We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Error Resilient H.264 Video Encoder with Lagrange Multiplier Optimization Based on Channel Situation

Jian Feng, Yu Chen, Kwok-Tung Lo and
Xu-Dong Zhang

Additional information is available at the end of the chapter

1. Introduction

Robust delivery of compressed video in wireless packet-switched networks is still a challenging problem. Video packets transmitted in wireless environments are often corrupted by random and burst channel error due to multi-path fading, shadowing, noise disturbance, and congestion in physical wireless channel.

To achieve an optimum transmission over a noisy wireless channel, both the source coding and network should be jointly adapted. An acceptable video quality in wireless environment can be obtained by the adjustment of parameters in video codec and wireless network. For the former, people have proposed many error resilient video encoding algorithms to enhance the robust performance of the compressed video stream in wireless networks. These algorithms can be divided into three categories: 1) error detection and error concealment algorithms used at video decoder of wireless receiver; 2) error resilient video encoding algorithms located at video encoder of wireless transmitter; 3) robust error control between video encoder and decoder based on 1) and 2). Fig.1 summarizes different techniques at different parts of a wireless video transmission system.

Since error concealment algorithms are only used at video decoder in wireless receiver, they do not require any modification of video encoder and channel codec. Hence, there is not any increase of coding computing complexity and transmission rate. Therefore, error concealment algorithms can be easily realized in present wireless video transmission system. However, since error concealment algorithms make full use of spatial and temporal correlation in video stream to estimate the corrupted region of video frames, when the correlation between corrupted region and correctly received frames is weak, error concealment algorithms

cannot achieve good effect so that there is apparent distortion in repaired reconstructed video frames. In addition, although error concealment algorithms can reduce the intensity of temporal error propagation, it cannot reduce the length of temporal error propagation. As we know, human visual system (HVS) is not very sensitive to short term obvious error propagation while long term even slight error propagation will annoy the observation of HVS impressively. Therefore, desirable error repaired effect should make the intensity and length of error propagation minimum simultaneously.
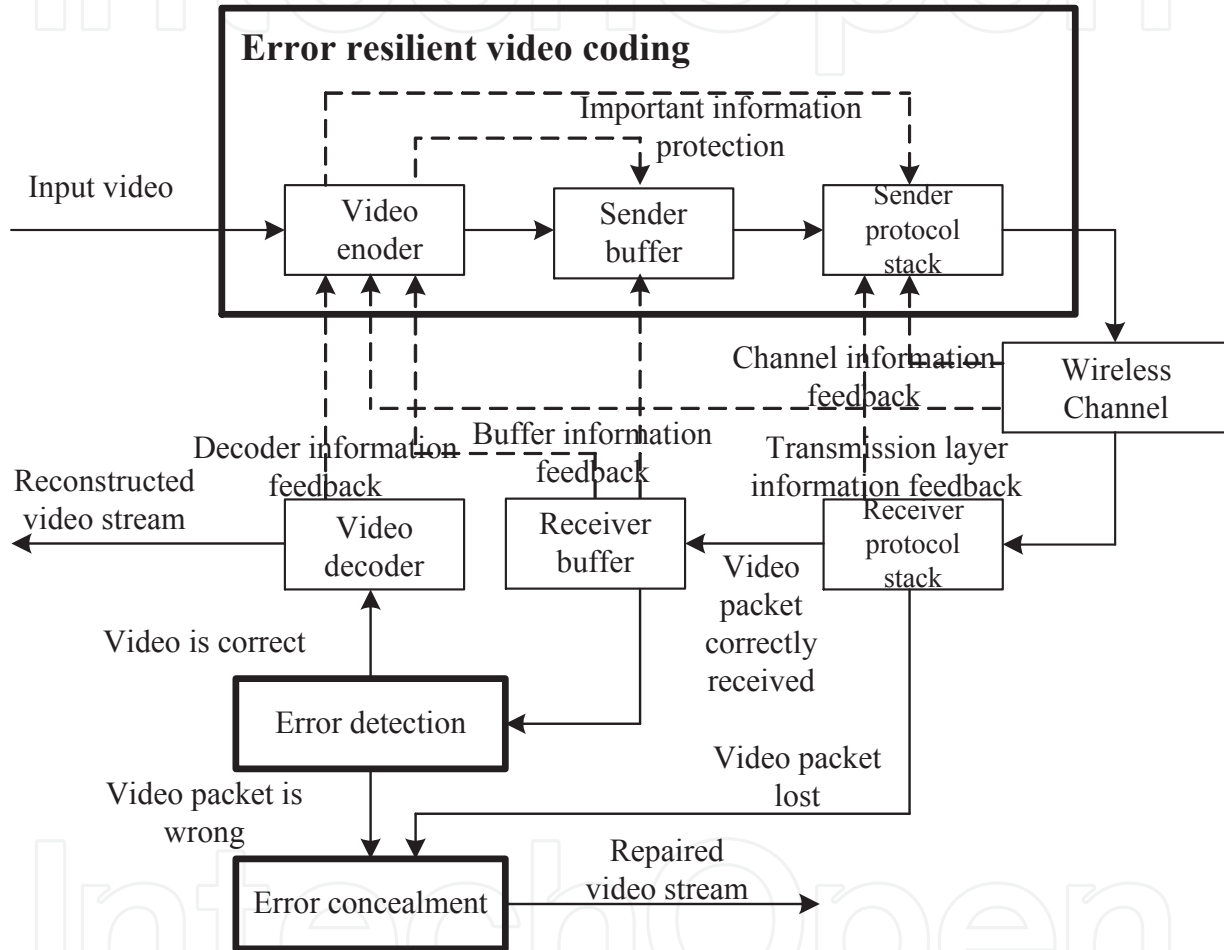

