

A Novel Fast Error-resilient Video Coding Scheme for H.264

Jiajun Bu, Linjian Mo, Genfu Shao, Zhi Yang, and Chun Chen

College of Computer Science

Zhejiang University

Hangzhou 310027, P. R. China

Email: {bjj, molin, andrew22, yangzh, chenc}@zju.edu.cn

Abstract—Both perceptually and statistically, compressed video with large or disordered motion is sensitive to errors. In this paper, we propose a novel fast error-resilient video coding scheme, which is based on significant macroblock (MB) determination and protection. The scheme takes three impact factors (inter-block mode, motion vector difference and SAD value) to build a statistical model. The model takes error concealment (EC) into consideration in advance and generates several parameters for further significant degree (SD) evaluation for MBs. During encoding, we build an SD table for each frame based on the parameters and pick up those MBs with the largest SD values as significant MBs (SMBs). Few additional computations are induced into SMB determination, thus make our scheme practical in real time video coding scenarios. Simulations show that the scheme has an acceptable SMB determination accuracy and the corresponding protection method can prevent errors effectively.

I. INTRODUCTION

H.264 is the latest and most advanced international video coding standard. It aims at not only improving coding efficiency, but also enhancing error robustness. With its outstanding performance compared with previous standards, H.264 has been widely used in Internet/wireless-based scenarios, such as video conference, video-on-demand, video surveillance and conversational services. Meanwhile, due to the extremely high demand of video data but the limited bandwidth for the underlying networks, packets that carry video data often suffer from loss. The lost packets affect the current frame as well as the subsequent frames due to error propagation, causing a degradation of video quality. So many error resilient schemes have been devised for H.264, e.g., intra block refreshing, unequal protection (UEP) and many error concealment schemes [1] [2] [3].

Region-based error resilient scheme has been widely adopted in static image communication to guarantee the high fidelity in the region of interest (ROI). It is newly applied in error resilience for video sequences recently. The idea behind error-resilient ROI video coding is that there exist one or more regions perceptually or determined more significant than the rest of the frame in many video scenarios like video surveillance, video conferencing and news broadcasting. A ROI-predefined error resilient scheme by employing a nonlinear transform that duplicates MBs inside the region of the ROI is proposed in [4]. This scheme can yield a high reconstruction of ROI without modification to the source codecs, but fails

to define ROI adaptively. Another method proposed by [5] introduces an adaptive ROI detecting. It exploits the perceptual importance of each MB by calculating mean absolute difference (MAD) between the pixels in the filtered version of current and latest encoded frames. A set of MBs corresponding to the ROI is then identified by a threshold. However, the scheme is restricted only for surveillance video which has most regions inactive.

Traditionally, ROI corresponds to the area with interesting contents by users' perception. However, in order to improve the overall video quality for general video scenarios with no obvious ROI, side information is usually exploited to determine significant data which will cause severe video quality degradation if corrupted. Then UEP-like methods will be enforced upon them. For error resilience purpose, the side information of MB's significance is usually estimated by considering video encoding with decoder's perspective. An SMB-determination-like method has been devised in [6]. Motion vectors (MVs) and concealment distortions have been used to evaluate the error propagation effect, and intra refresh protection is carried out on those MBs with the highest loss impact factors. This method is optimal to prevent error propagation, but it needs transcoding due to large computation. In [7], a similar significant MV (SMV) determination and protection method is proposed which selects SMVs by evaluating the absolute PSNR error decoding using only motion vector and decoding using error concealment. This method can be extended to protect not only SMVs but also significant residual errors.

Typically, for those MBs in terms of SMBs in this paper that will cause a large PSNR decrease, a large motion or a high visual complexity can be usually found on them. To evaluate the significant degree of an MB, three inter-predicted factors in a typical video coding standard including H.264/AVC [8], can be acquired after motion estimation (ME) and used as the contributing factors (CFs). The three CFs are block division mode, MVD and SAD, which are described as:

- Mode: representing the complexity of an MB;
- MVD: the difference between predicted motion vector (PMV) and the result MV estimated by ME;
- SAD: sum of absolute difference, which indicates the residual error between original frame and its reference frame.

