Adaptively Optimized Error Protection for H.264 Scalable Video Coding

Bin Zhang and Mathias Wien*

RWTH Aachen University, Aachen, Germany
{bzhang,wien}@ient.rwth-aachen.de
http://www.ient.rwth-aachen.de/cms/b_zhang

Abstract. We focus on the problem of providing robust video streaming under varying channel conditions in error-prone networks. We propose to apply unequal error protection based on priority encoding transmission scheme and Reed-Solomon codes to H.264/AVC scalable video (SVC) bit streams. We developed a fast algorithm to optimize the allocation of channel bit rate to different network abstraction layer (NAL) units on the GOP basis. To achieve this, we determined the real utility of each NAL unit in the SVC bit stream. Our simulations demonstrated the good performance of our algorithm. Good video quality can be guaranteed at the receiver side over a large range of packet loss rates by applying our algorithm to SVC bit stream. The results achieved with our fast algorithm are very close to the results obtained from time-consuming extensive search algorithm.

Keywords: SVC, UEP, Video Streaming.

1 Introduction

Services based on video streaming are becoming more and more important, due to the latest development in video compression technology and the enhanced capacity of mobile and wired networks. Although networks are more reliable than ever before, packet loss is still unavoidable because of factors such as network congestion, channel fading, link layer outage. To deal with the packet loss effects, it is necessary to provide error protection for video bit streams. As an extension of the most advanced video coding standard H.264/AVC, scalable video coding (SVC) can provide highly efficiently compressed scalable bit streams with only about 10% overhead with optimized encoding setting [1] compared to bit streams compressed with the single layer coding (SLC) method specified by the standard. The excellent performance of SVC and the great demand of multimedia streaming services motivated us to develop a new mechanism to provide a robust transmission of video bit streams over lossy networks using SVC.

The network abstraction layer (NAL) units of a SVC bit stream have inherently different importance. By applying the unequal erasure protection (UEP)

* This work has been supported by the UMIC Research Centre, RWTH Aachen University.

J. Rodriguez, R. Tafazolli, C. Verikoukis (Eds.): MOBIMEDIA 2010, LNICST 77, pp. 515–529, 2012.

technique to SVC bit streams, the so called graceful degradation can be achieved under varying channel conditions. Different layers shall be protected with channel codes of various strength.

Channel coding can be applied on the link layer or on the application layer. The advantage of application layer UEP (AL-UEP) is that no additional modification on the existing network infrastructure is needed. In our work, we adopt the AL-UEP approach and aim at developing an algorithm with low complexity to determine the optimal assignment of the protection strength, adapted to the varying channel conditions, for different NAL units.

Many previous works have explored the UEP approach for different video coding techniques. [2] focused on developing an algorithm to find an approximately optimal allocation of the protection strength for 3D-SPIHT bit stream. [3] applied UEP to H.263 bit stream by reconstructing it into an embedded bit stream. Simple linear group of picture (GOP) structure consisting of I and consequent P pictures (denoted as I-P-P) was used. [4] proposed a distortion model to determine the utility of video packets in I-P-P GOP for H.263 bit stream. However, the distortion model is only applicable for I, P and non-reference B frames. [5] applied UEP to H.264 bit stream with I-P-P GOP. The authors determined the utility of the packets by emulating error-concealment in the encoding process. The propagation distortion was, however, ignored. The utility/cost ratio of the packets was then used as an indicator for the importance. Recent works on the UEP for H.264 SVC include [6] and [7]. [6] proposed a layer-aware protection approach to bit stream with spatial scalability in multicast applications. The protection strength for the base layer (BL) and the enhancement layers (ELs) is not adaptive to the varying channel conditions. [7] provided a limited adaptive scheme for spatial BL and EL by dividing the NAL units into a few priority classes and only one Reed-Solomon (RS) code is used. In [8], the authors proposed a model to estimate the distortion of each NAL unit to the video sequence, if it is absent. However, the author didn't discuss the complexity to build to model.

To the best of our knowledge, adaptive UEP scheme using the SVC video coding to enhance the robustness of video streaming has not been well studied yet. In this work, we propose a novel UEP framework based on medium grained scalable (MGS) video coding. The enhanced robustness is achieved through the hierarchical GOP structure, the layered coding for each frame and a highly adaptive UEP allocation algorithm. Compared to I-P-P GOP structure discussed frequently in previous works, the drift error cause by one lost frame only propagates to B-frames at higher temporal levels within the hierarchical GOP. To avoid distortion propagation, we must consider the inherent dependency of the NAL units in the GOP when determining their utilities and impose a constraint on the level of protection strength for NAL units at different temporal and quality levels. Since no accurate distortion model is available for the B-frames in the hierarchical GOP, we determine the utilities of NAL units by decoding bit streams containing selected subsets of the NAL units within the GOP multiple times. Error concealment is performed if the BL NAL unit for one frame is lost.

The rest of the paper is outlined as follows: Section 2 describes the important procedures and components in the UEP framework. Section 3 formulates the optimization problem. In Section 4, we discuss about the solution algorithm. Section 5 provides the simulation results. This paper ends with a conclusion in Section 6.