Figure 1. Error resilient methods used in packet-switched wireless networks

In a practical wireless video transmission system, one entire frame is normally encapsulated into one video packet in order to make full use of limited wireless bandwidth. In this situation, any loss of one video packet would degrade image quality of successive frames in video decoder apparently since existing video standards utilize inter-frame prediction to make high compression efficiency. Hence, many error resilient methods have been developed to reduce the impacts of errors and improve the video quality in wireless video transmission in recent years [1-4]. However, most of the previously developed algorithms mitigate coding efficiency by adding redundancy to the video stream to enhance error resilient performance.

As mentioned, real time wireless video applications are very sensitive to the increase of coding overhead in [5], which may not only result in additional delay that makes correctly received video packets invalid, but also deteriorate the quality of service in wireless environment especially in ad hoc networks [6]. Therefore, it is necessary to make compressed video stream more resilient to errors at minimum expense of coding overhead.

In order to overcome the error propagation effect caused by video packet losses, long term memory motion-compensated prediction [7] is a reasonable way to suppress error propagation in the temporal domain at the cost of reducing the coding efficiency. In [8], the selection of reference frame in long-term motion compensated prediction is proposed for H.263 video with referring to the rate-distortion optimization (RDO) criteria. As a further work of [8], based on the original RDO model in error free condition, an error robust RDO (ER-RDO) method has been proposed in [9] for H.264 video in packet lost environment by redefining the Lagrange parameter and error-prone RD model. However, the ER-RDO method still requires a very high computational complexity to accurately determine the expected decoder distortion. To reduce the computational burden, Zhang et al. [10] developed a simplified version of the ER-RDO method by making full use of block-based distortion map to estimate the end to end distortion. Since the selected Lagrange parameters in these two methods are not precise enough to make corresponding rate distortion optimization, their cost for coding overhead for real time wireless video communication system is not desirable.

In the periodic frame method [11], a periodic frame is only predicted by previous l reference video frame, which is the previous periodic frame. l is the frame interval between neighboring periodic frames. When the frames between two periodic frames are lost, second periodic frame is still decoded correctly, so error propagation can be suppressed efficiently. However, the coding overhead of periodic frame increases obviously when the correlation between neighboring periodic frames is not high. To alleviate the heavy burden on wireless channel resulted by periodic frame, Zheng et al. also proposed the periodic macroblock (PMB) method [11] to reduce the increase of coding overhead by selecting only certain number of important MBs to be predicted by previous l reference video frame. PMB can effectively control the coding overhead with the sacrifice of the error reconstruction effect. Another effective way to constrain error propagation is to insert intracoded MBs. Compared to long term reference frame prediction, it needs more redundancy by adopting the intracoded mode. To obtain a better trade-off between the coding efficiency and error resilient performance, the methods based on accurate block-based distortion estimation model [12] [13] were developed for MPEG4 and H.261/3. The end-to-end approach in [12] generalized the RD optimized mode selection for point-to-point video communication by taking into account both the packet loss and the receiver's concealment method. In [13], the encoder computes an optimal estimate of the total distortion at decoder for a given rate, packet loss condition, and the concealment method. The distortion estimation is then incorporated within an RD framework to optimally select the coding mode for each macroblock. Both methods achieved better error resilient performance. However, their computational complexity and implementation cost are too high.

In this chapter, we develop a new channel based rate distortion (RD) model for error resilient H.264 video codec, which aims at minimizing the coding overhead increase while maintaining a good error resilience performance. In the new RD model, the practical channel conditions like packet lost rate (PLR) and packet lost burst length (PLBL), error propagation and error concealment effects in different reference frames are taken into consideration in analyzing the expected MB-based distortion at encoder. Moreover, for each reference frame, its corresponding Lagrange parameter is adjusted according to the variation of the channel based RD model, which can more accurately describe the relationship between coding rate and expected distortion at decoder in the sense of packet lost environment than other existing methods. Moreover, in our proposed new RD model, a proper intra-coded mode for error resilient performance is also considered. Therefore, more appropriate reference frame and encoding mode can be selected for each MB with the proposed method.

In the following of this chapter, a brief review on the error-robust rate-distortion optimization (ER-RDO) method is given in Section 2. The derivation of our proposed error resilient rate distortion (RD) optimization will be described in the same section. In section 3, the error resilient performance of the proposed method and some existing methods will be evaluated using computer simulations on H.264 video codec. Finally, some concluding remarks will be given in Section 4.

2. The proposed error resilience optimization method

As the latest video coding standard, H.264 has supreme coding performance by adopting lots of advanced techniques [14]. With the rate distortion optimization (RDO) operation, H.264 achieves a very good coding efficiency and a high PSNR simultaneously in error free condition. For encoding m^{th} MB in n^{th} frame, the RDO operation can find its most proper coding mode and reference frame by minimizing the cost as follows:

$$J_{org}(n,m,r,o) = D_s(n,m,r,o) + \lambda R(n,m,r,o) \tag{1}$$

where $D_s(n,m,r,o)$ and $R(n,m,r,o)$ are the source distortion and the coding rate when the MB is predicted by r^{th} reference frame and encoded with mode o. In an error free environment, the Lagrange parameter can be determined by the quantization parameter Q as follows [15]

$$\lambda = \begin{cases} 0.85 \times Q^2 & : (H.263) \\ 0.85 \times 2^{Q/3} & : (H.264) \end{cases} \tag{2}$$

However, the cost in (1) doesn't consider the distortion caused by error propagation and error concealment. Therefore, it cannot be directly used for finding the best reference frame

and encoding mode in an error prone wireless packet-switched network if the channel con-
dition is taken into consideration.

