

# Dynamic Sub-GOP Forward Error Correction Code for Real-Time Video Applications

Jimin Xiao, Tammam Tillo, *Member, IEEE*, Chunyu Lin, and Yao Zhao, *Senior Member, IEEE*

**Abstract**—Reed-Solomon erasure codes are commonly studied as a method to protect the video streams when transmitted over unreliable networks. As a block-based error correcting code, on one hand, enlarging the block size can enhance the performance of the Reed-Solomon codes; on the other hand, large block size leads to long delay which is not tolerable for real-time video applications. In this paper a novel Dynamic Sub-GOP FEC (DSGF) approach is proposed to improve the performance of Reed-Solomon codes for video applications. With the proposed approach, the Sub-GOP, which contains more than one video frame, is dynamically tuned and used as the RS coding block, yet no delay is introduced. For a fixed number of extra introduced packets, for protection, the length of the Sub-GOP and the redundancy devoted to each Sub-GOP becomes a constrained optimization problem. To solve this problem, a fast greedy algorithm is proposed. Experimental results show that the proposed approach outperforms other real-time error resilient video coding technologies.

**Index Terms**—Dynamic Sub-GOP, error resilience, H.264/AVC, real-time transmission, systematic Reed-Solomon.

## I. INTRODUCTION

THE H.264/AVC [1] video coding standard provides higher coding efficiency and stronger network adaptation capability in comparison to all the previously developed video coding standards. However, as with previous video compression standards, it is based on a hybrid coding method, which uses transform coding with motion compensated prediction (MCP). Therefore, when the video bitstream is transmitted over packet-loss networks, it suffers from error propagations and this leads to the well-known drifting phenomenon [2], [3].

Many error-resilient techniques have been developed to meet the requirements of video communication over lossy

networks [4]. These techniques include intra macroblock (MB) refreshment [5], [6], Automatic Repeat reQuest (ARQ) [7], [8] and feedback-based reference picture selection (RPS) [9], redundant picture coding with equal or lower quality [10], [11], multiple description coding (MDC) [12], [13], and forward error correction (FEC) coding [14]–[16]. Among those error resilient approaches, intra macroblock refreshment, redundant picture coding and MDC cause no additional delay, making them suitable for delay constrained applications. However, for the Intra MB refreshment approach, since the coding efficiency of intra mode is typically several times lower than inter mode, the coding efficiency is compromised dramatically. For the redundant picture coding and MDC, when the redundant version is used to replace the primary one or some of the descriptions are lost during transmission, there would be mismatch error and the mismatch error will propagate all over the GOP. ARQ and RPS usually cause long delay because of the network round-trip time, and consequently they cannot be employed for real-time applications. For the FEC approaches, the delay depends on the channel coding block size. In [14], the Reed-Solomon (RS) coding block includes the whole GOP, and one GOP of delay is caused. In [15], the RS coding block contains one block of packets (BOP) generated from different frames, and unequal loss protections are allocated for different packets based on both the frame position in the GOP and the data partition to which it belongs. For this approach, one BOP of delay is caused, and the delay depends on the length of the BOP. Furthermore, as the packets from one BOP are divided into two RS coding blocks based on the data partitioning type, the performance of the RS code is compromised. In [16], the RS code is at frame level, and no FEC coding delay is created. However, for the frame level FEC approach, usually the source packet number is not big enough for the FEC code to be efficient.

In this paper, a Dynamic Sub-GOP FEC Coding (DSGF) approach is proposed, and in this approach systematic Reed-Solomon erasure code is used to protect the video packets in real-time mode, while allowing to provide an error-free version of the reference frame to stop the propagation error. In order to enlarge the RS coding block size, all frames in one Sub-GOP are used as RS coding block. The length of the Sub-GOP is dynamically tuned, according to the Sub-GOP position, the probability of packet loss, and other encoding parameters, so as to minimize the expected total distortion of this GOP. On the encoder side, for the systematic RS code, the data is left unchanged and the parity packets are appended; therefore, there is no encoding delay. Meanwhile, at the receiver end, to decode and display one frame in the Sub-GOP, the video decoder only needs packets belonging to this frame. If some packets of this frame get lost during transmission, error concealment is applied to conceal the lost packets. In this

Manuscript received June 28, 2011; revised November 29, 2011; accepted April 01, 2012. Date of publication April 09, 2012; date of current version July 13, 2012. This work was supported by the National Natural Science Foundation of China (No.60972085), the Sino-Singapore JRP (No.2010DFA11010), the National Science Foundation of China for Distinguished Young Scholars (No.61025013), and the National Basic Research Program of China (Grant No. 2012CB316400). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Zhihai (Henry) He.

J. Xiao is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, L69 3GJ, U.K., and also the Department of Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University, Suzhou, China (e-mail: jimmin.xiao@liverpool.ac.uk).

T. Tillo is with the Department of Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University, Suzhou, China (e-mail: tammam.tillo@xjtlu.edu.cn).

C. Lin and Y. Zhao are with the Institute of Information Science, Beijing Jiaotong University, Beijing Key Laboratory of Advanced Information Science and Network Technology, Beijing, China (e-mail: yuailian@gmail.com; yzhao@bjtu.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMM.2012.2194274

manner, the decoder does not need to wait for all the packets belonging to this Sub-GOP. Therefore, there is no delay on the decoder side. Later, when the transmission of all packets of this Sub-GOP is finished, the systematic RS decoder would try to recover the lost packets. If enough packets are received, the RS decoder will be able to recover all the lost packets of this Sub-GOP, and the video decoder will re-decode this Sub-GOP with all the received and recovered packets, updating the reference frame, so the concealment distortion would not propagate to later frames. It is worth noticing that in our preliminary work [17], the Sub-GOP based FEC approach was proposed for real-time video streaming, whereas in this paper, the system model is generalized and analyzed in a detailed way, where the network packet loss model is extended to both random and burst cases; moreover, the greedy allocation algorithm is newly proposed to lower the computational complexity, which makes the proposed scheme more feasible.

The rest of the paper is organized as follows. A brief review of systematic RS code is provided in Section II. In Section III, firstly the frame level Evenly FEC approach is introduced; this approach is used as a benchmark for the real-time FEC coding. Later the proposed DSGF approach is presented in detail. In Section IV, some simulation results validating the proposed approach are given. Finally, some conclusions are drawn in Section V.

## II. SYSTEMATIC REED-SOLOMON ERASURE CODE

In this section, we will briefly recall some concepts and notations about systematic Reed-Solomon (RS) erasure code, which will be used throughout this paper. The systematic RS erasure code has been widely used as FEC code to protect data packets against losses in packet erasure networks. In RS  $(N, K)$  code, for every  $K$  source packets,  $(N - K)$  parity packets are introduced to make up a codeword of packets. As long as a client receives at least  $K$  out of the  $N$  packets, it can recover all the source packets. If the received packet number is less than  $K$ , the received source packets can still be used, because they have been kept intact by the systematic RS encoding process. In general, for the same code rate  $K/N$ , increasing the value of  $K$  would enhance the performance of RS code.

One important parameter for the systematic RS code that we will need in this paper is the remaining packet loss rate after the RS correction,  $p'$ . So, for example, for the Bernoulli i.i.d packet loss model, this parameter is determined by the value of  $N$ ,  $K$ , and the average network packet loss rate,  $p$ . Whereas, for the Markov burst packet loss model,  $p'$  is also impacted by other parameters of the Markov model, i.e., the average burst loss length.

