PERCEPTUAL QUALITY FEEDBACK BASED PROGRESSIVE FRAME-LEVEL REFRESHING FOR ROBUST VIDEO COMMUNICATION

Liang Cheng, Magda El Zarki

School of ICS, University of California, Irvine
Email: lcheng61@ics.uci.edu, elzarki@uci.edu

ABSTRACT

In this paper, we propose an encoder side error resilience technique, namely the progressive group of picture (PGOP), which is well suited for rate controlled real-time video transmission over wireless channels. The PGOP scheme can be seen as an alternative approach to the group of picture (GOP) that periodically uses intra-frames (I-frames) for low bit-rate channels. This is because: (1) the PGOP scheme is shown to completely stop error propagation in a timely manner; (2) the PGOP scheme avoids the bit rate fluctuation problem associated with I-frames by using a progressive intra-updating scheme. We use the PGOP scheme in a perceptual quality feedback framework, and apply a novel online algorithm that helps shape the bit rate and explores the maximum recovery capability of PGOP. The performance of our scheme is demonstrated in a wireless LAN environment.

1. INTRODUCTION

Wireless networks experience high bit error rates due to long and short fades, shadowing and environmental noise, none of which are conducive to the transmission of time sensitive highly compressed data. Channel coding, such as forward error correction (FEC) and automatic repeat request (ARQ), can help protect and recover the data [1]. These schemes, however, add overhead and delay to the transmission and cannot guarantee 100% recovery. The consequence of a packet loss is exacerbated in the case of predictive video coding schemes because the prediction loop, especially motion compensation, propagates errors.

The most intuitive and effective method to stop these propagating errors caused by the imperfections of the network is to periodically insert intra-frames (I-frames) and to form the group of picture (GOP). The propagation of errors terminates at the beginning of each GOP with the I-frame. For low-bandwidth (especially wireless) video transmission (e.g. $\leq$ 128kbit/s) environments, the use of I-frames is, however, highly undesirable: I-frames incur higher delay because they have a larger number of bits due to the lack of predictive coding. Therefore, in most encoders designed for wireless use, only one I-frame is used at the beginning of the sequence, followed by all P-frames. Unless we can guarantee that no errors will occur, some form of data refreshing has to take place continually to halt the error propagation.

Over the past few years, researchers have proposed a variety of approaches to increase the robustness of low bit-rate video communications[2]. Selectively forcing the intra coding of a number of macroblocks (henceforth referred to as forced intra MBs, updated MBs, or refreshed MBs) in the P-frames is well recognized as an effective mechanism to mitigate the propagation effect of interframe prediction[3] [4]. However, none of these proposed schemes provide a refresh pattern to completely eliminate the propagation effect of the existing errors in a practical and timely manner. For any such update scheme, we are faced with three questions: (1) how to allocate the intra-MBs in P-frames? (2) how to accurately detect the instantaneous occurrence of an error and reflect it in the intra/inter mode selection? (3) how to avoid the bit rate spikes caused by the insertion of intra-MBs?

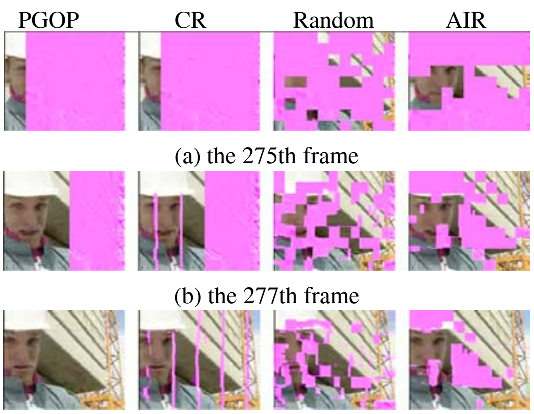
To address the first question, we design a method that is able to completely refresh errors existing in the prediction loop in a progressive and systematic manner, which we refer to as progressive group of picture (PGOP). The second problem is solved by applying PGOP in a quality feedback enabled framework. The third problem is solved by limiting that the bit rate fluctuations to stay within the designated range.

The rest of the paper is divided as follows. In section 2, we describe the PGOP algorithm. In section 3, we discuss the PGOP in a quality feedback enabled setting. In section 4, we present the simulation results to demonstrate the performance of our scheme. In section 5, we conclude and outline future work.