In this paper, we propose a novel fast error-resilient video coding scheme for H.264. A Mode, MVD and SAD-based SMB determination model is built. The off-line training is adopted to get parameters of the model. The parameters include: 1) impact degrees (IDs), which represent how a possible value of a CF contributes to SD evaluation; 2) weights, which indicate how each CF contributes to SD evaluation. During encoding, three stages are then introduced for each frame to determine and protect SMBs: 1) SMB determination stage, 2) SMB protection (SMBP) stage and 3) SMB recovery stage. In SMB determination stage, a SD table is first built to evaluate the significant degree of each MB. Then we pick up those MBs with the largest SD values as SMBs. After that, many UEP schemes can be adopted to protect SMBs, such as intra MB refreshing and layered coding. Recent literatures have indicated multiple description coding (MDC) as a promising approach to handle video transmission errors [9]. In our proposed method, a simple but error-robust SMB-based polyphase down-sampler multiple descriptions (PDMD) scheme is adopted as SMB protection scheme [10]. It should be noted that only inter data of SMBs are protected in our proposed MDC scheme for channel bitrate consideration. For those insignificant MBs, only one description is allocated. Thus a few redundancies induced by PDMD achieve an acceptable EC ability even with some channel noises. Moreover, the proposed SMB determination method induces very few additional computations, thus make it possible to be adopted in most real-time scenarios and portable devices.

The remainder of this paper is organized as follows. In section 2, The parameters training of SMB determination model based on several different sequences is presented. In section 3, we give the detailed description for SMB determination and protection. Section 4 and 5 give the experimental results and conclusions, respectively.

II. PARAMETERS TRAINING OF SMB DETERMINATION MODEL

Subjectively, we have observed that if an MB has large or disordered motion, it is usually not easy to be concealed when corrupted. In this section, we demonstrate this subjective perception with statistics based on mode, MVD and SAD which are all ready after ME. Since PSNR can well depict the video quality which we concern most, we use it as the statistical measure in our model. The off-line training is adopted to get the parameters including ID values and weights.

For H.264 inter MB, there are seven different block division modes: $16*16$, $16*8$, $8*16$, $8*8$, $8*4$, $4*8$, and $4*4$. In our model, the seven modes are divided into two groups: 1) the first four modes are for $16*16$ MB while 2) the other three are for $8*8$ sub-block, as the two groups correspond to different block sizes and affect the PSNR degradation in different behaviors. Thus we get four CFs, $16*16$ MB mode, $8*8$ sub-block mode, MVD and SAD. If the sub-block division is detected for an MB, we then take the smallest division as the value of CF of sub-block mode. JM10.2 [11], the recommended H.264 reference software, is adopted in our

SMB determination model. The frame-level PSNR difference between in lossless environment and in condition that the evaluating MB is concealed is calculated as the PSNR decrease value. EC scheme here is modified from the default frame copy and MV copy EC scheme in JM10.2 to make it for MB level. During the parameters training of our SMB determination model, the result PSNR decrease value of each MB is taken as the input of the target value and the values of four CFs are the other inputs, as Fig.1. Estimated ID values and weights are output as the parameters for future use. The rest of this section describes the concrete method to define ID values and weighting factors.

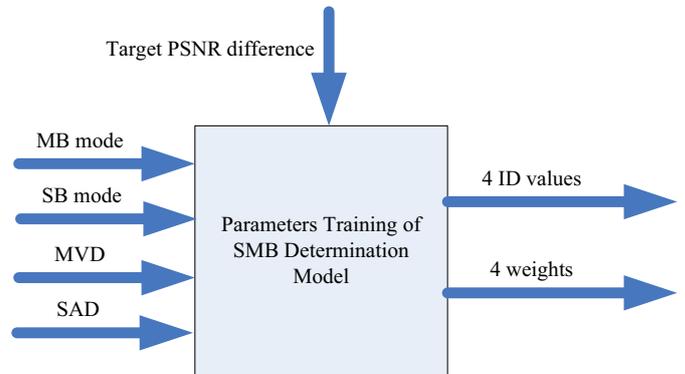


Fig. 1. Training of SMB determination model.

