

Layer-Aware Unequal Error Protection for Scalable H.264 Video Robust Transmission over Packet Lossy Networks

Yifei Sun, Xingjun Zhang, Feilong Tang, Scott Fowler, Huali Cui and Xiaoshe Dong

Linköping University Post Print

N.B.: When citing this work, cite the original article.

©2011 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Yifei Sun, Xingjun Zhang, Feilong Tang, Scott Fowler, Huali Cui and Xiaoshe Dong, Layer-Aware Unequal Error Protection for Scalable H.264 Video Robust Transmission over Packet Lossy Networks, 2011, Network-Based Information Systems (NBIS), 2011 14th IEEE International Conference on, 628-633.

<http://dx.doi.org/10.1109/NBiS.2011.106>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-74131>

Layer-aware Unequal Error Protection for Robust Scalable H.264 Video Transmission over Packet Lossy Networks

Yifei Sun*, Xingjun Zhang*, Feilong Tang[†], Scott Fowler[‡], Huali Cui*, Xiaoshe Dong*

*Department of Computer Science & Technology, Xi'an Jiaotong University, China

Email: xjzhang@mail.xjtu.edu.cn

[†]Department of Computer Science & Engineering, Shanghai Jiaotong University, China

Email: tang-fl@cs.sjtu.edu.cn

[‡]Department of Science & Technology, Linköping University, Sweden

Email: scott.fowler@liu.se

Abstract—The Scalable Video Coding (SVC) amendment of the H.264/AVC standard is an up-to-date video compression standard. The various scalable layers have different contribution to the quality of the reconstructed video sequence due to the use of hierarchical prediction and the drift propagation. This paper proposes a novel trapezoidal-unequal error protection (UEP) scheme which significantly reduces the redundancy but rarely decreases the performance by taking into account the characteristics of the video coding and the adoptive forward error correction (FEC) sufficiently. In order to optimally distribute FEC codes, the paper then proposes a layer-aware distortion model to accurately estimate the decrement of video quality caused by the loss of quality enhancement layers, drift propagation and error concealment in the scalable H.264/AVC video. Experimental results show that the proposed trapezoidal UEP scheme has better robustness and in the meanwhile reduces the coding redundancy greatly in different channel circumstance compared with the traditional UEP scheme.

Index Terms—channel coding, forward error correction, unequal error protection, scalable video coding

I. INTRODUCTION

Along with the development of network technology, there is more attention about the video application on the Internet due to the fact users are demanding the use of video. However, a majority of Internet is still working in the unstable networks due to the fluctuation of channel bandwidth, high bit-error rate as well as cyber intrusion [1] without the guarantee of Quality of Service (QoS), this presents on problem for video application since users expect high quality video. Therefore, how to provide robust, reliable, high effective and a scalable real-time video transmission in an unreliable network is a significant challenge. Recently, the newly developed Scalable Video Coding amendment (SVC) of the H.264/AVC standard [2], [3], [4] provides a superb coding efficiency, high bitrate adaptability, and low decoder complexity. Normally, the source video will be coded with a base layer and a group of enhancement layers with different levels of significance. The quality of the received video is greatly affected when packets are

missing due to hierarchical prediction and the drift propagation of H.264/SVC.

Adopting the UEP forward error correction (FEC) scheme makes full use of the different importance of each part in scalable video stream to reduce the error rates. In recent years, much research [5], [6], [7], [8], [9], [10], [11] has been performed on protecting different categories of layered video coding using an UEP scheme. For example, the paper [6] describes the MPEG-4 fine grange scalable (FGS) compressed video and adopts the unequal protection strategy by using rate-distortion information in each layer, which only considers the quality dimension in the SVC stream. In [8] and [9], the different importance between the temporal dimension and the quality dimension in the SVC stream is considered and then the two-dimensional protection scheme which uses the Reed-Solomon (RS) code to protect the video stream is proposed. However, the computational load of the distortion model in [8] is too high to compute in practice. And the distortion model and the performance metric proposed in [9] do not fully represent the distortion caused by loss of temporal layers and quality layers.