2. FRAME LEVEL PROGRESSIVE REFRESH

In order to maximally scatter the intra-MBs and help the rate control mechanism (which usually changes the quantization scale of each MB in a scan order) allocate the bit budget more uniformly, we refresh intra-MBs on a column-by-column basis from left to right. We note that errors may propagate across the column being refreshed from what we call the unrefreshed area of the frame to the refreshed area. What was a refreshed MB may be affected by error propagation again if it uses part of the unrefreshed area as its reference for motion estimation. To solve this problem, we propose to augment the refresh process to trap these error propagations by refreshing the affected MBs. We call this *stride back*.

Our implementation of PGOP is based on the MPEG-4 baseline encoder [5]. We setup a bit-map table to keep the status of every pixel in the whole frame. For one MB (which has 256 pixels), we need 32 bytes. Each bit represents the status of one pixel, i.e., if that pixel is affected or not. When half-pixel motion estimation is used, some extra pixel interpolation operations are needed. In the simulation throughout this paper, we use half-pixel motion estimation.



(a) the 275th frame

(b) the 277th frame

(c) the 280th frame

Fig. 1. Comparison of PGOP and conventional refresh schemes

Fig. 1 compares the refresh capability of PGOP (2 column/frame), column based refresh without *stride back* (CR), random intra update (Random), and adaptive intra refresh (AIR) [5]. Note that AIR updates a specified number of MBs that are calculated to have the highest motion (highest difference). For both the random update and the AIR scheme, 24 MBs are intra-coded per frame. We assume that the whole 274th frame of sequence Foreman is lost without any error concealment. We can see that the *stride back* function enables PGOP to eliminate the propagating errors, which affect neighboring MBs for the CR scheme. In Fig. 1-c, we notice that by the 280th frame our proposed PGOP scheme has completely recovered from all existing errors as well as their propagation effects. Thus, the total length of PGOP illustrated in Fig 1 is 6 frames. Using average the same number of intra-MBs, none of other refresh schemes are able to achieve the goal of frame-level error elimination within 6 frames.

Of course, we cannot accurately locate the lost MBs. When applying PGOP, if some MBs are in motion but are not in error, the regular column based refresh and the *stride back* will also refresh those MBs, which is actually unnecessary. The technique that will be introduced in the next section will help to decrease the needless intra-coding by informing the encoder of the occurrence of channel loss and the subsequent requirement of MB refresh.

3. FEEDBACK ENABLED PROGRESSIVE FRAME REFRESHING

In this section, we apply PGOP in a perceptual quality feedback framework. The major difference between “quality feedback” and conventional feedback schemes (i.e., packet loss rate, lost MB’s address) is that the quality feedback reflects the distortion caused by the channel loss in a cumulative way. That is, the received quality feedback shows the quality distortion jointly determined by all the losses before it, no matter whether or not previous feedbacks (if any) were received. It also takes into account the effect of the applied error concealment scheme. The Institute for Telecommunications Sciences (ITS) in [6] developed a spatial-temporal distortion metric. Instead of using pixel-comparison, the ITS model based quality assessment algorithm calculates the quality features of processed spatial-temporal (S-T) regions. Thus it is suitable for low band-width in-service video quality monitoring. In our implementation, we extract the quality features of the decoded video, send them back to the encoder, and compare them to the features of the encoder side reconstructed video. Once the encoder realizes any inconsistencies, which indicates the occurrence of channel loss, the encoder will send one PGOP to quickly remove the errors jointly caused by all previous channel losses and propagation effects thereof. For error free periods, nevertheless, the highest possible coding efficiency can be preserved as no MBs will be force-updated.

There is a tradeoff between the error recovery capability (i.e., PGOP period) and the bit rate fluctuation. On the one hand, it is desirable to set the PGOP period as short as possible. I-frame is an extreme instantiation: I-frame has the maximum error recovery capability because it can stop error propagation within one frame. On the other hand, we could lower the bit rate fluctuation if we increase the length of PGOP, i.e., decrease the number of intra-MBs in each frame. When the PGOP is sporadically inserted into the compressed stream, the bit rate fluctuation may happen due to the mismatch of the quantization scale and the percentage of intra-MBs in the first frame of the PGOP. The percentage of the intra-MBs in the first frame of the PGOP is usually larger than its previous frames, which are regular P-frames. The rate control mechanism (which is independent of the PGOP design in this paper) is not able to increase the quantization scale promptly to compensate for the higher number of bits generated by the force-updated MBs. After the first frame of the PGOP, the rate control mechanism can adapt to the large intra-MBs percentage of the remaining frames. The bit rate will converge quickly to the desired value.