A. Impact Degrees Determination

1) *Average PSNR degradation calculation*: Six commonly used QCIF sequences, that is, "foreman" (400 frames), "akiyo" (90 frames), "silent" (300 frames), "news" (300 frames), "carphone" (90 frames) and mother&daughter (400 frames) are employed in our statistics. All of them are encoded with H.264 baseline profile, with QP=28, and fps=30. To evaluate how each CF contributes to determining SMBs, we operate statistical experiments on average PSNR degradation of each factor's value for these sequences. Three steps are needed. 1) We threshold all the possible values for each factor, based on both perceptual observation and statistical experiments. Then the classification for each group can be acquired. 2) For each classified value, we average the corresponding PSNR degradation. 3) In view of CF of $8*8$ sub-block mode, the average PSNR degradation of possible values should be reduced by the $8*8$ value of MB mode to avoid overestimating, because the sub-block modes are based on $8*8$ MB mode. Fig. 2 shows the average PSNR degradation of the classified values for each CF based on the six video sequences.

2) *Unify and Define Impact Degrees*: The average PSNR degradation statistical result above demonstrates how each possible value of the corresponding CF impact the PSNR degradation. However, we have also observed that, given different sequences, the same value generates different effects. Further more, in order to make our model compatible with all types of videos, we unify the average PSNR degradation

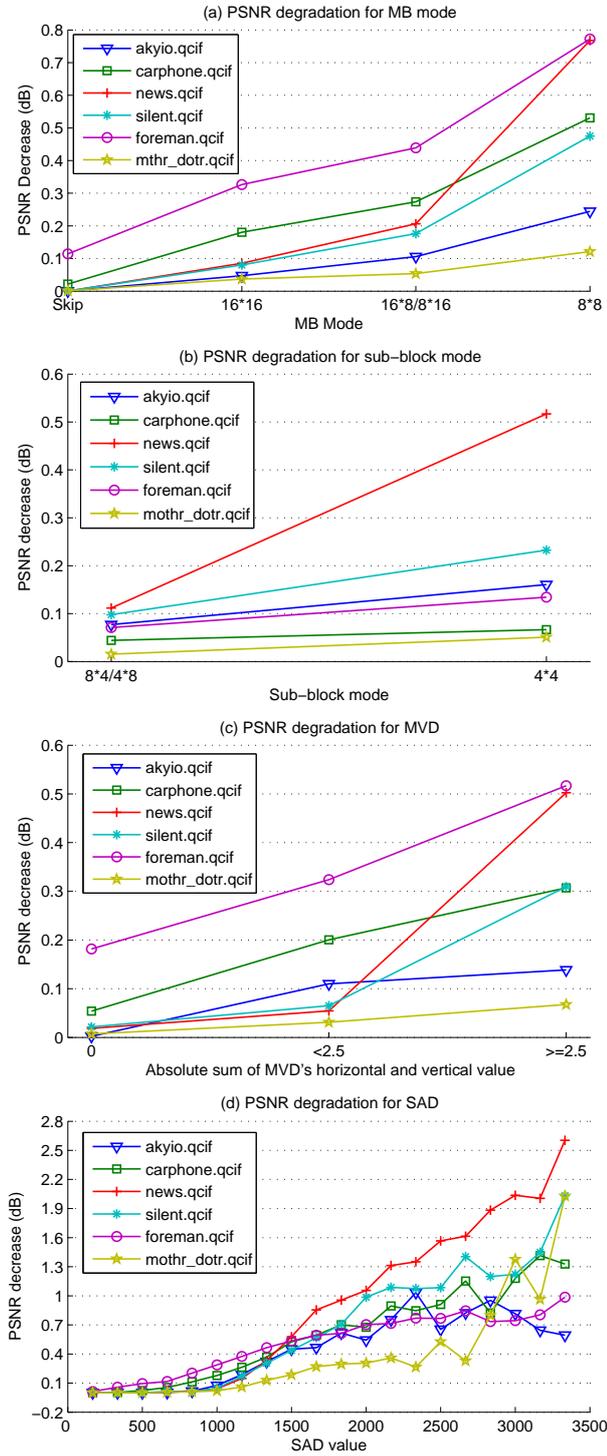


Fig. 2. The average PSNR degradation of for each CF, based on the six considered QCIF sequences: (a) for MB mode (The 16*8 and 8*16 modes are treated as the same value of MB mode); (b) for sub-block mode which has been reduced by average PSNR degradation of 8*8 MB mode (The 8*4 and 4*8 modes are treated as the same value of SB mode); (c) for MVD; (d) for SAD.

