PAPER

New Error Resilience Technique Using Adaptive FMO and Intra Refresh for H.264 Video Transmission

Tien HUU VU[†], *Nonmember*, **Supavadee ARAMVITH**[†a)], *Member*, **and** **Yoshikazu MIYANAGA**[††], *Fellow*

SUMMARY   In this paper, we propose an error resilience scheme for wireless video coding based on adaptive flexible macroblock ordering (FMO) and intra refresh. An FMO explicit map is generated frame-by-frame by using prior information. This information involves estimated locations of guard and burst sections in the channel and estimated effect of error propagation (EEP) from the previous frame to the current frame. In addition, the role of the current frame in propagating an error to the next frame is also considered. A suitable intra refresh rate which is adaptive to the channel state is used to reduce the dependence between frames and thus can stop the EEP. The results in experiments show that the proposed method gains some improvements in terms of peak signal-to-noise rate (PSNR) as compared with some other methods that have not considered the channel condition and the error propagation in generating an FMO map.
key words: H.264, FMO, error propagation, intra refresh, error resilience

1. Introduction

As a new tool of H.264/AVC, the FMO scheme [1] enables an image to be divided into regions called slice groups. Each slice group can be divided into several slices and a slice can also be decoded independently. An identification number for each macroblock (MB) is given by a macroblock allocation map (MBAmap) to specify which slice group that MB belongs to. The number of slice groups is limited to 8 for each picture to avoid complex allocation schemes. A suitable MBAmap disperses the important MBs into some slice groups. Because of independence between slice groups, if a slice group is lost during transmission, the important MBs in the other slice groups are not affected.

There are seven types of FMO map in which six types are default and a type called explicit type is defined by users. Although the explicit type comes with a higher overhead cost. It helps spread out burst errors and can result in reducing the number of error MBs in a frame while incurring the same amount of overhead bits as the method using forward error correction code (FEC). In addition, the explicit type is more flexible in generating the FMO map comparing to the other FMO types. Thus the explicit type is used in this work.

To design slice group maps, previous approaches use an indicator to evaluate the importance of an MB. These indicators can be bit-count [2], distortion-from-error conceal-

ment [3], MB impact factor [4] or spatial temporal indicator [5]. After that, the MBs are sorted in descending order of importance and arranged consequently to 8 slice groups. As another approach to generate an FMO map, the MBs firstly are divided into two slice groups as type 1 of an FMO map. Then these slice groups are split further into smaller slice groups according to the impact factor of the MBs [6]. By dispersing the important MBs, the number of lost MBs is reduced. However, the important MBs are not thoroughly protected from the errors. Because of variable length coding, if an MB is lost, the following important MBs in the same slice group are also affected. Therefore, using only the FMO is not enough to reduce the number of lost important MBs.

In [3], a channel state is considered in order to combine with the FMO. If a channel state is bad, the FMO and interleaving are used. In [7], redundant slices are added depending on the characteristic of the wireless channel. In this method, the ROI slices are transmitted during good state of the channel. During bad state, redundant slices are transmitted. Nevertheless even in the good state, there may be some errors in the channel. Thus, the important MBs in good state may still be affected by error. Furthermore, by adding redundant slices, compression efficiency of video codec is reduced.

When transmitting video signals over error-prone channels like wireless channels, beside the error caused by the transmission channel, the error propagated between frames is also taken into account. To stop the error propagation, intra refresh algorithm is used. However, selecting a suitable intra refresh rate is a problem that needs to be considered. The intra MBs can effectively stop the error propagation, but the number of intra MBs in a frame is limited by target bit rate. Because of bit-consuming in intra mode, a coarser quantization process is used to achieve a given target bit rate. This results in the degradation of video quality. Moreover, with a limited target bit rate allocated for a group of pictures (GOP), the frame having too many intra MBs consumes a large number of target bits and thus affects to the target bit of the next frames in the same GOP [8]. Thus, it is necessary to balance the benefit of reducing effect of error propagation and the drawback of using a number of intra coded MBs in a frame. In [9], a fixed number of MBs having highest distortion in the current frame are coded in intra mode. In addition, there have been some researches based on joint source-channel rate-distortion optimization to decide intra refresh rate [10]–[13]. In these methods, the relationship
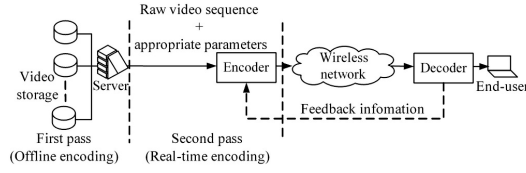
Fig. 1 Video streaming application over wireless channel.