2 Unequal Erasure Protection Framework

Fig. 1 illustrates the UEP framework. The media server in the framework encoded and stores a number of precoded SVC bit streams with different BL and EL bit rates denoted as R_{BL}, R_{EL}. The total source bit rate is $R_{\text{src}} = R_{\text{BL}} + R_{\text{EL}})$ for the coded bit stream. The utility and cost of the BL and EL NAL units are analyzed and stored as meta data for each bit stream on the server. The bit stream to be transmitted is selected by the server based on the channel bit rate R_{ch} and R_{src} at the beginning of the transmission. The EL NAL units are intended to provide an enhanced quality for occasions with good channel conditions and the BL NAL units prevent frequent frame loss for occasions with bad channel conditions, ensuring an acceptable quality.

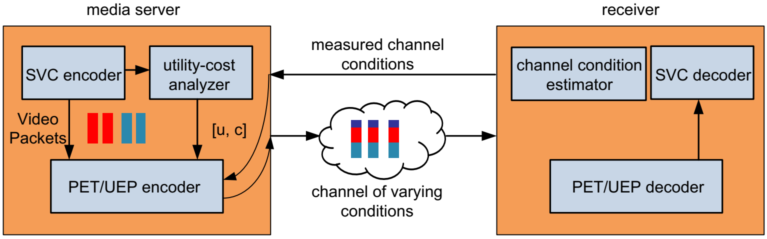

Fig. 1. UEP framework

During the transmission, the receiver sends some parameters about the channel status, i.e., channel bit rate R_{ch}, packet loss rate π and packet loss correlation ρ as feedback through a secure signaling channel to the media server on regular basis. The BL and EL NAL units are protected with Reed-Solomon (RS) codes of different strength, which are determined at the server side after it gets knowledge of the current channel status from the receiver. In this way, the error protection for different NAL units is adaptive to the varying channel status. We allow the server to drop the EL NAL units at poor channel conditions to strengthen the protection for BL NAL units using the saved bit rate from video source.

The transmission packets are constructed from the NAL units using priority encoding transmission (PET) scheme. With PET scheme, we can guarantee that NAL units protected with stronger channel codes are never lost prior to those protected with weaker codes independent of the channel realizations. In this framework, the server adapts the protection strength on a GOP basis by maximizing the expected end-to-end GOP utility at the receiver. In Section 2.1, 2.2, 2.3, we will go into details about the UEP framework.

2.1 Utility and Cost

The cost c of one NAL unit is measured in its size in terms of bytes. The utility u of one NAL unit is defined by the reduction of the distortion summed over all frames influenced by the NAL unit in the GOP. Compared to the wavelet-based video bit stream or bit stream with I-P-P GOP structure, the frames in the hierarchical GOP do not have a unique order according to which the video quality is progressively refined. Therefore, it is very important to reorder the NAL units in the bit stream to make the sequence progressively refinable before determining u of each NAL unit.

Since the drift effect deteriorates the video quality very adversely, we place the NAL units belonging to a frame f_a before all NAL units belonging to the other frames that depend on f_a. This indicates that the following order should be applied: 1) The BL NAL units are ordered prior to EL NAL units. 2) NAL units at lower temporal layer are ordered prior to those at higher temporal layers. 3) For NAL units at the same temporal and quality level, they are ordered firstly according to their displaying order. Let η_l^{tq} denote the l^{th} NAL unit on the t^{th} temporal and the q^{th} quality layer. Fig. 2 shows two hierarchical GOPs of a bit stream with GOP size of 4 consisting of one BL and one EL. The NAL units in each GOP are ordered independently. The sorted set of the NAL units in both GOPs would be: $\mathcal{N} = \{\eta_0^{00}, \eta_0^{10}, \eta_0^{20}, \eta_1^{20}, \eta_0^{01}, \eta_0^{11}, \eta_0^{21}, \eta_1^{21}\}$. In the following, we denote $|\mathcal{N}|$ as Q. The q^{th} NAL unit in $\mathcal{N}$ is denoted as η_q and the q^{th} subset of $\mathcal{N}$ is $\mathcal{N}_q = \{\eta_1, \ldots, \eta_q\}$. Note that $\mathcal{N}_0 = \emptyset$. We define a function $\mathrm{Dec} : \{\mathcal{N}_1, \ldots, \mathcal{N}_Q\} \mapsto \log(\mathrm{SSE})$, which outputs the logarithmic sum of squared error (SSE) of the GOP decoded with the specified subset of NAL units. If the BL NAL unit of one frame is not included in the subset, then error concealment is performed by replacing the missing frame with the temporally nearest decoded frame. The utility of η_q is defined as:

$$u_q = \mathrm{Dec}(\mathcal{N}_{q-1}) - \mathrm{Dec}(\mathcal{N}_q). \tag{1}$$

With the calculated u_q and c_q for all NAL units, we reorder the NAL units at the same temporal and quality level in the original set $\mathcal{N}$ in decreasing order of u_q/c_q. For example, if the u_q/c_q ratio associated with η_1^{21} is greater than that associated with η_0^{21}, then the final sorted set of NAL units is $\mathcal{N}^* = \{\eta_0^{00}, \eta_0^{10}, \eta_0^{20}, \eta_1^{20}, \eta_0^{01}, \eta_0^{11}, \eta_1^{21}, \eta_0^{21}\}$.