Fully considering of video coding characteristics as well as the FEC code characteristics, we propose a novel two-dimensional trapezoidal unequal error protection scheme to achieve the robust transmission of video stream. And then the optimal FEC assignment scheme can be formulated as the minimum distortion model solution problem under restrictive conditions of the available bandwidth and the different importance between temporal layers and quality layers in SVC stream. However, the distortion model and the performance metric in [9] are too simple and cannot fully indicate the reasons of drift distortion caused by the quality layers loss and error concealment brought by temporal layers loss, respectively. To overcome the aforementioned shortcomings, we present a model that considers the overall distortion caused by the enhancement layers truncation, propagation of drift and error concealment to accurately approximate the expected

distortion in each frame of one GOP. Then we use the genetic algorithm (GA) to solve the optimization problem and achieve a fast channel bit allocation.

The rest of this paper is organized as follows. In Section II, we provide an overview of the proposed 2-D trapezoidal UEP scheme and describe the problem formulation. In Section III, we present the method of calculating the expected distortion and the proposed UEP assignment algorithm. Experimental results are shown in Section IV and the final conclusions are drawn in Section V.

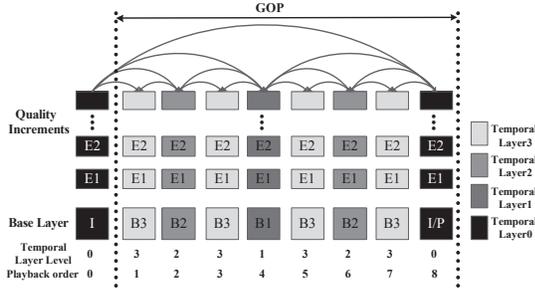


Fig. 1: Structure of a single resolution SVC bit stream for temporal and quality layers

II. PROPOSED 2-D TRAPEZOIDAL UEP SCHEME AND PROBLEM FORMULATION

A. System Model

The SVC extension of H.264/AVC supports spatial, temporal and quality scalabilities. The scalability adopts the layered approach. Each video sequence can be coded as a new layer and based on each layer the temporal scalability and quality scalability can be achieved. To improve the coding efficiency, inter-layer prediction technique is exploited. In the SVC extension of H.264/AVC, different with the previous standards, the temporal scalability is achieved through the coding structure of hierarchical B-pictures, and this coding method inherently provides the temporal scalability [12]. The enhancement layer picture is coded as hierarchical B-picture. The quality scalability may be regarded as the peculiar case when the picture's size is the same between enhancement layers and the base layer and can be realized by essential method of the spatial scalable code, which is also called the coarse grain scalability (CGS) code.

The division structure of the temporal layer and quality layer in the SVC stream is shown in Fig.1. The temporal scalability can be inherently realized by the hierarchical coding order on the temporal layer. The decode procedure is similar with the encode process, decode of the frame on the low temporal layer does not rely on the high temporal layer. By adding reference frames will improve the coding efficiency, however, this kind of improvement becomes smaller as the increase of the number of frames in one GOP.

As shown in Fig.1, each frame belongs to a particular temporal layer and each temporal layer contains several quality layers. The scalable unit (SU) is defined as each quality layer

in each frame of one GOP. If the frame is confirmed, and then its respective temporal layer is also determined.

Due to the hierarchical prediction structure and the dependence between the quality layers, packets loss may have serious impact on the reconstructed video quality. Thus, the effective robustness may be required to ensure that the video quality will not be excessive decline under the influence of unreliable channels in the video applications. In order to guarantee the reliability of SVC video data when it transmits in the error-prone channels, some measures need to be carried out for error control with the purpose of minimizing the influence of packets loss.

Forward error correction (FEC) is one kind of error correction technologies which is widely applied in communication system. It improves the robustness of the channels by adding redundant data. The receiver can recover the lost source data packets by using the redundant packets when the number of the loss packets does not surpass the error correction ability. As is well known, the larger the FEC code redundancy is, the more powerful its error correction ability is. Yet, the more redundancy means that the redundant data takes more bandwidth and lessens bandwidth utilization. Also, FEC may increase the delay because both the sender and receiver have to wait some time to execute FEC encoding and decoding respectively.

