One-pass Explicit FMO Map Generation for H.264/AVC Wireless Video Transmission

Jantana Panyavaraporn* and Supavadee Aramvith*
Department of Electrical Engineering
Chulalongkorn University, Bangkok 10330, Thailand
E-mail: j.panyavaraporn@gmail.com, supavadee.a@chula.ac.th Tel: +66-822-18-6909

Abstract—Flexible Macroblock Ordering (FMO) is a new error-resilient tool at the encoder in H.264/AVC. In this paper, a method to generate explicit FMO map using one-pass encoding is proposed. To generate explicit FMO map, coded bit-count which acts as spatial information and distortion measure based on the concealment error which acts as temporal information are used. Our simulation results performed under slow and fast fading channels show that the proposed technique using one-pass encoding saves some encoding delays while still can retain the same level of PSNR.

I. INTRODUCTION

The problem of wireless video transmission systems arises from the bursty error nature caused by the time-varying channel itself. Variable Length Coding (VLC) schemes used on video codec provide good compression, however a single bit error in the coded bit stream can propagate to the next synchronization point resulting in unacceptable video quality. The amount of degradation depends on the location of error in the bit-stream. Due to the use of motion compensated prediction the error can propagate in the spatial and temporal direction. The bandwidth of channel is limited, so the use of efficient coding scheme to compress the video signal is necessary. H.264/AVC [1] is the best candidate codec for wireless multimedia application.

H.264/AVC standard also supports enhanced error resilience feature. One features of the standard is an adoption of a robust error resilience tool at encoder know as Flexible Macroblock Ordering (FMO) [2]. FMO can significantly enhance the robustness of data losses by managing the spatial relationship between regions that are coded in each slice and a macroblock level interleaving tool used to spread consecutive burst error.

Recently, some works in FMO studied different methods on how to quantify the importance of a MB. In [3], the error sensitivity of a MB is determined by computing a MB PSNR parameter. A MB impact factor is computed which depends on some information derived from the used pixels in [4]. In [5], number of macroblock coded bit count has been investigated as an indicator for a choice of FMO map of each frame. The distortion measure based on the concealment error as indicators for a choice of macroblock-address-map (MBAmap) of each picture in [6]. However, all previously methods used a two-pass encoding scheme which is more suitable for encoding non real time and transmit video from video server. In the first pass, spatial and temporal information are collected to generate explicit FMO maps. In the second pass, the video sequence is encoded using the explicit FMO maps. To be able to support the encoding and transmission of real time live video, in this paper, we propose to generate explicit map using one-pass encoding using spatial and temporal information.

The paper is organized as follows. Some background on FMO and wireless channel characteristic are presented in Section II. Section III describes our proposed method of one-pass explicit FMO map using bit count information and distortion measure. In Section IV, the simulation results are discussed. Conclusion and future works are given in Section V.

II. FMO UNDER CHANNEL ERROR

In this section, we briefly review FMO in H.264/AVC and wireless channel model.

A. FMO in H.264/AVC

Flexible Macroblock Ordering (FMO) in H.264/AVC allows flexible means of macroblock grouping. The order of coding macroblock for each picture can be done in a non-raster-scan order. The groups of macroblocks, called slice, can be altered differently for each picture and are independently intra-predicted. With the current implementation of H.264, each macroblock can be mapped to a particular slice through the MBAmap data structure. However, the macroblock order within a slice must be in ascending order. The FMO scheme supports six different slice group map types; fives of them are predefined macroblock mapping that can be specified through the picture parameter sets (PPS), FMO type 6, i.e., explicit slice group map type, is the most general type where the entire MBAmap is actually coded into a picture parameter set as shown in Fig.1.

Type 0 (interleave): uses run lengths which are repeated to fill the frame. Therefore only those run lengths have to be known to rebuild the image on the decoder side.

Type 1 (dispersed): also known as scattered slices; it uses a mathematical function, which is known in both the encoder and the decoder, to spread the macroblocks. The distribution

---

10-0110521057©2010 APSIPA. All rights reserved.

Proceedings of the Second APSIPA Annual Summit and Conference, pages 1052–1057,
Biopolis, Singapore, 14-17 December 2010.

1052
in the figure, in which the macroblocks are spread forming a chess board, is very common.

Type 2 (foreground and Background): is used to mark rectangular areas, so called regions of interest. In this case the coordinates top-left and bottom-right of the rectangles is saved in the MBAnap.

Type 3-5 (Box-out, Raster and Wipe): are dynamic types that let the slice groups grow and shrink over the different pictures in a cyclic way. Only the growth rate, the direction and the position in the cycle have to be known.

Type 6 (Explicit): is the most random one and allows full flexibility to the user. All the other ones contain a certain pattern.

B. H.264/AVC in Wireless Channel

To analyze the effects of FMO for wireless video transmission, the modification of reference software and wireless channel simulator are needed. In the study, the H.264 JM codec version 9.2 reference software is used. No modification of the encoded bit-stream is needed. However, the encoder in this version does not fully support encoding of explicit slice-group-map (i.e. slice-group-map-type = 6). We modify the source code such that it can read the entire slice group configuration file and it can resend the updated picture parameter set (PPS) information for each encoded picture. Note that, the encoded bit-stream can be decoded by the decoder. However, some modifications needed at the decoder to discard undecodable macroblock and to continue the decoding process.

