

# Error Resilient Video Coding Using Unequally Protected Key Pictures

Ye-Kui Wang<sup>1</sup>, Miska M. Hannuksela<sup>2</sup>, and Moncef Gabbouj<sup>3</sup>

<sup>1</sup> Nokia Mobile Software, Tampere, Finland

<sup>2</sup> Nokia Research Center, Tampere, Finland

<sup>3</sup> Tampere University of Technology, Finland

**Abstract.** This paper proposes the use of unequally protected key pictures to prevent temporal error propagation in error-prone video communications. The key picture may either be an intra-coded picture or a picture using a long-term motion compensation reference picture through reference picture selection. The inter-coded key picture uses previous key pictures as motion compensation reference. Key pictures are better protected than other pictures using forward error correction in either source or transport coding. Simulation results show significantly improved error resiliency performance of the proposed technique compared to conventional methods.

## 1 Introduction

In order to gain maximal compression efficiency, video codecs make use of predictive coding. However, predictive coding is vulnerable to transmission errors, since an error affects all pictures that appear in the prediction chain after the erroneous position. Therefore, a typical way to make a video transmission system more robust against transmission errors is to weaken the prediction chains. Methods to weaken prediction chains include insertion of intra pictures or macroblocks (MBs) [1][2], video redundancy coding [3], reference picture selection (RPS) based on feedback information [1], and intra picture postponement [4].

Insertion of intra MBs can avoid coding of large-sized intra pictures. However, because of the remaining inter-coded MBs for each picture, temporal error propagation cannot be completely prevented. RPS was adopted as an interactive error resilient coding tool in ITU-T H.263 and ISO/IEC MPEG-4 part 2. To apply RPS, the decoder side transmits information about corrupted decoded areas and/or transport packets to the encoder side. The communication system must include a mechanism to convey such feedback information. After receiving the feedback information, the encoder encodes the subsequent picture using appropriately selected reference pictures that are available in both sides. However, since feedback information cannot be used in many applications, such as broadcast or multicast to a huge number of receivers, the use of RPS is limited.

Another way to improve error resilience is to use unequal error protection (UEP). UEP refers to techniques that protect part of the transmitted bit-stream better than the rest. In order to apply unequal error protection, video bit-streams have to be organized in portions of different importance in terms of user experience in visual quality. Techniques achieving this goal include data partitioning [1] and scalable video coding [5]. UEP can be done in either transport coding. Examples of applicable UEP techniques in transport coding include application-layer selective retransmission [6], forward error correction (FEC, e.g. RFC 2733 [7]), guaranteed network Quality of Service (e.g. QoS architecture of Universal Mobile Telecommunication System [8]), and Differentiated Services (DiffServ) [9]. Examples of UEP methods in source coding include repetition, FEC or adding other redundancies in portions of the bit-stream.

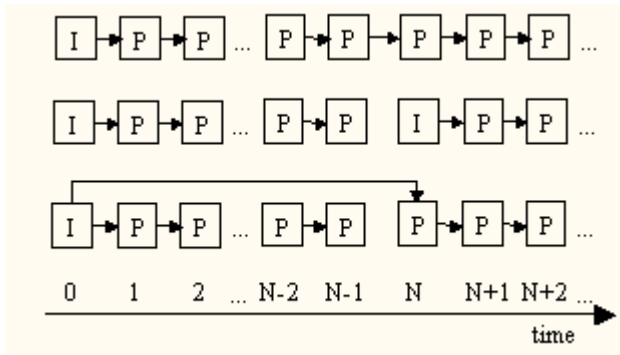
In this paper, we propose a novel error resilient video coding method combining insertion of intra pictures, use of RPS and application of UEP. A concept of key picture, which may be an intra-coded picture or a picture using long-term motion compensation references, is introduced. By correctly decoding a key picture, temporal error propagation from earlier coded pictures may be completely prevented even when some previous pictures contain errors. Therefore, key pictures are more important than other pictures. To improve the probability of completely preventing temporal error propagation, key pictures are better protected than other pictures. For inter-coded key pictures, the motion compensation reference can only be earlier coded key pictures. The novelty of the method lies in two aspects. An idea using RPS with UEP and data partitioning was proposed in [10]. The main difference between our proposal and [10] is that in our method key pictures can also be intra-coded. Furthermore, we propose to consider the underlying UEP method in encoding when loss-aware MB mode selection is in use.

