An Adaptive Error Resilient Scheme for Packet-Switched H.264 Video Transmission

Jian Feng[1], Yu Chen[2,3], Kwok-Tung Lo[2] and Xudong Zhang[3]
[1]*Dept. of Computer Science, Hong Kong Baptist Uni.,*
[2]*Dept. of Electronic and Information Eng., Hong Kong Polytechnic Uni.,*
[3]*Dept. of Electronic Eng., Tsinghua Uni., Beijing,*
[1,2]*Hong Kong*
[3]*China*

1. Introduction

When applying conventional standard video codecs in wireless video applications, error resilience and coding efficiency are the two main issues need to be considered. Since it is difficult to corroborate robust quality of service (QoS) in wireless networks, transmitted video packets are sometime lost or corrupted due to fading and shadowing effect of wireless channel. Providing robust video transmission in wireless packet-switched networks is therefore a challenging task as compressed video is very vulnerable to channel error.

In recent years, many error resilient tools have been proposed to enhance the robust performance of video transmission in wireless environment (Chen et al., 2008)(Stockhammer et al., 2003) (Stockhammer et al., 2005) (Vetra et al., 2005) (Wang et al., 2000) (Wiegand et al. 2003). Coding efficiency is one of the important issues to be taken into account for the limited bandwidth of wireless networks (Etoh and Yoshimura, 2005). To achieve a robust video transmission over wireless channels, the video codec on one hand should have supreme error resilient performance, on the other hand, it should also maintain a good coding efficiency by limiting the overhead information introduced by the error resilient tools. Hence, a good compromise between the error resilience performance and coding efficiency should be made.

Interactive error control is one of the effective error resilient techniques adopted in video codec. In this category, some error resilient techniques based on data hiding are proposed (Zeng, 2003) (Yilmaz and Alatan, 2003) (Kang and Leou, 2005). In such techniques, the important information for error concealment is extracted and embedded into video bitstream at the video encoder. When some packets are lost or corrupted, their corresponding embedded data at proper positions can enhance the error concealment effect at the video decoder. Although data hiding methods can obtain satisfied error resilient effect, their notable increase of bits overhead is disadvantageous for coding efficiency. Since the principle of embedding important information (Yin el al., 2001) they adopted modifies the original AC coefficients, not only video quality is degraded, but also coding overhead will be increased significantly. In wireless channels, as the transmission rate is limited, an obvious increase on coding overhead results in inevitable delay. Moreover, in wireless packet-switched networks, when a packet arrives at receiver beyond the maximum system waiting time, the receiver will consider this as packet lost [11]. Hence, embedded information should be essential and refined.

In order to simultaneously obtain better error resilient performance and preserve original coding efficiency in the video stream, a low redundancy error resilient scheme for H.264 video transmission in packet-switched environment is proposed in this chapter. The proposed method firstly utilizes content analysis to classify macroblocks (MBs) in a P frame into four categories with different protection measures. Each MB will then be protected by inserting proper information in next frame, which is determined by its protection type. Considering limited bandwidth of wireless channel, the inserted redundancy is selected as concise as possible while it can still facilitate error concealment to obtain better reconstruction effect. Finally, with the feedback from receiver, an adaptive transmission strategy of video packet is developed to effectively mitigate the required transmission rate especially in low packet loss rate (PLR) environments. Simulation results on H.264 JM 8.2 codec in different PLRs show that the proposed method can obviously outperform some reference methods in both PSNR and subjective quality. And it just brings little increase in coding overhead.

In the following of this chapter, a review on various error resilient techniques for wireless video transmission will be reviewed in section 2. A new error resilient scheme for H.264 video will then be described in section 4. Simulation results will be presented and discussed in section 4. Finally, some concluding remarks will be given in section 5.

2. Error resilient video coding

With the rapid development of wireless communications technologies, the demand for transmission of various video contents over wireless environments has been greatly increasing in recent few years. Therefore, providing robust video transmission in wireless environments draws much people's attention from different communities. However, it is a challenging task to make video information robust in wireless transmission. First, the quality of service (QoS) of wireless channel is hardly reliable for its high bit error rate and limited transmission bandwidth. Second, as techniques like predictive coding and variable length coding are generally adopted in most of the existing video codecs, it will cause not only spatial error propagation in present frame, but also temporal error propagation in successive frames. Hence, the visual quality at the receiving end will be greatly reduced.

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 (Wang and Zhu, 1998) (Villasenor et al, 1999) (Wang et al., 2000) (Chen et al., 2008). 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). Figure 1 summarizes different techniques at different parts of a wireless video transmission system.