2.2 Channel Model and Reed-Solomon Codes

In this section we discuss the channel model we used for the wired and wireless networks. We assume that the channel is approximately stationary between successive feedbacks from the receiver during the transmission. Therefore, we can use a two-state Markov model as proposed in [9] to capture the channel behavior in that temporal gap. A Good (G) and a Bad (B) state are defined in this model. In the G state, packets are transmitted error-free and timely, while in the B state, the transmitted packets are either lost or arrive too late at the receiver.

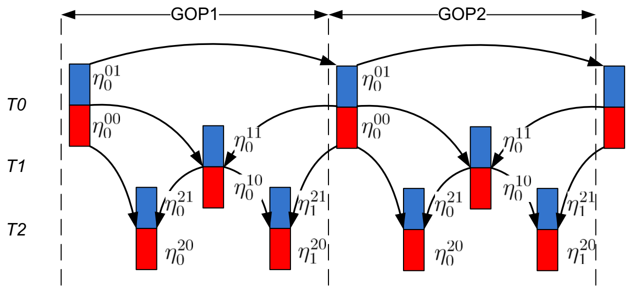

Fig. 2. Two GOPs with GOP size 4

The channel can be totally characterized by the transition probability p_{GB} from G state to B state and by p_{BG} from B state to G state. The transition probability matrix is as follows:

$$P = \begin{bmatrix} p_{GG} & 1 - p_{GG} \\ 1 - p_{BB} & p_{BB} \end{bmatrix}. \tag{2}$$

The probabilities of being in states G and B are $\pi_G = \frac{1-p_{BB}}{2-p_{GG}-p_{BB}}$ and $\pi_B = \frac{1-p_{GG}}{2-p_{GG}-p_{BB}}$, respectively.

In the proposed framework, we assume that the receiver measures the packet loss rate π and the packet loss correlation ρ and sends them as feedback to the server. The server can derive p_{GG} and p_{BB} from π and ρ as follows:

$$p_{GG} = 1 - \pi + \pi\rho, \tag{3}$$

$$p_{BB} = \rho + \pi - \pi\rho. \tag{4}$$

We use RS codes to protect the SVC video bit stream. If the source data consists of k symbols, we can generate $n - k$ parity symbols by applying the RS code denoted as (n, k). For RS codes (n, k), $1 \leq k \leq n$, the protection strength increases with decreasing k. Since RS codes can be categorized as maximum distance separable codes, if at least k symbols out of the n encoded symbols are received by the receiver, the source data can be reconstructed. With given channel condition parameters π and ρ, the effectiveness curve associated with all applicable error protection modes (indexed by $1 \leq t \leq n + 1$) can be determined as in Fig. 3, where $2 \leq t \leq n + 1$ indicates the usage of RS codes with increasing protection strength and $t = 1$ indicates discarding of the NAL unit. $p(t)$ is the function giving the probabilities of successful transmitting a NAL unit with the error protection mode t. More details about the calculation of $p(t)$, please refer to [9]. $r(t)$ is the corresponding channel coding overhead factor, i.e., reciprocal value of the code rate function, which is defined as $r(t) = \min(t - 1, 1)\frac{n}{n+2-t}$.

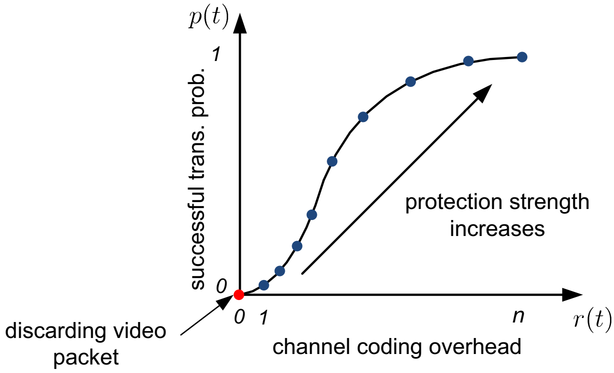

Fig. 3. Effectiveness curve of error protection modes

2.3 Priority Encoding Transmission

The PET scheme was first introduced in [10]. A PET transmission block is a byte matrix with n columns (see Fig. 4). For example, we assume that the PET scheme assigns the i^{th} NAL unit η_i a RS code (n, k_i). Then, each time k_i bytes are taken from η_i and supplied with $n - k_i$ parity bytes generated with the specified RS code. The total n encoded bytes are then put into one line of the transmission block. The process is repeated until all data of η_i are written into the block, then the data of next NAL unit will be processed. The NAL unit η_i occupies $\left\lceil \frac{c_i}{k_i} \right\rceil$ lines in the block.