The paper is organized as follows. Mathematical analysis of temporal error propagation is given in Section 2. The scheme of unequally error protected key pictures is presented in Section 3. Section 4 describes the method of applying loss-aware MB mode selection when UEP is in use. Simulation results comparing the proposed method with conventional error resilient video coding methods are shown in Section 5. Finally conclusions and discussion are given in Section 6.

## 2 Analysis of Temporal Error Propagation

In order to introduce the concept of key pictures, we first give an analysis on temporal error propagation in video communications. Fig. 1 shows three coding schemes for real-time low-latency applications under error-prone transmission channels. The top one is the normal predictive coding (denoted as the normal case), where the very first picture is intra coded and each subsequent picture is inter-coded with the previous coded picture as its motion compensation reference. To obtain acceptable image quality in the decoder side, insertion of intra MBs, e.g. the loss-aware rate-distortion optimized MB mode selection (LA-RDO) [2], should always be applied in this coding scheme. The middle (denoted as the intra case) and bottom (denoted as the RPS case)

schemes make use of intra picture insertion and RPS, respectively. In error-prone transmission, intra MB insertion should also be applied in these two coding schemes.



**Fig. 1.** Video coding schemes using normal predictive coding (top), insertion of intra pictures (middle), and reference picture selection (bottom). Arrows indicate motion compensation prediction relationships.

Without loss of generality, assume that the probability of a picture being hit by transmission error is  $p_1$ , and transmission errors of different pictures are independent. In the normal case, the probability of the  $N^{\text{th}}$  picture being hit by temporal error propagation from previously coded pictures is

$$p_{\text{normal}} = 1 - (1 - p_1)^N$$

In the intra and the RPS cases, the probabilities are

$$p_{\text{intra}} = 0$$

$$p_{\text{RPS}} = 1 - (1 - p_1)^1 = p_1$$

We can see that for  $N > 1$ ,  $p_{\text{normal}} > p_{\text{RPS}} > p_{\text{intra}}$ . In other words, the normal case has the worst performance in preventing temporal error propagation; while the intra case is the best (it can prevent temporal error propagation completely).

When transmission error occurred, an error concealment process is usually applied. In the intra case, if previously decoded pictures are utilized in the error concealment process, temporal error propagation will occur; if only spatial information is utilized, the quality will normally be much worse than utilizing also previously decoded pictures [11]. Considering the above reason for all the three coding schemes, the probabilities of picture  $N$  being hit by temporal error propagation from previously coded pictures are

$$p_{\text{normal}} = 1 - (1 - p_1)^{N+1}$$

$$p_{\text{intra}} = 1 - (1 - p_1)^1 = p_1$$

$$p_{\text{RPS}} = 1 - (1 - p_1)^2$$

If the coding scheme as shown in Fig.1 is periodical with period  $N$ , then each inter-coded key picture uses previously decoded key pictures as its motion compensation reference. In this case, the three probabilities of picture  $N \times k$  are

$$p_{\text{normal}}(k) = 1 - (1 - p_1)^{N \times k + 1}$$

$$p_{\text{intra}}(k) = p_1$$

$$p_{\text{RPS}}(k) = 1 - (1 - p_1)^{k \times 1}$$

Another factor should be considered is that intra-coded picture often has a larger size in bits than inter-coded picture. The larger size generally results into more packets for transport, which means higher error probability. Therefore, in practice,  $p_{\text{intra}}(k)$  and  $p_{\text{RPS}}(k)$  is closer than as shown above.

### 3 Key Picture with Unequal Error Protection

According to the analysis in previous section, the probabilities of completely preventing temporal error propagation in picture  $N \times k$  in Fig.1 for the normal, intra and RPS coding schemes are  $1 - p_{\text{normal}}(k)$ ,  $1 - p_{\text{intra}}(k)$  and  $1 - p_{\text{RPS}}(k)$ , respectively. If  $N=15$ ,  $k=5$  and  $p_1=5\%$ , then the values of the probability are about 2%, 95% and 74%, respectively. Therefore it is likely that temporal error propagation is stopped after decoding of the key picture in the intra and RPS cases, while almost impossible in the normal case.