### A. Bernoulli i.i.d Packet Loss Model

In the Bernoulli i.i.d packet loss model,  $p'$  could be evaluated as

$$p' = \frac{\sum_{i=1}^K i p_{rs}(i)}{K} \quad (1)$$

with  $p_{rs}(i)$  representing the probability of still having  $i$  unrecoverable source packets after RS correction. From now on, we will refer to those packets as unrecoverable lost packets. To

evaluate  $p_{rs}(i)$ , let us use  $p_s(n)$  and  $p_r(n)$  to denote the probability of losing  $n$  packets before decoding the RS code among the source packets and parity packets, respectively:

$$p_s(n) = \binom{K}{n} (1-p)^{K-n} p^n \quad (2)$$

$$p_r(n) = \binom{N-K}{n} (1-p)^{N-K-n} p^n. \quad (3)$$

Since having  $i$  unrecoverable lost packets is caused by losing  $i$  source packets, and at the same time losing more than  $N - K - i$  RS parity packets, then the probability of this event is

$$p_{rs}(i) = \begin{cases} p_s(i) P_r(N - K - i + 1) & \text{for } i \leq N - K \\ p_s(i) & \text{for } i > N - K \end{cases} \quad (4)$$

where  $P_r(j)$  is used to denote that no less than  $j$  RS parity packets are lost. This could be evaluated based on (3) as

$$P_r(j) = \sum_{n=j}^{N-K} p_r(n). \quad (5)$$

### B. Burst Packet Loss Model

For the burst loss model, we will use the Gilbert two-state model because it is one of the most common models used for multimedia transmission simulation. For this model, the formula to calculate the remaining packet loss rate,  $p'$ , after the RS correction was presented in [18]. It is worth noticing that, in Gilbert two-state model,  $p'$  is not only determined by the value of  $N$ ,  $K$ , and  $p$ , but also influenced by the average burst length, and  $p'$  can be evaluated as follows:

$$p' = \frac{p}{K} \sum_{i=1}^K i R(i, K) \sum_{j=\max(0, N-K+1-i)}^{N-K} R(j+1, N-K+1) + \frac{1-p}{K} \sum_{i=1}^{K-1} (K-i) S(i, K) \sum_{j=0}^{K-i-1} S(j+1, N-K+1) \quad (6)$$

where  $R(m, n)$  denotes the probability that  $m - 1$  consecutive packet losses occur following a packet loss, and  $S(m, n)$  denotes the probability that  $m - 1$  consecutive packets arrive following one packet arrival. For the detailed procedure of calculating  $R(m, n)$  and  $S(m, n)$ , please refer to [18].

## III. REAL-TIME FEC VIDEO TRANSMISSION APPROACHES

Since our objective is to design FEC video transmission system for real-time applications while minimizing the delay caused by the encoding stage, therefore B-frame will not be used, so we will use the IPPP GOP structure. It is also important to note that the most commonly used applications for real-time system are video telephony with low latency requirements. This application typically uses the baseline profile of H.264/AVC, where only I-frames and P-frames are used [1]. To make the RS code efficient, fixed length slice scheme, in terms of byte, is used to create slices. The slice length is decided by the maximum transmission unit (MTU) of the underlying networks. With this method, as many MBs are put into one slice as possible under the constraint that the slice length is no more than the target length, therefore, the length of all the

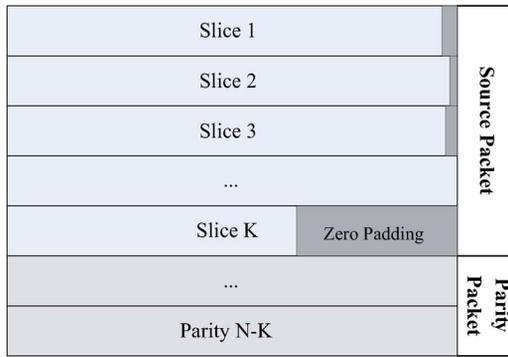


Fig. 1. Packet padding method for the RS coding, H.264 fixed slice length method is used, the slice length is nearly the same except for the last one.

slices except the last ones in each frame is very close to the target slice length. As shown in Fig. 1, for the slices other than the last slice, only very few zero bytes are padded, whereas for the last slice of one frame, usually more dummy zero bytes are padded. The length of each RS protection packet (i.e., parity packet) is the same as the target slice length. In this paper, the term packet and slice are used interchangeably, as one packet per slice packetization method is adopted.

#### A. Frame-Level Evenly FEC

For real-time FEC video packet protection, one common approach is to perform RS coding in frame level, which means that the RS coding block contains data packets from the same video frame. Under this constraint, RS coding does not introduce any additional delay. Let us assume the GOP length is  $L$  frames, and the  $i$ th frame has  $K(i)$  source packets and  $R(i)$  RS parity packets. If we want to provide even protection for all the GOP frames, and taking into account that, in general,  $K(i)$  does not change largely. Then  $R(i)/K(i)$  needs to be almost constant over all the GOP's frames. However, taking into account that  $R(i)$  should be an integer and that  $K(i)$  may vary from frame to frame, so we can write  $R(i)$  as

$$R(i) = \begin{cases} \lceil \mu K(1) \rceil & \text{if } i = 1 \\ \lceil \mu \sum_{k=1}^i K(k) \rceil - \sum_{k=1}^{i-1} R(k) & \text{if } i > 1 \end{cases} \quad (7)$$

where  $\mu = (N - K)/K$  is the parity packet rate of RS coding, and operation  $\lceil X \rceil$  is used to get the minimum integer number greater than or equal to  $X$ . From now on, we name this approach as Evenly FEC.

#### B. Dynamic Sub-GOP FEC Coding

For a probability of packet loss  $p$ , and transmitting packet number  $N = K + \mu K$ , to be able to recover  $pN$  losses, we need to insert  $\mu K$  redundant packets, with  $\mu K \geq pN$  or equivalently  $\mu \geq p/(1-p)$ . In the limit case, according to the law of large number, when  $N \rightarrow \infty$ , then  $\mu K$  could be as small as  $pN$ , namely  $\mu K \approx pN$ , which means that the inserted redundancy could be as small as  $\mu = p/(1-p)$ . In practical situations,  $N \rightarrow \infty$  is impossible; in this case with the same parity packet rate  $\mu = (N - K)/K$ , the larger the value of  $K$  is, the higher the performance of RS code can be. Motivated by this fact, we propose to encompass packets from a Sub-GOP of frames to one RS coding block to enlarge the value of  $K$ . Fig. 2 shows one

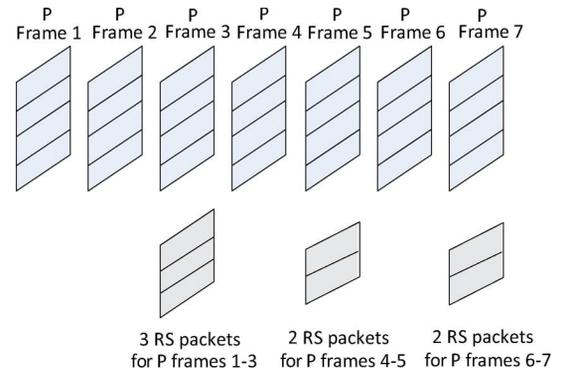


Fig. 2. One example of RS parity packets allocation for the dynamic Sub-GOP FEC coding approach.

example of how to generate Sub-GOPs and allocate RS parity packets at the end of each Sub-GOP. In addition, in order to meet the real-time constraint, we use the systematic RS code, so the source packets are intact in the RS coding process. Therefore, at the receiver side, the video decoder only needs packets belonging to one frame to decode and display that frame. If some packets of this frame are lost during transmission, error concealment is used to conceal them.

For easy illustration, let us take one example. We use the second Sub-GOP in Fig. 2, i.e., Frames four and five. In this case, the Sub-GOP contains two frames and each frame generates four packets. The redundancy due to the RS coding is 25%, which means for the eight source packets of this Sub-GOP, (10, 8) RS code is applied. In one Sub-GOP, when the first frame is encoded by the video encoder, immediately the encoded four packets are transmitted over the network. Due to network failure, let us assume two packets among these four packets are dropped. Then upon receiving the two packets, the video decoder will decode and display this frame, and for the lost packets, error concealment is used. In this manner, no additional delay is introduced. Later, the second frame is encoded, generating another four source packets. Meanwhile, as the second frame is the last frame in this Sub-GOP, two RS parity packets are generated for the eight source packets of this Sub-GOP. Suppose this time the four source packets and the two RS parity packets successfully arrive at the receiver side without any loss. In this case, at the receiver side, the (10,8) RS decoder will be able to recover the two lost packets, so the video decoder will re-decode the first frame of this Sub-GOP with the two received source packets and the two recovered packets, and update the reference frame buffer. In this case there will be no distortion in the second frame, and no error would propagate to the incoming frames. However, if the Evenly FEC approach is adopted, with the same amount of redundancy, for each frame (5,4) RS code will be used. The RS code would not be able to recover the two lost packets of the first frame; eventually the concealment distortion propagates to all the following frames, and severely degrades the video quality at the receiver side. In this illustrative simplified example, both the Evenly FEC and DSGF approach use the same amount of RS parity packets, and in both cases, no additional delay is introduced. However, the advantage of the DSGF approach is obvious, because this approach is able to restrict concealment distortion in very few frames.