2.1. ER-RDO model

To take into account the packet lost effect, an error robust RDO (ER-RDO) method was de-
veloped in [11] by redefining the Lagrange parameter and error-prone RD model based on
the practical wireless channel situation and potential decoded MB corrupted distortion. In
the ER-RDO model, the expected overall distortion of mth MB in n^{th} frame is determined as

$$D_e(n,m,r,o) = (1-p)p_c D_s(n,m,r,o) + pD_{ec} + (1-p)(1-p_c)D_{ep} \qquad (3)$$

where D_{ec} is the error concealment distortion if this MB is lost, and D_{ep} represents the expect-
ed error propagation distortion in the case that this MB is received correctly but the refer-
ence frames are erroneous. p is the current wireless channel packet loss rate (PLR), and p_c is
the probability that all reference frames are correct, which is computed by

$$p_c = (1-p)^k \qquad (4)$$

where k is the number of reference frames in the encoder buffer.

If we assume high-resolution quantization, the source distortion D_s depends on the rate (R)
as follows [9]:

$$D_s(R) = \beta \times 2^{-\alpha R} \qquad (5)$$

where α and β parameterize the functional relationship between rate and distortion [13]. If
uniform quantization is used, then we have

$$D_s(\Delta) = \Delta^2 / 12 \qquad (6)$$

where Δ is the quantization step size.

Referring to (5) and (6), the selected Lagrange parameter in ER-RDO model is computed as

$$\lambda_{ER-RDO} = -\frac{dD}{d\Delta}\frac{d\Delta}{dR} = (1-p)p_c\lambda \qquad (7)$$

With λ_{ER-RDO}, (3) and (4), the best reference frame r^* and encoding mode o^* for m^{th} MB in n^{th}
frame selected as in [11] are determined as follows.

$$(r^*, o^*) = \arg\min(p_c D_s(n,m,r,o) + (1-p_c)D_{ep}(n,m,r,o) + p_c \lambda R(n,m,r,o)) = \arg\min(p_c J_{org}(n,m,r,o) + (1-p_c)D_{ep}(n,m,r,o)) \tag{8}$$

From (4) and (7), we can find that the selected λ_{ER-RDO} in each reference frame is identical when the number of reference frame and PLR is known. That is to say, the correlation between the coding rate and expected overall distortion for all reference frames is equal. However, as we know, when the distance between the selected reference frame and the present encoding frame turns to be longer, the probability of correct reconstruction of this frame at receiver is higher with the degradation of the coding efficiency. Therefore, the term $(1-p_c)D_{ep}$ in (8) is not accurate enough (a comprehensive interpretation will be given in next subsection). So in the sense of error resilience, the correlation between the coding rate and expected overall distortion for each reference frame at decoder should be different and be varied according to not only PLR and the range of reference frame, but also the distance between the selected reference predicted frame and the present encoding frame.

2.2. The proposed channel based RDO model

To overcome the problems of the ER-RDO model for H.264 video, we propose a new channel based RDO model to more accurately trade-off the coding efficiency and error resilient performance. For n^{th} video frame to be encoded, there are k reference frames in encoder buffer, namely $n-1$, $n-2…$ $n-k$, as illustrated in Fig.2.
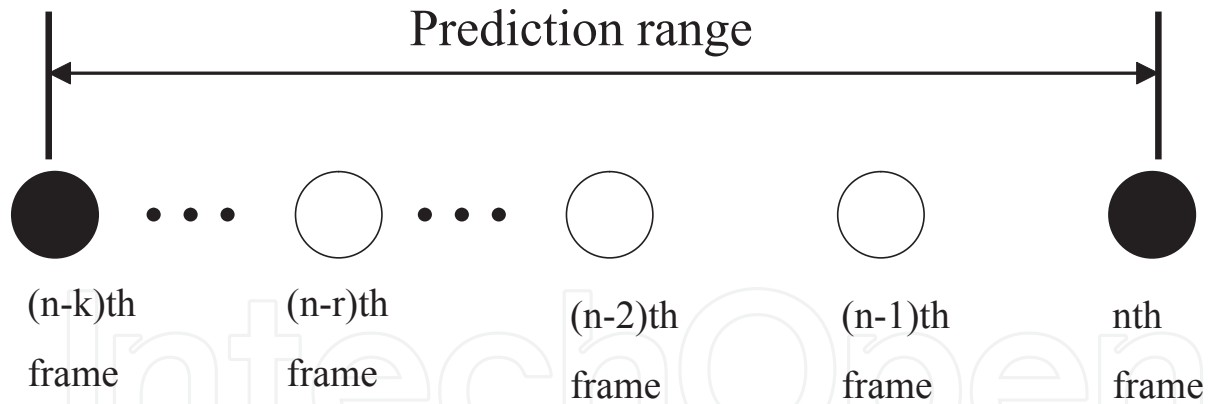