Fig. 2 Packet sequence for an error channel.

between end-to-end distortion and intra refresh rate is established. Then, the encoder has to try all coding modes for each MB to choose a suitable intra refresh rate to achieve minimum distortion.

In this work, the proposed scheme is applied for the video streaming applications such as video-on-demand and video news archive. In these applications, the video sequences are encoded in advance to extract motion vector and bit-count of MBs in every frame of video sequences. These parameters are stored in a database. When an end-user requests a video sequence, the encoder retrieves the requested raw video sequence and appropriate data from a database, as shown in Fig. 1.

The main contribution of our work is to propose an error resilience framework considering channel prediction to generate an adaptive FMO map and adopting a suitable intra refresh rate for H.264 video transmission. Initially, a three-state Markov model is used to locate the error bursts of channel. Then, a technique is introduced to evaluate the importance of MBs based on the error propagation between frames. Finally, an explicit FMO map of the current frame is generated as the rule: The MBs with low importance are arranged in slice groups which are transmitted in the error burst sections and the MBs with high importance are arranged in slice groups which are transmitted in the error-free sections. Furthermore, based on the estimated channel state, a suitable intra refresh rate is computed to stop error propagation.

This paper is organized as follows. Three-state Markov model and a method proposed to locate burst and guard sections are introduced in Sect. 2. Section 3 introduces the method to generate an explicit FMO map. Section 4 shows simulation results and discussions. Finally, conclusions are given in Sect. 5.

2. Locating Burst and Guard Sections over Wireless Channel

2.1 Three-State Markov Model [14], [15]

There are many different methods used to model the wireless channel. In [16], channel is modeled by two-state Markov model whose transition probabilities are functions of the channel characteristic. However, using two-state model is inadequate in locating guard and burst sections of channel. As an extension of two-state model, finite-state Markov model is analyzed in [17]. The authors indicated that the higher state of Markov model is used, the more accurate in capturing the error burst nature of channel. In addition,
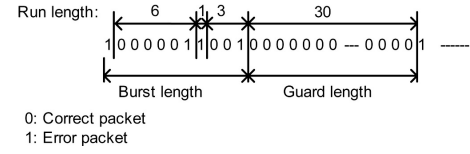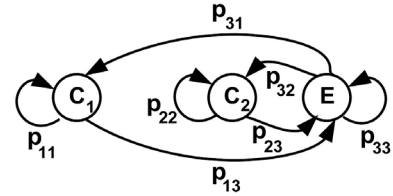

Fig. 3 Three-state Markov model.

the complexity level of model is directly proportional to the number of states in Markov model. With the limitation of two-state model and the cost in computing of higher state models, three-state Markov model is selected to estimate channel state.

In this section, three-state Markov model in [14] and [15] is introduced. However, instead of using model at bit level, we apply three-state Markov model at packet level to estimate the position of the error bursts in wireless channel. Figure 2 shows an example of a packet sequence in an error channel. Similarly to [14] and [15], we define the following definitions at packet level. A guard section is defined as a duration in which all packets are error-free. A burst section is defined as a duration sandwiched between guard sections. From now on, the section means guard or burst section. Minimum guard length is the minimum number of error-free packets a guard section should have. In this system, the minimum guard length is chosen to be 30 packets for computing the transition probability. Thus, each guard section is longer than or equals to 30 consecutive error-free packets. The run length is defined as the length from an error packet to the next error packet excluding the first error packet. The first return probability $P(i)$ is defined as the occurrence probability of each run length i. Figure 3 shows the transition probabilities of a three-state Markov model where C_1 and C_2 show the error-free states which are state 1 and state 2, and E shows the error state which is state 3. C_1 shows the error-free state in a guard section while C_2 and E show the error-free and the error state in a burst section. p_{nm} is the probability of transition from state n to state m. p_{nn}^{i-2} is the probability of the case in which there are $(i-2)$ consecutive transitions from state n to state n. The first return probabilities are computed by