After all Q NAL units are encoded with the corresponding RS codes and put into the block, the data in the columns are used as the payload of the RTP transmission packets. Therefore, if one packet is lost, one encoded byte is erased from the n encoded bytes in each line for all NAL units. The PET scheme has one important property, that the receipt of any k out of n transmission packets guarantees that all data protected with a RS code stronger than RS code (n, k) can be restored. This property indicates that stronger protection should be assigned to the NAL unit used by other frames for prediction within the PET transmission block. While applying the PET scheme to the SVC bit stream, we keep the NAL units belonging to one GOP in the same PET transmission block. In this way, we guarantee that frames protected with stronger RS codes within the GOP will not get lost before the frames protected with weaker RS codes. It should be noted that the drift effect across GOPs cannot be totally eliminated through this arrangement: If the key picture of one GOP gets lost, the GOPs that use the lost key picture for reference will be affected.

3 Optimization Problem

We maximize the expected utility of one GOP by finding the indices of optimal error protection modes for all NAL units in $\mathcal{N}^*$. The indices are represented by

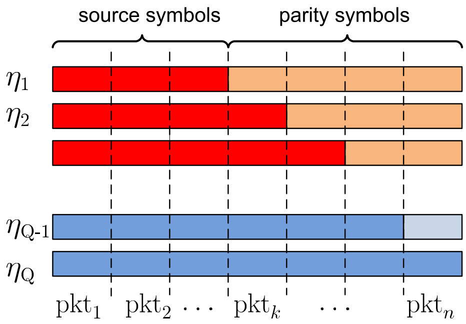

Fig. 4. Priority encoding transmission scheme

a vector $\mathbf{t} = [t_1, \ldots, t_Q]$, where $1 \leq t_q \leq n+1$, $1 \leq q \leq Q$. t_q is the index of the error protection mode for the q^{th} NAL unit. We denote the vector space of $\mathbf{t}$ as $\mathbf{T}$. The optimization problem can be formulated as follows:

$$\bar{\mathbf{t}} = \arg\max_{\mathbf{t}\in\mathbf{T}} \sum_{q=1}^{Q} u_q p(t_q), \tag{5}$$

$$\text{subject to}: \quad \sum_{q=1}^{Q} c_q r(t_q) \leq C_{\max}, \tag{6}$$

$$1 \leq t_Q \leq t_{Q-1} \leq \cdots \leq t_1 \leq e+1, \tag{7}$$

where $C_{\max}$ is the cost budget for the current GOP in analysis.

Two constraints must be met: 1) The total cost of the GOP should not exceed the total cost budget. 2) NAL units with smaller index in the sorted set $\mathcal{N}^*$ should be protected with stronger codes. To make the discussion in the following clearer, we denote the expected GOP utility $\sum_{q=1}^{Q} u_q p(t_q)$ as $U_{\text{GOP}}(\mathbf{t})$ and the total GOP cost $\sum_{q=1}^{Q} c_q r(t_q)$ as $C_{\text{GOP}}(\mathbf{t})$, respectively.

4 Solution Algorithm

The objective function (5) with the constraints (6) and (7) comprises a non-linear constraint optimization problem. To obtain the globally optimal solution, dynamic programming may be used. However, it is too complex for practical real-time applications. To keep the complexity low, we propose an improved Lagrangian relaxation method by introducing $\lambda > 0$, i.e.,

$$\hat{\mathbf{t}}^{\lambda} = \arg\max_{\mathbf{t}\in\mathbf{T}} \sum_{q=1}^{Q} u_q p(t_q) + \lambda \left(C_{\max} - \sum_{q=1}^{Q} c_q r(t_q) \right), \tag{8}$$

$$= \arg\max_{\mathbf{t}\in\mathbf{T}} \sum_{q=1}^{Q} u_q p(t_q) - \lambda \sum_{q=1}^{Q} c_q r(t_q) + \lambda C_{\max}. \tag{9}$$

Since $\lambda C_{\max}$ is a constant, we can remove it from (9). The objective function is then reduced to:

$$\hat{\mathbf{t}}^{\lambda} = \arg\max_{\mathbf{t}\in\mathbf{T}} \sum_{q=1}^{Q} u_q p(t_q) - \lambda \sum_{q=1}^{Q} c_q r(t_q). \tag{10}$$

To solve the problem in (10), we can decompose it into Q sub-problems for each NAL unit in $\mathcal{N}^*$ as follows:

$$\hat{t}_q^{\lambda} = \arg\max_{1\leq t_q \leq n+1} p(t_q) - \lambda_q r(t_q), \tag{11}$$

where $\lambda_q = \frac{c_q}{u_q}\lambda$.

We can find $\hat{t}_q^{\lambda}$ by checking all available protection modes. Actually, only points on the convex hull of channel coding effectiveness curve need to be checked. Furthermore, we notice that if the utility/cost ratio does not increase in the sorted set $\mathcal{N}^*$, i.e., $\frac{u_1}{c_1} \geq \ldots \frac{u_q}{c_q} \ldots \geq \frac{u_Q}{c_Q}$, the constraint (7) is guaranteed automatically after solving the sub-problems. However, since the costs of I picture and P picture are manifold higher than that of B pictures compared to their relative gain on utility against B pictures, the constraint (7 may not hold for the original utility/cost ratios. Therefore, we may need to increase the utility/cost ratios for NAL units of I and P pictures to make them no less than the the ratio of the NAL unit that comes immediately after them in the sorted set $\mathcal{N}^*$.

