when channel noise is very high, which degrades the reconstructed video quality. However, the use of Forward Error Correction (FEC) or channel coding presents an interesting solution for error control in video transmission over wireless network. In this paper, we use the H.264/AVC data partitioning tool (DP) to make difference between video packets and using of unequal error protection (UEP) [3][15] which makes the third class of [7]. In addition, we take into account the packet size to allocate bandwidth to different video packets without exceeding the channel capacity. Channel coding in MAC layer (S-FEC) using Reed Solomon codes is applied to the video bitstream taking into account the packet priority. The unrecoverable S-FEC blocks can be balanced by a Packet FEC (P-FEC) in the application layer for more robust error protection. Thus, our cross-layer scheme combines the application layer with the MAC layer. The performance of the proposed cross-layer unequal error control is demonstrated over wireless network by performing simulations under several channel conditions. Simulation results are compared with an equal error protection (EEP) results.

The rest of this paper is organized as follows: section 2 provides a brief overview of scalability in H.264/AVC, data partitioning and a review of error-control techniques. In section 3, we give a detailed description of the proposed Cross-Layer video transmission. In section 4 we report on performances of our application, using parameters like image quality, bit error rate(BER), packet loss rate (PLR) and PSNR (Peak Signal-to-Noise Ratio). Finally, a conclusion and further works are provided in section 5.

2 - Related Work

The growing importance of wireless video communications which are faced with channels that are both noisy and band limited, required a joint optimization of source coding and channel coding JSCC, where the goal usually is to improve the quality of video transmission for a given channel. This idea has drawn more and more attention recently [2][10][3][11]. The majority of the proposed works use the scalable extension of the H.264/AVC [14] video coding design (SVC) which demonstrates a superb adaptability in video communications. Such design makes a distinction between a Video Coding Layer (VCL) and a Network Abstraction Layer (NAL) [12], the output of the encoding process is VCL data which are mapped to NAL units prior to transmission. Each NAL units (NALU) makes up a packet where it contains some number of bytes including a header and a payload. The header specifies the type of each NALU and the payload contains related data. In VCL, picture frames are divided into macroblocks (MBs). An integer number of MBs are further grouped to form a slice which can be encoded to fit the size of one or more separate NALU that can independently decodable.

The H.264/AVC places the data that makes up a slice in three separate Data Partitions with different importance in terms of decoded video quality. This property in conjunction with unequal error protection UEP, belonging to joint source channel coding and cross-layer protection strategies, is especially useful in many transmission scenarios by using a stronger protection of the more important information. The most important partition in an H.264/AVC bitstream that named A, contains the slice header and disparity compensated prediction vectors. This part of data needs the highest level of error protection. Unlike partition C which requires least protection since information in partition A and B can be used to decode data in partition C using the error concealment tool EC offered by the H.264/AVC standard. An excellent review of the existing error concealment mechanisms is given by Yanling and Yuanhua [16]. We must consider that the EC can be accomplished only if the partition A is received correctly. Note that the concealment, or recovery, of a lost NAL unit is made easier if the NAL unit is small. Moreover, a small NAL unit contains less information, and its loss is less damaging. However, it is difficult to recover a long
partition which is considered in [3].

In [9] an adaptive JSSC over wireless channel based on a new rate-quality (R-Q) model of the H.264/AVC and the error protection characteristics of the turbo code is proposed. In [3], Azni et al proposed a new method of rate adaptation to the allowed maximum channel transmission rate which does not assume source rate control in the source encoder. Then, unequal error protection is applied at the network abstraction layer to different partitions with RCPC codes. This application needs considerable time coding which is on opposite with the requirements of real-time systems. For robustness, they used joint optimization of forward error correction at the transmitter and error concealment at the source decoder.