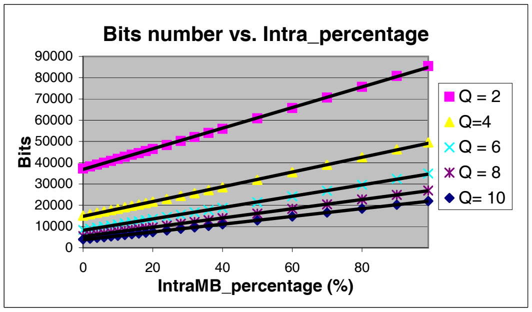


Fig. 2. Linear relation between average bit number and intra-MB's percentage for the total 400 frames of *Foreman*.

Therefore, we determine the number of intra-MBs (or say intra-MB columns) of each frame in the PGOP, which can maximally expedite the frame-level refresh period and keep the output bit rate fluctuation within a designated range. From empirical studies we conducted on several video sequences (i.e., *Foreman*, *Suzie*, and *MotherDaughter*), we know that the relationship between the percentage of the forced intra MBs and the average number of bits per frame is linear, as shown in Fig.2. Of course a significant scene change or high motion may cause this relation to be non-linear because those forced intra-MBs may overlap with the intra-MBs that have already been determined by the coding loop. As a matter of fact, the overlapping of some intra-MBs reduces the bit rate fluctuation.

Therefore, we propose an algorithm to minimize the impact of PGOP on the bit rate fluctuation. Before we describe our algorithm, we define some parameters in Table. 1.

Table 1. Parameters definition.

$\bar{Q}_{pre}$	Average quantization scale of previous frame
R	Bit rate of the video sequence
F	Frame rate of the video sequence
C	Total number of MB-columns in one frame
α	Percentage that the bit number can exceed the average bit allocation
β	Percentage of intra-MBs for the first frame of PGOP
Col_{pgop}	Maximum number of intra-columns for the first frame of PGOP

Thus, the linear relationship between the percentage of intra-MBs and the bit number is described as:

$$Bits = a \times \beta + b. \tag{1}$$

Where b represents the number of bits for a predictive frame with zero intra-MB percentage. Then $(a + b)$ represents the number of bits for a predictive frame with 100% intra-MB percentage (i.e., $\beta = 1$).

From Fig. 2, we also notice that the number of bits in a frame with a certain percentage of intra-MBs relies on the quantization scale. Thus, we design an online algorithm to determine the values of a and b by pre-coding the current frame once with 0% intra-MBs and once with 100% intra-MBs. We use $\bar{Q}_{pre}$ for all the MBs in the current frame. Note that a MB based rate control mechanism will increase the quantization scale if it is aware of a sudden increase in bits, which makes our assumption fairly conservative.

The algorithm is as follows:

(1) The last quality feedback shows that channel loss has occurred and one PGOP is requested.

(2) Get $\bar{Q}_{pre}$ and prepare to apply it to every MB of the current frame, i.e., first frame of PGOP.

(3) Encode the current frame in the regular predictive mode and assign the encoded number of bits to parameter b.

(4) Encode the current frame with intra mode. Assign the encoded number of bits to $(a + b)$ to obtain the linear relationship denoted by Equation (1).

(5) Calculate the maximum number of bits (including the fluctuation range) that can be used in the current frame, which is

$$Bits = \frac{R}{F} \times (1 + \alpha). \tag{2}$$

(6) Calculate the maximum intra-MB percentage that can be applied in the current frame by using Equation (1) and Equation (2):

$$\tfrac{R}{F} \times (1 + \alpha) = a \times \beta + b,$$

from which we get

$$\beta = \frac{(1 + \alpha) \times (R/F) - b}{a}. \tag{3}$$

(7) Calculate the number of intra-columns for the PGOP.

$$Col_{pgop} = \lfloor C \times \beta \rfloor. \tag{4}$$

Note: $\lfloor \cdot \rfloor$ denotes rounding to the lowest integer. If Col_{pgop} is negative, set it to zero.

Fig. 3 shows the exploited maximal number of intra columns. Of course, the proposed algorithm increases the encoding complexity as the encoder must determine the value of a and b; the first frame of the PGOP is encoded three times. However, a lot of coding routines (such as motion estimation, packetization, I/O, etc.) do not need to be repeated. Hence, our scheme only needs moderate additional complexity at the encoder when a perceptual quality drop is reported.

