AN EFFICIENT RESYNCHRONIZATION TECHNIQUE FOR PERCEPTUAL QUALITY ENHANCEMENT FOR ROBUST VIDEO TRANSMISSION

Lap-Pui Chau and Tao Fang

School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore

ABSTRACT

In this paper, using results from a simplified macroblock (MB)-based segmentation algorithm, we propose a framework called content-based resynchronization (CBR) for the effective positioning of resynchronization markers such that the image quality of foreground can be improved at the expense of sacrificing unimportant background. We do this because, in applications such as video telephony and video conferencing, foreground is typically the most important image region for viewers. Experimental results demonstrate that this scheme significantly improve the subjective quality of video sequence for robust video transmission.

1. INTRODUCTION

Since the amount of bandwidth available on current communication channels is limited, low or very low rate compression algorithms have to be applied to the video data before it can be transmitted. When compressed video data is transmitted over error-prone channels, the effect of channel errors on the video can be extremely severe. Thus, the video applications will have to provide sufficient robustness to ensure that the quality of the decoded video is not overly affected by the channel unreliability. Much effort has been invested in building error resilience into the compressed bitstream [1]-[3]. Among the state-of-art error-resilient techniques, resynchronization has been proved to be a very effective tool.

Previous video coding standards such as H.261 and H.263 (Version 1) logically partition each of the images to be encoded into rows of MBs called Group Of Blocks (GOBs) and resynchronization markers are allowed to occur only at the left edge of the images. Since the resynchronization markers are likely to unevenly spaced in the bitstream, some areas of the picture will be more susceptible to errors. In MPEG-4 encoder, resynchronization markers can be inserted in the bitstream at approximately constant intervals. Thus, the MB interval between the resynchronization markers is a lot closer in the high activity areas and a lot farther apart in the low activity areas. In the presence of a short burst of errors, the decoder can quickly localize the error to within a few MBs in the important high activity areas of the image and preserve the image quality in these important areas [3]. Besides the conventional approaches mentioned above, there are still many literatures considering the insertion of resynchronization markers [4]-[6]. Yoo [4] proposes an adaptive resynchronization marker positioning (RMP) method for rapidly resynchronizing the VLC decoding synchronization. Taking into account the channel condition and the error concealment method, Cote, *et al* [5] optimize placement of synchronization markers in the compressed bitstream. Yang, *et al* [6] present simpler but efficient techniques to optimize placement of synchronization markers considering the impact of the data partitioning in error concealment. In this paper, we present a content-based resynchronization method using segmentation results. That is, after pictures are segmented into foreground and background at the MB level, we design a novel resynchronization scheme to improve the perceptual quality of video sequences.

The rest of this paper is organized as follows. In Section 2, we first give a description of revised Partial Backward Decodable Bit Stream (PBDBS) and perceptual impact of error on video quality, after which, we propose a resynchronization approach using content-based segmentation results. Section 3 gives the simulation results and Section 4 concludes the paper.

2. CONTENT-BASED RESYNCHRONIZATION

In the beginning of this section, PBDBS, which is proposed by Gao and Tu [7], will be revised for using in the proposed resynchronization method. After that, we will analyze the perceptual impact of error on video. Finally, the content-based resynchronization method will be proposed.

2.1. Revised PBDBS

The structure of revised PBDBS is shown in Fig. 1. After encoding some MBs of one slice, the bit position is

