ERROR RESILIENCE TRANSCODING USING PRIORITIZED INTRA-REFRESH FOR VIDEO MULTICAST OVER WIRELESS NETWORKS[+]

Chih-Ming Chen[1,3], Yuh-Ruey Lee[2], Chia-Wen Lin[2], and Yung-Chang Chen[1]

[1]Department of Electrical Engineering, National Tsing Hua University,
Hsinchu, Taiwan 30013, R.O.C., ycchen@ee.nthu.edu.tw
[2]Department of Computer Science & Information Engineering, National Chung Cheng University,
Chiayi, Taiwan 621, R.O.C., cwlin@cs.ccu.edu.tw
[3]Chunghwa Telecom Co., Ltd.

ABSTRACT

In this paper, we propose a two-pass intra-refresh transcoding scheme for inserting error-resilience features to a compressed video at the media gateway of a three-tier streaming system. The proposed transcoder can adaptively vary the intra-refresh rate according to the video content and the channel's packet-loss rate to protect the most important macroblocks (MBs) against packet loss. In this work, we consider the problem of multicast of video to multiple clients having disparate channel loss profiles. We propose a minmax loss-rate estimation scheme to select a single intra-refresh rate for all the clients. Experimental results show that the proposed method can effectively mitigate the error propagation due to packet loss, and its fairness for multicast.

1. INTRODUCTION

Transmitting video data over error prone networks can be very unreliable due to packet-loss, and 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 client terminals for decoding and playback. There are several existing video coding techniques developed to compress video sequences into bitstreams to reduce the data sizes. These video encoding techniques exploit spatial and temporal redundancy to achieve a high compression ratio, while making the compressed data very sensitive to transmission error. This packet-loss problem may lead to serious video quality degradation, which not only affects the quality of current frame, but also leads to error propagation to subsequent frames due to the motion-compensated prediction technique used in standard video codecs. Furthermore, the heterogeneity of client networks also makes the encoder very difficult to adapt the video contents to a wide degree of different client channel conditions, especially for wireless client terminals. In order to achieve error robustness for transmitting video over wireless networks, the server must be able to adapt or transcode the non-error-resilient compressed video streams into error-resilience-capable streams at the intermediate network node. To serve this purpose, a video transcoder [1-3] can be placed in a network node connected to a high-loss network to insert error-resilience features into the video bitstream to achieve robust video transmission over wireless channels [4-7]. In our previous works [8], a novel two-pass error-resilient transcoding scheme by using prioritized intra-refresh was proposed. Like as adaptive intra refresh (AIR) in the MPEG4 standard, it does not need to make any change for standard video decoders, which is important in terms of cost and convenience for many practical applications.

In this paper, we propose a method that adopts a minmax penalty criterion and the prioritized intra-refresh strategy to solve the problem of multicast of a single stream to multiple receivers having diverse channel loss characteristics.

The remainder of this paper is organized as follows. Our previously proposed two-pass error-resilience transcoding scheme is briefly reviewed in Section 2. A minmax-based error resilience transcoding strategy for video multicast in heterogeneous environments is proposed in Section 3. Section 4 show experimental results. Finally the conclusion is drawn in Section 5.

2. ERROR RESILIENCE TRANSCODING USING PRIORITIZED INTRA-REFRESH

2.1. Two-Pass Error-Resilience Transcoding

At the first-pass front-end encoding of the proposed two-pass error-resilience transcoder architecture, the encoder utilizes the motion vectors and the estimated concealment distortion to estimate the error-propagation effect at the MB and frame levels within a group of pictures (GOP) as the side information which is stored in the streaming server to be used as hints to guide the error-resilience transcoding operation. In the second-pass transcoding process, the transcoder uses the side information received from the streaming server to determine the intra-refresh rate according to the channel statistics to determine the intra-refresh allocation for each frame of a GOP, and then performs intra-fresh on a number of high-priority MBs with highest loss-impact factors based on the intra-refresh allocation. In the proposed scheme, most of the computation is done in the first-

[+] This work was supported in part by Ministry of Economic Affairs, R.O.C., under grant 93-EC-17-A-02-S1-032

pass front-end encoding, which usually does not need to be done in real-time for prestored video applications. Only a small amount of computation is left to the second-pass transcoding, which usually has to meet the real-time requirement.