$$
\begin{aligned}
P(1) &= p(E/E) = p_{33} \\
P(2) &= p(C_1, E/E) + p(C_2, E/E) \\
&= p_{31}p_{13} + p_{32}p_{23} \\
&\cdots \\
P(i) &= p(C_1, \ldots, C_1, E/E) + p(C_2, \ldots, C_2, E/E) \\
&= p_{31}p_{11}^{i-2}p_{13} + p_{32}p_{22}^{i-2}p_{23}
\end{aligned}
\tag{1}
$$

where $p(\alpha/\beta)$ means that β is the first state and then the sequence of α occurs. The transition probabilities are computed by

$$
\begin{aligned}
p_{33} &= P(1) \\
p_{32} &= \sum_{i=2}^{L_{min}} P(i) \\
p_{23} &= \frac{p_{32}}{\sum_{i=2}^{L_{min}} (i-1)P(i)} \\
p_{31} &= 1 - p_{33} - p_{32} \\
p_{13} &= \frac{1}{\sum_{i=L_{min}+1}^{\infty} \frac{(i-1)P(i)}{p_{31}} - L_{min}} \\
p_{22} &= 1 - p_{23} \\
p_{11} &= 1 - p_{13}
\end{aligned}
\tag{2}
$$

where L_{min} is the minimum guard length. According to [14], the average guard length, L_G, and average burst length, L_B, are computed as

$$
\begin{aligned}
L_G &= L_{min} + \frac{1}{p_{13}} \\
L_B &= \frac{\frac{p_{31}}{p_{13}} + \frac{p_{32}}{p_{23}} + 1}{p_{31}p_{11}^{L_{min}-1}} - \left(1 + \frac{1}{p_{13}}\right)
\end{aligned}
\tag{3}
$$

2.2 Proposed Method to Locate the Error Burst Positions

Assume that before encoding the current frame (n), the encoder receives channel feedback information containing position of error packets of frame $(n-2)$. In this work, a roundtrip delay of less than 100 ms is assumed enabling the feedback delay of one frame. Because of the error propagation, the longer feedback delay affects the decrease in the PSNR value. In fact, the average PSNR drop due to higher value of feedback delay for slow fading case is around 0.6 dB while for the fast fading case is 0.1 dB when feedback delay is varied experimentally with values from 1 to 4 frames. Thus, one frame for feedback delay is selected in the system.

Based on feedback information, the encoder uses Eq. (3) to estimate values L_G and L_B of channel. Then values of L_G and L_B are updated when the encoder receives the next feedback information. To estimate the location of the first burst or guard section in frame $(n-1)$, the encoder uses feedback information to locate burst and guard sections in frame $(n-2)$. If a border between frame $(n-2)$ and frame $(n-1)$ is in a guard section (as shown in Fig. 4), distance from the last burst in frames $(n-2)$ to the first burst in frame $(n-1)$ is L_G packets. Otherwise, if the border is in a burst section, the distance from the last guard in frame $(n-2)$ to the first guard in frame $(n-1)$ is L_B packets. The next sections in frame $(n-1)$ and frame (n) are estimated from the position of the first section in frame $(n-1)$.
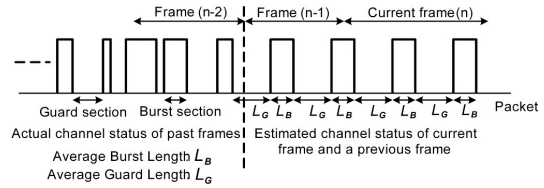