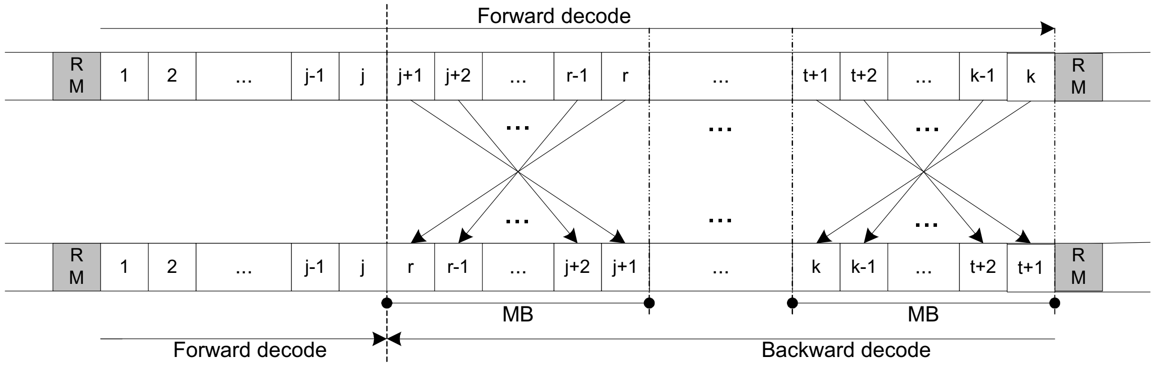


Fig. 1. Encoding procedure: H.263 baseline (top) and revised PBDBS (bottom).

memorized, which is denoted as j. After bit j, all the bits in each following MB will be reversed bit by bit. When decoding, if there is no error happening, the decoder will carry a reverse procedure of the encoding one. That is, after the decoder decodes to the bit j, it will search the next resynchronization marker and decode in a backward direction from bit $t+1$ to bit r. On the other hand, when the error happens in the forward decodable stream, i.e., before bit j, the decoder will stop decoding and jump to the next synchronization marker to decode in a backward direction.

2.2. Perceptual Impact of Error on Video Quality

Regardless of the resynchronization scheme being used, the foreground data may generally distribute in one slice (i.e., between two resynchronization markers) in four cases: in the beginning, in the middle, in the end, one part in the beginning and the other part in the end. Then we classify the location of error in two cases: in the forward decodable stream and in the backward one. Thus, there will be 4×2 different impacts on perceptual video quality according to the different positions where the foreground data distributes and where the error happens (see Fig. 2). Considering the best protection of the foreground data, we can summarize the perceptual impact of error on video quality as follows:

1. Best case: one part of foreground in the beginning and one part in the end. As we can see from Fig. 2 (d) and (h), when the error happens in the beginning or in the end, little foreground data will be lost.

2. For other cases, that is, foreground in the beginning, in the end or in the middle, wherever the error happens in the bitstream of one slice, part of the foreground will be affected. The effect depends on the location where error happens. This is shown in Fig. 2 (a)-(c), (e)-(g).

It is now clear that the impact of error on the perceptual video quality depends tremendously on the location of the