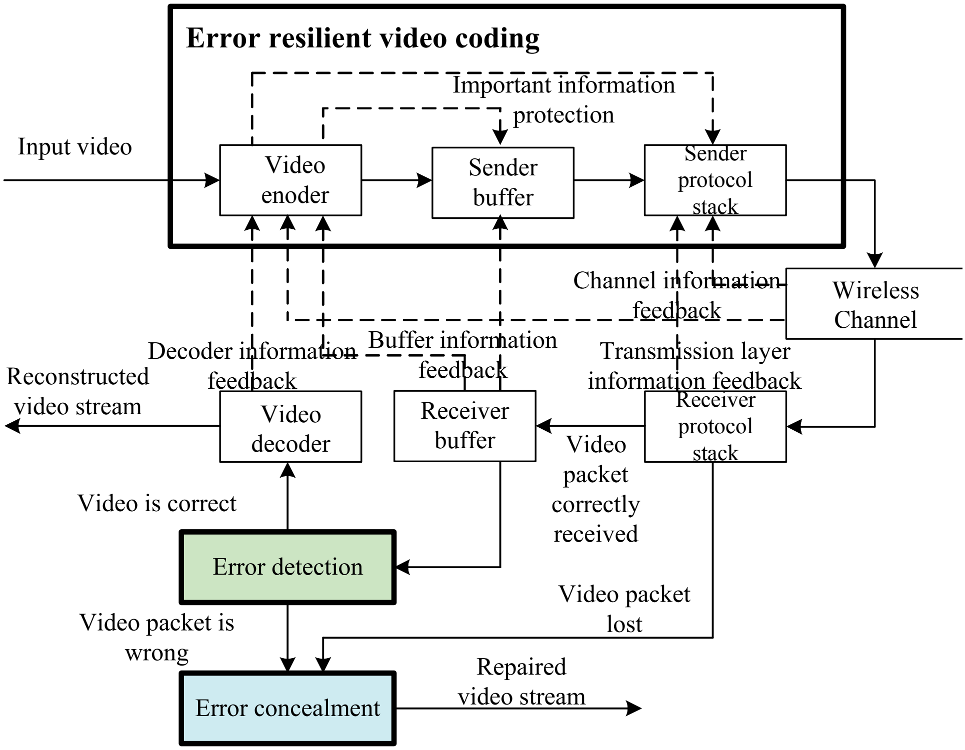


Fig. 1. Error resilient methods used in packet-switched wireless networks

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.

In order to compensate the defects of error concealment algorithms, a number of error resilient video encoding algorithms had been developed in the last decade to make the compressed video stream be accustomed to wireless transmission environment. These algorithms can be divided into five categories as discussed in the following.

The first category is concerned on the location of error protection information. Its main purpose is to reduce the length and intensity of spatial and temporal error propagation. In this category, four kinds of representative algorithms are developed based on resynchronization mode, adaptive intra coding, flexible reference frame and multiple reference frames.
The second category of error resilient algorithms utilizes data partition scheme to aggregate same type of syntax information, such as the aggregation of motion vector, header and texture information. When channel error appears in this type of video stream, all of information in the same region is not simultaneously wrong, and there is some correct information left for corrupted region. So with residual correct information, coarse reconstruction effect is still achieved at video decoder, which is always more satisfactory than that of error concealment.
The redundant error resilient video encoding algorithms can efficiently improve the performance of robust decoding for their inserted redundancy to mitigate the corrupted probability of video stream. Reversible Variable Length Coding (RVLC) (Takishima et al., 1995) can effectively reduce the range of spatial error propagation by reversely decoding from the position of next resynchronization mode with expense of apparent increase of encoding overhead. Multiple Description Coding (MDC) (Wang et al., 2005) divides conventional compression video stream into several pieces of sub-video stream, and each of them has same priority for transmission. When any of them is corrupted or lost in transmission, residual correctly received pieces of video stream can still be used to reconstruct coarse picture. Flexible Graphic Scalable (FGS) coding (Jens-rainer, 2005) is another type of error resilient algorithms to adopt multiple layers coding for video compression. In FGS coding, there are depending associations among base layer and enhanced layers that if only base layer is correctly decoded, the other enhancement layers can be decoded.
The fourth category is developed to compensate the defects of existing error concealment algorithms. For the spatial and temporal correlation in video stream is not always high, and the correct data used as reference by error concealment is not always enough, practical prediction effect of error concealment is not precise, whose final repaired effect is not better than direct replacement and weighted interpolation. In order to avoid this, some essential verification information are necessary to add into original video stream in order to improve the preciseness of error concealment prediction effect.
The last category is the wireless channel based error resilient video coding algorithms (Stockhammer et al. 2002). With respect to original rate distortion optimization (RDO) model in conventional video codec, these algorithms are designed to get better video quality and compression efficiency simultaneously. This type of RDO model may not be best suited to the wireless transmission environment. The distortion caused by channel error should be taken into RDO model so that the corresponding optimization parameters in the RDO model can be adjusted according to varied channel parameters, such as, packet lost rate (PLR), bit error rate (BER) and burst error average length (BEAL).

3. Content-based error resilient scheme

In a generic packet-based video transmission system (Katsaggelos et al., 2005), both the source coding and network will be jointly adapted to achieve an optimum transmission over a noisy channel.

A formulation is given in (Katsaggelos et al., 2005) to minimize the cost required to send video stream confined to a desirable level of distortion and tolerable delay. This is suitable for mobile terminal where transmission power, computational resource and wireless bandwidth are limited. The optimization for this formulation is

$$C_{tot}(S^*,N^*) = \min C_{tot}(S,N) \quad \{S^* \in S, N^* \in N\} \tag{1}$$

$$s.t.: D_{tot}(S^*,N^*) \le D_0 \quad \text{and} \quad T_{tot}(S^*,N^*) \le T_0$$

where D_0 is the maximum allowable distortion, and T_0 is the end-to-end delay constraint. Here, S denotes the set including available error resilient source coding types, and N represents the network parameters and transmission strategy to be controlled. S^* and N^* are the best selection of S and N to obtain minimum coding overhead min C_{tot} (S^*,N^*) while their end-to-end distortion $D_{tot}(S^*,N^*)$ is not beyond D_0, and their delay $T_{tot}(S^*,N^*)$ does not exceed T_0.