The layered optimization leads to sub-optimal results in multimedia performance. Alternatively, under adverse conditions, wireless stations need to optimally adapt their compression parameters and transmission strategies jointly across the protocol stack in order to guarantee a predetermined quality at the receiver. [13] presents a cross-layer framework for jointly analyzing, selecting, and adapting the different strategies available at the various OSI layers in terms of multimedia quality, consumed power, and spectrum utilization. In [6] Amin Abdel Khalek et al. proposed an APP/MAC/PHY cross-layer optimization framework that allows the channel to adapt at a faster time scale than the video codec. Unlike the framework proposed by Jeong-Yong Choi and Jitae Shin [4] which combines cross-layer error protection techniques composed of error correction code in the link/MAC layer, erasure code in the application layer, and ARQ scheme with low-overhead across the link/MAC layer and application layer.

This paper is mainly motivated by [4]. We combine cross-layer error protection techniques composed of error correction code (reed-solomon) in the link/MAC layer and erasure code in the application layer with the code rate allocation algorithm proposed in [3].

3 - System Overview

In this section, we develop our cross-layer design, which combines packet FEC at the application layer with symbol adaptive FEC at the data link layer. Consider the single-transmit single-receiver in an IEEE 802.11.e wireless network. Figure 1 represents the layer structure system on both the transmitter and the receiver side. The system input is a digital video content that has been compressed using the most recent H.264/AVC compression standard. H.264/AVC introduces the concept of parameter sets, which contain information that can be used to decode a large number of encoded video sequences.

![Figure 1: The Proposed Cross Layer Architecture](image-url)
In the proposal, we consider that the sequence parameter set is transmitted out of band and error-free because its integrity is a critical constraint in the source decoding. The non-VCL units can be transmitted without channel coding for carrying enhancement information which are not necessary for source decoding. The remaining VCL units are protected using an optimal JSCC scheme where a variable code rate is calculated for each NALU in one frame. The code rate selection is warranted by a rate allocation algorithm [3] described in Section 3.1; which uses both type and size of NALU as constraints. Source packets are then sent to the application layer FEC Reed Solomon (RS) encoder which generates M additional packets for N input source ones. The frames that have even single bit errors cannot be sent to the application layer FEC decoder. In such a case, applying FEC along each source packet would mean that in case of link layer frame loss, the complete source packet would be lost, and application RS decoding would thus fail. In order to cope with this problem, we propose a combined S-FEC in the MAC layer using a shortened RS code, based on the remaining channel code rate from the source channel allocation. The unrecoverable packets will be passed to the P-FEC decoder process which will try to recover them instead of passing the unreliably decoded packets to the source decoder.

3 – 1 JSCC Rate Allocation algorithm

The first important feature of this work is that the source rate does not change across the transmission. By using this idea, a considerable reduction in the processing requirement can be achieved at the encoder. The incoming compressed bitstream is first encoded at the application layer. This is accomplished using H.264/AVC standard that includes the data partitioning mode. As a result, each source frame is encoded as a sequence of 3N NALUs, where N is the frame slice number. The JSCC algorithm [3] calculates the optimal channel code rate for each partition and adapts the bit rate to the capacity of the given line. So, this algorithm is responsible for optimally allocating the available channel rate between the source and FEC with respect to their importance to minimize the video distortion at the decoder side.

![Figure 2: Frames NALUs's Classification](image)

Here, source packet consists of a SVC NAL unit portrayed as a row in Figure 2. This figure shows all the source packets included for transmission in one frame which are ordered according to their subjective importance (type and size) and divided in two subsets. The first one contains the A partitions which have a same code rate. The second one is the concatenation of three subsets, which are B partitions, C partitions and parity packets respectively; where the partitions of each subset are listed in descending order of size. Therefore, we need to sort the corresponding code rate vector starting by an initial one. Figure 2 illustrates this correspondence, and it depicts the percentage of bandwidth allocated to the source and channel coding respectively. The
dark areas correspond to the additional percentage of the bandwidth that is used for channel coding. The bigger part of the bandwidth is needed for the first subset channel coding where all NALUs have the same percentage of parity bits. However, the percentage vector of the second subset is proportional to partitions size. The complete JSCC allocation algorithm is described in [3].