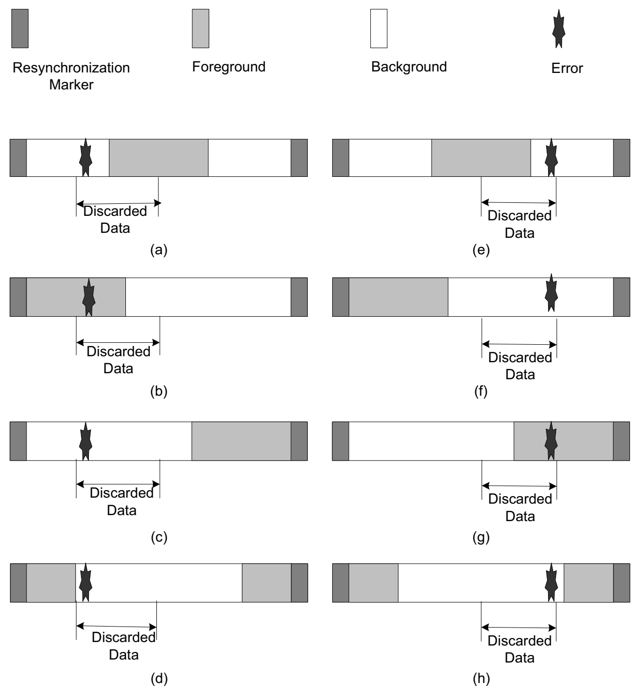


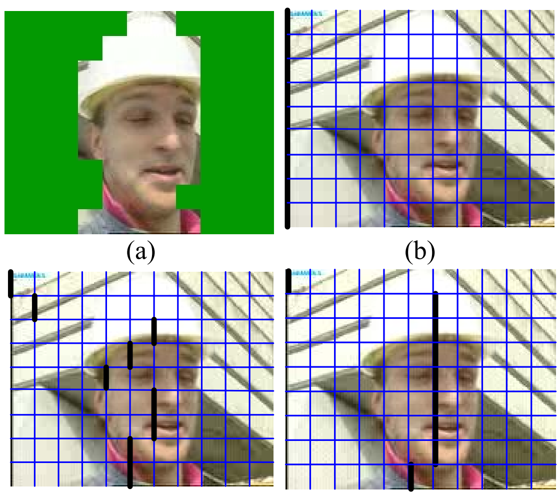
Fig. 2. Impact of error location on perceptual video quality according to different location of foreground in one slice.

foreground data in one slice. Below using the result of MB-based segmentation, which is shown in Fig. 3 (a), we propose a novel resynchronization scheme with which perceptual quality of video will be significantly improved. Note that the segmentation result is based on an MB-based algorithm. Normally, the segmentation is carried based on pixel level. Based on the result of pixel-level segmentation, our MB-based segmentation algorithm works like this: If the number of foreground pixels in one MB is noted as N, the decision for a MB to be either in foreground or in background can be made as follows:

$$MB_{i,j} = \begin{cases} foreground\ MB, if\ \ N \geq T \\ background\ MB, if\ \ N < T \end{cases}$$

Table 1
Pseudo-code for the proposed content-based resynchronization scheme.

```
RM_1 = 1; // RM_i is the position of resynchronization marker in the ith row.
for (i = 2; i ≤ I; i++) // I is the MB's max index in a column.
    for (j = 1; j ≤ J; j++){// J is the MB's max index in a row.
        if (MB_{i,j} = foreground MB)
            FG_beginning_i = j; // The index of 1st MB in foreground.
        if (MB_{i,j} = background MB){
            FG_end_i = j; // The index of 1st MB after foreground.
            RM_i = 1/2 (FG_beginning_i + FG_end_i);
            break;
        }
    }
}
```



(a) (b) (c) (d)

Fig. 3. (a) MB-based segmentation result and the spatial position of resynchronization markers of (b) H.263 (version 1) packet approach, (c) MPEG-4 periodic approach and (d) this work.

where i and j are MB indexes. T is the threshold and in our experiment we set it to $\frac{3}{4}\times16\times16$.

2.3. Content-based Resynchronization Scheme

From Fig. 3 (b), we can see that for H.263 (version 1) packet approach, almost all the slices belong to the case that foreground data is in the middle of one slice. That means that error may have a bad effect on the foreground data, which leads to an impact on the perceptual quality of the video sequence. The reason has been analyzed in the previous subsection. On the other hand, for MPEG-4 periodic approach, which is shown in Fig. 3 (c), all the cases (best case and other cases in the previous subsection) will happen when error occurs.

Intuitively, if we want the best case happen most of time, the simplest way is to put the resynchronization markers just in the middle of the foreground, which is shown in Fig. 3 (d). In such a way, the foreground will distribute in a slice as illustrated in Fig. 2 (d), (h). Thus, wherever error happens, we can get the best case for all the slices. (Note that for the first slice, the resynchronization marker is put at the beginning of one picture instead of in the middle of foreground data because the resynchronization marker in the first slice is also supposed to be the resynchronization marker of the whole frame.) The algorithm's brief pseudo-code is given in Table 1.

3. EXPERIMENTAL RESULTS

In the experiment, 80 frames of the Foreman QCIF sequence at frame rate of 15 fps were encoded to test our proposed scheme. We use both subjective and objective performance measures to evaluate the performance of the algorithm. Besides the results from CBR, the results of the packet approach of H.263 and the periodic approach of MPEG-4 are also given for comparison. To make the comparison fair, the bitrate for the three approaches was all set to 140 kbps. Due to the random nature of losses and their impact on the compressed video, 50 different runs of experiments were conducted.

The subjective quality of Foreman QCIF is shown in Fig. 4. The figures clearly show that with the proposed method, the area of interest is much improved, which is the result of inserting the resynchronization markers right in the middle of foreground data as described in our proposed scheme. The degradation in the background region of the results from the proposed scheme is hardly noticeable since we concentrate on the foreground when looking at the whole picture.