### C. Optimal Sub-GOP Size and RS Packet Allocation

As described above, using the RS protection in Sub-GOP level could be better than in frame level, but the problem of how to divide frames into Sub-GOPs and how to allocate the RS parity packets among all the Sub-GOPs will be addressed in the following. To tackle this problem we have to note that, on one hand, if the Sub-GOP includes too few frames, the value of  $K$  for the RS code will not be large enough to make the RS code efficient. On the other hand, if the Sub-GOP includes too many frames, as the RS correction codes will not be available until the last frame of this Sub-GOP, the quality of those frames before the last frame would degrade significantly. Consequently, the Sub-GOP length should be properly tuned to increase the efficiency.

In general, I-frame generates much more bits than P-frame, and therefore more source packets are produced for I-frame. In our DSGF approach, for the I-frame we provide RS protection in frame level, the same as Evenly FEC approach, whereas for the P-frame we allocate RS parity packets in Sub-GOP level. Our objective is to optimally allocate Sub-GOP and RS parity packets and minimize the expected total distortion of this GOP. For easy access, all the notations, used in the optimal allocation process, are listed in the table provided in the Appendix. To do the optimal allocation, we need to know the detailed information of this GOP, including the slice number in each frame, the concealment distortion caused by losing each slice, and how the distortion propagates. However, those information are not available for real-time on-the-fly transmission system. In light of such circumstance, we established a model to represent these information. The model parameters include the number of P-frames in one GOP,  $L$ , and the average number of slices in each P-frame,  $S$ . For simplicity, let us assume the value of  $S$  is unchanged. This assumption is more accurate for low motion video with little change in the content. The expected concealment distortion of losing one packet is  $\bar{d}$ , the distortion in current frame propagates to the following frames, and the attenuation function of the distortion is  $f(n)$ . This means if the concealment distortion of one slice is  $\bar{d}$ , it will propagate to the following frames and become  $f(n)\bar{d}$  after  $n$  frames. For the sake of simplicity, the function  $f(n) = \alpha^{n-1}$  ( $0 < \alpha < 1$ ) is employed. This expression approximates, at low levels of attenuation, the function  $f(n) = (1 + \lambda_1 n)^{-1}$  and  $f(n) = (1 + \lambda_2 n)^{-1/2}$  reported in [19] and [20], respectively,  $\alpha$ ,  $\lambda_1$ , and  $\lambda_2$  being parameters to be selected. Let us assume that distortion caused by losing slices are uncorrelated, and in this case, the total expected distortion for the whole GOP is the sum of all the expected distortions caused by individual slices. The assumption on slice concealment distortion uncorrelation is reasonable. In fact, concealment distortions can be considered as uncorrelated with the pixel values; then concealment distortions caused by losing different slices can also be considered as uncorrelated. The additive distortion model has been verified experimentally in [10]. For the P-frames in one GOP, the total allocated RS parity packet number is  $R = \mu SL$ ; here  $\mu$  is the parity packet rate of RS coding. We use  $R(i)$  to denote the number of RS parity packet for P-frame  $i$ :

$$\sum_{i=1}^L R(i) \leq R. \quad (8)$$

Fig. 2 shows one example of how RS parity packets are allocated. We assume there are totally  $t$  positions where we insert RS parity packets, with frame number  $r_1, r_2, \dots, r_t$ , whereas other positions have no RS parity packets. The number of RS parity packets are  $R(r_1), R(r_2), \dots, R(r_t)$ . In the example of Fig. 2, we allocate RS parity packets in the three positions ( $t = 3$ ), the 3 positions are  $r_1 = 3, r_2 = 5, r_3 = 7$ , and RS packet number is  $R(r_1) = 3, R(r_2) = 2, R(r_3) = 2$ . The RS parity packets allocated under the Frame  $r_{m+1}$  are used to protect the frames from  $r_m + 1$  to  $r_{m+1}$ . Therefore, the parameters of RS  $(N, K)$  code for this Sub-GOP are  $N = (r_{m+1} - r_m)S + R(r_{m+1})$  and  $K = (r_{m+1} - r_m)S$ . We use  $\bar{D}(r_m + 1, r_{m+1})$  to denote the expected distortion caused by losing packets from Frame  $r_m + 1$  to  $r_{m+1}$ . It is important to note that  $\bar{D}(r_m + 1, r_{m+1})$  does not only include the distortion in frames from  $r_m + 1$  to  $r_{m+1}$ , but also accounts for the propagated distortion in subsequent frames:

$$\bar{D}(r_m + 1, r_{m+1}) = \bar{D}_i(r_m + 1, r_{m+1} - 1) + \bar{D}_p(r_{m+1}). \quad (9)$$

As described in (9), the distortion  $\bar{D}(r_m + 1, r_{m+1})$  is caused by two parts:  $\bar{D}_i(r_m + 1, r_{m+1} - 1)$  is the expected distortion within frames  $r_m + 1$  to  $r_{m+1} - 1$ . For those frames, the recovering capability of RS code cannot be used by the time when those frames are decoded and displayed. The subscript  $i$  means that  $\bar{D}_i(r_m + 1, r_{m+1} - 1)$  accounts for the *internal* distortion and their propagated distortion only within frames  $r_m + 1$  to  $r_{m+1} - 1$  due to the eventual losses in the frames. The error propagation to frames outside this Sub-GOP is not accounted in this term. Whereas,  $\bar{D}_p(r_{m+1})$  is the sum of expected distortion in Frame  $r_{m+1}$  and it also accounts for the propagated distortion in the subsequent frames. The subscript  $p$  denotes that this term includes the *propagated* distortion to frames beyond this Sub-GOP. Note that by the time of decoding and displaying Frame  $r_{m+1}$ , the RS code would try to recover the lost source packets within this Sub-GOP, but when the packet loss number in this Sub-GOP is beyond the recovery capability of RS code, the RS code would not be able to recover the lost packets.

For the term  $\bar{D}_i(r_m + 1, r_{m+1} - 1)$ , as from Frame  $r_m + 1$  to  $r_{m+1} - 1$ , the expected number of lost packets for each frame is  $pS$ , and by taking the error propagation inside this Sub-GOP into consideration, we get

$$\bar{D}_i(r_m + 1, r_{m+1} - 1) = \sum_{i=1}^{r_{m+1} - r_m - 1} \phi(i) pS \bar{d} \quad (10)$$

where  $\phi(i) = \sum_{n=0}^{i-1} f(n)$ .

Now let us evaluate the term  $\bar{D}_p(r_{m+1})$ . To do this, we have to note that from Frame  $r_m + 1$  to Frame  $r_{m+1}$ , there are  $(r_{m+1} - r_m)S$  source packets, while the RS parity packet number is  $R(r_{m+1})$ . Firstly, let us evaluate the expected distortion in Frame  $r_{m+1}$ ,  $\bar{D}_l(r_{m+1})$ , taking into account the error propagation from previous frames inside this Sub-GOP:

$$\bar{D}_l(r_{m+1}) = \frac{\bar{\Gamma}}{r_{m+1} - r_m} \phi(r_{m+1} - r_m) \bar{d} \quad (11)$$

with  $\bar{\Gamma}$  being the expected number of unrecoverable lost packets among Frame  $r_m + 1$  to Frame  $r_{m+1}$ , and this is  $\bar{\Gamma} = (r_{m+1} - r_m)p'S$ , where  $p'$  is the remaining packet loss rate after the RS correction. The detailed process to calculate  $p'$  has been given in

Section II for both i.i.d and burst model. Therefore, the expected number of unrecoverable lost packets in each frame would be  $\bar{\Gamma}/(r_{m+1} - r_m)$ . Given that the distortion in Frame  $r_{m+1}$  will propagate to the end of this GOP, therefore

$$\bar{D}_p(r_{m+1}) = \bar{D}_l(r_{m+1})\phi(L - r_{m+1} + 1). \quad (12)$$

At this point,  $\bar{D}_i(r_m + 1, r_{m+1} - 1)$  and  $\bar{D}_p(r_{m+1})$  can be used to calculate the expected distortion caused by each Sub-GOP using (9). By adding up the expected distortion caused by each Sub-GOP, we evaluate the total expected distortion for the whole GOP,  $\bar{D}_{total}$ , as in (13), shown at the bottom of the page where condition  $r_t \equiv L$  means that the last Sub-GOP has RS parity packets, whereas  $r_t < L$  means there are no RS parity packets for it.