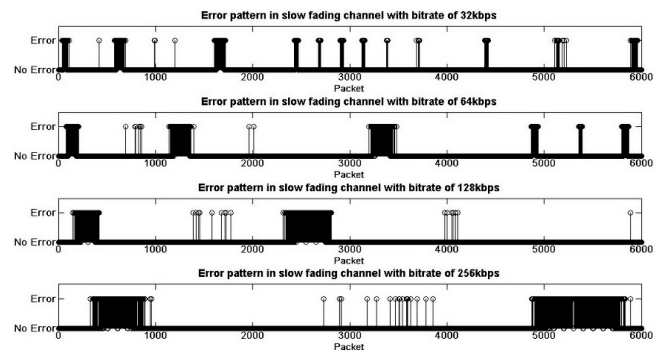**Fig. 4** Locating burst and guard section for the current frame.

Fig. 5 Error patterns of wireless channel simulator in slow fading at different bit rates.
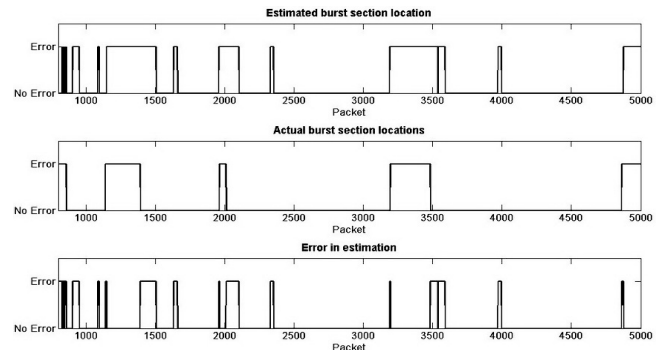

Fig. 6 Estimated burst locations of wireless channel for the Akiyo sequence in slow fading, *bitrate* = 64 kbps.

Table 1 Percentage difference (P_d) between estimated and actual locations of burst sections.

Bitrate	Akiyo		Foreman		Claire		Carphone	
	Slow	Fast	Slow	Fast	Slow	Fast	Slow	Fast
32 kbps	5.7	8.0	5.6	10.5	5.5	6.8	5.1	8.7
64 kbps	5.3	7.8	4.5	9.0	4.4	4.6	5.5	7.0
128 kbps	5.9	11.2	4.2	7.8	4.8	7.7	4.9	6.8
256 kbps	5.5	7.8	4.6	7.9	4.2	7.9	4.4	7.9

In this work, to validate the accuracy in locating burst sections of wireless channel, the estimated burst locations are compared with the actual burst locations of wireless channel which are generated from wireless channel simulator [18]. Figure 5 shows actual burst locations of wireless channel at different bit rates. Figure 6 illustrates an example of estimated burst locations when transmitting 100 frames of the "Akiyo" sequence over slow fading channel at bit rate of 64 kbps.

Table 1 shows the percentage difference of four video

sequences "Akiyo", "Foreman", "Claire" and "Carphone" in both slow and fast fading cases. The percentage difference is calculated by

$$P_d = \frac{e}{T}.100\% \qquad (4)$$

where P_d is the percentage difference. e is the number of errors in estimation. T is total number of packet of a video sequence.

3. Adaptive Explicit FMO Map Generation

In this section, the importance of an MB is estimated from distortion caused by the error propagation. After that, an explicit FMO map of the current frame is generated by mapping the high important MBs and the low important MBs into slice groups which are transmitted in guard and burst sections respectively.

3.1 The Estimation of MB Importance

To estimate the importance for an MB in the current frame, firstly, we estimate the distortion at that MB caused by error MBs in the past frame. Secondly, we measure the EEP caused by that MB to the next frame. The sum of distortion and the EEP are considered as the importance of that MB. This importance and the estimated burst locations are taken into account to generate an FMO map and to decide inter/intra coding mode for that MB.

Figure 7 describes the error propagation from the past frame $(n-2)$ to the current frame (n) and to the next frame $(n+1)$.

Step 1: Compute the distortion at the frame $(n-1)$ caused by an error pixel in the frame $(n-2)$:
Assume that a pixel j in the frame $(n-1)$ refers to a pixel s in the frame $(n-2)$.