3 -2 Symbol-Level FEC (S-FEC)

In link/MAC layer, the MAC Protocol Data Unit (MPDU) format defined in the draft specification of IEEE 802.11.e is RS encoded using the remaining rate $r_i$ from the channel allocation where $r_i$ is the NALU of the picture, the data IEEE 802.11.e frame is illustrated in [1]. Basically, a shortened Reed-Solomon code, defined in Galois Field, $\text{G}(256)$ is used, while a (36,48) shortened RS code is used for MAC header, and CRC-32 is used for the Frame Check Sequence (FCS). It is important to note that any RS block can correct up to byte-errors. The MPDU integrity is verified using the FCS field, if the FCS is correct the receiver skips the RS decoding process. Usually, the unrecoverable S-FEC frame is discarded and will not pass to the upper layers which make the application parity packets useless. To cope with this problem, we propose a technique which allows the receiver to keep a copy of all the corrupted NALUs mapped into each MAC frame in a buffer and allows the MAC layer to pass the unrecoverable packets to the UDP transport layer which in turn relies on the application layer to handle the corrupted data using the application RS decoding. By using this idea, the unrecoverable S-FEC blocks can be compensated by P-FEC for more robust error protection.

3 – 3 Packet-Level Forward Error Correction (P-FEC)

In the application layer we implement P-FEC mechanism based on erasure code. We utilize RS codes since it has good error correction properties and is widely used in FEC schemes. The fundamental concept of erasure code is described in [5]. In general a RS code contains $n$ source packets and $m$ parity packets. The key idea behind erasure codes is that any subset of $n$ encoded blocks suffices to reconstruct the source data knowing that it's correction capacity is $m/2$. In the current implementation, we use systematic codes where the first $n$ of the $n+m$ encoded packets are identical to the $n$ source packets. In our experiments we utilize $RS(11,9)$ ($n = 9, m = 2$) codes where we generate 2 parity packets for 9 source packets which form one picture. As long as the number of lost packets is less than the corrector capacity, all original video packets can be decoded successfully. When the loss exceeds the FEC correction limit, only the received video packets are put into the decoded video files which will contain mostly the A partitions since they are the most protected by the S-FEC unlike the B and C partitions. In such a case, the last problem that the decoder has to cope with is the recovery of the B and C partitions which are poorly protected by our technique. This function is warranted using the H.264/AVC error concealment tool which can recreate the B and C partitions using the most protected partitions.

4 - Experimental Results

4 – 1 Simulation Model

To illustrate the effectiveness of our proposed Cross-Layer scheme, we have performed experiments on two video sequences Tree and Mobile calendar (QCIF format), encoded using the JM10.0 standard. Since an encoded picture in a QCIF video sequence includes 99 MBs, we set the number of MB per slice at 33 to obtain pictures encoded in 3 slices each. We also chose the extended profile which allows us to encode each slice into 3 NALUs using the data partitioning mode. Therefore, each picture is encoded in 9 NALUs. The 9 NALUs will be RS(11,9) encoded in the application layer.
Note that the initial rate vector is set to \([k/85, k/80, k/75, k/70, k/65, k/60, k/55, k/50]\), and Symbol Size=50. The compared EEP method, using RS code, is fixed channel coding rate as follow:

\[
T_{moy} = \frac{1}{N} \sum_{n} \frac{P_{s}(n)}{R_{c}}
\]

where \(R_{c}\) is the channel rate and \(P_{s}(n)\) is the number of bits in the picture.