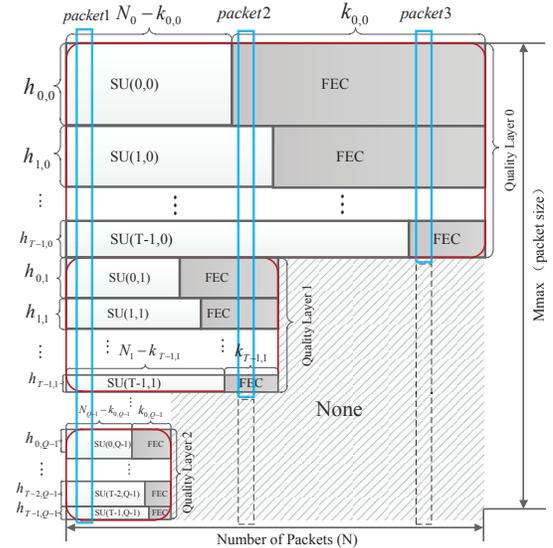


Fig. 2: Trapezoidal UEP scheme for scalable units of a GOP

B. Problem Formulation

On the basis of traditional rectangular UEP strategy [8], [9], a novel trapezoidal UEP strategy is proposed. Fig. 2 shows this novel UEP scheme for all scalable units. There are T temporal layers and each layer is divided into Q quality layers. The data of the quality base layer is the most important data and the quantity is larger than any other temporal layers and quality layers by means of the hierarchical video coding mechanism

and the forms as well as the contents of the source video. As the number of quality layers increases, the data quantity of each layer will decrease. In the most extreme situation, the quantity is less than 1% of quality base layer data. However the length of the FEC code source data have some restriction. Therefore, if continue to adopt the rectangular UEP strategy, we need to add massive data to source packets to perform the FEC code. These data are useless because they are not the parity packets arisen from the FEC code but only used to be padded to achieve the required length of source data for FEC. Thus, the network conditions may be more serious even the situation of more severe congestion and packets loss will happen.

The SVC supports up to 16 quality layers, but in the practical application, these quality enhancement layers may not be used totally. Therefore, we assume that there are 8 quality layers at most to be used. In this paper, famous Reed-Solomon code will be used as FEC. The Reed-Solomon code, i.e. RS(N, v) code, is one kind of FEC cyclic code which is defined in the Galois Field(2^m) [13]. From this we need to change v original data packages to $N - v$ redundant data packages by code transformation and guarantee that any v subset of data packets in v original data packets and $N - v$ redundant data packets, N data packets in all, could be recovered to the original data packages. Also, it permits for losing at most $N - v$ data packets at the transmission. Since the RS code is based on the GF(2^m) and the code length N is $2^m - 1$, the different Galois Field decides the length of RS code.

Upon this theory and the characteristics of video coding as well as RS code, the trapezoidal UEP strategy is proposed. The quality base layer uses RS code on GF(2^8) and the first quality enhancement layer uses RS code on GF(2^7), by parity of reasoning.

SU(t, q) is defined as the scalable video data in each unit (t, q), where t represents the temporal layer and its scope is from 0 to $T - 1$, similarly, q represents the quality layer and its scope is from 0 to $Q - 1$. The data length and height of FEC code are separately defined as $k_{t,q}$ and $h_{t,q}$ in each SU(t, q). $N_q - k_{t,q}$ is the length of the source data in each SU(t, q), N is the number of data packets in each SU(t, q) and M_{max} is the maximum length of data packets. Thus, the height of FEC code in each SU(t, q) can be calculated by the following formula.

$$h_{t,q} = \left\lceil \frac{R_{t,q}}{m(N_q - k_{t,q})} \right\rceil \quad (1)$$

where $R_{t,q}$ is the number of source data bits for SU(t, q) and m is the length of a symbol.