If the pixel s in the frame $(n-2)$ is error, the decoder will copy the pixel s of the frame $(n-3)$ while assuming that the non-motion compensated error concealment method is used. Therefore, the distortion at the pixel j in the frame $(n-1)$, $D(s, j, n-1)$, is computed as

if j is inter coded:
$$D(s, j, n-1) = \Big\| \big| f(j, n-1) - f(s, n-2) \big| \\ - \big| f(j, n-1) - f(s, n-3) \big| \Big\| \qquad (5)$$
if j is intra coded:
$$D(s, j, n-1) = 0$$

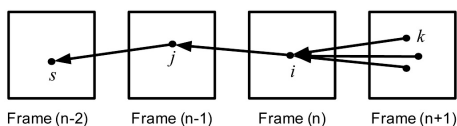where $f(x, y)$ is reconstructed value of the pixel xth in the frame yth.

Fig. 7 Error propagation from the past frame to the next frame.

If the pixel s in the frame $(n-2)$ is error-free, the distortion at the pixel j in the frame $(n-1)$ is computed in Eq. (6).

$$D(s, j, n-1) = 0 \qquad (6)$$

Step 2: Compute distortion at the frame (n) caused by an error pixel in the frame $(n-2)$:
Assume that the pixel i in the frame (n) refers to the pixel j in the frame $(n-1)$.

If the pixel j is inter coded, the distortion at the pixel i is computed as:

if j is error:
$$D(j, i, n) = \Big\| \big| f(i, n) - f(j, n-1) \big| \\ - \big| f(i, n) - f(j, n-2) \big| \Big\| \qquad (7)$$
if j is error-free:
$$D(j, i, n) = D(s, j, n-1)$$

In conclusion, in the case pixel j is inter coded, the distortion at the pixel i is computed as

$$D(j, i, n) = q(j) \Big\| \big| f(i, n) - f(j, n-1) \big| - \big| f(i, n) \\ - f(j, n-2) \big| \Big\| + (1 - q(j)) D(s, j, n-1) \qquad (8)$$

where $q(j)$ is error probability of pixel j. The $q(j)$ depends on the error probability of packets and the length in packets of the MB containing pixel j.

If the pixel j is intra coded, the distortion at the pixel i is computed as

if j is error:
$$D(j, i, n) = \Big\| \big| f(i, n) - f(j, n-1) \big| \\ - \big| f(i, n) - f(j, n-2) \big| \Big\| \qquad (9)$$
if j is error-free:
$$D(j, i, n) = 0$$

In conclusion, in the case j is intra coded, the distortion at pixel i is computed as shown in Eq. (10).

$$D(j, i, n) = q(j) \Big\| \big| f(i, n) - f(j, n-1) \big| \\ - \big| f(i, n) - f(j, n-2) \big| \Big\| \qquad (10)$$

Step 3: Compute the EEP at the frame $(n+1)$ caused by pixel i in the frame (n):
To compute the EEP from the past frame to the next frame through the current frame, all MBs in the current frame are coded in inter mode in the first pass. Assume that the pixel i in the current frame is referred by a pixel k in the next frame. The distortion caused by error propagation for the pixel k is computed by

$$D(i, k, n+1) = q(i) \Big\| \big| f(k, n+1) - f(i, n) \big| \\ - \big| f(k, n+1) - f(i, n-1) \big| \Big\| \\ + (1 - q(i)) D(j, i, n) \qquad (11)$$

where $D(j, i, n)$ is computed as shown in Eq. (8) or Eq. (10)

depending on coding mode of the pixel j. The overall distortion propagated from pixel i in the current frame (n) to the next frame $(n + 1)$ is computed as

$$I(i, n) = \sum_{k \in \{N\}} D(i, k, n + 1) \tag{12}$$

where N is the number of pixels in the frame $(n + 1)$ which refer to the pixel i in the current frame (n).