Figure 2. Inter-coded prediction reference frame range

The estimated cost for n^{th} frame predicted by $n-r$ ($1 \leq r \leq k$) reference frame is

$$J_p(n, n-r) = D_p(n, n-r) + \lambda_r R(n, n-r) \tag{9}$$

where $R(n,n-r)$ is the coding overhead of nth frame predicted by n-r reference frame, and referring to (3), $D_p(n,n-r)$ is the expected overall distortion of nth frame at decoder in the proposed channel based RDO model with n-r reference frame. It is given by

$$D_p(n,n-r) = (1-p)^{k+1}(D_s(n,n-r)+D_{lep})+(1-p)(1-(1-p)^k)D_{ep_r}+pD_{ec} \qquad (10)$$

where $D_s(n,n-r)$ is the source distortion predicted by n-r reference frame in error free situation, D_{lep} is distortion caused by the long term error propagation when frames before reference frames are lost. And D_{ep_r} is the potential distortion caused by the frame loss in the range of reference frame when *n-r* frame is the reference frame, which can be computed as followed.

$$D_{ep_r} = \sum_{j=r}^{k} D_{r_{j+1}} q_{ep}(r,j)+q_s(r)D_s(n,n-r) \qquad (11)$$

For computing D_{ep_r} as in (11), it includes two parts: one is the error propagation distortion caused by *n-k*, *n-k+1*… *n-r* reference frame. The term $D_{r_{j+1}}$ in (11) is error concealment reconstruction distortion when *n-j* frame is lost ($r \le j \le k$), and its corresponding occurrence probability is

$$q_{ep}(r,j) = p(1-p)^{k-j} \qquad (12)$$

When the frames after present reference frame *n-r* are lost, present encoding frame *n* can still be decoded correctly, this occurrence probability is computed as

$$q_s(r) = (1-p)^{k-r}(1-(1-p)^r) \qquad (13)$$

So another part is the multiplying results of $q_s(r)$ and $D_s(n,n-r)$ as in (11).

With (9), (10) and (11), the final estimate cost for n^{th} frame predicted by *n-r* reference frame is

$$\begin{aligned}
J_p(n,n-r) &= (1-p)^{k+1}(D_s(n,n-r)+D_{lep})+(1-p)(1-(1-p)^{k+1})D_{ep_r}+pD_{ec}+\lambda_r R(n,n-r)\\
&= (1-p)^{k+1}(D_s(n,n-r)+D_{lep})+pD_{ec}+\lambda_r R(n,n-r)\\
&\quad +(1-p)(1-(1-p)^k)(\sum_{j=r}^{k} D_{r_{j+1}} q_{ep}(r,j)+q_s(r)D_s(n,n-r))
\end{aligned} \qquad (14)$$

Finally, $J_p(n,n-r)$ is computed as

$$J_p(n,n-r) = ((1-p)^{2k+1}+(1-p)^{k-r+1}-(1-p)^{2k-r+1})D_s(n,n-r)+(1-p)^{k+1}D_{lep}+(1-p)(1-(1-p)^k)\sum_{j=r}^{k}D_{r_{j+1}}p(1-p)^{k-r}+pD_{ec}+\lambda_r R(n,n-r) \qquad (15)$$

So with the derivatives of $J_p(n,n-r)$ for Δ as (7), the optimized Lagrange parameter for present encoding frame n predicted by reference frame n-r is obtained by

$$\lambda_r = -\frac{dD(n,n-r)}{d\Delta}\frac{d\Delta}{dR} = ((1-p)^{2k+1} + (1-p)^{k-r+1} - (1-p)^{2k-r+1})\lambda \tag{16}$$

where we assume that the buffer length of reference frame k is larger than real-time PLBL obtained from the feedback of wireless channel situation.

2.3. Implementation of reference frame and mode selection algorithm

With the results obtained before, we apply the proposed channel based RDO model to select the best reference frame and encoding mode in an H.264 encoder as follows. For one MB in P frame, it has two categories of encoding modes: intracoded and intercoded. Intracoded modes include direct coding, *intra_4×4* and *intra_16×16*; intercoded modes include *inter_16×16*, *inter_16×8*, *inter_8×16* and *inter_P8×8* mode (this mode is composed of *inter_8×8*, *inter_8×4*, *inter_4×8* and *inter_4×4 sub 8×8* block modes). For each intercoded mode, the best reference predicted frame r^* for m^{th} MB in n^{th} frame in coding mode o is selected by finding the minimum cost of interceded mode$J_p(n, m, o, r)$.

$$J_p(n,m,o,r^*) = \arg\min(J_p(n,m,o,r))$$
$$= \arg\min(D_p(n,m,o,r) + \lambda_r R(n,m,o,r))$$
$$= \arg\min(((1-p)^{2k+1} + (1-p)^{k-r+1} - (1-p)^{2k-r+1})D_s(n,m,o,r) \tag{17}$$
$$+ (1-p)^{k+1}D_{lep} + (1-p)(1-(1-p)^k)\sum_{j=r}^{k} D_{r_{j+1}} p(1-p)^{k-r} + pD_{ec} + \lambda_r R(n,m,o,r))$$

where r^* is best reference predicted frame in coding mode o. Since $(1-p)^{k+1}D_{lep}$ is same for any reference frame to predict n^{th} frame, andD_{ec} is independent of encoding modes and reference frame [8], (17) can be simplified to