values of each CF for all sequences and refine them to [0, 1]. The unified IDs have been presented in Table.1, termed

in this paper as *ID Table* which will be further referenced to determine SMBs during SMB determination and protection process.

TABLE I
ID Table: THE CLASSIFIED VALUES OF EACH CF AND THEIR CORRESPONDING ID VALUES WHICH HAVE BEEN UNIFIED FOR EACH SEQUENCE.

Contributing Factors	Possible values (Classified)	ID values (Unified)
MB Mode	Skip	0.000000
	16*16	0.234509
	16*8/8*16	0.416249
	8*8	1.000000
Sub-block mode	No sub-block division	0.000000
	8*4/4*8	0.383990
	4*4	1.000000
MVD	0	0.000000
	≤2.5	0.402905
	2.5	1.000000
SAD	0–199	0.000000
	200–399	0.008963
	400–599	0.019046
	600–799	0.029335
	800–999	0.055977
	1000–1199	0.095919
	1200–1399	0.165315
	1400–1599	0.249909
	1600–1799	0.339742
	1800–1999	0.400320
	2000–2199	0.465979
	2200–2399	0.499504
	2400–2599	0.598390
	2600–2799	0.645016
	2800–2999	0.625197
	3000–3199	0.714499
3200–3399	0.716955	
3400–3599	0.804658	
3600–3799	0.797420	
≥3800	1.000000	

B. Weights Determination

After impact degree determination, we have got the unified ID values, leaving only the four weights not defined. Based on the observation on Fig. 2 and Table. 1, a simple linear relation between PSNR decrease and the four CFs is exploited. The weights denoted as $cf_weight = [cf_w_1, cf_w_2, cf_w_3, cf_w_4]^T$ for MB mode, SB mode, MVD and SAD, respectively, can then be got by (1).

$$\begin{cases} \min \sum_j |Unify(\Delta PSNR_j) - \sum_{cf.i=1}^4 (\omega_{cf.i} * ID_{cf.i,j})|^2 \\ s.t. \sum_{cf.i=1}^4 \omega_{cf.i} = 1; \quad 0 \leq \omega_{cf.i} \leq 1; \end{cases} \quad (1)$$

where $Unify()$ is the function used to unify a series of data; $\Delta PSNR$ represents the PSNR decrease value; $ID_{cf.i}$ and $\omega_{cf.i}$ are ID value and weight for the CF which indexed by $cf.i$; j is the MB from large amount of considered sequences. After training we obtain the four weights as $cf_weight = [0.0250, 0.7968, 0.0106, 0.1676]^T$.

III. SMB DETERMINATION AND PROTECTION

In previous section, our proposed SMB determination model has been constructed. Now we will adopt it to pick up SMBs during encoding for UEP purpose. This section presents three detailed stages of SMB determination and protection.

A. Stage 1: SMB Determination

Now the ID_Table and determined weights defined by the proposed parameters training of SMB determination model can be used to evaluate the SD of each MB. For MB i in current frame, SD value can be evaluated as,

$$\begin{cases} SD(i) = \sum_{cf.i=1}^4 [cf \cdot \omega_{cf.i} * ID_{cf.i}(i)] \\ 0 \leq i \leq PicSizeInMBs \end{cases} \quad (2)$$

where ID values can be defined by ID_Table , and $PicSizeInMBs$ is the total MB number of a frame.

An SD table is then generated for each frame. Since the SD value of each MB corresponds to the estimated loss impact of that MB in our proposed method, we could just pick up those MBs with the largest SD values as SMBs, and protect certain amount of them based on the bandwidth available for SMBP.