Finally, the optimization problem can be formulated as the following constrained minimization:

$$\begin{cases} \min & \bar{D}_{total} \\ \text{subject to} & \sum_{i=1}^N R(i) \leq R. \end{cases} \quad (14)$$

#### D. Greedy Algorithm for Fast RS Parity Packet Allocation

It is computational prohibitive to get the global optimal solution for (13). When one GOP includes  $L$  P-frames, and the RS parity packet number for P-frames is  $R$ , there are totally  $\binom{L+R-1}{R}$  possible allocation solutions. For example, if the GOP length is  $L = 30$  and the number of RS parity packet is  $R = 40$ , there would be  $\binom{69}{40} = 2.39 \times 10^{19}$  possible allocation solutions. Obviously, calculating the value of  $D_{total}$  using (13) for all the  $2.39 \times 10^{19}$  allocation patterns is impossible. Since it is computational prohibitive to get the global optimal solution, we propose to use a greedy algorithm to get a sub-optimal RS parity packet allocation. In this algorithm, each time one RS parity packet is allocated, by trying to allocate this packet for all possible  $L$  positions, while calculating the value of  $D_{total}$  using (13) for all these positions, then the algorithm chooses to allocate the RS packet to the position which makes  $D_{total}$  minimum. Suppose adding the RS parity to Frame  $j$  would make  $D_{total}$  minimum, then  $R(j) = R(j) + 1$ . By iterating the previous steps  $R$  times, all the parity packets will be allocated. With this greedy algorithm, allocating  $R$  RS parity packets to  $L$  P-frames will have a computational complexity order of  $O(RL)$ , which is much less than  $\binom{L+R-1}{R}$ . The detailed procedure of greedy RS parity packet allocation algorithm is shown in Algorithm. 1.

---

#### Algorithm 1: Greedy RS parity packets allocation algorithm

---

**for**  $i = 1$  to  $L$  **do**

$$R(i) \leftarrow 0$$

**end for**

**for**  $j = 1$  to  $R$  **do**

$$index \leftarrow 0$$

$$distortion \leftarrow \infty$$

**for**  $i = 1$  to  $L$  **do**

$$R(i) \leftarrow R(i) + 1$$

calculate  $D_{total}$  using Equation (13)

$$R(i) \leftarrow R(i) - 1$$

**if**  $D_{total} \leq distortion$  **then**

$$index \leftarrow i$$

$$distortion \leftarrow D_{total}$$

**end if**

**end for**

$$R(index) \leftarrow R(index) + 1$$

**end for**

---

Fig. 3 shows two practical examples of how to divide P-frames into Sub-GOP and allocate the RS parity packets among all the Sub-GOPs with the greedy algorithm. These results have been obtained by assuming that one GOP has 30 P-frames, each P-frame includes 5 or 10 slices, the value of  $\alpha$  is 0.95, the packet loss rate is 5% with i.i.d model, and the parity packet rate is 20%. It is interesting to find some regular patterns behind the allocations. Firstly, in general, the P-frames at the beginning of the GOP have more RS parity packets than those at the end of the GOP. In Fig. 3(a), the first 2 Sub-GOPs have 4 RS parity packets for each Sub-GOP, the subsequent Sub-GOP has 3 RS parity packets, and so on. Whereas, for the last frame in the GOP, no RS parity packets are allocated. This is because any distortion in the front frames will propagate to the following frames, and usually losing one packet in the front frame would lead to more distortion for the whole GOP than losing one in the end. Therefore, it is reasonable to allocate more RS parity packets to the frames at the beginning of GOP. Secondly, it is important to note that at the beginning of the GOP, one Sub-GOP usually contains more frames than the Sub-GOP in the end of the GOP. In Fig. 3(a), the first 8 Sub-GOPs include 3 frames, the 9th Sub-GOP contains 2 frames, while the 10th and 11th Sub-GOPs contain only one frame. This is also because the distortion propagation paths in the frames at the beginning of a GOP are long. So putting more frames into one Sub-GOP can make the value of  $K$  large, which means that the RS code can recover the lost packets with higher probability, and eventually effectively cut down the error propagation. Thirdly, comparing results in Fig. 3(a) with Fig. 3(b), the average Sub-GOP length in Fig. 3(a) is larger than that in Fig. 3(b). This is because the

$$\bar{D}_{total} = \begin{cases} \bar{D}(1, i_1) + \sum_{m=1}^{t-1} \bar{D}(r_m + 1, r_{m+1}) & \text{for } r_t \equiv L \\ \bar{D}(1, i_1) + \sum_{m=1}^{t-1} \bar{D}(r_m + 1, r_{m+1}) + \bar{D}_i(r_t + 1, L) & \text{for } r_t < L \end{cases} \quad (13)$$

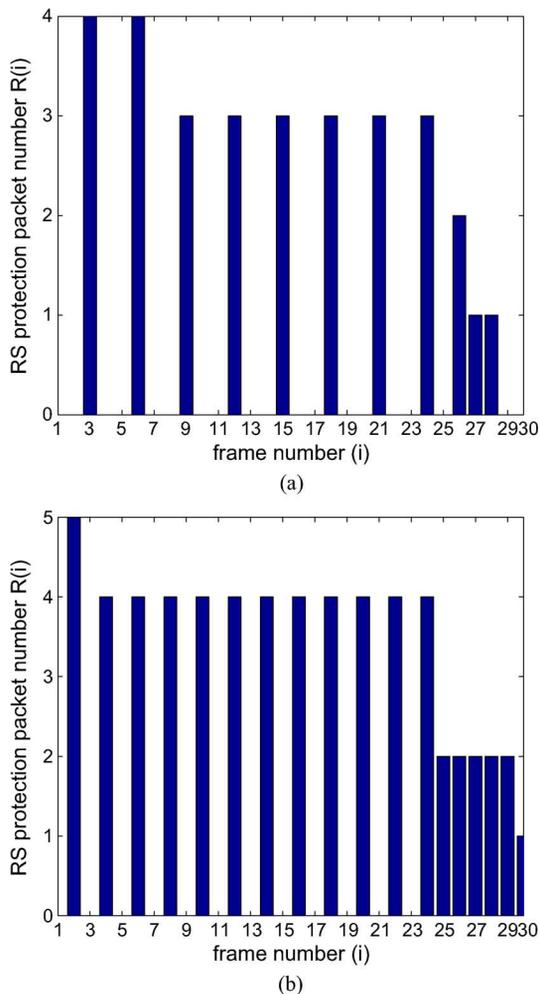


Fig. 3. RS allocation example with the greedy algorithm; packet loss rate  $p = 5\%$ ; one GOP has 30 P-frames; RS parity packet rate  $\mu = 20\%$ ; each frame includes  $S$  slices. (a)  $S = 5$ . (b)  $S = 10$ .

number of slices in each frame,  $S$ , is large in Fig. 3(b), and there is no need to put as many frames into one Sub-GOP as in Fig. 3(a).

#### E. Adaptive Estimation of the Attenuation Factor ( $\alpha$ )

As described above, in the Sub-GOP and RS packet allocation process, we need to know the distortion attenuation function  $f(n) = \alpha^{n-1}$  ( $0 < \alpha < 1$ ). However, for various video sequences, the attenuation factor  $\alpha$  is different; at the same time, for one specific video sequence, the parameter  $\alpha$  changes with time. Therefore one possible solution to determine  $\alpha$  would be to adaptively estimate its value at the end of each GOP, and to use this  $\alpha$  for the next GOP. The detailed process is as follows. While encoding the GOP, its slices will be decoded, and only one slice is assumed to be lost during the decoding process. So if we assume this slice is in Frame  $k$ , then, and because of the error propagation, the distortion due to this emulated loss will be  $d(k), d(k+1), d(k+2), \dots, d(L)$ ; here  $L$  is the length of the GOP. These distortions could be obtained by comparing the decoded sequence with the emulated loss with the error-free version. At the end of the GOP,  $\alpha$  could be evaluated as  $(1/L - k) \sum_{i=k+1}^L (d(i)/d(i-1))$ . Given the fact that the attenuation factor changes slowly, the estimated  $\alpha$  could be used to do the RS allocation of the following GOP. In addition, the

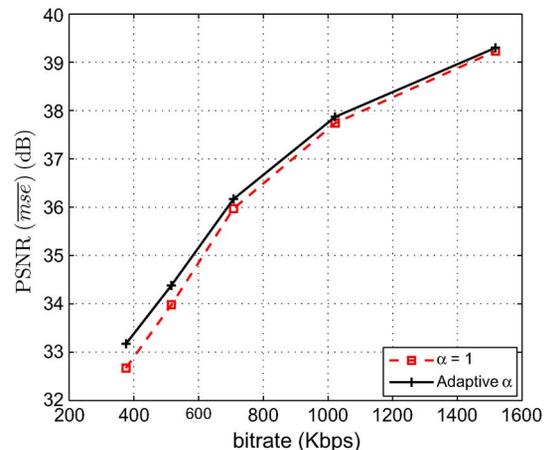


Fig. 4. PSNR versus bitrate for fixed  $\alpha$  value ( $\alpha = 1$ ) and adaptive  $\alpha$  for the DSGF approach; CIF Foreman sequence is used; GOP length is 30; packet loss rate is 5%; parity packet rate is 20%.

value of  $\alpha$  could be updated at the end of each GOP by running the previous procedure for each current GOP. More accurate estimation of  $\alpha$  could be obtained by emulating the loss of more than one slice.