We encode 4 video sequences using the baseline profile at level 3.0 in our simulation. IPPP... GOP structure is used for low-delay purpose. Each sequence is encoded for a total of 100 frames with frame rate of 10 frames/s. The rate-control is enabled at fixed bit rate of 32 kbps. The default encoder parameters [7] are used with the exception on the following FMO-related parameters.

Slice_group_map_type (6)
Num_slice_group_minus_one (0 to 7)
SliceGroupConfigFileNamet(proposed MBAmap)

Detailed description of H.264 JM codec parameter usage can be found in [8] and [9].

To investigate the benefits of using FMO on wireless channel, the Rayleigh fading wireless channel simulator is used in our simulation. The details of the simulator can be found in [10]. To simulate the effects of slow and fast fading channels, the maximum Doppler frequency parameter is set to 1 Hz, for slow-fading and 40 Hz, for fast-fading, respectively. We set up the simulation by assuming that errors may not attack the PPS header. The average bit error rate (BER) and average packet error rate (PER), for 80-bit packet [10], are 0.06 and 0.09, respectively, as shown in Table 1.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>WIRELESS CHANNEL PARAMETER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple Access</td>
<td>TDMA</td>
</tr>
<tr>
<td>Modulation</td>
<td>QPSK</td>
</tr>
<tr>
<td>Channel Rate</td>
<td>32 kbps</td>
</tr>
<tr>
<td>Maximum Doppler Frequency</td>
<td>1 Hz (slow), 40 Hz (fast)</td>
</tr>
<tr>
<td>Transmitted Signal Power</td>
<td>15 dB</td>
</tr>
<tr>
<td>Time Delay Spread</td>
<td>¼ of symbol period</td>
</tr>
<tr>
<td>Power Delay Profile</td>
<td>2-ray with equal power</td>
</tr>
<tr>
<td>Antenna Diversity</td>
<td>1</td>
</tr>
</tbody>
</table>

III. ONE-PASS EXPLICIT FMO MAP GENERATION

Wireless video transmission system is characterized by the burst error nature caused by the time varying channel. FMO can possibly alleviate the effect of data losses by managing the spatial relationship between regions that are coded in each slice and also serves as a macroblock level interleaving tool that can spread out consecutive burst error. In this section, we explain the proposed framework of one-pass explicit FMO map generation.

A. FMO map using Bit-Count and Distortion Indicators

From our previous works in [5], we proposed to use bit count as an indicator for the importance of the macroblock in
video frames. To generate macroblock-to-slice group maps, two pass encoding was used. In the first pass, the statistics of bit count information are collected to generate the corresponding map by the sorting of bit-count of each macroblock to different slice groups. In the second pass, the generated explicit map is used to encode the videos. The results from [5] showed that, using the proposed method can reduce the number of undecodable macroblocks. This proposed method based on spatial information. Nevertheless, our previously proposed work in [6], we proposed and analyzed another indicator, called distortion measure from concealment error, as the temporal information in video frames to indicate macroblock importance. Given a packet corresponded to area of pixels in macroblocks is error, the undecodable macroblock at the decoder will then be concealed using non-motion compensated error concealment method (MB copy). The MB copy method is the simplest non-motion compensated temporal error concealment method. Here the missing macroblocks in a frame are replaced by the spatially corresponding macroblocks in the previous frame. The distortion due to the error concealment based on the sum of absolute difference (SAD). The computed distortion measure per macroblock is then sorted in ascending order. The macroblocks with high distortion measures are assigned to different slice groups, in order to interleave the seemingly important macroblock into different slice groups, as shown block diagram of two-pass explicit FMO in Fig.1.

B. One-Pass FMO Map Generation

The downside of two-pass explicit FMO map generation is that it needs two-pass encoding which introduces extra delay in the encoding process. This encoding delay may make it impractical for live video transmission. In this paper, FMO explicit map using the indicators proposed in [5-6] has been extended to fit one-pass encoding framework. The block diagram of one-pass FMO map generation is shown in Fig.2.

We propose to use one-pass explicit FMO mapping using bit-count information at encoder of the previous frame is collected to generate explicit map of the current frame. The basic idea of our approach is that the macroblocks that use a higher number of bits are the more important macroblock due to the properties of motion-compensated prediction. Therefore, we try to interleave consecutively the macroblock with high bit-count to be in a different slice. The steps to generate one-pass explicit FMO map using bit count information is shown, as in Fig. 3. An example of the MBAmap of the carphone sequence using the slice group mapping technique is shown as in Fig.4.
We propose and analyze another indicator, called a distortion measure from concealment error, as the temporal information in video frames to indicate macroblock importance. To generate MBAmap, one-pass encoding is used. Pixel values of the previous frame are collected to generate explicit FMO map of each frame.