$$J_p(n,m,o,r^*) = \arg\min(((1-p)^{2k+1} + (1-p)^{k-r+1} - (1-p)^{2k-r+1})D_s(n,m,o,r)$$
$$+ (1-p)(1-(1-p)^k)\sum_{j=r}^{k} D_{r_{j+1}} p(1-p)^{k-r} + \lambda_r R(n,m,o,r))$$
$$= \arg\min(\alpha_r(D_s(n,m,o,r) + \beta_r\sum_{j=r}^{n} D_{r_{j+1}} p(1-p)^{n-j} + \lambda R(n,m,o,r))) \tag{18}$$
$$= \arg\min(\alpha_r(J_{org}(n,m,o,r) + \beta_r D_{referror}(r)))$$

$$a_r = \lambda_r / \lambda, \beta_r = \frac{(1-p)(1-(1-p)^k)}{\alpha_r} \tag{19}$$

$$D_{referror}(r) = \sum_{j=r}^{k} D_{r_{j+1}} p(1-p)^{n-r} \tag{20}$$

And then, the best encoding mode o^* in intercoded mode is

$$J_P(n,m,o^*,r^*) = \arg\min(J_P(n,m,o,r^*)) \tag{21}$$

For the best intracoded mode o^{**} for this MB, it can be determined as follows with the cost for intracoded mode $J_i(n, m, o, 0)$.

$$
\begin{aligned}
J_i(n,m,o^{**},0) &= \arg\min(J_i(n,m,o,0)) \\
&= \arg\min((1-p)D_s(n,m,o,0) + pD_{ec} + (1-p)\lambda R(n,m,o,0)) \\
&= \arg\min((1-p)J_{org}(n,m,o,0) + pD_{ec})
\end{aligned} \tag{22}
$$

As the final results, the best encoding mode $\hat{o}$ and its potential best reference predicted frame $\hat{r}$ in the sense of optimized error resilience for m^{th} MB in n^{th} frame are found as

$$J_{best}(n,m,\hat{o},\hat{r}) = \arg\min(J_p(n,m,o^*,r^*), J_I(n,m,o^{**},0)) \tag{23}$$

3. Experimental results

In this section, we evaluate the performance of the proposed channel based RDO model in terms of video quality and coding efficiency in wireless packet lost environment. In our experiments, we use H.264 JM 8.2 codec as test platform where video stream structure is selected as IPPP…. Three standard QCIF video sequences, namely Salesman, Susie and Foreman, are used in the simulations. The range of tested intracoded frames in these sequences is from 10th to 100th frame. Their QP is set as 28, their frame rate in H.264 JM8.2 is 30 fps, and their buffer of reference frames includes previous five frames. In order to make full use of wireless channel bandwidth, each compressed video frame is transmitted by a single packet. A simple error concealment method is used to make analysis of potential error propagation and error concealment effect at video encoder. When a MB is assumed to be lost, it

will be replaced by the MB at same position in the previous error free frame. As a comparison, we use the original H.264 JM8.2 codec, the periodic frame method, the PMB method [11] and ER-RDO method [8] as reference algorithms. In addition, for the PMB method, we use PMB (11%), and PMB (22%) and PMB (33%) to denote the corresponding performance when the proportions of periodic MB in video frame are 11%, 22% and 33% respectively.

We first look at the error resilience performance of the proposed method by considering the PSNR performance of the reconstructed video under a packet loss environment. Fig.3 shows the error reconstruction effect of three test sequences using different methods when PLR = 0.1 and PLBL < 5. At each point in Fig.3, it is an average PSNR result when any reference frame of present encoding frame is lost.

It is shown in Fig.3 that the proposed method always achieves the best reconstruction effect for the three test sequences when compared with other methods. In Fig.3 (a), for Salesman sequence with low motion scene, the proposed method outperforms H.264 JM8.2, PMB (11%), PMB (22%), PMB (33%) and ER-RDO with an average PSNR improvement of 1.18dB, 1.14dB, 1.04dB, 0.8dB and 0.2dB, respectively. In Fig.3 (b), for Susie sequence with moderate motion scene, the proposed method performs better than H.264 JM8.2, PMB (11%), and PMB (22%), PMB (33%) and ER-RDO with an average PSNR improvement of 2.48dB, 2.03dB, 1.43dB, 0.13db and 0.21dB, respectively. In Fig.3 (c), for Foreman sequence with high motion scene, the proposed method achieves better results than H.264 JM8.2, PMB (11%), and PMB (22%), PMB (33%) and ER-RDO with an average PSNR improvement of 3.61dB, 3.04dB, 2.45dB, 1.72db and 0.53dB, respectively. As a conclusion, the proposed method can achieve more robust error resilient performance in different video scenes.

For evaluating the coding efficiency of different methods, we consider their impacts on overall coding rate requirement and PSNR performance of reconstructed video in error free environment. The simulation results for the three test sequences are listed in Table 1, 2 and 3 respectively. It is seen that all of the error resilient methods have little effect on original video quality. For fair comparisons, the PSNR performance of the reconstructed video is more or less kept constant for different methods. We then compare the coding rate required for each method.

| Method	PSNR-Y (dB)	PSNR-U (dB)	PSNR-V (dB)	Bit rate (kb/s)	Increase (%)
H.264 JM 8.2	35.57	39.6	40.14	56.83	0%
Periodic Frame FFFrame	35.54	39.61	40.19	60.08	5.72%
PMB (33%)	35.54	39.59	40.15	57.48	1.14%
PMB (22%)	35.59	39.61	40.17	57.05	0.39%
PMB (11%)	35.59	39.59	40.17	57.03	0.35%
ER-RDO	35.61	39.66	40.23	60.4	6.28%
The proposed method	35.57	39.6	40.15	57.14	0.54%