Herein we define an intra-coded picture or an inter-coded picture using long-term motion compensation references through RPS as a key picture. By correctly decoding a key picture, temporal error propagation from earlier coded pictures may be completely prevented even when some previous pictures contain errors. Obviously key pictures are more important than any of other pictures in terms of preventing temporal error propagation. Therefore, it should be beneficial to apply better error protection, in either the source coding or the transport layer, for key pictures than other pictures in error-prone transmission. For example, each key picture can be repeated in source or transport coding. Key picture repetition in source coding level can be implemented using the so-called sync frames in ITU-T H.263 or the redundant slices feature in ITU-T H.264 (also know as MPEG-4 part 10 or AVC). In this case, the probabilities of picture  $N \times k$  being hit by temporal error propagation from previously coded pictures for the intra and RPS coding schemes are

$$p_{\text{intra}}(k) = p_1^2$$

$$p_{\text{RPS}}(k) = 1 - (1 - p_1^2)^{k+1}$$

If  $N=15$ ,  $k=5$  and  $p_1=5\%$ , then the probabilities of completely preventing temporal error propagation in picture  $N \times k$  for the two cases are 99.8% and 98.5%, respectively.

An intra-coded key picture is more likely to completely preventing temporal error propagation, while an inter-coded key picture generally has a smaller size in bits. In

practice, the two kinds of key pictures can be switched adaptively. For each key picture, the coding method resulting in better rate-distortion performance should be selected, taking into account the repetition or retransmission.

## 4 Loss-Aware MB Mode Selection under Unequal Error Protection

As stated earlier, insertion of intra MBs should be applied in error-prone transmission whatever the coding scheme is in use. Loss-aware MB mode selection is one of such methods. During encoding, the known or estimated error rate is utilized to help finding the best coding mode for each MB. However, when UEP is used, the well-protected parts would naturally have lower error rates than other parts. This factor should be considered in encoding when Loss-aware MB mode selection is in use.

We propose that the error rate of the well-protected picture or region is calculated according to the applied UEP method. For example, if repetition is used, then the final error rate is the square of the original value. The final error rate should be used instead of the original value in selection the coding mode of each MB in well-protected region.

## 5 Simulation Results

Five coding schemes were compared: 1) the normal case, 2) the intra case, 3) the intra case with UEP, 4) the RPS case and 5) the RPS case with UEP. The simulated UEP is repetition in source coding or in transportation. The target application is real-time multicast in Internet with a large number of receivers, where feedback should not be used. The simulations were based on the JVT joint model version 2.0b [12]. The common test conditions for packet-lossy environments [13] were applied. For the cases other than the 1st one, the key picture period is 1 second. LA-RDO [2] was applied in all the coding schemes with use of the idea described in Section 4. To produce best quality for a majority of the receivers, the assumed packet loss rate was 5%. The bit-streams were then decoded under packet loss rates of 0, 3, 5, 10 and 20 percent, respectively.

300-400 pictures of each designated sequence were used, to ensure that at least 100 pictures are encoded when the frame rate is at least 7.5 frames per second (fps). The slicing method was to let each coded slice has a maximum size of 1400 bytes. Each slice was encapsulated into one packet. We assumed that the packet containing parameter sets [14] is conveyed reliably (possibly out-of-band during the session setup), and therefore no error pattern was read from the error pattern file [13] for it.

The coded bitstream was decoded 20 times (each time is called a decoding run). The beginning loss position of the  $n+1$ th run continuously follows the ending loss position of the  $n$ th run. The overall average PSNR was obtained by averaging the average PSNR values of all decoding runs. The representative decoding run was se-

lected so that its average PSNR was the closest to the overall average PSNR. The decoded sequence of the representative run was stored for subjective quality evaluation.



**Fig. 2.** Snapshots of Foreman@144kbps, 7.5 fps. From top to bottom, the applied coding schemes are “normal”, “intra-1” and “RPS-1”, respectively.

Tables 1 and 2 show the average Y-PSNR values under different packet loss rates for Foreman@144kbps (7.5 fps) and News@144kbps (15 fps), respectively, where intra-1 represents the intra case with UEP and RPS-1 represents the RPS case with UEP. The largest PSNR value in each column is shown in italic bold font. Some typical snapshots under packet loss rate of 5% are shown in Fig.2.