The two-state Markov model[14] is widely used to model packet loss rate on the Internet for its simplicity and mathematical tractability. The Markov model can be calculated by $P(i, N)$ which indicates the probability of losing i packets among N packets. If the quantity of lost packets is greater than the redundancy packets, the source data cannot be recovered according to the principle of FEC. Therefore P_u which

means the loss probability of each packet in SU(t, q) can be formulated as

$$P_u = 1 - \sum_{i=0}^k P(i, N) \quad (2)$$

where k represents redundancy packets added to recover the original data.

According to the trapezoidal UEP strategy in Fig.2, for the packets in a GOP, the different positions of the packets loss may have different impact on the various scalable units. For example, when packet1 is lost, each SU(t, q) has lost data packets; when packet2 is lost, only the quality base layer and the first quality enhancement layer have data packets loss, however the SU(t, q) of the afterward quality layers do not have packets loss, such as the black dotted portion illustrated in Fig.2. Similarly, when packet3 is lost, only quality base layer has packets loss.

Compared with the traditional rectangular UEP strategy, the trapezoidal UEP protection strategy makes the useless FEC code redundancy of different SU(t, q) decrease greatly in each GOP. The diagonal region in Fig.2 shows the reduced useless redundancy. The values of (2) will be used as a part of the distortion model to calculate the expected distortion. Although increasing the FEC code redundancy can enhance the robustness of video transmission, there has to be limit to the amount of redundancy otherwise it may lead to over saturation of the network bandwidth resulting in less effect network conditions. Therefore, the optimal FEC assignment strategy is to be found under constraints of various restrictions for the purpose of minimizing of overall distortion while reconstructing the video sequence at the different network conditions. Thus, a distortion model is required to estimate accurately the impact on video quality due to packets loss caused by poor network conditions. In this paper, the PSNR is used to measure the amount of the video quality distortion. The overall expected distortion is calculated as follows:

$$E\{\bar{D}_{overall}\} = \sum_{n=1}^F \{E\{\bar{D}_n|BL\} + E\{\bar{D}_n|nBL\}\} \quad (3)$$

where F represents the amount of frames in one GOP. $E\{\bar{D}_n|BL\}$ means the distortion of receiving the quality base layer and losing some enhancement layers. Also, $E\{\bar{D}_n|nBL\}$ means the distortion of performing error concealment after losing the quality base layer i.e. temporal layer. The specific computational process of distortion model will be introduced in details in the next section. The transform of FEC assignment strategies could lead to the change of loss probability of various quality layers in each frame, i.e. the different SU(t, q). Thus, whether the SU(t, q) lose or not may influence directly the expected distortion of reconstructed video sequence. Therefore, the key of minimizing expected distortion is to find out the optimal assignment matrix \mathbf{K} .

$$\mathbf{K} = \begin{bmatrix} K_{0,0} & K_{0,1} & \cdots & K_{0,Q-1} \\ K_{1,0} & K_{0,1} & \cdots & K_{1,Q-1} \\ \cdots & \cdots & \cdots & \cdots \\ K_{T-1,0} & K_{T-1,1} & \cdots & K_{T-1,Q-1} \end{bmatrix} \quad (4)$$

At present, the problem of finding the optimized FEC assignment strategy is formulized:

$$\text{Min } E \{ \text{Doverall}(\mathbf{K}) \} \quad (5)$$

Subject to

$$k_{t,q} \geq k_{t+1,q} \quad t = 0, 1, \dots, T-2 \quad (6)$$

$$k_{t,q} \geq k_{t,q+1} \quad t = 0, 1, \dots, Q-2 \quad (7)$$

$$\sum_{t=0}^{T-1} \sum_{q=0}^{Q-1} h_{t,q} \leq M_{max} \quad (8)$$

$$\sum_{q=0}^{Q-1} \sum_{t=0}^{T-1} h_{t,q} \cdot N_q \leq B_{tot} \quad (9)$$

As is known, the lower the temporal layers and the quality layers that errors happen in is, the greater the impact on the reconstructed video quality is. So, the constraints of (6) and (7) mean adding more protection to the lower temporal layers and quality layers. The constraint of (8) means that the maximum length of produced data packets cannot surpass the length of stipulation. In (9), B_{tot} is the total number assigned to one GOP, which means that the total number of information source and channel coding cannot surpass the total number of bits assigned.

