

ERROR-RESILIENT TRANSCODING USING ADAPTIVE INTRA REFRESH FOR VIDEO STREAMING

Hong-Jyh Chiou, Yuh-Ruey Lee, and Chia-Wen Lin

Department of Computer Science and Information Engineering,
National Chung Cheng University, Chiayi, Taiwan, R.O.C.

E-mail: cwlin@cs.ccu.edu.tw

ABSTRACT

Transmitting video data over wireless networks can be very unreliable due to packet loss, leading to serious video quality degradation which is annoying to human perception. The error due to lost packets not only affects the quality of current frame, but also propagates to subsequent frames due to the motion compensation technique used in standard video codecs. Adding error-resilience to video bitstream for robust video delivery to users thus becomes a very important issue. In this work, we propose a two-pass transcoding scheme for enhance the capability of error-resilience in the media gateway of a three-tier streaming system. The transcoder can adapt intra-refresh rate according to a channel's packet loss rate, hence protect the most important MBs subjected to packet loss. Experimental results show that the proposed method can effectively mitigate the error propagation due to packet loss and improve the video quality significantly.

1. INTRODUCTION

With the rapidly growing demand and widely deployed infrastructure of wireless networks (e.g., GPRS, 3G, and wireless LANs), applications of streaming video over wireless links have attracted much attention in recent years. However, the packet erasure and bandwidth variation characteristics of wireless links still present a number of challenges to streaming video applications. In a video streaming system, a server pre-stores encoded video streams and transmits them to clients for decoding and playback. There are several existing video coding techniques developed, for example, H.26x, and MPEG-x, to compress video sequences to bitstreams. These video encoding techniques exploit spatial and temporal redundancy to achieve a high compression ratio, but they are usually unaware about the channel conditions during the coding process, especially for prestored video streaming applications. Therefore, transmitting video data over wireless networks can be very unreliable due to packet loss. This packet loss problem may lead to serious video quality degradation which is annoying to human perception. The lost packets not only affect the quality of current frame, but also lead to error propagation to subsequent frames due to the motion compensation technique used in standard video codecs. In order to achieve error robustness for transmitting video over wireless networks, one must be able to adapt or transcode the compressed video streams into error-resilience-capable streams at the intermediate network node.

We consider the application scenario of three-tier video streaming system for home networking, which involves a

streaming server, a media gateway (i.e., home server), and client terminals (e.g., information appliances). In a home network, the communication channels to client terminals may have different packet-loss rates and available bandwidths for different clients. The home server has to deploy different error-resilience features and regulate the bitrates of coded bitstreams to match different dynamic channel characteristics. A transcoder is usually located at the home server to serve this purpose. Using a transcoder to handle the different demands (e.g., bandwidth, resolution, frame-rate, and channel condition) from different information appliance (IA) devices can reduce the complexity and transmission cost from the ISP streaming server to the home receivers. We adopt the three-tier architecture with the error-resilient transcoder located at the home server for the addition of error robustness to video streams prior to delivering video data to the users.

There have been some research works on error-resilient transcoding. In [1], an error-resilient transcoding scheme is proposed for general packet ratio services (GPRS) mobile-access networks with the transcoding process taking place at a video proxy which can be located at the edge of two or more networks. Two error-resilience coding tools, the AIR (Adaptive Intra Refresh) and RFS (Reference Frame Selection) schemes with FCS (Feedback Control Signaling), are exploited to reduce error effects when users playback video. The method in [2] develops a rate-distortion framework with analytical models that characterize the error propagation of corrupted video bitstream subject to bit errors, thereby optimizing the combination of spatial error-resilience, temporal error-resilience, and transmission bitrate. However, this scheme is computationally very expensive for real-time video streaming applications.

In this work, we propose a two-pass error-resilient video transcoding scheme with a transcoder located at the media gateway of the three-tier architecture. Intra-refresh is adopted in our proposed transcoder as the error-resilience coding tool because it does not need to make any change for standard video decoders. The proposed transcoder can adaptively vary the intra-refresh rate according to the channel's packet-loss rate to protect the most important macroblocks (MBs) against packet loss. In the first-pass encoding, the encoder estimates the error propagation effect of each MB while getting lost, and then generates side information as transcoding hints for further use at the transcoder. In the second-pass transcoding, the error-resilient transcoder adaptively determines the intra-refresh rate and the locations of MBs to perform intra-refresh according to the transcoding hints.