2.2. Estimation of Loss-Impact

To estimate the error propagation effect of a lost MB, we first define the pixel-level loss-impact (*LI*) metric as the product of two parameters: *PRC* (Pixel Reference Count) and *PCE* (Pixel Concealment Error), to characterize the amount of pixel-wise error propagation as follows:

$$LI(x,y,n) = PCE(x,y,n) \times PRC(x,y,n) \quad (1)$$

where *PRC*(x,y,n) represents the frequency of pixel (x,y) of frame n being referenced by pixels in the following frames within a GOP in the motion-compensated prediction process. It can be calculated recursively by summing up the individual reference counts of pixels in frame n+1 which reference to pixel (x,y) of frame n in the reverse tracking order from the last frame to the first frame of a GOP as in (2). And $PCE(x,y,n)$, shown in (3), denotes the norm of concealment error of pixel (x,y) of frame n should this pixel be corrupted, where $f(x,y,n)$ is the pixel value of pixel (x,y) in frame n. In this work, the zero-motion error concealment scheme is adopted.

$$PRC(x,y,n) = \begin{cases} \sum\limits_{(x',y',n+1)\text{ ref. to }(x,y,n)} PRC(x',y',n+1) & 1 \le n < N_{\text{GOP}} \\ 1 & n = N_{\text{GOP}} \end{cases} \quad (2)$$

$$PCE(x,y,n) = \left| f(x,y,n) - f(x,y,n-1) \right|^2 \quad (3)$$

We then use the motion information to calculate the current-frame's MB-level error-propagation (from the previous frames) as follows:

$$EP_{\text{MB}}(m,n) = \sum_{(x,y)\in \text{MB}_m} LI(x+\text{MV}_x, y+\text{MV}_y, n-1) \quad (4)$$

where m is the MB index in a frame; (x,y) is the pixel coordinate; n is the time index; $(\text{MV}_x,\text{MV}_y)$ is the associated motion vector of pixel (x,y). Finally, all EP_{MB}'s in each frame are summed up to estimate the frame-level error-propagation as follows:

$$EP_n = \sum_{m=1}^{N_{\text{MB}}^{\text{F}}} EP_{\text{MB}}(m,n) \quad (5)$$

where N_{MB}^{F} is the number of MBs in a frame. After obtaining the above features in the first-pass front-end encoding, EP_{MB}'s of MBs and the frame-level EP_n are extracted and stored at the streaming server which will be sent to the intermediate transcoder as side information to enhance error resilience while streaming.

2.2. Transcoding Using Prioritized Intra-Refresh

In the second-pass transcoding, we propose a prioritized intra-refresh scheme to determine the intra-refresh rate and the intra-MB allocation strategy for each GOP so as to adapt the transcoded video to varying network conditions. 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 using (6), where $N_{\text{intra}}^{\text{GOP}}$ is the total number of MBs to be intra-refreshed in a GOP; N_{GOP} is the GOP size; PLR_{TC} is the channel packet-loss rate estimated at the transcoder by using the client feedback information; TH_{intra} is a scaling parameter.

$$N_{\text{intra}}^{\text{GOP}} = \frac{\dfrac{1}{N_{\text{GOP}}} \sum\limits_{n=1}^{N_{\text{GOP}}} EP_n \times PLR_{\text{TC}}}{TH_{\text{intra}}} \quad (6)$$

The intra-refreshed MBs are then distributed to a GOP using the following algorithm,

If $n = 2$ (i.e., the first P-frame in a GOP)

$$N_{\text{intra}}^{n} = \frac{EP_n}{\sum\limits_{i=n}^{N_{\text{GOP}}} EP_i} \times N_{\text{intra}}^{\text{GOP}} \quad (7)$$

else if $3 \le n \le N_{\text{GOP}}$

$$N_{\text{intra}}^{n} = \frac{EP_n}{\sum\limits_{i=n}^{N_{\text{GOP}}} EP_i} \times \left(N_{\text{intra}}^{\text{GOP}} - \sum_{i=2}^{n-1} N_{\text{intra}}^{i} \right) \quad (8)$$

end if

where $N_{\text{intra}}^{n} = \min(N_{\text{intra}}^{n}, k_{\text{MB}} N_{\text{MB}}^{\text{F}})$ is the number of MBs to be intra-coded in frame n; N_{MB}^{F} is the number of MBs in a frame; k_{MB} ($0 \le k_{\text{MB}} \le 1$) is a control parameter to constrain the number of intra-coded blocks in a frame not to exceed an upper limit. For the nth frame of a GOP, we select a total of $N_{\text{intra}}^{\text{F}}$ MBs with top-ranked EP_{MB} values to perform intra-refresh.