## IV. SIMULATION RESULTS

Our simulation setting is built on the JM14.0 [21] H.264 codec. CIF video sequence Foreman, Bus, and Stefan are used for the simulations. We select these three sequences because they represent different motion characteristics. Foreman has moderate movement and video texture, Bus has fast and translational movement, whereas Stefan has fast movement with different motion directions. The GOP structure is IPPP, and GOP length is 30 frames. The reference frame number is 1; in other words, only the previous frame is used for prediction. One slice is transmitted in one packet; taking the MTU of wireless network into account, we set the target slice length as 400 bytes. We use the average luminance PSNR to assess the objective video quality, which is denoted as  $\text{PSNR}(\overline{mse})$ ; this is obtained by evaluating the mean squared error (mse) over all the frames and over 200 trials. Then the value of  $\text{PSNR}(\overline{mse})$  is calculated based on the averaged mse. In order to have a fair comparison, we compare the proposed DSGF approach with RS-MDC [10] and Evenly FEC approach, because all those approaches meet the real-time constraint and cause no additional delay. In the following simulations, we assume the packet loss follows the i.i.d model, unless burst packet loss model is explicitly specified. In Fig. 4, we compare the effects of using adaptive  $\alpha$  value with fixed  $\alpha$  value. For the fixed  $\alpha$  case, we choose  $\alpha = 1$ , which means the distortion will propagate without attenuation. It is interesting to note the gap between the two curves is rather narrow (always less than 0.5 dB), especially in high bitrate. Taking into consideration the huge computational resource for adaptively estimating the  $\alpha$  value, we use  $\alpha = 1$  in all the following simulations for simplicity. It is important to point out that with the adaptive  $\alpha$  estimation method described in Section III-E, the performance of the proposed approach is expected to be further enhanced.

In the first set of simulations, we study the effects of allocating different parity packet rates for RS code. The network

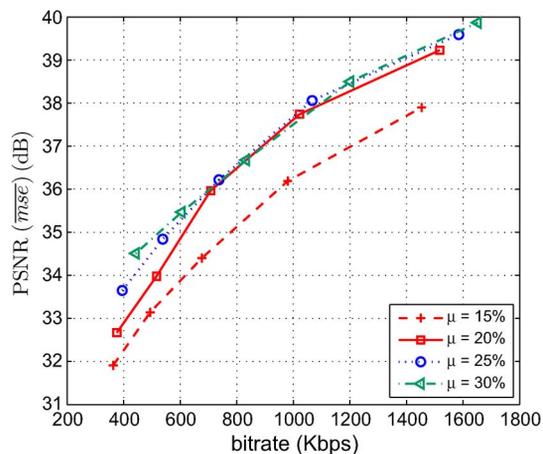


Fig. 5. PSNR versus bitrate for various parity packet rate  $\mu$ ; CIF Foreman sequence is used; packet loss rate is 5%; RS parity packet rate  $\mu$  includes 15%, 20%, 25%, and 30%.

packet loss is an i.i.d random process; for the same packet loss rate  $p = 5\%$ , we try different  $\mu$ , including 15%, 20%, 25%, and 30%. We do simulations with various quantization parameters (QP) to span a considerable bitrate range. Fig. 5 shows the PSNR versus bitrate with different RS parity packet rates  $\mu$ . The PSNR curve for parity packet rate 15% is much lower than other cases in all bitrate. In intermediate and high bitrate, the PSNR curves for  $\mu = 20\%$ , 25%, and 30% are very close, while in low bitrate, higher redundant rate can provide better PSNR. This is because in low bitrate, the slice number in each frame is small, which makes the performance of RS low, and high RS parity packet rate is required to compensate this. In general, the PSNR curves for redundant rate 20%, 25%, and 30% are similar; consequently, in the following simulations, we use parity packet rates 20% for 5% packet loss. Similarly, in later simulations, 40% RS parity packet rate is used for 10% packet loss.

Figs. 6 and 7 compare the performance of different approaches in term of PSNR versus bitrate and for random packet loss rate 5% and 10%, respectively. As mentioned above, for DSGF and Evenly FEC for the packet loss rate 5% and 10%, the RS coding redundancy is 20% and 40%, respectively, whereas for the RS-MDC, the redundancy is tuned as described in [10]. The proposed DSGF approach always outperforms RS-MDC and Evenly FEC in all the simulation environments. For all the Foreman, Bus, and Stefan sequences, the gain over Evenly FEC can be more than 2 dB in low bitrate when packet loss rate is 5%, and the gain over RS-MDC could be over 4 dB in high bitrate when packet loss rate is 10%. It is very interesting to note that the gap between the DSGF and Evenly FEC is larger in low bitrate than in high bitrate. This is because in high bitrate, more packets are generated for each frame, and consequently, for the frame level Evenly FEC approach, the value of  $K$  is relatively large to make the RS coding efficient.

In Fig. 8, for the Foreman and Stefan sequences, the PSNR of each frame in one GOP are plotted. In each Sub-GOP, the video quality degrades frame by frame gradually because of the random packet loss. However, at the end of each Sub-GOP, with high probability, the RS parity packets will be able to recover all the lost packets of this Sub-GOP, so the PSNR of the last frame of each Sub-GOP is higher than other frames in this

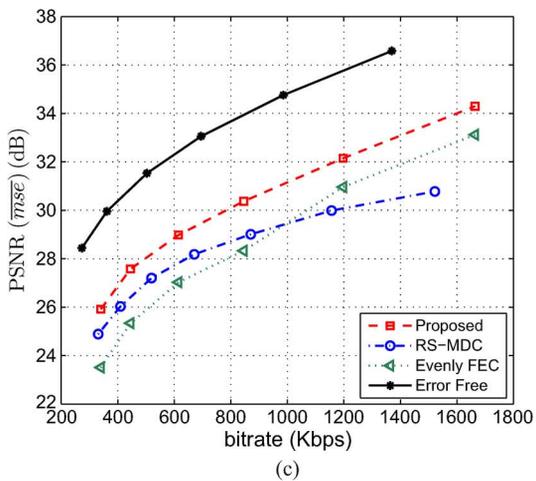
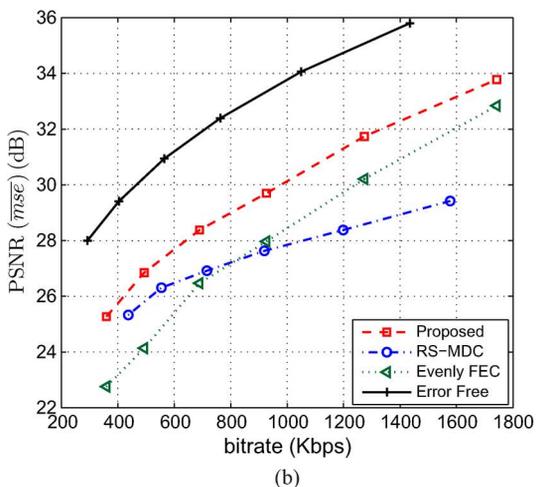
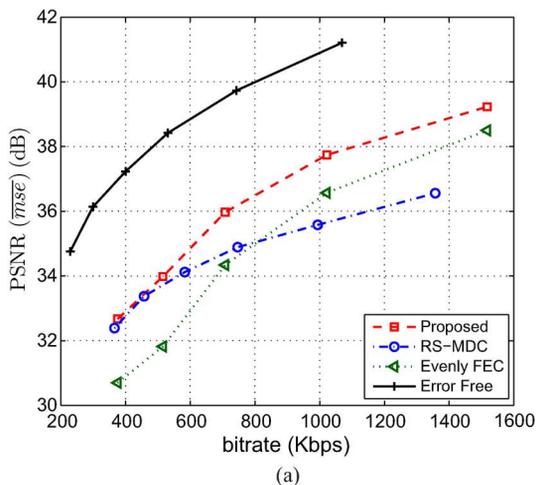


Fig. 6. PSNR versus bitrate curves; the network packet loss rate is 5% and the parity packet rate  $\mu$  is 20%. (a) Foreman sequence. (b) Bus sequence. (c) Stefan sequence.