2.3. Transcoding using adaptive intra-refresh

In the second-pass transcoding, we propose a priority-based intra-refresh scheme to determine the intra-refresh rate and the intra-block allocation strategy for each GOP to adapt to the varying network environments. One key issue of the intra-refresh algorithm is to determine the number of MBs to be intra-coded in a GOP, which is determined in our work using Eq. (5).

$$N_{\text{intra}}^{\text{GOP}} = \frac{1}{N_{\text{GOP}}} \sum_{n=1}^{N_{\text{GOP}}} FEP^n \times PLR \quad (5)$$

where $N_{\text{intra}}^{\text{GOP}}$ is the number of MBs to be intra-coded in a GOP; N_{GOP} is the GOP size; PLR is the packet-loss rate which can be obtained via the receiver RTCP reports from the clients through a feedback channel (see Fig. 2); T is the parameter which is used to characterize the relationship between $N_{\text{intra}}^{\text{GOP}}$ and the error-propagation effect in a GOP and can be obtained empirically. The intra-refreshed MBs are then distributed to a GOP using the following algorithm.

if $n = 1$

$$N_{\text{intra}}^n = \frac{FEP^n}{\sum_{i=1}^{N_{\text{GOP}}} FEP^i} \times N_{\text{intra}}^{\text{GOP}} \quad (6)$$

else if $2 \leq n \leq N_{\text{GOP}}$

$$N_{\text{intra}}^n = \frac{FEP^n}{\sum_{i=1}^{N_{\text{GOP}}} FEP^i} \times \left(N_{\text{intra}}^{\text{GOP}} - \sum_{i=1}^{n-1} N_{\text{intra}}^i \right) \quad (7)$$

endif

if $N_{\text{intra}}^n > N_{\text{MB}}^f$ then $N_{\text{intra}}^n = N_{\text{MB}}^f$

where n is the frame index in a GOP; N_{intra}^n is the number of MBs to be intra-coded in frame n .

For the n th frame of a GOP, we select a total of N_{intra}^n MBs with top-ranked MBEP values to perform intra-refresh according to the ranking information. Probability P is updated every GOP. This can ensure that our method can adapt to frequently changing network conditions.

3. EXPERIMENTAL RESULTS

Three 300-frame CIF (352x288) test sequences, ‘‘Foreman,’’ ‘‘Salesman,’’ and ‘‘Dancer’’ are pre-encoded at 30 fps and 384 Kbps. The group of picture (GOP) size is $(N_{\text{GOP}}, M) = (30, 1)$, where M is the number of B-frames between two I/P-frames. In this work, we use a two-state Markov model to simulate the channel conditions. We adopt a simplified Gilbert channel at the packet level [5,6] to generate the packet-loss patterns with four packet loss rates (PLR): 5%, 10%, 15%, and 20%, respectively. The proposed method is compared with random intra-refresh [7], regular intra-refresh [7] and CBERC [8]. Suppose the average number of intra-refreshed MBs in a frame is m , for the random intra-refresh scheme, the intra-refreshed positions are randomly selected independently for each frame. For regular intra-coding, the intra-refreshed positions are $1 \sim m$ in the first frame, $m+1 \sim 2m$ in the second frame, and so on. If all MB positions have been refreshed once, the first positions are intra-refreshed again. In the proposed method, the parameter of intra-refresh rate, T , in Eq.

(5) is determined empirically. Fig. 3 shows the frame-by-frame PSNR with different T for three test sequences. We can observe that $T = 1200$ almost achieves best performance for all sequences.