Table 1. Coding rate comparison of different methods in *Salesman* sequence
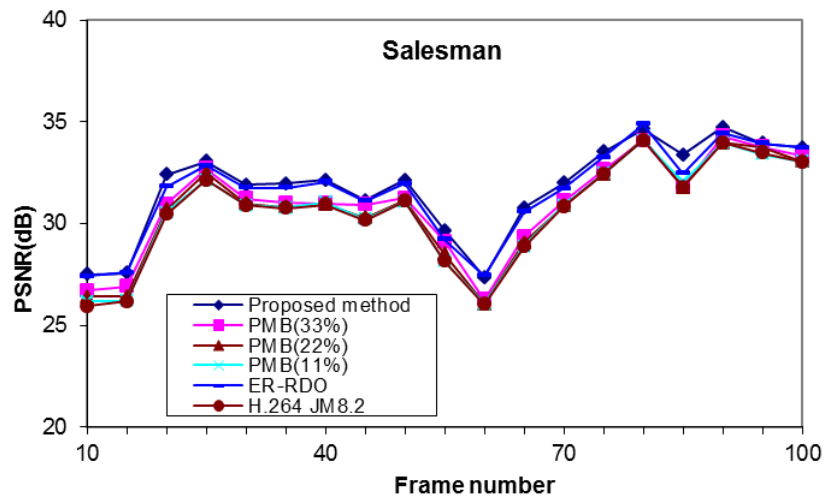
(a) Salesman
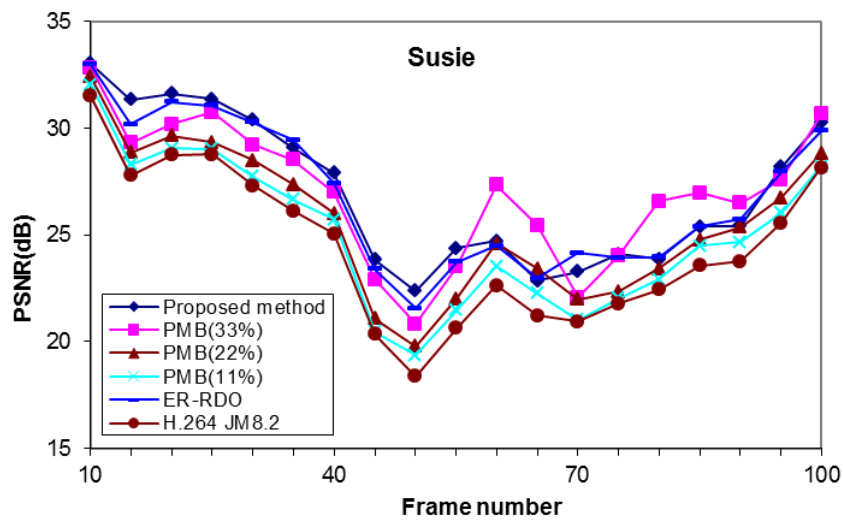
(b) Susie
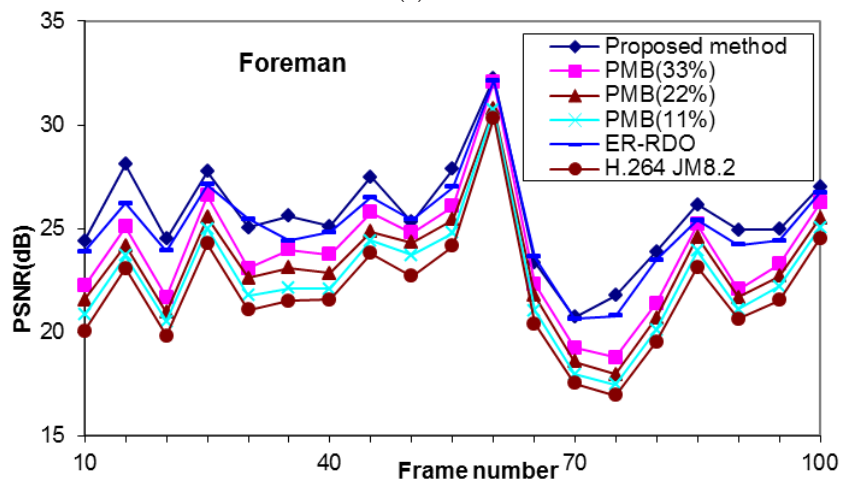
(c) Foreman

Figure 3. Reconstruction effect comparison of different methods when PLR = 0.1 and PLBL < 5

Table 1 shows the coding rate requirement of different methods for *Salesman* sequence, in which there is high correlation between reference frames and encoding frame. It is noted that the coding redundancy resulted in all methods is smallest among the three test sequences. The coding rate increase of ER-RDO method is not desirable as it needs more bits than the periodic frame method, while the PMB method in different level of long term predicted MB can obtain less rate increase. The coding overhead increase of the proposed method is not obvious as it is only slightly larger than PMB (11%) and PMB (22%) and apparently smaller than PMB (33%).