In this chapter, one effort is to find an optimized error resilient tool S^* from S that can guide error concealment to get the best reconstruction effect with minimum amount of redundancy. Another effort concerns an adaptive transmission strategy N^* for video packets based on the receiver's feedback to reduce burden on channel overhead. Here, we assume that network parameters, such as channel coding and network protocols, have been optimized.

3.1 New error resilient approach

To derive an efficient error resilient scheme, we consider the following issues for embedded video coding used in error resilience: (1) how to use correct evaluation tools to find MBs with important information that is crucial to human vision system (HVS), (2) how to extract these crucial information from the video stream and represent them with minimum bits, and (3) how to add extracted important information into the video stream at proper position with little coding overhead increase and modification on original video stream. Since slice video coding is resilient to wireless channel error, in this paper, we assume that each row of MBs in a P frame is a slice transmitted by a video packet.

To locate the important information in P frames, we consider using the pre-error concealment scheme (Kang and Leou, 2005). In this scheme, a process called pre-error concealment is performed to each MB and the mean absolute error (MAE) between the error-free MB and pre-concealed MB is calculated. Those macroblocks having large MAE values are regarded as important ones since missing of them will result in larger reconstruction error. More embedded protection information will then be allocated to them. However, this result may not be consistent with human visual system. Since mobile video phone and video conference are the common applications of current wireless video communication, the main content of such services is human portrait. According to (Wang et al., 2003), the region of human face is recognized as the foveation area of HVS. In this region, even a slight variation in face between neighboring frames such as blink and smile can draw HVS more attention than other MBs that may have a larger MAE in the background. In this regard, we propose to make a refinement of the pre-error concealment scheme with the concern of HVS by taking into account whether the MB is located in the foveation area or not. The weighted pre-error concealed distortion (WPECD) f_i of the i^{th} MB is defined as follows.

$$f_i = \begin{cases} q_i m_f : (MB_i \in \Gamma) \\ q_i m_{nf} : (MB_i \notin \Gamma) \end{cases} \tag{2}$$

where m_f (m_f >1) and m_{nf} (0< m_{nf} <1) are the weighted values assigned to the MBs in the foveation region and the background, respectively, which are determined by the network parameters and the size of the foveation region. Γ denotes the foveation region (removed), Here, we uses q_i to denote the original pre-error concealed distortion (PECD) value of the i^{th} MB obtained by pre-concealment as follows.

$$q_i = \frac{\sum_{x=1}^{16}\sum_{y=1}^{16}\left| m_{org_i}(x,y) - m_{prc_i}(x,y) \right|}{\sum_{x=1}^{16}\sum_{y=1}^{16} m_{org_i}(x,y)} \tag{3}$$

where $m_{org_i}(x,y)$ and $m_{prc_i}(x,y)$ are the original and pre-concealed pixel value of the luminance component of the i^{th} MB. In this paper, we adopt the error concealment method in (Tsekeridou and Pitas, 2000) to obtain those pre-concealed values.

After identifying the important information, we have to determine which information should be embedded for protecting the content. For the MB having a small PECD value based on (3), it means that the error concealment process can reconstruct the original information with little visual degradation especially in the foveation area (removed) because of the high temporal correlation between the reference MB and lost MB. But when the MB is in the moderate and high motion region, it will have a larger PECD value. If it is lost, existing error concealment methods may not accurately estimate its motion vector (MV), as well as its residual transform coefficients. To measure the deviation of the error concealment method in estimating MV and residual error, the following two parameters are defined.

$$q_{mc_i} = \frac{\sum_{x=1}^{16}\sum_{y=1}^{16}\left| m_{mc_i}(x,y) - m_{prc_i}(x,y) \right|}{\sum_{x=1}^{16}\sum_{y=1}^{16} m_{mc_i}(x,y)} \tag{4}$$

where q_{mc_i} denotes the deviation level between error concealed MB and original motion compensated MB without the residual distortion, $m_{mc_i}(x,y)$ is the original motion compensated pixel value of luminance component of i^{th} MB.

$$q_{res_i} = \frac{\sum_{x=1}^{16}\sum_{y=1}^{16}\left| m_{mc_i}(x,y) - m_{org_i}(x,y) \right|}{\sum_{x=1}^{16}\sum_{y=1}^{16} m_{org_i}(x,y)} \tag{5}$$

where q_{res_i} is the deviation level between original MB and original motion compensated MB. These two parameters reflect to a certain extend the degree of temporal and spatial correlation of subsequent frames. Based on these two factors, we classify MBs in a P frame into four categories (two extra bits are required for MB classification) as follows.

i. *Class I MB*: The MB has a small PECD value, i.e., $q_{mc} < T_{mc}$ or $q_{res} < T_{res}$, where T_{mc} and T_{res} are two thresholds. As mentioned before, the error-concealed effects of Class I MBs are very close to the original video quality. Hence, it is not necessary to insert protection information for them. Also, in most cases, the previous frame is the best predicted frame and there are little distortion among the reconstruction effects of 16×16 block mode and other block modes. Thus, we make a compromise that these MBs can be restored desirably with 16×16 mode and previous frame. Therefore, additional bits overhead for best mode and predicted frame for important MB in (Kang and Leou, 2005) are saved.

ii. *Class II MB:* The MB has a larger PECD value and $q_{mc} >> q_{res}$. In this situation, it is necessary to insert motion vector (MV) information to compensate the defect of error concealment. As in (Kang and Leou, 2005), 16 bits for entire MV of important MB is embedded in next P frame. Based on the method in (Lee et al., 2005), only 8 bits inserted into the video stream for residual MV is enough to regenerate the original MV.