From the results shown in Tables 1 and 2, the follow conclusions can be drawn: 1) Applying UEP can improve error resilience of both the intra and the RPS coding schemes. 2) The two cases with UEP outperform the normal cases. 3) For Foreman theintra case with UEP performs best while for News the RPS case with UEP per

**Table 1.** Average Y-PSNR (dB) of Foreman@144kbps under different packet loss rates

Coding Scheme	0%	3%	5%	10%	20%
normal	27.02	26.45	25.94	24.95	23.62
intra	<b>27.07</b>	<b>26.57</b>	26.08	25.15	23.85
intra-1	26.87	26.47	<b>26.13</b>	<b>25.36</b>	<b>24.17</b>
RPS	27.03	26.26	25.82	24.92	23.14
RPS-1	26.81	26.36	26.06	25.07	23.72

**Table 2.** Average Y-PSNR (dB) of News@144kbps under different packet loss rates

Coding Scheme	0%	3%	5%	10%	20%
normal	<b>35.85</b>	34.81	33.24	32.46	29.68
intra	34.47	33.79	33.46	32.36	30.76
intra-1	34.33	33.92	33.47	32.62	30.91
RPS	35.39	33.90	33.78	32.31	<b>28.99</b>
RPS-1	35.69	<b>35.21</b>	<b>34.00</b>	<b>33.28</b>	30.35

forms best. As shown by the snapshots, both of the two cases with UEP have perceivable subjective quality improvement over the normal coding scheme. These results verify the effectiveness of the proposed technique.

## 6 Conclusions and Discussion

A novel error resilient video coding technique combining insertion of intra pictures, use of reference picture selection and application of unequal error protection is proposed in this paper. The intra-coded pictures or the inter-coded pictures using long-term motion compensation prediction references are called key pictures and protected better than other pictures. Improved error resilience is shown by the simulation results. It should be pointed out that if adaptive switching between the two coding modes of key pictures were carried out, the error resiliency performance could be further improved. This is a prospective direction for future investigation.

## References

1. Y. Wang, S. Wenger, J. Wen and A. K. Katsaggelos, "Error resilient video coding techniques," *IEEE Signal Processing Magazine*, vol. 17, no. 4, pp.61–82, July 2000
2. T. Stockhammer, D. Kontopodis and T. Wiegand, "Rate-distortion optimization for H.26L video coding in packet loss environment," 12<sup>th</sup> International Packet Video Workshop (PV 2002), Pittsburg, PY, May 2002
3. S. Wenger, G. Knorr, Jörg Ott, and F. Kossentini, "Error resilience support in H.263+," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 7, pp. 867–877, Nov. 1998

4. M. M. Hannuksela, "Simple packet loss recovery method for video streaming", International Packet Video Workshop PV2001, Kyongju, South Korea, 30 April – 1 May 2001
5. W. Li, "Overview of fine granularity scalability in MPEG-4 video standard", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 3, pp. 301–317, Mar. 2001
6. J. Rey, D. Leon, A. Miyazaki, V. Varsa, R. Hakenberg, "RTP retransmission payload format," IETF Internet Draft draft-ietf-avt-rtp-retransmission-05.txt, February 2003
7. J. Rosenberg, H. Schulzrinne, "An RTP Payload Format for Generic Forward Error Correction" IETF RFC 2733, December 1999
8. 3GPP TS 23.107, "QoS Concept and Architecture (Release 4)", V4.6.0, December 2002, available [ftp://ftp.3gpp.org/specs/2002-12/Rel-4/23\\_series/23107-460.zip](ftp://ftp.3gpp.org/specs/2002-12/Rel-4/23_series/23107-460.zip)
9. S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services", IETF RFC 2475, December 1998
10. Y. Wang and M. D. Srinath, "Error resilient video coding with tree structure motion compensation and data partitioning," 12<sup>th</sup> International Packet Video Workshop (PV 2002), Pittsburg, PY, May 2002
11. Y.-K. Wang and M. M. Hannuskela, "Signaling of shot changes," Joint Video Team document JVT-D099, July 2002
12. ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC Draft Text, Joint Video Team document JVT-E133d37, Nov. 2002
13. S. Wenger, "Common conditions for wire-line, low delay IP/UDP/RTP packet loss resilient testing," ITU-T VCEG document VCEG-N79r1, Sep. 2001
14. ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC Joint Model, version JM-2.0b, May 2002