Method	PSNR-Y (dB)	PSNR-U (dB)	PSNR-V (dB)	Bit rate (kb/s)	Increase (%)
H.264 JM 8.2	37.26	43.54	43.28	95.56	0%
Periodic Frame	37.24	43.59	43.37	108.91	13.97%
PMB (33%)	37.23	43.59	43.29	102.41	7.17%
PMB (22%)	37.26	43.62	43.27	99.76	4.39%
PMB (11%)	37.25	43.55	43.32	97.78	2.32%
ER-RDO	37.29	43.64	43.28	100.82	5.50%
The proposed method	37.26	43.56	43.24	95.75	0.2%

Table 2. Coding rate comparison of different methods in *Susie* sequence

For *Susie* sequence where the correlation between reference frames and encoding frame is moderate, the coding overhead is in general more than that of *Salesman* sequence, as shown in Table 2. It is noted that the coding rate of the periodic frame method has increased about 14%, which is a heavy burden for wireless channel. The coding rate increase of ER-RDO is smaller than PMB (33%), while it is still more than PMB (11%) and PMB (22%). The coding rate of the proposed method is just 0.2% higher than that of H.264 JM 8.2 but smaller than all other methods.

Method	PSNR-Y (dB)	PSNR-U (dB)	PSNR-V (dB)	Bit rate (kb/s)	Increase (%)
H.264 JM 8.2	35.72	39.04	40.72	109.17	0%
Periodic Frame	35.74	39.17	40.84	129.07	18.23%
PMB (33%)	35.69	39.14	40.78	116.41	6.63%
PMB (22%)	35.7	39.06	40.7	113.79	4.23%
PMB (11%)	35.71	39.03	40.75	112.46	3.01%
ER-RDO	35.73	39.04	40.75	116.62	6.82%
The proposed method	35.72	39.05	40.76	111.28	1.93%

Table 3. Coding rate comparison of different methods in *Foreman* sequence

For *Foreman* sequence, as there is low correlation between reference frames and encoding frame, the required coding rate of all methods is largest in among the three test sequences, as shown in Table 3. Again, our proposed method achieves the best coding efficiency. The coding rate increase of the proposed method is only 1.93%, while that of PMB (11%), PMB (22%), PMB (33%), ER-RDO and the periodic frame method is 3.01%, 4.23%, 6.63%, 6.82% and 18.23%, respectively.

As a conclusion with the results of error resilient performance and the coding efficiency, the proposed method can obtain not only more satisfying video reconstruction effect but also smaller coding rate increase than the reference methods.
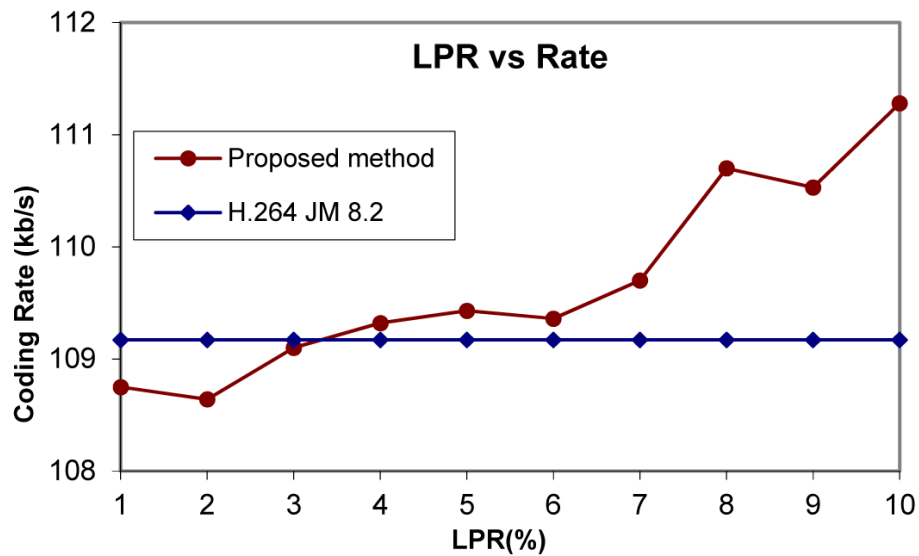

Figure 4. The coding rate (kb/s) of the proposed method with respect to original H.264 JM 8.2 codec in different PLR from 0.01% to 0.1%

Fig.4 shows the coding efficiency of the proposed method in different PLR from 0.01% to 0.1% of Foreman sequence. In Fig.4, we can find that the increase of coding rate using the proposed method is small when compared with that of H.264 JM 8.2 codec. Even in some instances of low LPR of Fig.4, the proposed method can achieve a slightly smaller coding rate than the original H.264 JM 8.2 codec.

As a further analysis on error resilient performance of the proposed method with respect to the PMB and ER-RDO method, Table 4, 5 and 6 give more detailed reconstruction PSNR (dB) effect comparison in Salesman, Susie and Foreman sequences when each of the reference frames in encoder buffer is lost. From the tables, we can find that the PMB method, especially PMB (33%) can achieve better results when the lost reference frame is far away from present encoding frame. On the contrary, ER-RDO can obtain better reconstruction effect when lost reference frame is near to present encoding frame. Our proposed method achieves a compromise between the two methods and obtains better average error reconstruction performance. In addition, it is always better than H.264 JM 8.2 when any reference frame in the encoder buffer is lost.