iii. *Class III MB:* The MB has a larger PECD value and $q_{mc} << q_{res}$. In this situation, these MBs are usually intra-coded, and they have stronger correlation with spatial neighboring MBs than that of temporal neighboring MBs. Hence, spatial error concealment is desirable for this category of MBs. We use Soble operator to extract the direction of edge in them as embedded information to facilitate spatial error concealment where 4 bits are used to denote potential 16 possible directions.

iv. *Class IV MB:* The MB has a larger PECD value and ($q_{mc} > T_{mc}$ and $q_{res} > T_{res}$). In this situation, both the spatial and temporal neighboring MBs have certain correlation with this category of MBs. Certainly, corresponding residual MV and edge direction information should be inserted. With respect to the characters of error concealment method in (Al-Mualla et al., 1999), we extract following temporal and spatial correlation parameters q_{t_i} and q_{s_i} of i^{th} MB to enhance its multi-hypothesis error concealment compensation effect.

$$q_{t_i} = 1 - q_{mc_i} \tag{6}$$

$$q_{s_i} = 1 - \frac{\sum_{x=1}^{16}\sum_{y=1}^{16}\left|m_{sperc_i}(x,y) - m_{org_i}(x,y)\right|}{\sum_{x=1}^{16}\sum_{y=1}^{16} m_{org_i}(x,y)} \tag{7}$$

where $m_{sperc_i}(x,y)$ is the spatial pre-error concealed pixel value of the luminance component of i^{th} MB.

$$s_{t_i} = \frac{q_{t_i}}{q_{s_i} + q_{t_i}} \tag{8}$$

$$s_{s_i} = 1 - s_{t_i} \tag{9}$$

where s_{t_i} and s_{s_i} represent the weighted values of temporal and spatial error concealed pixel value of luminance component of i^{th} MB for multi-hypothesis error concealment compensation. The final reconstructed pixel value $m_{mul_i}(x,y)$ of luminance component in i^{th} MB is obtained by

$$m_{mul_i}(x,y) = s_{s_i} m_{sperc_i}(x,y) + s_{t_i} m_{tperc_i}(x,y) \tag{10}$$

where $m_{tperc_i}(x,y)$ equals to above $m_{prc_i}(x,y)$. Here, 4 bits are used for one of these two weighted values inserted into the video stream.

As a result, the total number of inserted bits for these four categories of MBs is only two bits, ten bits, six bits and eighteen bits respectively.

Because the methods in (Yilmaz and Alatan, 2003) and (Kang and Leou, 2005) use the strategy in (Yin et al., 2001) to embed extracted information, video quality and coding efficiency are greatly degraded when the original AC coefficients are modified. In this paper, we directly insert extracted information at the end of each video packet in next P frame with fixed length coding after all of MBs of this video packet have been encoded.

3.2 New adaptive transmission scheme for wireless packet video transmission

In (Yilmaz and Alatan, 2003) and (Kang and Leou, 2005), embedded important information is transmitted in wireless channel without considering whether the protected MBs are received correctly or not. With respect to the practical loss environment mentioned in (Katsaggelos et al., 2005), most of video packets in practical situation can be correctly received. Therefore, when one video packet is correctly received, it is not necessary transmit its protecting information in next frame. Based on the feedback of receiver, an adaptive transmission strategy that determines whether it is necessary to transmit embedded information for video packet is proposed here.

Using QCIF P frame as an example to describe this new transmission strategy as shown in Fig.2 (We assume that one video packet includes one row of MBs in each P frame), when k^{th} P frame arrives at receiver, receiver will find whether some video packets are lost in this frame, and give transmitter a nine bits feedback to represent the receiving situation of nine video packets of k^{th} P frame. If j^{th} ($1 \le j \le 9$) video packet is lost, j^{th} bit in feedback is set as 1 to denote j^{th} packet is lost, and other bits are zero in feedback when their corresponding video packets are correctly received. In the buffer of wireless video transmitter, there are two types of encapsulation for same video packets of $(k+1)^{th}$ P frame to be transmitted, one is TYPE-I packet including original video stream only, the other is TYPE-II packet including both the original video stream and the extracted important information for the video packet of the same position in k^{th} P frame. Hence, when feedback from receiver is 100000000, it means that only first video packet is lost, and others arrive at receiver correctly. Therefore, in wireless video transmitter, only the first video slice is transmitted by the TYPE-II packet followed by other eight TYPE-I video packets. This strategy can effectively mitigate the volume of unnecessary transmitted redundant information in wireless channel especially in low LPR situation. Here, we consider k^{th} P frame is lost when its feedback verification information cannot arrive at transmitter in time. So we have to adopt TYPE-II packet to transmit all of video packets of $(k+1)^{th}$ P frame in this situation.