3. PROPOSED INTRA-REFRESH STRATEGY FOR VIDEO MULTICAST

In our proposed error-resilience transcoding scheme described above, according to the estimated channel-loss characteristics PLR_{TC}, the transcoder determines an appropriate intra-refresh rate to reach a good tradeoff between error robustness and coding efficiency to maximize the visual quality for a single client. However, how to determine the intra-refresh rate for a single multicast stream delivered to multiple clients with different channel-loss characteristics is a practical problem in video multicast applications, which was not addressed. That is, how to determine an appropriate PLR_{TC} to obtain a best intra-refresh rate using (6) for multiple clients with different channel packet-loss rates PLR_{ch}'s, if only a single video stream is desirable.

3.1. Minmax Penalty Criterion

To characterize the effect of adopting at the transcoder an estimated packet-loss rate (i.e., PLR_{TC}) which does not match the packet loss-rate for a certain channel, we define the following PSNR penalty measure:

$$\Delta PSNR_i(x \mid p_i) = PSNR(PLR_{\text{TC}} = p_i \mid p_i) - PSNR(PLR_{\text{TC}} = x \mid p_i) \quad (9)$$

where we assume the packet-loss rate of the ith client is p_i, whereas the transcoder uses a different $PLR_{\text{TC}} = x$ to determine the intra-refresh rate for the outgoing video stream based on (6).

Fig. 1 shows an example of *PSNR* penalty plot for thee channels with different packet-loss rates (e.g., PLR_{ch} = 5%, 10% and 15%, respectively). In Fig. 1, the symbol '×' marks the optimal PLR_{TC} value which leads to minimum PSNR penalty for each client. With the proposed error-resilience transcoding method, the optimal PLR_{TC} for one channel is very close to the channel's packet-loss rate. Fig. 1 shows that, if the transcoder adopts a PLR_{TC} deviated from the optimal value, a smaller PLR_{TC} will lead to more error propagation caused by packet loss, since the intra-refresh rate is less. On the other hand, the more PLR_{TC} is, the more the intra-refresh rate will be, leading to worse coding efficiency which cannot be compensated by the performance gain from the enhanced error resiliency.

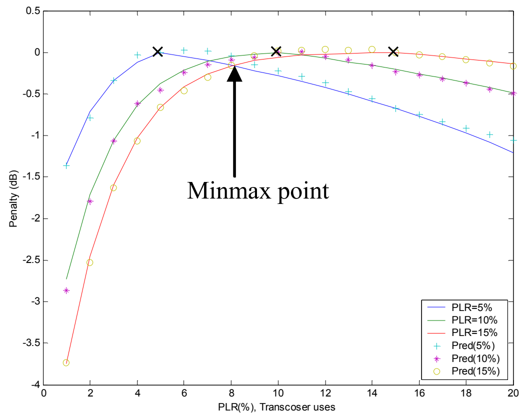

Fig. 1. Plot of PNSR penalty caused by using at the transcoder an estimated packet-loss rate which does not match the packet-loss rates of individual channels (PLR_{CH} = 5%, 10%, and 15%).

When multicasting a video stream to multiple clients with different loss characteristics, the transcoder should not just maximize the received visual quality for one certain client since it may lead to quality degradation for other clients. In the multicast scenario, we propose to determine PLR_{TC} based on the following minmax penalty criterion:

$$PLR_{TC} = \arg\min_{x} \max_{i} \{\Delta PSNR_i(x \mid p_i)\} \tag{10}$$

The transcoder then uses PLR_{TC} to determine the intra-refresh rate for the outgoing video stream according to (6). Such single intra-refresh rate will result in quality penalty $\Delta PSNR_i(x|p_i)$ for the *i*th channel due to the mismatch of channel-loss rates between PLR_{TC} and p_i. The selection PLR_{TC} of is optimized in the sense of minimizing the maximum penalty distortion that any client will suffer, thereby tending to reduce the distortion deviation among all clients to achieve fairness.