B. Stage 2: SMB Protection

MDC approach has been proved efficient to prevent errors during video transmission. In our proposed method, a simple but error-robust PDMD coding approach is introduced as our protecting scheme. Only those MBs determined as SMBs are protected in PDMD. The scheme has been presented in Fig.3.

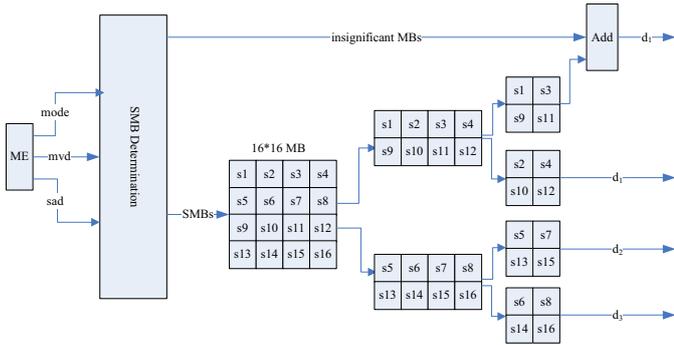


Fig. 3. The proposed polyphase down-sampler scheme for SMBP. Each SMB is partitioned into sixteen 4*4 sub-blocks, and protected in four descriptions. No vertical- or horizontal-adjacent 4*4 sub-blocks are in the same description so that the 4*4 sub-blocks of the lost descriptions can be easily interpolated from those well received ones.

C. Stage 3: SMB Recovery

When errors are detected at decoder side, EC tools are triggered to reduce the infection of errors. In this paper, EC is done based on the detection of whether the corrupted MB is an SMB or not:

- If the MB is not an SMB, the modified EC method is adopted.

- If the MB is an SMB, an inverse PDMD approach will help interpolate the corrupted 4*4 sub-blocks corresponding to lost descriptions spatially.

IV. EXPERIMENTAL RESULTS

Several QCIF video sequences at 30 fps are simulated in our experiments, including the six sequences we referenced in our SMB determination model, and a seventh sequence "salesman". The GOP structure is "IPPP.." and the test platform is JM10.2. First the accuracy of the SMB determination is evaluated under different SMB percentage consideration. Then, we present the PSNR gain based on our proposed SMB protection scheme.

SD values can be calculated using (2) and one SD table will be built for each frame to determine SMBs. Since our proposed SMB determination method is all-sequence compatible, the accuracy to determine SMBs for different types of videos is extremely important. There are two ways to evaluate the accuracy of our SMB determination method: 1) the percentage of match between SMBs determined by our proposed method and SMBs with the largest $\Delta PSNRs$; 2) the average of $\Delta PSNR$ of SMBs, compared with the average $\Delta PSNR$ of the rest MBs in the same frame. Fig. 4 (a) shows the average match percentage of SMB determination for the six tested sequences. The rule is exploited that for videos with low motion, such as "akiyo", "news" and "silent", the SMB determination gets its best accuracy of about 70% in [10%, 20%] proportion of SMBs protection in a frame, while for videos with high motion, the more proportion of SMBs we protect, the higher accuracy we get. The average difference between $\Delta PSNR$ of SMBs and $\Delta PSNR$ of insignificant SMBs is shown in Fig.4 (b). Since the PSNR degradation for sequences "akiyo" and "mthr_dotr" in Fig. 2 is low as well as "salesman", the average PSNR difference in Fig. 4 (b) is relevantly lower than other sequences.

At last, the performance of the proposed method in terms of PSNR gain is evaluated by testing on "foreman" and "silent". The modified EC scheme, the random MB protection (RMBP) by our proposed method, and random intra refresh (RIR) with different Intra MB proportion per frame are adopted to be compared with. For SMBP, RMBP and 10% MB RIR schemes, 10% MBs are protected. Another RIR simulation with 5% intra rate is also operated for future comparing. Fig. 5 plots the PSNR curves of "foreman" and "silent" with different MB loss rates. Compared with RMBP, the result for SMBP that average 0.07 dB (at loss rate 0.1) and 0.46 dB (at loss rate 0.4) PSNR gain in average for "foreman" and 1.67 dB (at loss rate 0.1) and 2.87 dB (at loss rate 0.4) for "silent" is observed to show the high accuracy of our SMB determination method. Though 10% MB RIR protection scheme achieves higher average PSNR than our proposed method a little but at no more than 0.9 dB in all cases, it scarifies with a much higher bitrate of about 2 times of SMBP. A more fairly comparison with 5% MB RIR whose bitrate is about as 1.5 times as SMBP is carried out. The result shows that at loss rate 0.1, our proposed method achieves a more or less average PSNR