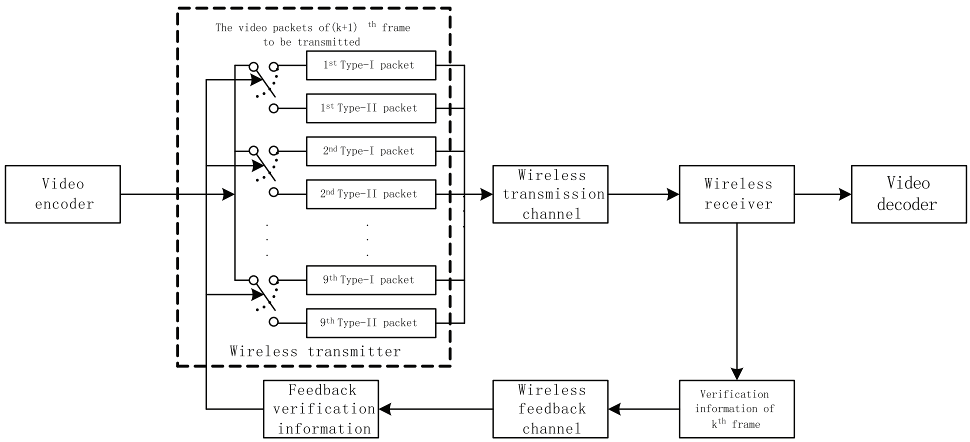
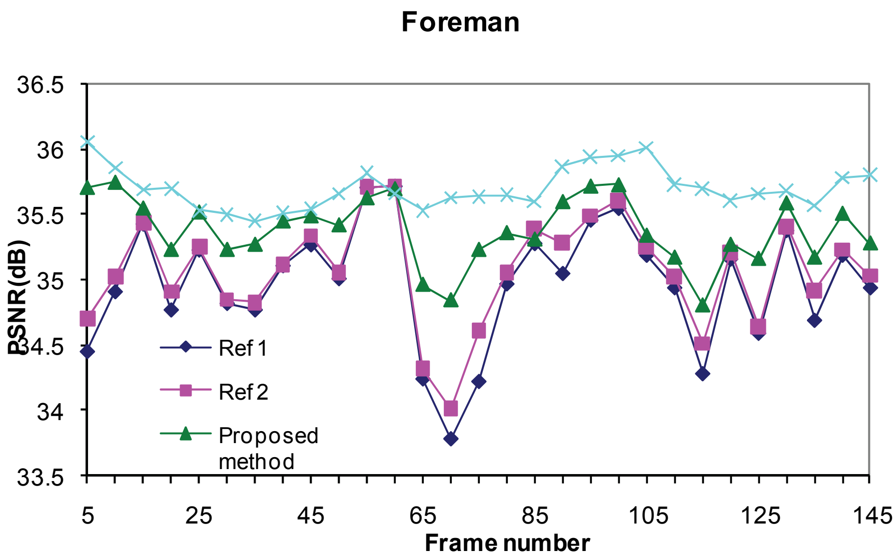


Fig. 2. Adaptive transmission strategy based on feedback information from receiver in QCIF video stream

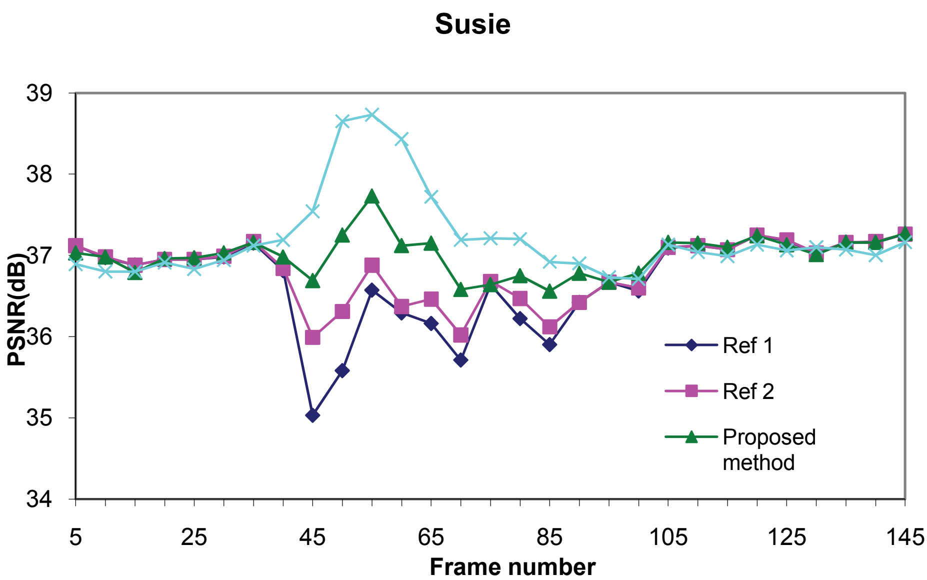
4. Simulation results

Computer simulations are performed in this section to evaluate the performance of the proposed error resilient scheme for video transmission in noisy environment. In our experiments, the H.264 JM 8.2 baseline codec is used, where the number of reference encoding frame is five. The frames from 5^{th} to 145^{th} frame in QCIF sequences "*Susie*" and "*Foreman*" are used for simulation. The video are encoded as IPPP…, and their QP is set at 28. The wireless channel adopted by simulation is random packet lost channel. We assume that each P frame is transmitted by nine video packets. For q_{mc} and q_{res}, they should be tuned according to real time video content. However, with respect to computation complexity, we assumed that they have fixed value. In simulation, both of them are set as 0.1. We compare our method with the data embedding based error resilient method (Kang and Leou, 2005) with Type-I data (denoted as Ref 1) and both Type-I and Type-II data (denoted as Ref 2).