Lost reference frame	Proposed method	PMB (33%)	PMB (22%)	PMB (11%)	ER-RDO	H.264 JM8.2
1	34.04	32.63	32.41	32.31	34.06	32.29
2	32.56	31.65	31.37	31.27	32.49	31.21
3	31.54	30.83	30.5	30.37	31.34	30.31
4	30.7	30.17	29.78	29.67	30.4	29.58
5	29.99	29.57	29.57	29.57	29.6	29.57

Table 4. Reconstruction PSNR (dB) comparison of different methods in *Salesman* sequence

Lost reference frame	Proposed method	PMB (33%)	PMB (22%)	PMB (11%)	ER-RDO	H.264 JM8.2
1	31.74	30.03	28.37	27.65	32.48	26.99
2	28.42	28.31	26.59	25.84	28.98	25.15
3	26.2	27.1	25.41	24.63	25.73	23.96
4	24.86	26.29	24.58	23.82	23.87	23.16
5	23.92	23.07	23.07	23.07	23.07	23.07

Table 5. Reconstruction PSNR (dB) comparison of different methods in *Susie sequence*

Lost reference frame	Proposed method	PMB (33%)	PMB (22%)	PMB (11%)	ER-RDO	H.264 JM8.2
1	31.81	27.34	26.51	25.8	32.45	24.98
2	27.78	25.26	24.31	23.58	27.95	22.77
3	24.7	23.83	22.86	22.1	23.37	21.32
4	22.63	22.73	21.77	21.03	21.3	20.3
5	21.06	20.23	20.23	20.23	20.24	20.23

Table 6. Reconstruction PSNR (dB) comparison of different methods in *Foreman* sequence

4. Conclusions

In this paper, an error resilient method based on the feedback of wireless channel condition is proposed for robust H.264 video stream transmitted in wireless packet lost environment. The proposed method can smartly adjust Lagrange parameter for each reference frame at encoder buffer by adopting proposed channel based RDO model. The modified Lagrange parameter can better reflect the association between the expected distortion and coding effi-

ciency of video streaming in the sense of error resilience in packet lost environments. Comprehensive experimental results show that the proposed method sufficiently absorbs the advantages of existing methods and achieves better error resilient performance with minimum increase of coding overhead.

Acknowledgements

Author details

Jian Feng[1], Yu Chen[2,3], Kwok-Tung Lo[3] and Xu-Dong Zhang[2]

1 Dept. of Computer Science, Hong Kong Baptist Uni., HK

2 Dept. of Electronic Eng., Tsinghua Uni., Beijing, China

3 Dept. of Electronic and Info. Eng., Hong Kong Polytechnic Uni., HK

References

[1] Wang Y., Wenger S., Wen J., and Katsaggelos K.K. Error resilient video coding techniques. IEEE Signal Processing Magazine 2000, 17(4), 61-82.

[2] Vetro A., Xin J., and Sun H.F. Error resilience video transcoding for wireless communication. IEEE Wireless Communication 2005., 12(4), 14-21.

[3] Stockhammer T., Hannuksela M.M., and Wiegand T. H.264/AVC in wireless environments. IEEE Trans. Circuits Syst. Video Technol. 2003, 13(7), 657-673.

[4] Hsiao Y.M., Lee J.F., Chen J.S. and Chu Y.S. H.264 video transmissions over wireless networks: Challenges and solutions. Computer Communications 2011, 34(14), 1661-1672.

[5] Katsaggelos A.K., Eisenberg Y., Zhai F., Berry R., and Pappas T.N. Advances in efficient resource allocation for packet-based real-time video transmission. Proceedings of the IEEE 2005, 93(1), 135-147.

[6] Zhu X.Q., Setton E., and Girod B. Congestion-distortion optimized video transmission over ad hoc networks. Signal Processing: Image Communication 2005, 20(8), 773-783.

[7]   Wiegand T., Zhang X., and Girod B. Long-term memory motion-compensated pre-diction. IEEE Trans. Circuits Syst. Video Technol. 1999, 9(2), 70–84.

[8]   T. Wiegand, Farber N., Stuhlmuller K., and Girod B. Error-resilient video transmis-sion using long-term memory motion-compensated prediction. IEEE Journal on Se-lected Areas in Communications 2000, 18(3), 1050–1062.

[9]   T. Stockhammer, D. Kontopodis and T.Wiegand, "Rate-distortion optimization for JVT/H.26L coding in packet loss environment," Proc. PVW, Pittsburgh, PY, April 2002.

[10] Zhang Y., Gao W., Lo Y., Huang Q.M. and Zhao D. Joint Source-Channel Rate-Dis-tortion Optimization for H.264 Video Coding Over Error-Prone Networks. IEEE Trans. Multimedia 2007, 9(3), 445-454.

[11] Zheng J.H. and Chau L.P. Error-resilient coding of H.264 based on periodic macro-block. IEEE Trans. Broadcasting 2006, 52(2), 223-229.

[12] Wu D., Hou Y.T., Li B., Zhu W., Zhang Y.Q. and Chao H.J. An end to end approach for optimal mode selection in Internet video communication. IEEE Journal on Select-ed Areas in Communications 2000, 18(6), 977-995.

[13] Zhang R., Regunathan S.L. and Rose K. Video coding with optimal inter/intra-mode switching for packet loss resilience. IEEE Journal on Selected Areas in Communica-tions 2000, 18(6), 966-976.

[14] Wigand T., Sullivan G.J., Bjntegaard G., and Luthra A. Overview of the H.264/AVC video coding standard. IEEE Trans. Circuits Syst. Video Technol. 2003, 13(7), 560-576.

[15] Wiegand T. and Girod B. Lagrange multiplier selection in hybrid video coder control. Proc. ICIP '01, 2001: 542-545.