Step 4: Estimate the importance of an MB in the current frame:

The importance of the lth MB in the current frame is computed as

$$D_l(n) = \sum_{i,j=1}^{M} D(j, i, n) + \sum_{i=1}^{M} I(i, n) \tag{13}$$

where $D(j, i, n)$ is computed in Eq. (8) or Eq. (10). $I(i, n)$ is computed in Eq. (12). M is the number of pixels in an MB. In this case, $M = 256$.

3.2 Adaptive Explicit FMO Map Generation

Based on the importance and the bit-count derived in the first pass, the MBs are arranged into slice groups of an FMO map according to the following criteria:

- The MBs considered as high importance (computed by Eq. (13)) are arranged into the slice groups which are transmitted in the guard sections of channel. The other MBs are arranged into the slice groups which are transmitted in the burst sections.
- The total number of packets of the MBs in the slice groups equals to the length in packet of burst or guard section in the channel over which those slice groups are transmitted.

Figure 8 illustrates the estimated burst locations of channel before encoding the frame 29th in the "Akiyo" sequence. As the estimation, the length in packet of frame is 34. In this duration of channel, there are two burst and two guard sections. Based on the bit-count and the importance of MBs, the encoder arranges the MBs in such a way that the highest important MBs are encoded and transmitted in the first guard section. The next are low important MBs which are transmitted in the first burst section. The second guard section contains the next high important MBs and the second burst section contains the rest of MBs of the current frame. The order of encoding and transmission of the MBs are shown in Fig. 8. Note that in a slice group, the encoding order is sorted in ascending order of the MB identification. As the result, a slice map for the frame 29th is generated as shown in Fig. 9.

Because of using variable length codes (VLC) in the entropy coding, the decoder relies on the length of previously decoded symbols to be able to decode the current symbol. Hence, an error MB can cause errors in the rest of MBs in the slice group. To reduce the effect of error propagation in a slice group, the MBs in each section are separated into
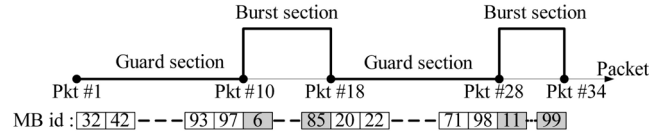

Fig. 8 Estimated burst locations for frame 29th Akiyo sequence.

Fig. 9 Explicit FMO map of frame 29th Akiyo sequence.

some slice groups. The number of slice groups in a section, N_{slg}, is computed by

$$N_{slg} = \left\lceil \frac{N_{mb}}{\gamma} \right\rceil \tag{14}$$

where $\lceil z \rceil$ is the smallest integer not less than z. N_{mb} is the number of MBs selected to be filled in a burst or a guard section. For the QCIF frames, to guarantee 99 MBs are fairly distributed in 8 slice groups, the number of MBs in a slice group, γ, is selected to be 13. As shown in Fig. 9, the number of MBs in the first guard section is 19. Therefore, the MBs in this section are mapped into two slice groups including the MBs from 32 to 96 and from 44 to 97.

In the guard section, the MBs having high importance are coded in intra mode. The number of intra MBs, N_{intra}, is computed as

$$N_{intra} = \lambda N_{slg} \tag{15}$$

To balance the compression efficiency and the effect of error propagation, λ is empirically selected to 2. Furthermore, these intra MBs are dispersed into slice groups to reduce the number of lost intra MBs in case channel prediction is not precise. In the frame 29th, with each slice group in the guard section, the encoder selects two MBs having highest importance for intra coding mode. Thus, there are eight intra MBs in two guard sections of the frame. According to Eq. (15), the number of intra MBs in frames varies from 0 (if there is no guard section in transmission duration of frame) to 16 (if the whole frame is transmitted in a guard section).

4. Simulation Results and Discussions

4.1 Experimental Set-Up

In the study, wireless channels are simulated for video transmission using 2-ray Rayleigh Fading channel [19]. The block diagram of the wireless channel simulator and the system parameters are shown in Fig. 10 and Table 2, respec-
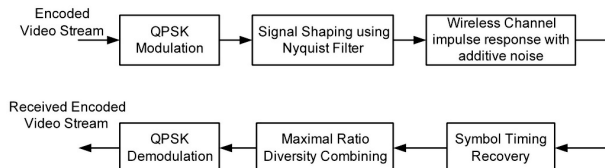
Fig. 10 Block diagram of wireless video transmission system.