Sub-GOP. All these factors make the video frame PSNR fluctuate, with a period the same as the Sub-GOP length. Nevertheless, for the majority of the frames in one GOP, PSNR of the proposed approach is higher than that of Evenly FEC approach and RS-MDC. In fact, among the 30 frames, only 6 and 3 frames have PSNR lower than that of the Evenly FEC approach for the Foreman and Stefan sequences, respectively; whereas

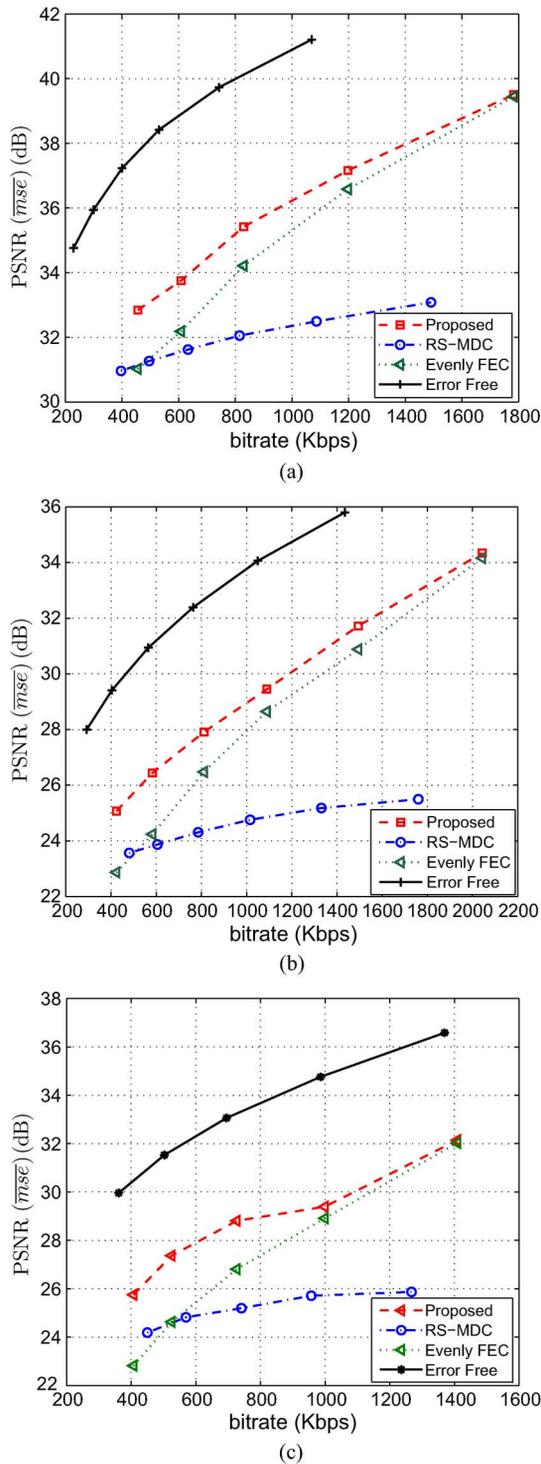


Fig. 7. PSNR versus bitrate curves; the network packet loss rate is 10% and the parity packet rate  $\mu$  is 40%. (a) Foreman sequence. (b) Bus sequence. (c) Stefan sequence.

almost all frames have better video quality than RS-MDC, although RS-MDC has some extra bitrate. It is worth noticing that for some video frames, PSNR of the proposed approach is more than 3 dB higher than that of the Evenly FEC approach, and for the second half of the GOP, our approach outperforms the Evenly FEC approach and RS-MDC significantly. Note that for the first frame in this GOP, which is I-frame, the video quality of

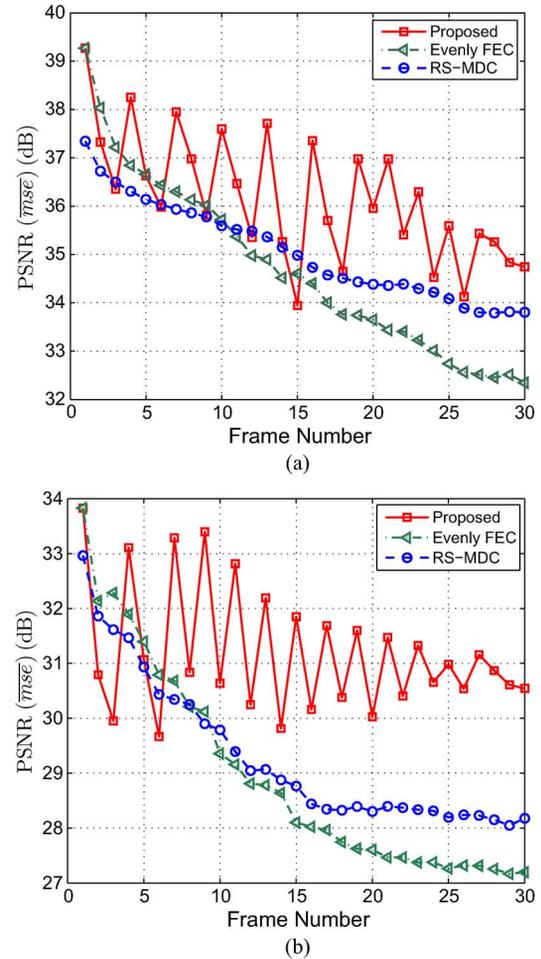


Fig. 8. Video quality versus frame number in one GOP with length 30; packet loss rate is 5%; parity packet rate  $\mu$  is 20%. (a) CIF Foreman sequence; QP = 26; bitrate for the proposed approach and the Evenly FEC approach is 707.9 Kbps, for RS-MDC is 746.5 Kbps. (b) CIF Stefan sequence; QP = 32; bitrate for the proposed approach and the Evenly FEC approach is 845.4 Kbps; for RS-MDC is 870.5 Kbps.

the proposed approach and Evenly FEC approach is the same, more than 0.8 dB better than RS-MDC. This is because the slice number in the I-frame is large, and that makes the RS code efficient, thereby providing higher PSNR than RS-MDC. It is worth noticing that similar results have been obtained for the Bus sequence. With the proposed approach, although the video quality rises and falls, this would not lead to inferior visual perception. To demonstrate that, two consecutive video frames after random packet loss are provided in Fig. 9 for the Foreman sequence, with the 10th frame has 38.27 dB and the 11th frame has 35.20 dB of PSNR. Despite the 3 dB PSNR gap, human eyes can hardly distinguish this quality fluctuation. In order to better visualize the video quality, the most damaged area in the 11th frame is zoomed in. We select this area because the slice that covers this area is lost, and concealment process is invoked; moreover, there is some motion in this area. However, even in this area, as we can see, the video quality is still acceptable.<sup>1</sup>

In all the previous experiments, the channel packet loss rate is assumed to be available at the video transmitter side. This can

<sup>1</sup>For the subjective visual quality, more results are provided in <http://www.mmtlab.com/Download.ashx>.