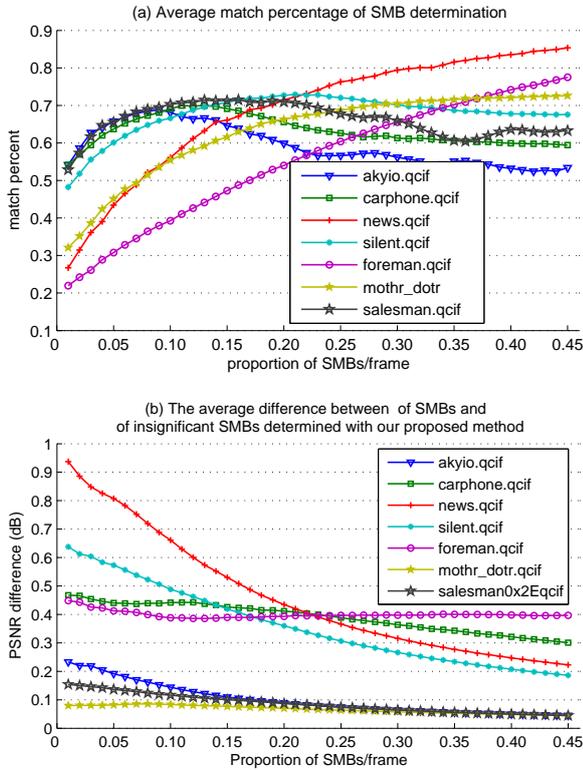


Fig. 4. Two ways to evaluate the accuracy of the proposed SMB determination method: (a) by the match percentage; (b) by $\Delta PSNR$ evaluation.

than RIR (5% MB), but outperforms it with 0.5 dB and 0.3 dB at loss rate 0.4, for "foreman" and "silent", respectively.

V. CONCLUSIONS

In this paper, we propose fast error resilient video coding scheme for H.264. We adopt the idea of considering from decoder's perspective by exploiting the side information at encoder side in an SMB determination model and describe how each CF affects to determine SMBs. Three steps are introduced in encoding and decoding process for SMB determination and SMBP. Experimental results show that SMB determination model can pick up those MBs with large PSNR degradation if concealed fairly well, thus the protection upon them can achieve great improvement on video quality. Future work will focus on whether we can take advantages of temporal correlation between current frame and previous frames to determine SMBs before ME, and thus make more UEP schemes practical in the protection stage.

ACKNOWLEDGMENT

The work was supported by National Natural Science Foundation of China (60573176), Key Technologies R&D Program of Zhejiang Province (2005C23047).

REFERENCES

[1] S. Kumar, L. Xu, M. K. Mandal, and S. Panchanathan, "Error resiliency schemes in h.264/avc standard," vol. 17, April 2006.