III. EXPECTED DISTORTION CALCULATIONS AND PROPOSED UEP ASSIGNMENT ALGORITHM

In scalable video coding, the impact on the video quality is not the same whether the quality base layer is lost or not. There are two cases to be discussed which are receiving the quality basic layer and performing error concealment while the quality base layer is lost. Also, the distortion model needs to be solved in detail. The basic idea is to quantify the impact on error propagation caused by packets loss rather than to calculate the exact video distortion caused by several conditions. In this way, computational complexity could be significantly reduced and the data packets of different layers could be measured accurately and effectively relatively.

A. Frames with Decodable Base Layer

In order to increase the adaptability and robustness as well as improve the coding efficiency of stream which covers multi-coding rate, SVC introduces the variant of CGS which is called the medium-grained quality of scalable video code (MGS). Fig.3 shows MGS video coding strategy. The MGS code introduces the concept of the key picture which means the picture of the temporal base layer. The quality base layers of key frames are used to predict the subsequent quality base layers of key pictures. The quality enhancement layers of key pictures and the lower temporal layers are used to predict quality base layers of the higher temporal layers. Except for the key pictures, the motion compensated prediction of all frames uses the reference pictures of the highest available quality layers. Therefore, while the refinement packets lose, propagation of drift should be considered. D_n^d represents

the distortion caused by propagation of drift. The $D_n^e(q)$ represents the errors caused by enhancement layers loss, which means that additional degradation of video quality is to be brought since quality increments of the frame n is lost.

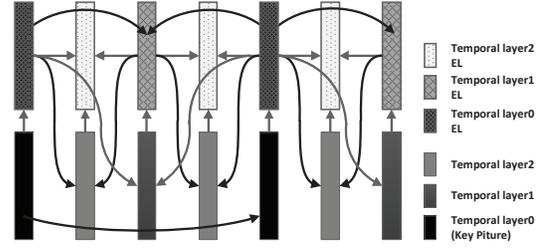


Fig. 3: The compromise between coding efficiency and drift

According to analysis above, for one GOP, the formula can be obtained:

$$E\{\bar{D}_n|BL\} = \sum_{q=1}^Q p_n^q D_n(q) \quad (10)$$

where p_n^q is calculated by (2), which means the unrecoverable probability of the q quality increments in frame n . $D_n(q)$ means the decrement of PSNR caused by propagation of drift and loss of enhancement layers while compared with the PSNR of video without packets loss. We suppose that $D_n(q)$ is the function of D_n^d and $D_n^e(q)$, i.e. $D_n(q) = F(D_n^d, D_n^e(q))$.

In a GOP, the decrement of PSNR caused by the loss of certain quality layers in a frame could be calculated by decoding the video sequences with some corresponding packets losing, also, it could be obtained by analyzing and calculating accurately in the process of video coding and decoding. However, considering that the computational procedure is extremely complicated and tedious, these two methods are not suitable. In order to overcome the complexity of computation, based on the performance metric proposed in [9], we propose a method to calculate simply and effectively the decrement of PSNR due to the packets loss.

According to the dependency among quality layers, the lower level quality layers which lose packets have a more serious impact on the reconstructed video quality. This is due to the fact that the contents of higher quality layers are dependent on lower quality layers. From the standpoint of improvement of PSNR, the quality base layer will improve the PSNR value more greatly than the quality enhancement layers. The data packets used in this paper satisfy directed acyclic graph (DGA) model proposed in [15], that is to say, in the condition of losing data packets of low quality layers, even if the data packets of high quality layers are received, they would be treated as useless packets and then discarded.

$D_n^e(q)$ can be modelled as the inverse function relation to the level of quality layers. It can be formulized as $(1+q)^{-C_2}$, where q represents the level of quality layer; C_2 represents the scale factor related to the complexity of video encoding.