Fig. 9. Two consecutive frame in one decoded sequence after random packet loss ( $p = 5\%$ ). (a) 10th frame, with PSNR 38.27 dB. (b) 11th frame, with PSNR 35.20 dB. (c) Zoom in the most damaged area in the 11th frame.

be implemented with the real time control protocol (RTCP) [22]. However, in practical situations, feedback packet loss rate information may be delayed from the video receiver. Therefore, the packet loss rate may not be exactly identical to the actual packet loss rate. To further evaluate the performances of the proposed DSGF approach in this scenario, we assume the packet loss rate is 10%; therefore the redundancy of the RS-MDC is tailored for 10% packet loss rate, as proposed in [10]. This will lead to 36.5% of redundancy with QP 26 and GOP length 30, and it is worth noticing that this amount of redundancy is optimal for RS-MDC for this specific transmission scenario. In order to have fair comparison with other approaches, this amount of redundancy and the same QP has been used for Evenly FEC and the DSGF approach. This will generate the same total bitrate for the three approaches. The actual packet loss rate is varied from 0 to 20%. In Fig. 10, the video qualities of the three approaches under different packet loss rates are plotted. As we can see, the PSNR of the proposed DSGF approach is the highest among the three approaches, which means when there is packet loss rate fluctuation, the proposed approach can still provide the best video quality. Meanwhile, the gap between the DSGF approach and Evenly FEC increases with packet loss rate; that is because in this case, the RS parity packet rate  $\mu$  is fixed, increasing the packet loss rate makes the redundancy relatively small comparing to the packet loss rate. In this case, it becomes

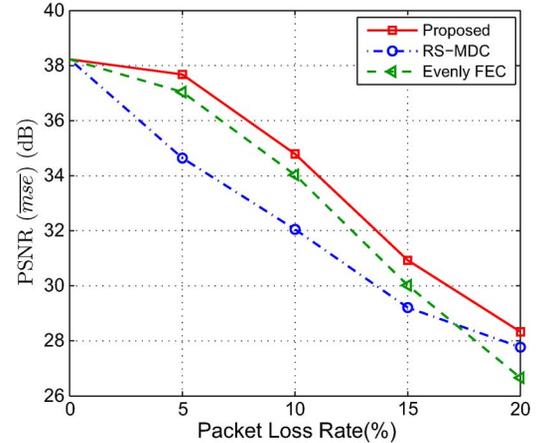


Fig. 10. Video quality versus different packet loss rate; CIF Foreman sequence is used; QP is 26; GOP length 30; the redundancy of the three approaches is  $\mu = 36.5\%$ .

more important to group frames together in order to increase the efficiency of RS coding.

In order to validate the performance of the proposed DSGF approach in different error distribution models, some additional results are provided for real Internet packet loss pattern and Gilbert burst loss pattern. In Fig. 11, the average PSNR versus bitrate for real Internet environments are presented, for the Foreman and Stefan sequences. The packet loss pattern for the Internet experiments specified in the file 10 of Q15-I-16r [23] is used to emulate the real Internet environments. This has an actual packet loss rate of 11.38%. From the results, we could note that, in the real Internet environments, the proposed DSGF approach outperforms RS-MDC and Evenly FEC, which is similar to the results obtained in the i.i.d packet loss environment. In Fig. 12, we compare the proposed DSGF approach with the other two approaches in Gilbert burst loss environment. As indicated in [24], we set the average burst length as two, and since burst packet loss usually requires higher redundancy comparing to i.i.d loss, 60% parity packets are inserted for the 10% average packet loss rate for both the DSGF and Evenly FEC approaches. In burst loss environment, the DSGF approach outperforms the Evenly FEC significantly, with an average gain of more than 2 dB, being higher than that in i.i.d case [Fig. 7(a)]. This is because, in burst loss environment, consecutive packets tend to be lost together; in this case, FEC coding at Sub-GOP level can mitigate this kind of loss more efficiently than Evenly FEC. It is also noted that, in burst loss environment, the gain over the RS-MDC is lower than that in i.i.d case [Fig. 7(a)]. This is due to the packet arrangement in RS-MDC. In fact, in RS-MDC, for each frame, redundant packets are grouped together and sent sequentially before the primary packets. In this case, the probability of losing both primary and redundant packets for the same video content becomes quite low. Nevertheless, in all the bitrate higher than 500 kbps, the DSGF approach provides much higher PSNR than RS-MDC, and the gap increases dramatically with bitrate.

## V. CONCLUSIONS

In this paper, a real-time FEC video transmission approach has been proposed. We firstly presented the general idea of this

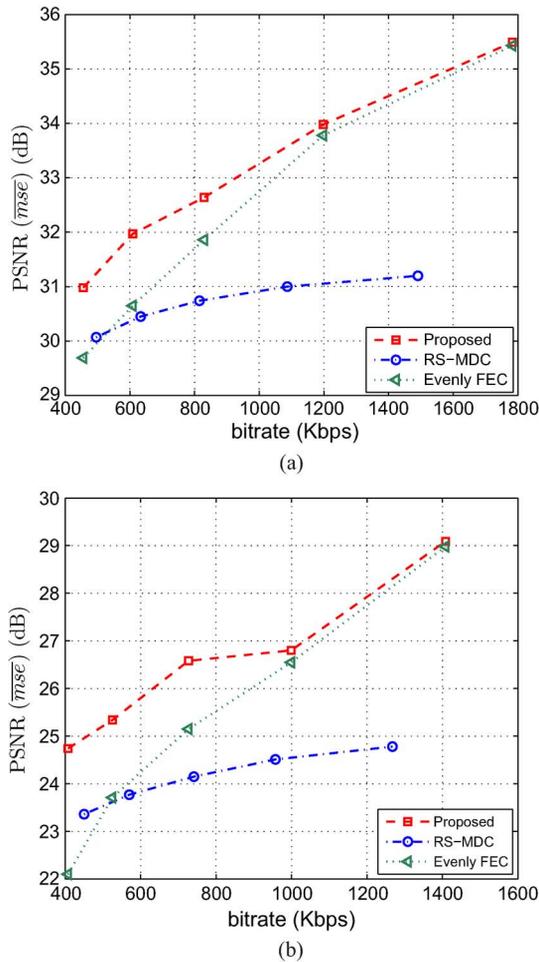


Fig. 11. PSNR versus bitrate curves; the network packet loss follows the pattern in file 10 of Q15-I-16r;  $\mu$  is 40%. (a) Foreman sequence. (b) Stefan sequence.

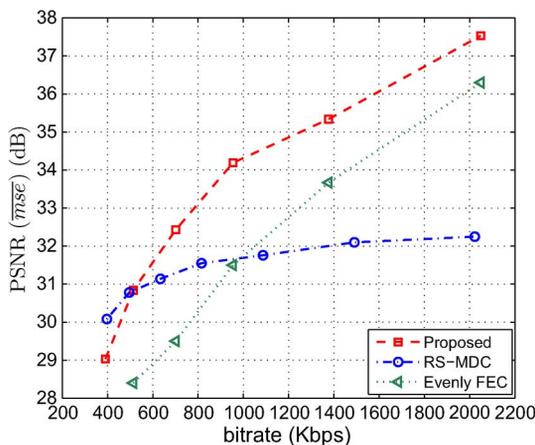


Fig. 12. PSNR versus bitrate curves for the Foreman sequence; the packet loss rate is 10%; average burst length is two;  $\mu$  is 60%.

approach; then the theoretical model for creating Sub-GOP and allocating FEC protection packets was given. With this model, the allocation problem became a constrained optimization problem. To resolve it, a fast greedy algorithm was proposed. In order to validate the proposed approach, its performance was compared with other real-time error resilient approaches, such

TABLE I  
ALL THE IMPORTANT NOTATIONS USED THROUGHOUT THIS PAPER

Notation	Description of the notation
$(N, K)$	$K$ is the Reed-Solomon source packet number, $N - K$ is the number of the parity packets
$p$	average packet loss rate before RS code correction
$p'$	remaining packet loss rate after RS code correction
$S$	average number of slices in each P-frame
$L$	P-frame number in one GOP
$d$	expected concealment distortion of losing one slice
$\mu$	parity packet rate
$R$	total number of parity packet for all the P-frames in one GOP
$f(n)$	distortion propagation attenuation function
$\alpha$	distortion propagation attenuation factor
$r_i$	the index of the $i$ -th P-frame where parity packets are inserted
$R(r_i)$	the number of parity packets for the frame with index $r_i$
$\bar{D}(r_m + 1, r_{m+1})$	the total expected distortion caused by frames $[r_m + 1, r_{m+1}]$
$\bar{D}_i(r_m + 1, r_{m+1} - 1)$	the expected distortion within frames $r_m + 1$ to $r_{m+1} - 1$ , caused by the current Sub-GOP
$\bar{D}_p(r_{m+1})$	the sum of expected distortion in Frame $r_{m+1}$ and the propagated distortion in the subsequent frames, caused by the current Sub-GOP
$\bar{D}_i(r_{m+1})$	the expected distortion for Frame $r_{m+1}$ , caused by the current Sub-GOP
$\bar{D}_{total}$	the total expect distortion of the GOP

as RS-MDC and Evenly FEC. Experimental results demonstrated that the proposed approach had considerable practical value for real-time applications.