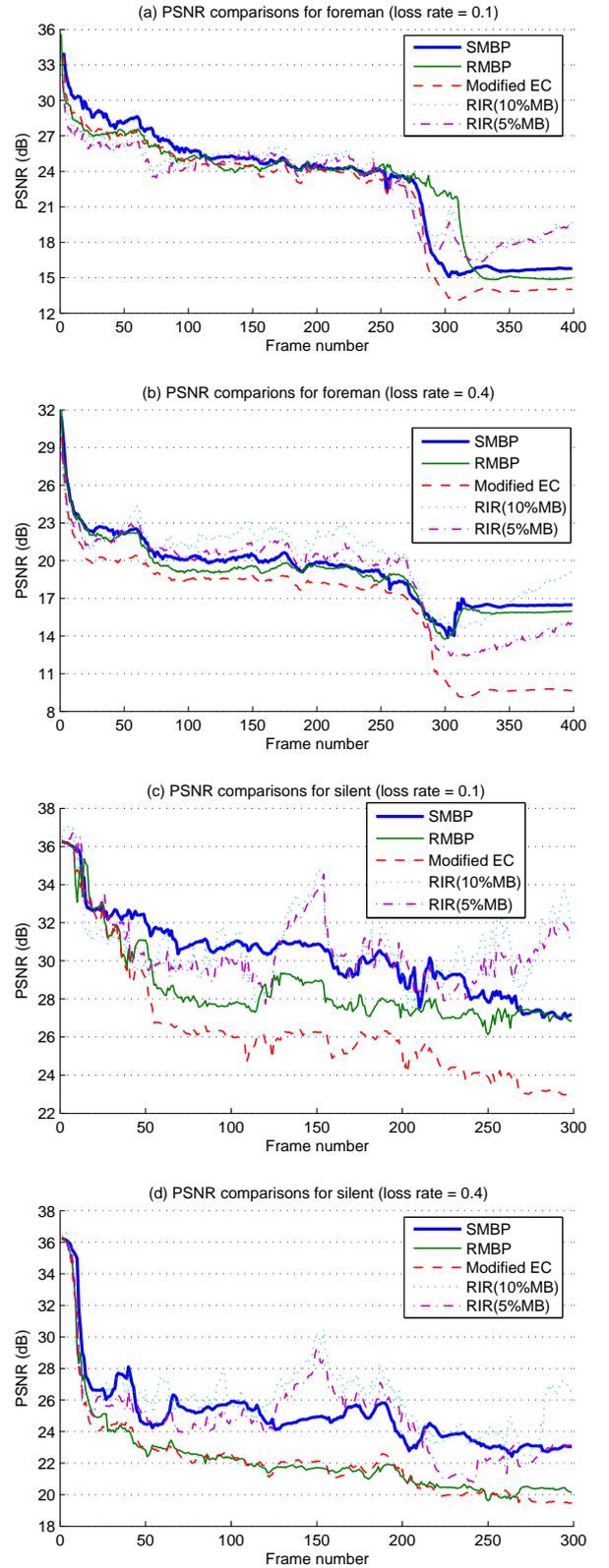


Fig. 5. Comparisons between different error control schemes for videos of "foreman" and "silent" with different loss rate.

[2] T. Halbach and S. Olsen, "Error robustness evaluation of h.264/mpeg-4 avc," vol. 5308, 2004, pp. 617–627.

- [3] L. Su, Y. Zhang, W. Gao, Q. Huang, and Y. Lu, "Improved error concealment algorithms based on h.264/avc non-normative decoder," 2004, pp. 1671–1674.
- [4] A. Jerbi, J. Wang, and S. Shirani, "Error-resilient region-of-interest video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, pp. 1175–1181, Sept. 2005.
- [5] P. Baccichet, X. Zhu, and B. Girod, "Network-aware h.264/avc region-of-interest coding for a multi-camera wireless surveillance network," April 2006.
- [6] H. Chiou, Y. Lee, and C. Lin, "Error-resilient transcoding using adaptive intra refresh for video streaming," vol. 3, May 2004, pp. 777–780.
- [7] J. Chen, C. S. Lu, and K. C. Fan, "A significant motion vector protection-based error-resilient scheme in h.264," 2004, pp. 287–290.
- [8] J. V. T. J. of ISO/IEC MPEG and I.-T. VCEG, "Draft itu-t recommendation and final draft international standard of joint video specification (itu-t rec. h.264 - iso/iec 14496/10 avc-jvt-g050r1.doc)," *ISO/IEC MPEG and ITU-T VCEG*, 2003.
- [9] Y. Wang, A. R. Reibman, and S. Lin, "Multiple description coding for video delivery," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 57–70, Jan 2005.
- [10] N. Franchi, M. Fumagalli, R. Lancini, and S. Tubaro, "Multiple description video coding for scalable and robust transmission over ip," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 15, no. 3, pp. 321–334, May 2005.
- [11] "<http://iphome.hhi.de/suehring/tml/download/jm10.2.zip>."