As is known SVC uses the MGS code, and then all the frames except for the key frame use the highest available

quality layer for motion compensated prediction. Thus, the loss of data in quality enhancement layer may lead to the drift propagation of data packets which are predicted by it. From the video predictive coding structure shown in Fig.1, if the loss of quality enhancement layer happens in the picture B2 in temporal layer 2, it may impact the $2^{4-2} - 2 = 2$ pictures predicted by it, these pictures are two neighbor pictures ($B3$). And then, the following formula can be obtained

$$E\{D_n|BL\} = \sum_{q=1}^Q p_n^q \frac{1}{(1+q)^{C_2}} (2^{(T-C_1t)} - 2) \quad (11)$$

where T represents the total number of temporal layers; t is the temporal layer level of the frame n ; q is the level of the lost quality layer and C_1 is the factor related to the propagation of drift.

B. Frames without Base Layer

The base quality NAL units may be lost at the transmission or be damaged in the channel and therefore become unavailable to the decoder. In this case, all subsequent NAL units belong to this frame will be given up at the decoder side and perform the error concealment. In order to confirm the impact on the overall quality of the video sequence when the frame is lost, the distortion value of frame loss after error concealment needs to be calculated.

In practice, the decoder usually has a specific error concealment strategy. For example, Frame Copy (FC or Picture Copy, PC) is one of the most common error concealment strategies. Under FC concealment strategy, each lost frame is replaced by the previous temporal picture in the higher hierarchical level [16]. So the frame loss in different position of the hierarchical structure will introduce different concealment distortion, which depends on the scalable level the lost frame belongs to (mainly the temporal layer level). For instance, if the frame ($B1$) in Fig.1 is lost, then frame I will substitute ($B1$) to conceal the error. Obviously, all the B-picture in this GOP will suffer from quality degradation. If i indicates the frame which is used to perform the error concealment, then the distortion of frame n after error concealment can be expressed as $D_{n,i}^{con}$. To do this the formula is as follows:

$$E\{\bar{D}_n|nBL\} = p_n^0 D_{n,i}^{con} \quad (12)$$

where p_n^0 means the probability of losing the quality base layer in frame n , i.e. probability of frame n loss. Through analyzing the error concealment strategy and coding method, the conclusion can be drawn that when the frame is lost and the error concealment is performed, it will impact on the current frame and simultaneously it may impact on the predicted frames due to the hierarchical video coding structure. As a result, $D_{n,i}^{con}$ can be calculated approximately by the following:

$$D_{n,i}^{con} = (\mu + \nu D_i)(2^{(T-C_3t)} - 2) \quad (13)$$

where D_i is the distortion value of the frame which are used to perform error concealment; and t is the temporal layer of the frame n ; μ , ν and C_3 is the scaling factor related with the

error concealment. Then, the formula can be obtained through discussed before:

$$D_{n,i}^{con} = (\mu + \nu E\{\bar{D}_i|BL\})(2^{(T-C_3t)} - 2) \quad (14)$$

Therefore, through (11) and (14), the value of expected overall quality distortion can be calculated in the case of packets loss, i.e. the value of formula (3).

From the distortion model, the method of exhaustion to enumerate all the possible solutions and find the optimal solution is unacceptable because the scale of the problem is greatly vast. Then, genetic algorithm (GA) is a suitable method to solve addresses the large solutions space.

IV. EXPERIMENTAL RESULTS

In this section, we will show the experimental results of the performance of the proposed trapezoidal UEP allocation algorithm using the expected distortion model. In the experiment, we test two QCIF sequences: Bus with 75 frames and Football with 130 frames, which have the same frequency of 15Hz and are encoded by the JSVM codec tools. One spatial layer is encoded to one quality base layer and two quality enhancement layers. Quality enhancement layers are encoded by MGS coding. With the encoding condition of key pictures with GOP size of 8 frames, there are four temporal layers. Table I represents the scaling factors of different video sequences respectively.