One direction for future work would be to take the network transmission delay into consideration in the Sub-GOP and protection packet allocation process. For such a system, certain amount of end-to-end delay would be allowed for video packet transmission. In this case, how to allocate the Sub-GOP and protection packets becomes another constraint optimization problem. Moreover, as certain amount of delay is allowed, introducing the B-frames into such system is also worth investigating, since typically B-frames provide higher compression performance comparing to the I-frames and P-frames.

## APPENDIX

Table I lists all the important notations used throughout this paper.

## ACKNOWLEDGMENT

The authors would like to thank professor M. Elia at Politecnico di Torino for guidance in Reed-Solomon code, and S. Yu in Duke University for English language correction.

## REFERENCES

- [1] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the h.264/avc video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.
- [2] S. Wenger, "H.264/avc over ip," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 645–656, Jul. 2003.
- [3] T. Stockhammer, M. Hannuksela, and T. Wiegand, "H.264/avc in wireless environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 657–673, Jul. 2003.
- [4] W. Yao, S. Wenger, J. Wen, and A. Katsaggelos, "Error resilient video coding techniques," *IEEE Signal Process. Mag.*, vol. 17, pp. 61–82, Jul. 2000.
- [5] R. Zhang, S. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 966–976, Jun. 2000.

- [6] Y. Zhang, W. Gao, Y. Lu, Q. Huang, and D. Zhao, "Joint source-channel rate-distortion optimization for h.264 video coding over error-prone networks," *IEEE Trans. Multimedia*, vol. 9, no. 3, pp. 445–454, 2007.
- [7] S. Soltani, K. Misra, and H. Radha, "Delay constraint error control protocol for real-time video communication," *IEEE Trans. Multimedia*, vol. 11, no. 4, pp. 742–751, 2009.
- [8] P. Chou and Z. Miao, "Rate-distortion optimized streaming of packetized media," *IEEE Trans. Multimedia*, vol. 8, pp. 390–404, Apr. 2006.
- [9] S. Lin, S. Mao, Y. Wang, and S. Panwar, "A reference picture selection scheme for video transmission over ad-hoc networks using multiple paths," in *Proc. IEEE Int. Conf. Multimedia and Expo (ICME 2001)*, 2001, pp. 96–99.
- [10] T. Tillo, M. Granello, and G. Olmo, "Redundant slice optimal allocation for h.264 multiple description coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 59–70, Jan. 2008.
- [11] C. Zhu, Y.-K. Wang, M. Hannuksela, and H. Li, "Error resilient video coding using redundant pictures," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 1, pp. 3–14, Jan. 2009.
- [12] I. Radulovic, P. Frossard, Y.-K. Wang, M. Hannuksela, and A. Halpapuro, "Multiple description video coding with h.264/avc redundant pictures," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 1, pp. 144–148, Jan. 2010.
- [13] B. A. Heng, J. G. Apostolopoulos, and J. S. Lim, "End-to-end rate-distortion optimized md mode selection for multiple description video coding," *EURASIP J. Appl. Signal Process.*, vol. 2006, no. 1, p. 12, 2006.
- [14] E. Baccaglini, T. Tillo, and G. Olmo, "Slice sorting for unequal loss protection of video streams," *IEEE Signal Process. Lett.*, vol. 15, pp. 581–584, 2008.
- [15] X. Yang, C. Zhu, Z. G. Li, X. Lin, and N. Ling, "An unequal packet loss resilience scheme for video over the internet," *IEEE Trans. Multimedia*, vol. 7, no. 4, pp. 753–765, 2005.
- [16] N. Thomos, S. Argyropoulos, N. Boulgouris, and M. Strintzis, "Robust transmission of h.264/avc video using adaptive slice grouping and unequal error protection," in *Proc. IEEE Int. Conf. Multimedia and Expo*, 2006, pp. 593–596.
- [17] J. Xiao, T. Tillo, C. Lin, and Y. Zhao, "Real-time forward error correction for video transmission," in *Proc. 2011 IEEE Visual Communications and Image Processing (VCIP)*, Nov. 6–9, 2011, pp. 1–4.
- [18] P. Frossard, "FEC performance in multimedia streaming," *IEEE Commun. Lett.*, vol. 5, pp. 122–124, Mar. 2001.
- [19] N. Farber, K. Stuhlmuller, and B. Girod, "Analysis of error propagation in hybrid video coding with application to error resilience," in *Proc. 1999 Int. Conf. Image Processing (ICIP 99)*, 1999, vol. 2, pp. 550–554.
- [20] B. Girod and N. Farber, "Feedback-based error control for mobile video transmission," *Proc. IEEE*, vol. 87, no. 10, pp. 1707–1723, Oct. 1999.
- [21] HHI Fraunhofer Institute, H.264/AVC Reference Software. [Online]. Available: <http://iphome.hhi.de/suehring/tml/download/>.
- [22] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A transport protocol for real-time applications," in *Internet Engineering Task Force—RFC 3550*, Jul. 2003.
- [23] Y.-K. Wang, S. Wenger, and M. M. Hannuksela, "Common Conditions for SVC Error Resilience Testing," *JVT document P206*, Aug. 2005.
- [24] D. Loguinov and H. Radha, "End-to-end internet video traffic dynamics: Statistical study and analysis," in *Proc. IEEE INFOCOM '02*, Jun. 2002, pp. 723–732.



**Jimin Xiao** was born in Suzhou, China. He received the B.S. and M.Eng. degrees in telecommunication engineering from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2004 and 2007, respectively. Since 2009, he has been pursuing the Ph.D. degree at the University of Liverpool, Liverpool, U.K.

From 2007 to 2009, he worked as a software engineer at Motorola Nanjing Center, and later as a system engineer at Realsil (Realtek) Semiconductor Corp. His research interests are in the areas of video

streaming, image and video compression, and 3-D video coding.



**Tammam Tillo** (M'05) was born in Damascus, Syria. He received the degree in electrical engineering from the University of Damascus in 1994, and the Ph.D. degree in electronics and communication engineering from Politecnico di Torino, Torino, Italy, in 2005.

From 1999 to 2002, he was with Souccar for Electronic Industries, Damascus, Syria. In 2004, he was a visiting researcher at the EPFL (Lausanne, Switzerland), and from 2005 to 2008, he worked as a Post-Doctoral researcher at the Image Processing Lab of Politecnico di Torino, and for a few months, he was an Invited Research Professor at the Digital Media Lab, SungKyunKwan University, Suwon, Korea. Since August 2008, he has been with Xi'an Jiaotong-Liverpool University, Suzhou, China. His research interests are in the areas of robust transmission, image and video compression, and hyperspectral image compression.

Prof. Tillo has been a member of the technical program committees for several international conferences.



**Chunyu Lin** was born in LiaoNing Province, China. He received the B.S. degree from Bohai University, Jinzhou, China, in 2002 and the M.S. degree from Yanshan University, Qinhuangdao, China, in 2005. In 2011, he received the Ph.D. degree from Beijing Jiaotong University, Beijing, China.

From 2009 to 2010, he was a visiting researcher at the ICT group of Delft University of Technology, Delft, The Netherlands. From 2011 to 2012, he was a postdoc in MMLab, Ghent University, Ghent, Belgium. His research interests are in the areas of image/

video compression and robust transmission, 2D-to-3D conversion, and 3D video coding.



**Yao Zhao** (M'06–SM'12) received the B.S. degree from Fuzhou University, Fuzhou, China, in 1989 and the M.E. degree from the Southeast University, Nanjing, China, in 1992, both from the Radio Engineering Department, and the Ph.D. degree from the Institute of Information Science, Beijing Jiaotong University (BJTU), in 1996.

He became an Associate Professor at BJTU in 1998 and became a Professor in 2001. From 2001 to 2002, he worked as a senior research fellow in the Information and Communication Theory Group, Faculty of Information Technology and Systems, Delft University of Technology, Delft, The Netherlands. He is now the Director of the Institute of Information Science, Beijing Jiaotong University. His research interests include image video coding, fractals, digital watermarking, and content based image retrieval. Now he is leading several national research projects from 973 Program, 863 Program, the National Science Foundation of China.

Prof. Zhao was the recipient of the National Outstanding Young Investigator Award of China in 2010.