3.2. Mismatch Distortion Models

In order to obtain an optimal PLR_{TC}, we propose to model the channel mismatch distortion as follows:

$$\Delta PSNR(x \mid p) = \begin{cases} G_0 \cdot (p-x) \cdot \dfrac{e^{m\cdot(p-x)}}{e^{m\cdot p}} & , \ x < p \\ G_1 \cdot (x-p) \cdot \dfrac{e^{n\cdot(x-p)}}{e^{n\cdot(100-p)}} & , \ x \geq p \end{cases} \tag{11}$$

where $G_i = c_i - k_i \times p^{1/3}$.

As mentioned above, the penalty, $\Delta PSNR(x \mid p)$, is mainly caused by error propagation when $x < p$, and by coding efficiency sacrifice when $x \geq p$. They depend on the diverged distance of x from p. The parameters G_0 and G_1 are decreased by a scale k_i from c_i to indicate the slope of decay. $e^{m\cdot(p-x)}/e^{m\cdot p}$ and $e^{n\cdot(x-p)}/e^{n\cdot(100-p)}$ are used to fine tune the smoothness. Fig. 1 illustrates the penalty models for three clients with channel packet-loss rates of 5%, 10%, and 15%, respectively, for the *Foreman* sequence. We use a fixed set of parameters for modeling each video stream which can be analyzed beforehand and stored as side information. For example, the set of model parameters used for *Foreman* are c_0 = 0.53, k_0 = 0.01, m = 0.35, c_1 = 3.29, k_1 = 1.15 and n = 0.035, respectively. The results of modeling for the *Salesman* and *Coastguard* sequences are also shown in Fig. 2.

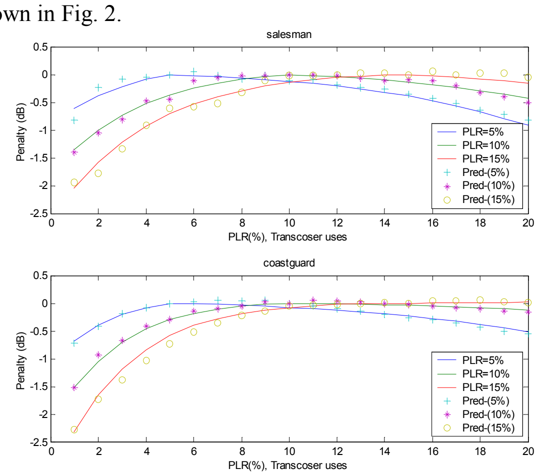

Fig. 2. PSNR penalty function models for the *Salesman* and *Coastguard* sequences for PLR_{CH} = 5%, 10%, and 15%.

4. EXPERIMENTAL RESULTS

In our experiments, several CIF (352x288) test sequences are pre-encoded at 30 fps and 384 Kbps. The group of picture (GOP) size is (N_{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 to generate various packet-loss patterns with three packet loss rates (*PLR*): 5%, 10%, and 15%, respectively. We apply the penalty model functions in (11) to compute the optimal PLR_{TC} which meets the minmax criterion for the case which includes six receivers with different channel loss rates as listed in Table 1.

Table 1 also shows the numerical results of the penalty distortion $\Delta PSNR_i(x|p_i)$ for each user with channel loss rate p_i, where "Average" refers to $x = (\sum_{i=1}^{6} p_i)/6$, and $x = \max\{p_i\}$ for "Worst", $x = \min\{p_i\}$ for "Best". Fig. 3 shows that the proposed minmax penalty criterion yields the best visual quality in terms of the mean and the variance of PSNR penalty values among the four methods.

Table 1. Comparison of penalty in distortion for six users under different case.

User	p_i (%)	Minmax	Average	Worst	Best
1	3	0.03	0.01	0.54	0
2	3	0.01	0.06	0.36	0
3	3	0.05	0.03	0.60	0
4	5	0.03	0	0.22	0.34
5	5	0.01	0	0.28	0.43
6	10	0.24	0.45	0	1.07

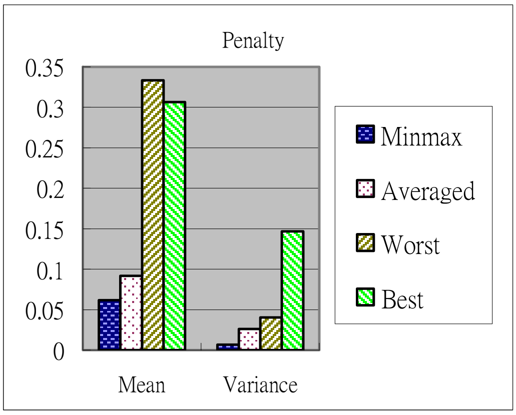

Fig. 3. Mean and variance of penalty of multiple users with disparate channel loss characteristics.