The Lagrangian relaxation algorithm alone cannot guarantee the identification of the globally optimal solution in some cases. While solving the problem (10) for different λ with Lagrangian relaxation, we are actually checking only the solutions on the upper convex hull of all possible solutions $\mathbf{T}$ in the space of $[U_{\mathrm{GOP}}\ C_{\mathrm{GOP}}]$. However, the best solution $\hat{\mathbf{t}}$ on the convex hull is not guaranteed to be the optimal solution $\bar{\mathbf{t}}$ for (5). Two cases are illustrated in Fig. 5. In Fig. 5(a), $\hat{\mathbf{t}}$ on the convex hull achieves obviously $U_{\max}$ with the cost budget $C_{\max}$. Therefore, it is the optimal solution $\bar{\mathbf{t}}$. In Fig. 5(b), $\hat{\mathbf{t}}$ can not utilize the cost budget sufficiently. Instead, the solution depicted as the red point, which is not on the convex hull, is the optimal solution. In this case, we can use $\hat{\mathbf{t}}$ as a solution seed and start a local-search algorithm to find out $\bar{\mathbf{t}}$.

As summarized in Algorithm 1, the solution algorithm consists of two stages. In the first stage (Line 2 to 16), we solve Problem (10) to get $\hat{\mathbf{t}}^{\lambda}$ for a given λ. Then we verify the validity of $\hat{\mathbf{t}}^{\lambda}$ with constraints (6) and (7). If the validity holds, $\hat{\mathbf{t}}^{\lambda}$ is a feasible solution for Problem (10). We need to update λ iteratively until the feasible solution converges or the maximal iteration limit is exceeded.

Let ΔC be the remaining channel bit rate after the first stage. The best solution $\bar{\mathbf{t}}$ that we get from this stage is the one minimizing ΔC. $\Delta C = 0$ indicates that $\bar{\mathbf{t}}$ is the globally optimal solution. If ΔC is significantly smaller than $C_{\max}$, $\bar{\mathbf{t}}$ should be a good solution although it may not be the globally optimal one. Otherwise, we need to refine the solution iteratively in the second stage as shown from Line 18 to 42.

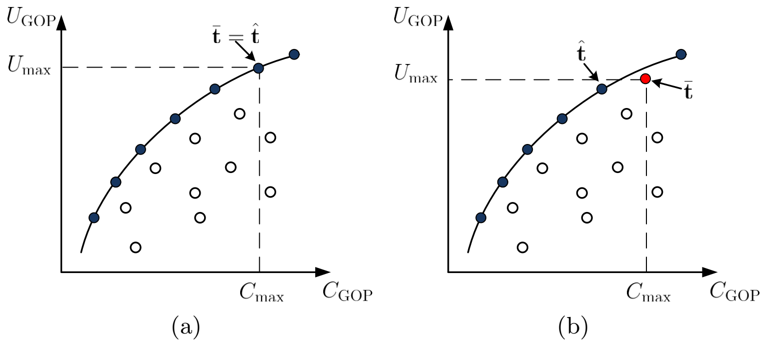

Fig. 5. Two cases of Lagrangian relaxation: (a) The optimal solution is on the convex hull; (b) The optimal solution is not on the convex hull

In this stage, the increase on the effective utility du_q and the transmission cost dc_q by incrementing the RS code index for all NAL units are first calculated. The set $\mathcal{D}$ shall contain the indices of the NAL units whose RS codes can not be strengthened any more due to the violation of the constraint (6). Inside the outer loop, $\mathcal{A}$ contains the indices of the NAL units whose RS codes can be potentially strengthened. Inside the inner loop, candidate NAL units that may receive a stronger RS code is searched for in $\mathcal{A}$, starting from the one with the largest ratio $uc_q = du_q/dc_q$ to that with lowest uc_q ratio. As soon as one proper candidate is found, its RS code is strengthened and its uc_q is updated. The outer loop will be broken, when ΔC lies within the defined tolerance or no candidate NAL units can be found in $\mathcal{A}$ without violating (6) and (7). Finally, the remaining ΔC is added to the $C_{\max}$ for the optimization process in the next GOP.

5 Experiments and Discussions

To verify the proposed algorithm, we carried out simulations with various sequences. Note that, the applicability of the proposed UEP scheme and solution algorithm is not constrained by the number of quality layers and NAL units in each quality layer. Actually, as long as the utilities of the NAL units in a GOP are determined, the optimal RS code for each NAL unit can be found with the proposed algorithm in Section 4. In the following, we will show some simulation results of the Foreman and Mobile (CIF@30 Hz, 300 frames) sequences. The bit streams, containing one BL ($QP_{\mathrm{BL}} = 36$) and one MGS EL ($QP_{\mathrm{EL}} = 30$), are coded with the SVC reference software (JSVM 9.14). Each GOP consists of 8 pictures with key pictures intra-coded.