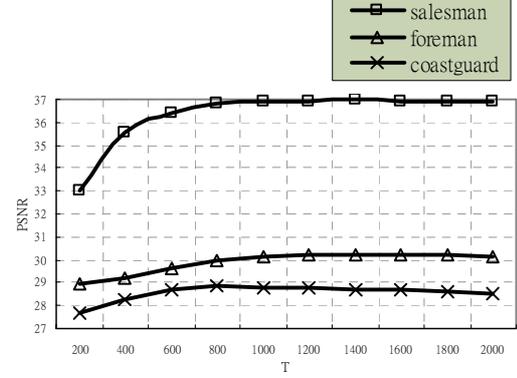
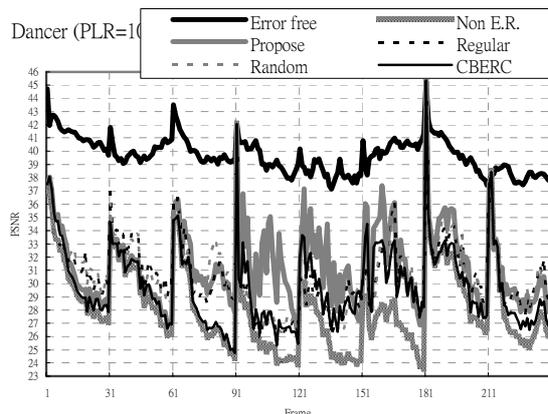


Fig. 3. Effect of changing the intra-refresh parameter T ($PLR = 10\%$).

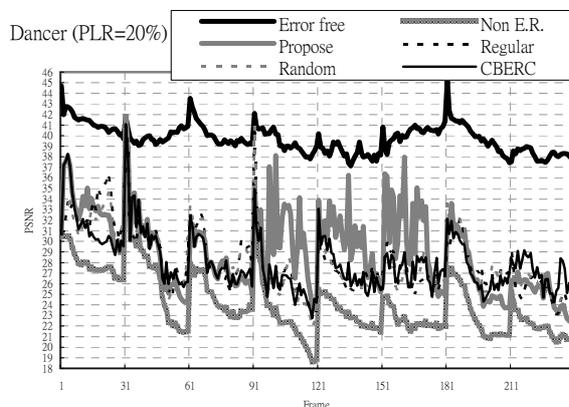
Table 1. Average PSNR comparison between different intra-refresh schemes for three test sequences (in dB)

Sequence	Method	PLR			
		5%	10%	15%	20%
Foreman	Error Free	35.80			
	Non-E.R.	29.73	26.63	25.37	22.95
	Proposed	31.51	30.25	29.21	28.46
	Regular	30.30	28.88	27.84	26.99
	Random	30.29	28.78	27.73	26.85
	CBERC	31.14	30.12	29.20	28.47
Salesman	Error Free	39.81			
	Non-E.R.	37.74	35.76	34.56	32.98
	Proposed	37.85	36.94	36.15	35.14
	Regular	37.61	36.28	34.89	33.92
	Random	37.54	36.27	34.79	33.82
	CBERC	37.79	36.74	35.84	35.10
Dancer	Error Free	39.71			
	Non-E.R.	30.87	28.25	26.51	24.2
	Proposed	32.09	31.30	30.29	29.12
	Regular	32.00	30.56	29.33	28.08
	Random	31.94	30.62	29.07	28.02
	CBERC	31.23	29.79	28.60	28.11

The average PSNR performance comparison with different methods and channel conditions is shown in Table 1. Fig. 4 shows the frame-by-frame PSNR performance comparison of the proposed method with the other three intra-refresh methods for the ‘‘Dancer’’ sequence. Since I-frames are important in terminating all error propagation, a corrupted I-frame may cause serious error propagation. Furthermore, MBs closer to the I-frame in a GOP are usually more important than other farther MBs. This is because, if these important MBs are corrupted, the resultant error propagation will be more serious, since the drifting error will propagate to more frames until reaching the next I-frame. This kind of effects can be observed obviously in Fig. 4, where large PSNR drops can be observed because many of MBs of more importance are corrupted. The proposed method can mitigate the error propagation due to the loss of important packets (e.g., packets close to I-frames) more effectively than the other methods.



(a) $PLR = 10\%$



(b) $PLR = 20\%$

Fig. 4. PSNR performance comparison using four intra-refresh methods under two different channel conditions (Dancer).