We first look at the objective performance of error concealed effects in terms of the PSNR value of reconstructed pictures with a PLR as 0.1, which is illustrated in Fig.3. The PSNR result generated by an error free H.264 codec is also plotted in the figure, which is denoted as H.264 JM 8.2. It is seen that the proposed method always has a higher PSNR value than the other methods. For sequence "*Foreman*", which contains rather high motions, the proposed method has an obvious PSNR improvement with respect to Ref 1 and Ref 2, and its reconstruction effect is also more robust than other methods. In this scene, the average PSNR achieved by Ref 1, Ref 2, our proposed method and H.264 JM 8.2 is 34.97dB, 35.06dB, 35.37dB and 35.77dB respectively. We can find that the achieved PSNR distance between the proposed method and H.264 JM 8.2 is not very apparent. For sequence "*Susie*", which contains mostly medium motions, the proposed method can still obtain more desirable effect in the middle of selected range of video sequence than other methods, where there is apparent motion of human portrait. While at the beginning and end of selected range of video sequence, where high temporal correlation is in neighboring video frames, there are no apparent difference in reconstruction effect between the proposed method and reference



(a)



(b)

Fig. 3. Objective error concealed effect comparison between the proposed method and reference methods.

methods. Though the proposed method cannot always outperforms the reference methods in this video sequence, the average PSNR achieved by Ref 1, Ref 2, proposed method and H.264 JM 8.2 is 36.68dB, 36.79dB, 37.01dB and 37.21dB respectively. We can find that the PSNR value of the proposed is very close to that of H.264 JM 8.2. Hence, a better error concealment performance is obtained by our method as it can accurately extract visually important information to be protected.
After evaluating the error concealed effect in Fig.2, we then examine the coding rate requirement to see how many coding overhead is imposed by different error resilient schemes. The results for the two test sequences are shown in Table 1. In the table, only Ref 1 for the method in [9] is shown because it always needs fewer bits than Ref 2 (less data are embedded). It is seen that the coding rate increases for the proposed method are only 9% and 5%, while those of Ref 1 are about 23% and 16 %, for sequence "*Susie*" and "*Foreman*" respectively. It reveals that our method can obtain better coding efficiency as it directly inserts concise important data into the video stream.

Method	Susie	Foreman
H.264 JM8.2	104.73	142.32
Ref 1	128.79	164.78
Proposed method	114.24	149.98

Table 1. Coding rate (kb/s) comparison with the proposed method and comparison methods.

In the following, we apply the proposed adaptive transmission scheme in sending the video in noisy environment. We assume that both the forward channel and feedback channel are random pocket lost channels. We vary PLR in the forward channel and the feedback verification information loss probability (FVILP) in the feedback channel to see the effectiveness of the proposed scheme in saving coding overhead for the error resilient video stream. The results for the two sequences are shown in Table 2 and Table 3, for different PLR and FVILP settings. It is shown that the required transmission rate of the error resilient stream in different PLR can be significantly mitigated by the proposed video packet transmission strategy even in the loss of feedback verification information.

Method\PLR	0.05	0.1	0.15	0.2
H.264 JM8.2	104.73	104.73	104.73	104.73
Ref 1	128.79	128.79	128.79	128.79
Propose method (FVILP = 0)	106.16	106.64	107.12	107.59
Propose method (FVILP = 0.05)	106.56	107.02	107.48	107.92
Propose method (FVILP = 0.1)	106.97	107.4	107.83	108.25
Propose method (FVILP = 0.15)	107.37	107.78	108.19	108.59
Propose method (FVILP = 0.2)	107.77	108.16	108.54	108.92

Table 2. Required transmission rate (kb/s) comparison with the proposed method and comparison methods at different FVILP and PLR in Susie sequence.

Method\PLR	0.05	0.1	0.15	0.2
H.264 JM8.2	142.32	142.32	142.32	142.32
Ref 1	164.78	164.78	164.78	164.78
Propose method (FVILP = 0)	143.66	144.05	144.43	144.81
Propose method (FVILP = 0.05)	143.98	144.35	144.71	145.07
Propose method (FVILP = 0.1)	144.29	144.64	144.99	145.35
Propose method (FVILP = 0.15)	144.61	144.94	145.26	145.59
Propose method (FVILP = 0.2)	144.92	145.24	145.54	145.84

Table 3. Required transmission rate (kb/s) comparison with the proposed method and comparison methods at different FVILP and PLR in Foreman sequence.

And then, we look at the actual visual quality of the reconstructed pictures by different error resilient schemes. The reconstructed pictures of the 22nd frame of "*Foreman*" sequence for different methods are shown in Fig.4. It is seen that the fourth row MBs are lost, with respect to original picture shown by Fig.4(a), the proposed method in Fig.4(d) has more vivid subjective effect in human eyes area than Ref 1 and Ref 2 shown by Fig.4(b) and (c).

Finally, as we know, the proposed algorithm needs q_i , q_{mc} and q_{res} to classify MB. The increasing encoding time for this process is listed in Table 4 for H.264 JM 8.2 baseline profile. In our experiments, the tested hardware platform is DELL Latitude D820 Notebook PC (Intel Core2 Duo 2GHz, 1024 Memory), and the total encoding time of first 150 frames (only first frame is I frame) in *Susie* and *Foreman* video sequence is given. The additional encoding time for these three parameters is limited. The total encoding time increase for them in tested video sequences is only less than 1%.