Given a packet corresponding to an area of pixels in macroblocks that have an error, the undecodable macroblock at the decoder will then be concealed using non-motion compensated error concealment method. The distortion due to the error concealment, $D_{CE}$, based on the sum of absolute difference (SAD), can be computed, as shown in eq. (1),

$$D_{CE} = \sum_{(x,y) \in \text{damaged area}} |f_{k-1}(x,y) - f_{k-2}(x,y)|$$

(1)

where a frame $k-1^{th}$ and $k-2^{nd}$ are the previous frame and the second previous frame, respectively, $f_{k-1}(x,y)$ is the reconstructed pixel value at the coordinate $(x,y)$ of a previous frame, $f_{k-2}(x,y)$ is the reconstructed pixel value at the coordinate $(x,y)$ of a second previous frame and $L$ is the damaged area due to the error packet.

We propose and analyze another indicator, called a distortion measure from concealment error, as the temporal information in video frames to indicate macroblock importance. To generate MBAmap, one-pass encoding is used. Pixel values of the previous frame are collected to generate explicit FMO map of each frame.

Given a packet corresponding to an area of pixels in macroblocks that have an error, the undecodable macroblock at the decoder will then be concealed using non-motion compensated error concealment method. The distortion due to the error concealment, $D_{CE}$, based on the sum of absolute difference (SAD), can be computed, as shown in eq. (1),

$$D_{CE} = \sum_{(x,y) \in \text{damaged area}} |f_{k-1}(x,y) - f_{k-2}(x,y)|$$

(1)

where a frame $k-1^{th}$ and $k-2^{nd}$ are the previous frame and the second previous frame, respectively, $f_{k-1}(x,y)$ is the reconstructed pixel value at the coordinate $(x,y)$ of a previous frame, $f_{k-2}(x,y)$ is the reconstructed pixel value at the coordinate $(x,y)$ of a second previous frame and $L$ is the damaged area due to the error packet.

$$D_{CE} = \sum_{(x,y) \in \text{damaged area}} |f_{k-1}(x,y) - f_{k-2}(x,y)|$$

(1)

where a frame $k-1^{th}$ and $k-2^{nd}$ are the previous frame and the second previous frame, respectively, $f_{k-1}(x,y)$ is the reconstructed pixel value at the coordinate $(x,y)$ of a previous frame, $f_{k-2}(x,y)$ is the reconstructed pixel value at the coordinate $(x,y)$ of a second previous frame and $L$ is the damaged area due to the error packet.
No FMO and [6] respectively. While in the case of fast fading, the 1P_FMO_bitCnt can reduce, on the average, the number of undecodable macroblocks of up to 77.6%, 3.39% compare to No FMO and [5] respectively. And the 1P_FMO_Dist can reduce, on the average, the number of undecodable macroblocks of up to 77.8%, 0.54% compared to No FMO and [6] respectively.

For the 16th frame of carphone sequence, a comparison of the decoded image quality of the proposed method under slow and fast fading are shown in Figs 6(a)-(h) carphone sequence has low motion content. Therefore in Fig 6(h) the decoded image can be very close to the original image. Comparison of PSNR in table 1 shows that our proposed method obtains the best objective quality. For subjective quality, as shown in the Figs. 6 the face of carphone is clearer as compared with the other methods.

Overall, using one-pass explicit FMO from bit-count and distortion indication can improve the PSNR with the number of undecodable being comparable. This is because of the unequal importance of each MB. Every frame is divided to 8 slice group, so in some case PSNR and undecodable MB is very close. Previous researches in [5-6] used two-pass encoder. In the first pass, the collection of parameters for FMO map generation takes place. In the second pass, the video sequence is encoded according to generated explicit FMO map. Thus, the two-pass method in [5-6] is more suitable for encoding non real-time video transmission. The additional complexity to the encoder is considered negligible, because only the FMO map generation process is changed, the encoder structure is the same as the JM reference encoder. This has the minimal impact when in comparison to the other more time consuming task in the decoder, such as motion compensation. In summary, this method is suitable for low resolution video sequence in real time transmission such as mobile communication and video conference.

ACKNOWLEDGMENT

This work has been supported in part by CHE-AUN/Seed-Net Scholarship and the Collaborative Research Project titled Wireless Video Transmissions.

V. CONCLUSIONS

In this paper, we proposed explicit FMO map is used one-pass encoding based on bit count information which acts as spatial information and distortion measure based on the concealment error which acts as temporal information for a choice of macroblock-address-map of each picture. Our simulation results performed under slow and fast fading channels confirm that the proposed technique can reduce the number of undecodable macroblock can reduce up to 67.66% and 80.58%, respectively. The PSRN improvements are up to 4.47 dB and 6.23 dB, respectively when compared to the method of using No FMO. Future works involve the extension of integrating the one-pass explicit FMO and error concealment method for improve the better video quality.
Fig. 6 16th carphone sequence (slow fading)  (a) Original, (b) No FMO, (c) FMO using bit count, (d) FMO using distortion measure, (e) One-pass FMO using bit count, (g) One-pass FMO using distortion measure

REFERENCES