4 – 2 Simulation Results and Discussion

In this section, the simulation results for two video transmission techniques are evaluated; UEP and EEP FEC coding. We quantify the performance of our method by calculating the Bit Error Rate (BER) at the output of channel decoder using both protection techniques EEP and UEP. This is one of the most important parameters of this study since it can quantify the coding efficiency. Video sequences are compressed using H.264/AVC encoder in DP mode. The video encoded stream is protected by a rate compatible punctured RS codes depending on the UEP and EEP scheme. Therefore, to compare effectiveness of the two correctors we add noise to both encoded sequences and analyze the bit error rate results in both cases, before and after channel decoding. In each group of experiments, we change the Signal Noise Ratio (SNR) from 3 to 10 in step size of 0.5. In both cases, the BER is improved through the correction. Indeed, the errors before channel decoding can't be fully corrected. The obtained results show that the error rates in both cases are almost identical; the two correctors have almost the same performance with a small advantage to the EEP scheme. However, it is worth noting that the EEP has an advantage over UEP regarding computational complexity. It requires less computation to allocate equal code rates to all NALUs, and its performance is slightly better than UEP against the bit error rate. On the other side, its major problem is that it can't take into account the video stream semantic and it is not able to effectively compensate the loss of high priority packets. In this case, the UEP corrector can be more useful because of its high performance against the peak-signal-to-noise ratio (PSNR) and visual quality, which will be shown later.

4 – 2 – 1 Packet Loss Rate

Now let's evaluate the effect of EEP and UEP on packet loss rate of the three partitions. This factor is defined as a ratio of the number of lost packets to the total number of transmitted packets. Note that a packet is considered lost if at least one bit is erroneous. This parameter has an immediate effect on the multimedia transmission quality (voice and video) and an indirect effect on data transfer applications which use typically TCP. Figure 3 shows the comparison results of the three partitions A, B and C respectively under different channel SNR, decoded using EEP and UEP schemes. Each result is obtained as the average of 10 runs.
In all observed cases, the packet loss rate decreases when SNR increases. Thus, asymptotically, we can see that the FEC using the UEP have an unequal protection capability of the three types of NAL units. In Figure 3(red), the UEP provides enhanced protection of partition A compared with the second technique whatever the value of the SNR, unlike partitions C, which are more protected by EEP than UEP (Figure 3(bleu)). Partitions B are better protected by the UEP between 0 and 4.5dB (Figure 3(black)), 4.5dB above, the protection offered by the UEP decreases and gives lower quality compared to the EEP. In Figure 3(black), the UEP scheme yields a better protection to large NALUs than small ones which make their recovery more efficient when the channel is too noisy. However, when the channel become less noisy, the average code rate of EEP scheme can recover both large and small NALUs. On the other hand, small packets suffer neglect by the unequal error protection which gives rise to their losses. This explains why the EEP is more effective at high SNR region and UEP gives minimal PLR values at low SNR.

4 – 2 – 2 Peak-Signal-to-Noise Ratio (PSNR)

To see the impact of our technique to the reconstructed quality, we can calculate the PSNR of received video sequences compared to the original sequences in both scenarios (UEP and EEP). The comparison results of the two schemes with different channel SNR is shown in Figure 4. Typical values for the PSNR in lossy image and video compression are between 30 and 35 dB, where higher is better. Acceptable values for wireless transmission quality loss are considered to be about 20 dB to 25 dB. It can be seen clearly from figure 4 that compared with the fixed channel coding rate scheme, our proposed one can achieve a higher reconstructed quality of 32db at the receiver side. It can recover most of A partitions which makes the concealment of partitions B and C more efficient by the source decoder.
5 - Conclusion and Future Work

In this paper, we have used the DP approach in the NAL unit with Cross-Layer protection scheme for H.264/AVC encoded video sequences. In addition, we have considered UEP Reed-Solomon coding in the MAC layer and EEP channel coding schemes in the application layer to protect the H.264/AVC video transmission over a wireless network. The UEP scheme was designed based on the type of NALUs and their size to allocate jointly the channel code rate. Our proposal, by using strong protection to the most important packets, has been proven effective over the equal protection against the PSNR and visual quality by up to 4dB when SNR=5.5db. However, it has both a BER greater than or equal to EEP and greater computational complexity.

We plan to balance the MAC Forward Error Correction by a low-overhead ARQ. When an uncorrectable error is detected, a selective ARQ system requests retransmission only for the uncorrectable A packets which reduces retransmission rate as well. This combination can provide higher reliability than a FEC system alone and higher throughput than a system with ARQ only.

Bibliography