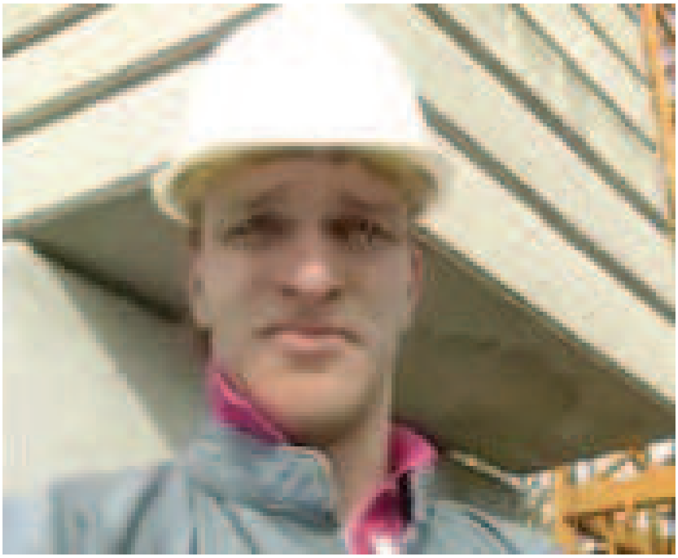
Sequence	Original	Proposed method	Increase %
Susie	111.58	112.48	0.81%
Foreman	109.17	109.98	0.74%

Table 4. Encoding time (s) comparison between original H.264 JM8.2 and proposed method

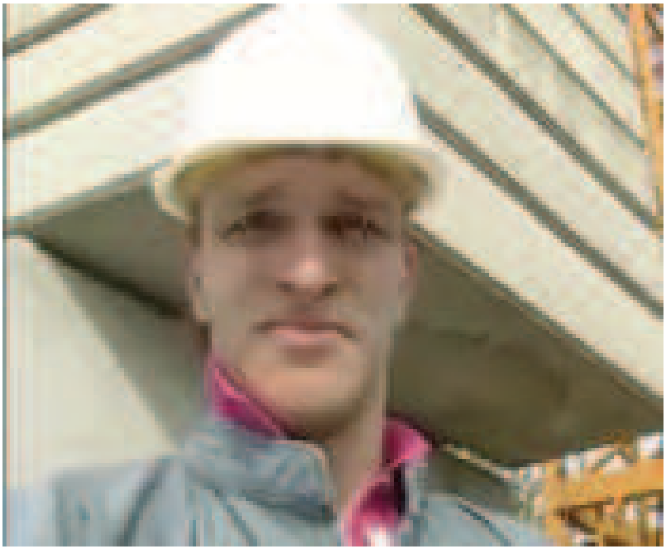
5. Conclusions

In this chapter, a review on various error resilient video coding techniques for wireless packet video transmission is given. Then a new error resilient method based on content analysis together with an adaptive transmission scheme are proposed for compressed video transmission over noisy environments. Simulation results show that the proposed method can obtain good error concealment effect with low redundancy by carefully extracting and inserting essential information in video encoder. It is also shown that the proposed adaptive transmission scheme can help further reduce the coding overhead.

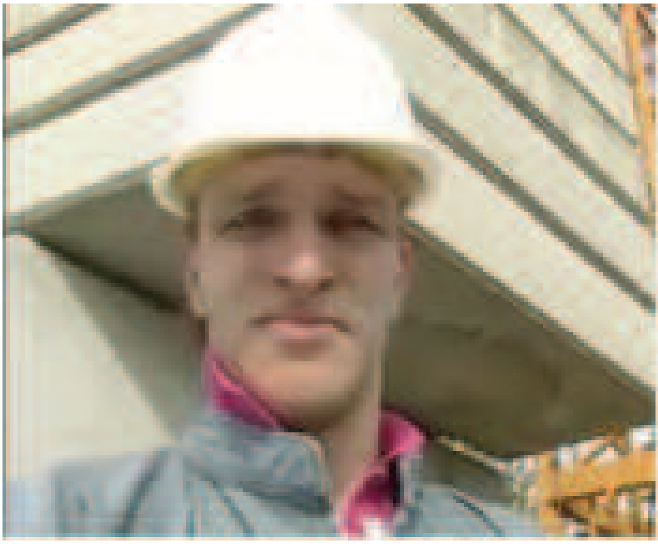
(a)



(b)



(c)



(d)

Fig. 4. The error concealment subjective effect comparison between the proposed method and comparison methods. (a) No error, (b) Ref 1, (c) Ref 2, (d) Proposed method.

6. Acknowledgements

The work of K.-T.Lo was supported by the Hong Kong Polytechnic University under Grant A/C 1-BB9G. The work of J. Feng was supported by the Hong Kong Baptist University under Grant No. FRG-06-07-II-06.

7. References

Al-Mualla, M., Canagarajah, N., and Bull, D.R. (1999). Temporal error concealment using motion field interpolation. *IEE Electronics Letters*, Vol.35, No.4, pp.215-217.

Chen, Y., Feng, J., Lo, K.T., and Zhang, X.D. (2008). Wireless Networks: Error Resilient Video Transmission. *Encyclopedia of Wireless and Mobile Communications*, pp.1434-1441, edited by B.Furht, CRC Press, ISBN 9781420043266, USA.

Etoh, M., and Yoshimura, T. (2005). Advances in wireless video delivery. *Proceeding of the IEEE.*, Vol.93, No.1, pp.111-122.

Jens-rainer, O. (2005). Advances in scalable video coding. *Proceedings of the IEEE*, Vol.93, No.1, pp.42-56.

Kang, L.W., and Leou, J.J. (2005). An error resilient coding scheme for H.264 video transmission based on data embedding. *Journal of Visual Communication and Inage Representation*, Vol.16, No.1, pp.93-114.