Table 2 shows the run-time analysis of the first-pass encoding and second-pass transcoding on an Intel Pentium-III 1 GHz PC. The encoder and transcoder are implemented based on MPEG-4 public domain software. With the proposed loss-impact estimation method, the first-pass encoding consumes significantly more time than the original one. On the other hand, the proposed method does not increase the computational complexity of the second-pass transcoding operation. Actually sometimes it consumes even less time than the original transcoder for two reasons. First, the computation for intra-refresh decision in Eqs. (5)-(7) in the second-pass transcoding is almost negligible compared to the whole transcoding process. Second, the error-resilient transcoding will increase the number of intra-coded MBs, thereby reducing the computation since the computational cost for intra-coding is much lower than that for inter-coding.

Table 2. Run-time analysis of the first-pass encoding and second-pass transcoding

Sequence	Encoding time		Transcoding time	
	original	proposed	Non-error resilient	Error-resilient
Foreman	11.0 s	23.7 s	15.1 s	14.5 s
Coastguard	11.1 s	23.7 s	14.9 s	14.7 s
Dancer	11.0 s	23.7 s	12.2 s	12.2 s
Salesman	11.2 s	23.7 s	15.2 s	15.0 s

5. CONCLUSIONS

In this paper, we proposed a novel two-pass error-resilient transcoding scheme by using a priority-based intra-refresh strategy. The first-pass encoding process of the proposed method estimates the pixel-wise mismatch error using an error-tracking technique. The estimated pixel-wise mismatch error is subsequently used to estimate the amount of error propagation (the loss-impact factor) caused by a lost MB from a frame to the following frames. After sorting the impact factors of all MBs in a GOP, the rank of loss-impact of each MB as well as its impact factor is stored in the streaming server as side information for the future transcoding.

In the second-pass transcoding, the home server first determines the intra-refresh budget according to the prestored information of frame-level loss-impact values and the packet-loss rate of the client, and subsequently use the side information to distribute the intra-refresh budget into each frame in a GOP. The extra computational complexity required for this computation is almost negligible, thereby making it suitable for real-time applications.

The proposed algorithm can effectively mitigate the error propagation due to packet-loss and improve the quality significantly. The degree of error resilience can be dynamically adjusted to adapt to a channel with time-varying error characteristics which can be estimated using the statistics collected through the feedback channel.

REFERENCES

- [1] S. Dogan, A. Cellatoglu, M. Uyguroglu, A. H. Sadka, A. M. Kondo, and A. M. Kondo, "Error-resilient video transcoding for robust internetwork communications using GPRS," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 453-464, June 2002.
- [2] G. de los Reyes, A. R. Reibman, S. F. Chang, and J. C. I. Chuang, "Error-resilient transcoding for video over wireless channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1063-1074, June 2000.
- [3] T. Warabino, S. Ota, D. Morikawa, M. Ohashi, H. Nakamura, H. Lwashita, and F. Watanabe, "Video transcoding proxy for 3Gwireless mobile internet access," *IEEE Commun. Mag.*, Vol. 38, pp. 66-71, Oct. 2000.
- [4] W. M. Lam, A. R. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Apr. 27-30, 1993, pp. V417-V420, Minneapolis, MN.
- [5] M. Zorzi, R. R. Rao, and L. B. Milstein, "ARQ error control for fading mobile radio channels," *IEEE Trans. Veh. Technol.*, Vol. 46, pp. 445-455, May 1997.
- [6] E. N. Gilbert, "Capacity of a burst-noisy channel," *Bell Syst. Tech. J.*, Vol. 39, pp. 1253-1265, Sept. 1960.
- [7] G. Cote and F. Kossentini, "Optimal intra coding of blocks for robust video communication over the internet," *Signal Processing: Image Commun., Special Issue on Real-time Video over Internet*, vol. 15, pp. 25-34, Sep. 1999.
- [8] W. J. Chen and J. N. Hwang, "The CBERC: a content-based error-resilient coding technique for packet video communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp. 974-980, Aug. 2001.