Algorithm 1. Fast algorithm for searching the optimal solution $\bar{\mathbf{t}} = [\bar{t}_1 \ \bar{t}_2 \cdots \bar{t}_Q]$

Require: NAL units on different quality layers and temporal level are sorted according to their dependency and the utility-cost ratio; $C_{\max} = \frac{R_{\text{ch}}}{R_{\text{src}}} * \sum_{q=1}^{Q} c_q + \Delta C_{\text{lastGOP}}$.

Ensure: $|\frac{\Delta C}{C_{\max}}| < \varepsilon, \ \varepsilon > 0$.

Stage 1: Solve the Lagrangian optimization problem.
$\lambda_l \leftarrow 0, \ \lambda_r \leftarrow \lambda_{\max}, \ \lambda \leftarrow \frac{1}{2}(\lambda_l + \lambda_r)$
while $|\frac{\Delta C}{C_{\max}}| > \varepsilon$ **do**
  Solve (10) to get $\hat{\mathbf{t}}^{\lambda}$, $\Delta C = C_{\max} - \sum_{q=1}^{Q} c_q r_m(\hat{t}_q^{\lambda})$
  $iter = iter + 1$
  if $iter > iter_{\max}$ **then**
    break
  end if
  if $\Delta C < 0$ **then**
    $\lambda_l \leftarrow \lambda$
  else
    $\lambda_r \leftarrow \lambda$
    $\bar{\mathbf{t}} \leftarrow \hat{\mathbf{t}}^{\lambda}$
  end if
  $\lambda \leftarrow \frac{1}{2}(\lambda_l + \lambda_r)$
end while
Stage 2: Allocate the remaining channel bit rate.
if $|\frac{\Delta C}{C_{\max}}| > \varepsilon$ **then**
  $\mathcal{D} \leftarrow \emptyset$, $\mathcal{Q} \leftarrow \{1, \cdots, Q\}$
  $du_q \leftarrow u_q \left(p(\bar{t}_q + 1) - p(\bar{t}_q)\right), q \in \mathcal{Q}$
  $dc_q \leftarrow c_q \left(r(\bar{t}_q + 1) - r(\bar{t}_q)\right), q \in \mathcal{Q}$
  $uc_q \leftarrow du_q / dc_q, q \in \mathcal{Q}$
  repeat
    $\mathcal{A} \leftarrow \mathcal{Q} \backslash \mathcal{D}$
    while $\mathcal{A} \neq \emptyset$ **do**
      Find q with the largest uc_q in $\mathcal{A}$
      if $\bar{t}_q + 1 < \bar{t}_{q'}$ for any $q' < q$ **then**
        if $dc_q < \Delta C$ **then**
          $\bar{t}_q \leftarrow \bar{t}_q + 1$
          $\Delta C \leftarrow \Delta C - dc_q$
          Update the values of du_q, dc_q and uc_q
          break
        else
          $\mathcal{D} \leftarrow \mathcal{D} \cup \{q\}$
          $\mathcal{A} \leftarrow \mathcal{A} \backslash \{q\}$
        end if
      else
        $\mathcal{A} \leftarrow \mathcal{A} \backslash \{q\}$
      end if
    end while
  until $\mathcal{A} = \emptyset$ or $|\frac{\Delta C}{C_{\max}}| \leq \varepsilon$
end if

The channel transmission bit rate was set according to the bit rate of the video bit streams, such that $R_{\mathrm{src}}/R_{\mathrm{ch}} = 1/1.4$. In the different channel settings, we varied the packet loss rate π from 2% to 40% with a step size of 2%. The packet loss correlation factor ρ was set to 0 or 0.20, respectively, to emulate channels with random packet loss and burst packet loss effect. Error concealment algorithm is applied to frames, whose BL NAL units are lost. For the RS codes, we set $n = 63$. All 63 different RS codes plus the option that a NAL unit is not dropped by the sender, are used as available error protection modes. The result for each channel setting was evaluated with the averaged Y-PSNR value of all frames in the reconstructed video sequence. To make the averaged PSNR value statistically representative, 200 simulations in each experiment were conducted under different channel realizations for each channel setting.

First, we want to evaluate the performance of our algorithm (denoted as LagOpt) on finding the optimal error protection. We compared the results of LagOpt against the results achieved by exhaustive search (denoted as FullSearch) using differerent sequences and channel settings. FullSearch represents the globally optimal solution of error protection. In Fig. 6, we show the results on the Foreman SVC bit stream, which are quite representative for other bit streams. In both cases, i.e., $\rho = 0$ and $\rho = 0.20$, the LagOpt algorithm performs as well as the FullSearch algorithm, with much lower complexity. The average run time of the FullSearch algorithm for each GOP is about 1.30 s on MATLAB platform, while the LagOpt algorithm needs only about 0.023 s on average on the same platform. Therefore, the LagOpt is applicable for real-time video streaming (30 Hz).