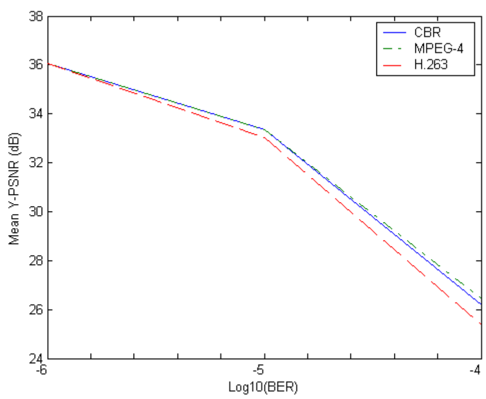
(a)

(b)

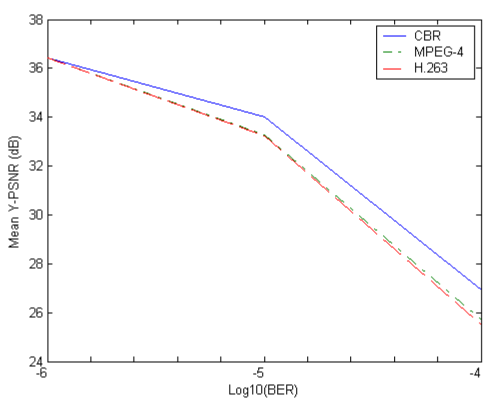
(c)

(d)

Fig. 4. Decoded Frame 55 of Foreman QCIF with BER=10^{-4}. (a) without channel error, (b) this work, (c) MPEG-4 approach and (d) H.263 approach.



(a)



(b)

Fig. 5. Foreman QCIF average PSNR comparison of (a) overall picture and (b) foreground.

Plots displaying the average PSNR achieved from 50 runs of experiments are provided in Fig. 5. The results of this work, MPEG-4 approach and H.263 approach are given with BER from 10^{-6} to 10^{-4}. Fig. 5 (a) shows the overall PSNR comparison of three. The results revealed that our scheme and the MPEG-4 scheme achieve higher PSNR values than the H.263 scheme. Fig. 5 (b) shows the average PSNR comparison of foreground. Comparing with the other two schemes, our scheme achieves a significant PSNR improvement in foreground. Furthermore, the improvement increases with the increase of BER.

4. CONCLUDING REMARKS

In this paper, we proposed a content-based resynchronization scheme so that we can have better protection on the areas of interest, i.e., foreground, resulting in better perceptual video quality. The simulation results indicated that our proposed scheme encodes the video sequences with higher quality on the foreground as compared to that of the conventional approaches over error-prone channels with BER from 10^{-5} to 10^{-4}.

5. REFERENCES

[1] Y.Wang and Q. Zhu, "Error control and concealment for video communication: a review", *Proc. IEEE*, vol. 86, pp. 974-997, May 1998.

[2] J. D. Villasenor, Y. -Q. Zhang and J. Wen, "Robust video coding algorithms and systems", *Proc. IEEE*, vol. 87, pp. 1724-1733, Oct. 1999.

[3] I. Moccagatta, S. Soudagar, J. Liang and H. Chen, "Error-resilient coding in JPEG-2000 and MPEG-4", *IEEE J. Select. Areas Commun*, vol. 18, pp. 899-914, Jun. 2000.

[4] K. Yoo, "Adaptive resynchronization maker positioning method for error resilient video transmission," *IEE Electronics Letters*, vol. 34, pp. 2084-2085, Oct. 1998.

[5] G. Cote, S. Shirani and F. Kossentini, "Optimal mode selection and synchronization for robust video communication over error-prone networks," *IEEE J. Select. Areas Commun*, vol. 18, pp. 952-965, Jun. 2000.

[6] K. H. Yang, D. W. Kang and A. F. Faryar, "Efficient intra refreshment and Synchronization Algorithms for robust transmission of video over wireless networks," *IEEE ICIP*, Oct. 2001, vol. 1, pp. 938 –941.

[7] S. Gao and G. Tu, "Robust H.263+ video transmission using partial backward decodable bit stream (PBDBS)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 182-187, Feb. 2003.