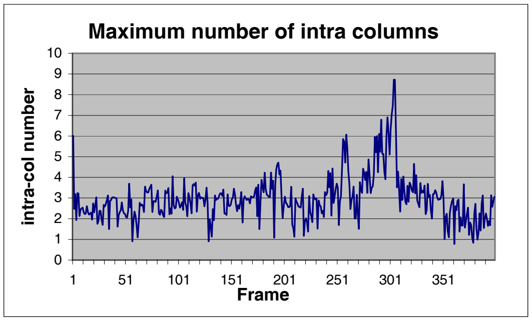


Fig. 3. Simulation result of Equation (4). Bitrate: 64kbps; Sequence: *Foreman*; Video format: QCIF;.$\alpha = 1$; $F = 10$.

4. PERFORMANCE

In Fig. 4, we demonstrate the performance of feedback-based PGOP for a wireless LAN packet loss pattern (Fig. 4-a), which reflects the pedestrian movement and moderate traffic congestion around the ICS school building at the University of California, Irvine. In Fig. 4-b, we notice the comparison of the feedback decoding quality and the reconstructed quality (i.e., ITS feedback). It shows that the encoder figures out correct positions to insert PGOPs. In Fig. 4-c, the number of intra-columns of PGOP is different and adapted to the video content and error locations. In Fig. 4-d, the bit number fluctuation is within the range (i.e. for $\alpha = 1$, the upper bound of the fluctuation range is 12.8kbits/frame if the average bit rate is 6.4kbits/frame). In Fig. 4-e, we see that the PGOP can effectively and promptly recover the decoded video from errors without influencing the decoded video quality in the error free period, when more advanced optimizations are applicable.

5. CONCLUSION

In this paper, we propose a frame-level error resilient scheme named PGOP. We then incorporate PGOP into a quality feedback framework, discuss its effect on the bit rate fluctuation, and evaluate its performance. The proposed methods are independent from any other encoder/decoder side control mechanisms (i.e. rate control, channel coding, etc.). Further optimization is possible if these control mechanisms are taken into consideration.

6. REFERENCES

[1] H. Liu et al., "Performance of h.263 video transmission over wireless channels using hybrid arq," *IEEE J. Select. Areas Commun.*, vol. 15, no. 9, Dec. 1997.

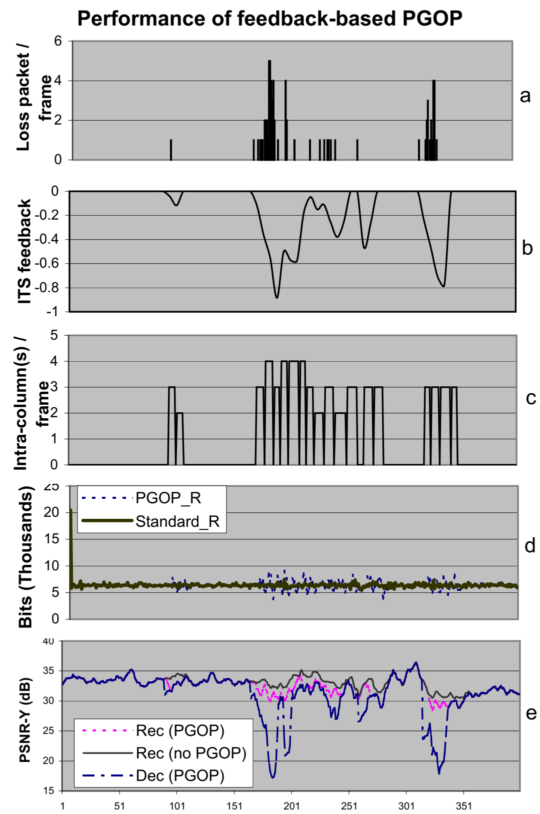


Fig. 4. Performance of feedback-based PGOP scheme in the wireless LAN environment. Bit rate: 64 kbps; packet size: 200 bytes; error concealment: direct copy from corresponding MBs in the previous frame; feedback interval: 6 frames; $\alpha = 1$.

[2] Y. Wang et al., "Error control and concealment for video communication: A review," *Proc. IEEE*, vol. 86, May 1998.

[3] P. Haskell et al., "Resynchronization of motion compensated video affected by atm cell loss," *Proc. Int. Conf. Acoust., Speech, Signal Processing*, vol. 3, May 1992.

[4] B. Girod et al., "Feedback-based error control for mobile video transmission," *Proc. of the IEEE, special issue on video for mobile multimedia*, vol. 97, no. 10, Oct. 1999.

[5] "Iso/iec 14496-5: Information technology - coding of audio-visual objects part 5: Reference software," 2001/Amd 1: 2002.

[6] S. Wolf et al., "Spatial-temporal distortion metrics for in-service quality monitoring of any digital video system," *Proc. of SPIE Int. Symp. on Voice, Video, and Data Commun.*, Sep. 1999.