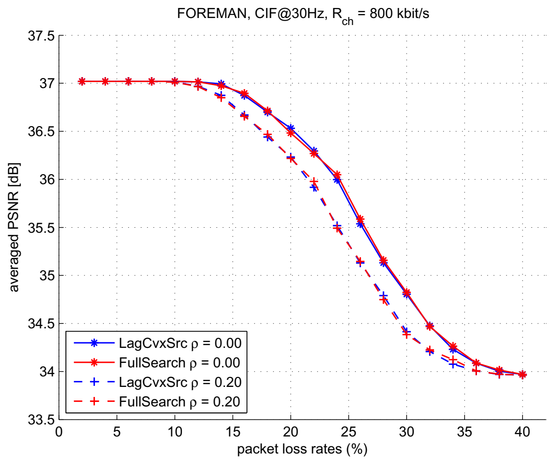

Fig. 6. Comparison of the UEP performance between the fast algorithm and the algorithm of exhaustive search on the SVC bit stream of Foreman

Fig. 7 shows the statistics of the received and lost NAL units for the Foreman SVC bit stream for $\rho = 0$. The bit stream contains 600 NAL units altogether. For $\pi \leq 12\%$, the available channel bit rate is still sufficient to protect all NAL units against the packet loss, therefore, all NAL units were correctly received and the video quality remains constantly at about 37 dB. With increasing π from 12%, the algorithm adjusts the protection scheme, tending to protect BL NAL units more strongly by reducing the protection for EL NAL units or drop part of them. The part of lost EL NAL units (in dark blue), which are transmitted by server but not received by client, is kept to a small amount. This reduces effectively waste of channel bit rate. As we can observe, there are no BL NAL units lost for the whole range of packet loss rates due to the adaptively determined error protecion modes. This guarantees an acceptable quality at high packet loss rate.

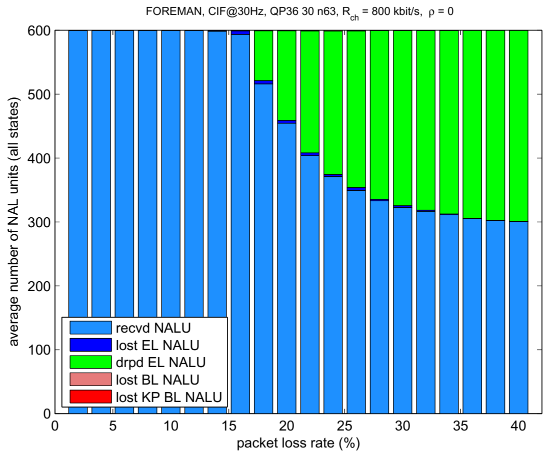

Fig. 7. Statistics of lost and received NAL units for the SVC bit stream of Foreman sequence. recvd NALU: NAL units are received; lost EL NALU: EL NAL units are lost; drpd EL NALU: EL NAL units are dropped at sender; lost BL NALU: BL NAL units are lost; lost KP BL NALU: BL NAL units belonging to key picture are lost.

Fig. 8 and Fig. 9 provides the simulation results for the Mobile SVC bit stream.

In Fig. 10, the proposed LagOpt algorithm was applied to two SVC and two SLC bit streams encoded with the Foreman sequence. The QP values are annotated in the plot. ρ was set to 0.20. As we can seen, SLC (QP 28) provides a very high quality ($\geq$ 38dB) at very low packet loss rates. However, its quality degrades rapidly for $\pi \geq 8\%$ because of frequent occurrence of lost frames. SLC (QP 30) provides good video quality up to $\pi = 20\%$. For $\pi \geq 20\%$, the reserve of the channel bit rate is not sufficient to protect SLC (QP 30) against loss of frame, resulting in rapid degradation of the video quality. On the other hand, both SVC bit streams provide a comparable quality to that of SLC (QP 30)

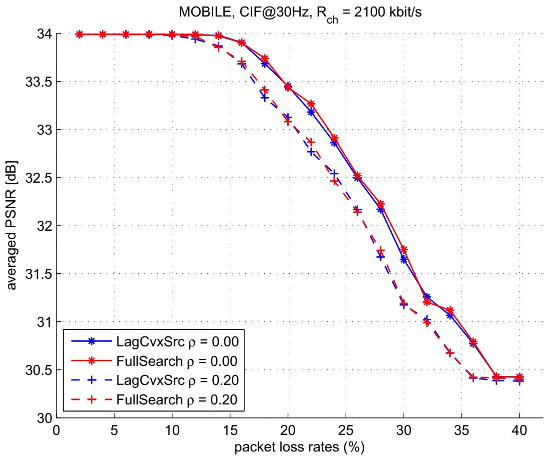

Fig. 8. Comparison of the UEP performance between the fast algorithm and the algorithm of exhaustive search on the SVC bit stream of Mobile