TABLE I: SALING FACTORS OF DIFFERENT VIDEO SEQUENCES

<i>YUV sequence</i>	C_1	C_2	C_3	μ	ν
Football	0.36	4.03	0.758	0.09	0.02
Bus	0.35	3.61	0.760	0.17	0.01

The performance of our proposed scheme is compared with four other schemes: 1) fixed unequal protection ration on the 2-D units without considering different packet loss rates; 2) UEP implements in temporal layers without considering the different importance of quality layers; 3) UEP carries out on quality layers without considering the different importance of the temporal layers; 4) UEP applies on 2-D units with performance metric LW-EZEP. The comparison result is showed in Fig.4 as a function of packet loss rates. Owing to the randomness of such a channel, 100 different runs of each video transmission experiment are executed with different original PLRs from 5% to 30%. To show the performance results, we use two sequences as mentioned above.

The performance of the proposed UEP scheme in Fig.4 shows obvious superiority over other four UEP methods. When the packet loss rate is 5%, the declining extent of PSNR is not very obvious, and no matter which kind of UEP methods is adopted, the PSNR of reconstructed video is closer. This is because when the packet loss rate is lower the fewer packets are lost at the transmission and various kinds of UEP methods protect the quality base layer. Therefore, the quality base layer almost does not lose packets, and the

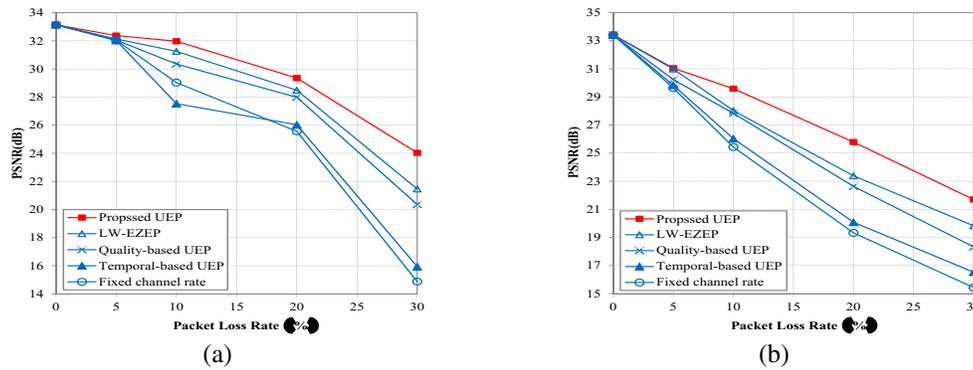


Fig. 4: Average PSNR comparison with five UEP schemes on different video sequences: (a) bus (b)football.

loss of quality enhancement layers may be insignificant for the quality of video. However, as the increase of packet loss rate, the declining of PSNR values becomes more evidently. When the PLR is larger than 10%, the PSNR values drop obviously for fixed channel coding rate and temporal-based UEP method while the proposed scheme still gives graceful degradation of PSNR in the packet loss range from 10% to 30%. And we can observe that the proposed UEP scheme still gives up to 1.87dB improvements for test sequences at higher packet loss rate compared to the UEP method which uses the performance metric LW-EZEP. The reason of this phenomenon is that the evaluation of importance of scalable units will be more accurate by using the modified expected distortion model and trapezoidal UEP scheme. And then, the quality base layer will be added more redundancy and the number of performing error concealment may be reduced. As is known, the impact on the error concealment for the video quality is greater than the impact caused by the quality enhancement layers loss.

V. CONCLUSIONS

This paper proposes a novel trapezoidal unequal protection scheme for H.264 SVC at video transmission over packet-lossy networks. This scheme shows reduction of the channel code redundancy at the transmission under the premise of unchanging the error correction ability and prevents the degradation of network environment caused by adding more redundancy. In order to find the optimal assignment of FEC code, we present a more comprehensive and accurate distortion model by considering the factors of the enhancement layer truncation, drift propagation and error concealment simultaneously. Afterward, the GA is applied to solve the problem of optimal allocation in the distortion model. Experimental results show that the proposed UEP scheme obtains better robustness and flexibility compared to other methods.