Katsaggelos, A.K., Eisenberg, Y., Zhai, F., Berry, R., and Pappas, T.N. (2005). Advances in efficient resource allocation for packet-based real-time video transmission. *Proceedings of the IEEE*, Vol.93, No.1, p.135-147.

Lee, J.S., Lim, J.W., Lee, K.H., and Ko, S.J. (2005). Residual motion coding method supporting error concealment in low bitrate video coding. *IEE Electronics* Letters, Vol.41, No.7, pp.406-408.

Stockhammer, T., Kontopodis, D., and Wiegand, T. (2002). Rate-distortion optimization for JVT/H.26L coding in packet loss environment. Proceedings of 2002 Packet Video Workshop, Pittsburgh, USA.

Stockhammer, T., Hannuksela, M.M., and Wiegand, T. (2003). H.264/AVC in wireless environments. *IEEE Trans. Circuits and Systems for Video Technol.*, Vol.13, No.7, pp. 657-673.

Stockhammer, T., and Hannuksela, M.M. (2005). H.264/AVC video for wireless transmission. *IEEE Wireless Communications,* Vol.12, No.4, pp.6-13.

Takishima, Y., Wada, M., and Murakami, H. (1995). Reversible variable length codes. *IEEE Transaction on Communications*, Vol.43, No.2, pp.158-162.

Tsekeridou, S., and Pitas, I. (2000). MPEG-2 error concealment based on block matching principles. *IEEE Trans. Circuits and Systems for Video Technol.*, Vol.10, No.6, pp.874-879.

Vetro, A., Xin, J. and Sun, H.F. (2005). Error resilience video transcoding for wireless communication. *IEEE Wireless Communications,* Vol.12, No.4, pp.14-21.

Villasenor, J., Zhang, Y.Q., and Wen, J. (1999). Robust video coding algorithms and systems. *Proceedings of the IEEE*, Vol.87, No.10, pp. 1724-1733.

Wang, Y., and Zhu, Q.F. (1998). Error control and concealment for video communication: a review. *Proceedings of the IEEE*, Vol. 86, No.5, pp.974-997.

Wang, Y., Wenger, S., Wen, J., and Katsaggelos, A.K (2000). Error resilient video coding techniques. *IEEE Signal Processing Magazine*, Vol.17, No.4, pp.61-82.

Wang, Y., Reibman, A.R., and Lin, S.N. (2005). Multiple descriptions coding for video delivery. Proceedings of the IEEE, Vol.93, No.1, pp.57-70.

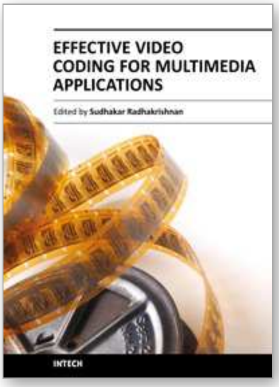
Wang, Z., Lu, L., and Bovik, A.C. (2003). Foveation scalable video coding with automatic fixation selection. *IEEE Trans. Image Processing*, Vol. 12, No.2, pp.243-254.

Wiegand, T., Sullivan, G.J., Bjontegaard, G., and Luthra, A. (2003). Overview of the H.264/AVC video coding standard. *IEEE Trans. Circuits and Systems for Video Technol.*, Vol.13, No.7, pp.560-576.

Yilmaz, A., and Alatan, A.A. (2003). Error concealment of video sequence by data hiding. *Proceedings of IEEE ICIP'2003,* Vol.3, pp.679-682, Barcelona, Spain.

Yin, P., Wu, M., and Liu, B. (2001). Error concealment using data hiding. *Proceedings of IEEE ICASSP'2001*, pp.1453-1456, Salt Lake City, Utah, USA.

Zeng, W. (2003). Spatial-temporal error concealment with side information for standard video codecs. *Procedings of ICME2003*, Vol.2, pp.113-116, Baltimore, Maryland, USA.



Effective Video Coding for Multimedia Applications

Edited by Dr Sudhakar Radhakrishnan

ISBN 978-953-307-177-0
Hard cover, 292 pages
Publisher InTech
Published online 26, April, 2011
Published in print edition April, 2011

Information has become one of the most valuable assets in the modern era. Within the last 5-10 years, the demand for multimedia applications has increased enormously. Like many other recent developments, the materialization of image and video encoding is due to the contribution from major areas like good network access, good amount of fast processors e.t.c. Many standardization procedures were carrried out for the development of image and video coding. The advancement of computer storage technology continues at a rapid pace as a means of reducing storage requirements of an image and video as most situation warrants. Thus, the science of digital video compression/coding has emerged. This storage capacity seems to be more impressive when it is realized that the intent is to deliver very high quality video to the end user with as few visible artifacts as possible. Current methods of video compression such as Moving Pictures Experts Group (MPEG) standard provide good performance in terms of retaining video quality while reducing the storage requirements. Many books are available for video coding fundamentals.This book is the research outcome of various Researchers and Professors who have contributed a might in this field. This book suits researchers doing their research in the area of video coding.The understanding of fundamentals of video coding is essential for the reader before reading this book. The book revolves around three different challenges namely (i) Coding strategies (coding efficiency and computational complexity), (ii) Video compression and (iii) Error resilience. The complete efficient video system depends upon source coding, proper inter and intra frame coding, emerging newer transform, quantization techniques and proper error concealment.The book gives the solution of all the challenges and is available in different sections.

Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821