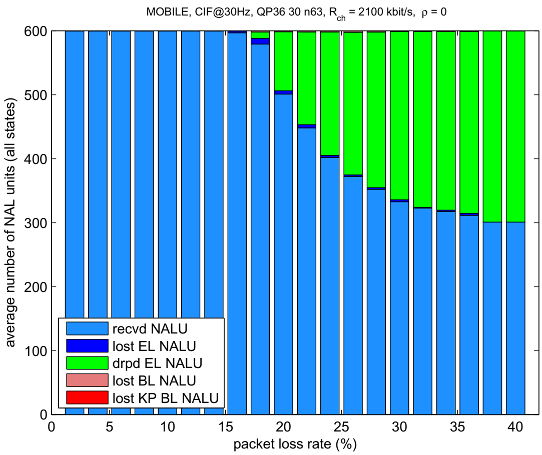

Fig. 9. Statistics of lost and received NAL units for the SVC bit stream of Mobile sequence

at low packet loss rates. For both SVC sequences, with increasing π, the server adaptively optimizes the error protection among different EL and BL NAL units on the GOP level. At high packet loss rates, the server starts to drop EL NAL units, starting from B pictures at high temporal levels, and spends more bit rates to protect the BL NAL units at lower temporal level. At extremely high packet loss rates, the entire ELs across the whole sequence are dropped and only the BL NAL units are sent. While SVC (QP 33 30) suffered slightly from frame loss for $30\% \leq \pi \leq 40\%$, SVC (QP 36 30) experiences no frame loss due to the lower

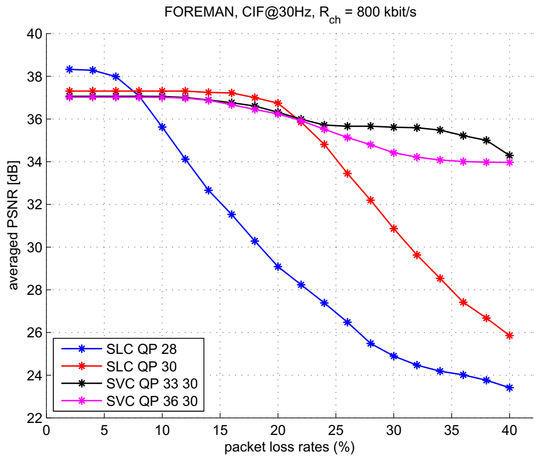

Fig. 10. Comparison of the fast algorithm applied to SVC and SLC bit streams

BL bit rate with $QP_{\mathrm{BL}} = 36$. The subjective impression for SVC (QP 36 30) at higher packet loss rates is better than that for SVC (QP 33 30), although the averaged PSNR values of SVC (QP 33 30) is higher for those packet loss rates.

6 Conclusions

In this paper, we proposed a framework based on H.264/AVC MGS SVC and UEP to provide adaptive robust video streaming for unicast applications. We proposed a method to determine the utility and cost of the NAL units in the hierarchical GOP structure. Based on that, we proposed an algorithm with low complexity to determine the optimized allocation of error protection modes to different NAL units on the GOP basis. The results achieved with our fast algorithm are very close to the results from exhaustive search. We applied our algorithm to SVC and SLC video bit streams. The simulation results demonstrated good performance with the combination of SVC bit stream and our algorithm in error-prone networks.

References

1. Wien, M., Schwarz, H., Oelbaum, T.: Performance Analysis of SVC. IEEE Trans. on Circuits Syst. Video Technol. 17(9), 1194–1203 (2007)
2. Riskin, E.A., Mohr, A.E., Ladner, R.E.: Approximately optimal assignment for unequal loss protection. In: Proc. IEEE Int. Conf. on Image Processing, Vancouver, Canada, vol. 1, pp. 367–370 (September 2000)
3. Goshi, J., Mohr, A.E., Ladner, R.E., Riskin, E.A., Lippman, A.: Unequal loss protection for H.263 compressed video. IEEE Trans. on Circuits Syst. Video Technol. 15(3), 412–419 (2005)

4. Fang, T., Chau, L.-P.: A novel unequal error protection approach for error resilient video transmission. In: Proc. IEEE Int. Symp. on Circuits and Systems, Kobe, Japan, vol. 4, pp. 4022–4025 (May 2005)
5. Baccaglini, E., Tillo, T., Olmo, G.: Concealment driven smart slice reordering for robust video transmission. In: Proc. IEEE Int. Conf. on Multimedia and Expo., Hannover, Germany, pp. 1173–1176 (June 2008)
6. Hellge, C., Schierl, T., Wiegand, T.: Receiver driven layered multicast with layer-aware forward error correction. In: Proc. IEEE Int. Conf. on Image Processing, San Diego, USA, pp. 2304–2307 (October 2008)
7. Nguyen, D.T., Hayashi, M., Ostermann, J.: Adaptive error protection for Scalable Video Coding extension of H.264/AVC. In: Proc. IEEE Int. Conf. on Multimedia and Expo., Hannover, Germany, pp. 417–420 (June 2008)
8. Maani, E., Katsaggelos, A.K.: Optimized bit extraction using distortion modeling in the scalable extension of H.264/AVC. IEEE Transactions on Image Processing 18(9), 2022–2029 (2009)
9. van der Schaar, M., Chou, P.A.: Multimedia over IP Wireless Networks, Compression, Networking and System. Academic Press (2007)
10. Albanese, A., Blomer, J., Edmonds, J., Luby, M.: Priority encoding transmission. IEEE Trans. Inf. Theory 42(6), 1737–1744 (1996)