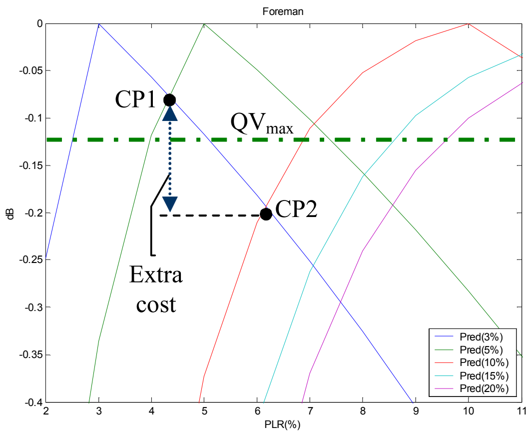

Fig. 4. Illustration of minmax point selection by excluding receivers with temporarily unreliable channel conditions.

When multicasting a video stream to heterogeneous receivers with varying channel characteristics, it is usually undesirable to sacrifice the visual quality of users with good channel conditions by compromising on some users with significantly poor channel characteristics, especially for wireless LAN environments where client mobility will temporarily result in channels with rather unstable channel conditions. In order to constrain the quality variation for those receivers staying at stable conditions, we will consider the cost to decide whether to account for the users with temporarily unstable or relatively diverse channel characteristics.

As an example, Fig. 4 shows that the minmax point is CP1 when we only consider the users with packet-loss rates of 3% and 5%. The minmax point will move to CP2 should the receiver with 10% loss rate be taken into account. The receiver who has 3% loss rate has the maximum penalty distortion and will incur some extra cost when it is taken into account in the intra-refresh rate selection. If the penalty distortion of a user is more than the quality variation constraint (QV_{max}), we will exclude the user from the minmax rate determination process.

5. CONCLUSION

In this paper, we extended our previously proposed two-pass error-resilience transcoding scheme to cope with more general video multicasting involving heterogeneous clients with diverse channel conditions. We have proposed a minmax loss-rate estimation scheme so as to determine an appropriate intra-refresh rate for all the clients. Simulation results show that the proposed scheme can effectively reduce the mean and variance of penalty distortion of all users. We have also discussed how to constrain the quality variation for applications with temporarily unreliable users such as video multicast in wireless LANs.

REFERENCES

[1] T. Warabino *et al.*, "Video transcoding proxy for 3G wireless mobile internet access," *IEEE Commun. Mag.*, vol. 38, pp. 66-71, Oct. 2000.

[2] T. Shanableh and M. Ghanbari, "Heterogeneous video transcoding to lower spatio-temporal resolutions and different encoding formats," *IEEE Trans. Multimedia*, vol. 2, no. 2, pp. 101-110, June 2000.

[3] J. Xin, C.-W. Lin, and M.-T. Sun, "Digital video transcoding," *Proc. IEEE*, vol. 93, no. 1, pp. 84-97, Jan. 2005.

[4] 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.

[5] S. Dogan, A. Cellatoglu, M. Uyguroglu, A. H. Sadka, and A. M. Kondoz, "Error-resilient video transcoding for robust internetwork communications using GPRS," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 453-464, June 2002.

[6] T.-C. Wang, H.-C. Fang, and L.-G. Chen, "Low delay and error robust wireless video transmission for video communication," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 12, pp. 1049-1058, Dec. 2002.

[7] A. Vetro, J. Cai and C.-W. Chen, "Rate-reduction transcoding design for wireless video streaming," *Journal of Wireless Communications and Mobile Computing*, vol. 2, no. 6, pp.549-552, Sept. 2002.

[8] H.-J. Chiou, Y.-R. Lee, and C.-W. Lin, "Error-resilient transcoding using adaptive intra-refresh for video streaming," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 3, pp. 777-780, May 2004, Vancouver, Canada.