VI. ACKNOWLEDGMENTS

This work was supported by the State 863 project of China (Grant No. 2009AA01A135), the NSFC projects (Grant Nos. 61073148 and 60773089), and the Natural Science Foundation of Shaanxi Province, China (Grant No. 2009JM8002-5). Part of the work was supported by the XJTU multi-disciplinary project under grant No. 2009xjtujc30.

REFERENCES

- [1] C. Fung, "Collaborative intrusion detection networks and insider attacks," *Journal of Wireless Mobile Networks, Ubiquitous Computing and Dependable Applications*, vol. 2, pp. 63–74, 2011.
- [2] ISO/IEC, "Pdam 19 reference software for svc, joint video team (jvt) of iso-iec mpeg & itu-t vceg," ISO/IEC, 8652(E) Ed. 3, Annex E Distributed Systems, 2001.
- [3] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the h.264/avc standard," *IEEE Transactions on Information Theory*, vol. 17, pp. 1103–1120, 2007.
- [4] ISO/IEC, "Fdam 3 scalable video coding, joint video team (jvt) of iso-iec mpeg & itu-t vceg," ISO/IEC, 8652(E) Ed. 3, Annex E Distributed Systems, 2005.
- [5] A. Albanese, J. Blomer, J. Edmonds, and M. Luby, "Priority encoding transmission," *IEEE transactions on Information Theory*, vol. 42, pp. 1737–1744, 1996.
- [6] L. Cheng, W. Zhang, and L. Chen, "Rate-distortion optimized unequal loss protection for fgs compressed video," *IEEE Trans. Broadcasting*, vol. 50, pp. 126–131, 2004.
- [7] X. Yu, J. W. Modestino, R. Kurceren, and Y. S. Chan, "A model-based approach to evaluation of the efficacy of FEC coding in combating network packet losses," *IEEE/ACM Trans. Netw.*, vol. 16, no. 3, pp. 628–641, 2008.
- [8] Y. Wang, T. Fang, L. P. Chau, and K. H. Yap, "Two-dimensional channel coding scheme for MCTF-based scalable video coding," *IEEE Transactions on Multimedia*, vol. 9, no. 1, pp. 37–45, Jan. 2007.
- [9] H. Ha and C. Yim, "Layer-weighted unequal error protection for scalable video coding extension of h.264/avc," *IEEE Transactions on Consumer Electronics*, vol. 2, pp. 736–744, 2008.
- [10] J. Liu, Y. Cho, Z. Guo, and C.-C. J. Kuo, "Bit allocation for spatial scalability coding of H.264/SVC with dependent rate-distortion analysis," *IEEE Trans. Circuits Syst. Video Techn.*, vol. 20, no. 7, pp. 967–981, 2010.
- [11] X. Zhang, X. Peng, S. Fowler, and D. Wu, "Robust h.264/avc video transmission using data partitioning and unequal loss protection," in *CIT*. IEEE Computer Society, 2010, pp. 2471–2477. [Online]. Available: <http://doi.ieeecomputersociety.org/10.1109/CIT.2010.423>
- [12] H. Schwarz, D. Marpe, and T. Wiegand, "Analysis of hierarchical b pictures and mctf," in *IEEE Int. Conf. Multimedia Expo*, vol. 17, no. 2, 2006, pp. 1929–1932.
- [13] S. Gravano, *Introduction to Error Control Codes*. Oxford University Press, 2001.
- [14] X. Yang, C. Zhu, Z. Li, X. Lin, and N. Ling, "An unequal packet loss resilience scheme for video over the internet," *IEEE Transactions on Multimedia*, vol. 7, no. 4, pp. 753–765, 2005.
- [15] P. A. Chou and Z. Miao, "Rate-distortion optimized streaming of packetized media," *IEEE Transactions on Multimedia*, vol. 8, no. 2, pp. 390–404, 2006.
- [16] H. Mansour, P. Nasiopoulos, and V. Krishnamurthy, "Modeling of loss-distortion in hierarchical prediction codecs," *ISSPIT*, vol. 1, pp. 536–540, 2006.