Table 2 Wireless channel and air interface parameters [18].

Multiple access	TDMA
Modulation	QPSK
Channel rate	32, 64, 128, 256 kbps
Maximum Doppler frequency	1 Hz (slow fading), 40 Hz (fast fading)
Transmitted signal power	15 dB
Time delay spread	$\frac{1}{4}$ of symbol perriod
Power delay profile	2-ray with equal power
Antenna diversity	1

Table 3 Comparison of the average PSNR (dB) at 64 kbps, 10 fps.

PSNR (dB)	Akiyo		Foreman		Claire		Carphone	
	Slow	Fast	Slow	Fast	Slow	Fast	Slow	Fast
No FMO	30.6	26.2	19.1	17.2	26.9	24.6	23.4	20.7
Bit-count	34.5	30.6	23.3	19.4	30.6	27.6	25.6	22.8
STI-FMO	33.8	31.8	22.6	**21.5**	31.8	29.4	24.9	23.3
Proposed method (no IR + no CP)	34.7	32.5	23.6	21.3	33.1	29.0	25.1	24.5
Proposed method (no IR + CP)	**35.3**	**33.3**	**24.1**	20.2	**33.7**	**30.6**	**27.2**	**25.7**

Table 4 Comparison of the average PSNR (dB) at 256 kbps, 30 fps.

PSNR (dB)	Akiyo		Foreman		Claire		Carphone	
	Slow	Fast	Slow	Fast	Slow	Fast	Slow	Fast
No FMO	31.0	27.9	22.2	20.3	27.7	26.2	25.8	22.6
Bit-count	34.0	32.4	24.4	22.4	33.3	31.5	26.0	24.6
STI-FMO	34.1	33.0	25.1	**24.4**	3 3.5	**32.2**	26.6	24.7
Proposed method (no IR + no CP)	35.8	33.3	25.2	23.2	34.0	31.0	27.2	25.7
Proposed method (no IR + CP)	**36.5**	**34.2**	**25.4**	23.5	**34.5**	31.4	**29.2**	**25.9**

tively. One important wireless channel simulator's parameter is Doppler frequency. In our study, we are interested to test our technique in both slow fading channel, i.e., Doppler frequency = 1 Hz, and fast fading channel, i.e., Doppler frequency = 40 Hz. We used H.264 reference software JM 9.2 [20] with baseline profile in simulation. The video sequences in the QCIF format (176×144 pixels/frame) are coded at 32, 64, 128, and 256 kbps with frame rates of 10, 20, and 30 fps. The following video sequences are used in the experiment: Akiyo, Foreman, Claire, and Carphone. At the decoder, the non-motion compensated error concealment is used.

4.2 Simulation Results

In this experiment, to evaluate the efficiency of using the EEP as an indicator for the FMO as well as the effectiveness of channel prediction, the proposed method using channel prediction (CP) and without channel prediction (no CP) are compared with some other methods using different indicators, including bit-count [2] and spatial-temporal indicator (STI) [5]. In the no CP method, the step of estimating the locations of bursts and guards section in channel is skipped. To generating an explicit FMO map, the MBs in a frame are arranged consequently into slice group according to the EEP values. Moreover, to validate the method of selecting the MBs to be intra coded and computing intra refresh rate, the proposed method is compared with methods using fix intra refresh rate (FIR) [9] and random intra refresh rate (RIR) [20]. In comparisons, we use both subjective and objective measures. For objective measure, the PSNR is used as the performance metric in quantifying the effectiveness of methods.

Tables 3 and 4 show the average PSNR of video sequences in the scenario of slow and fast fading channels. The simulation results show that the proposed method without intra refresh and channel prediction (no IR + no CP) gains a higher average PSNR than the conventional methods. Especially, if compared with the method without the

FMO, the improvement of the average PSNR is up to 7 dB. However, in some cases, the PSNR of the new method is lower. This is because the quality of measurements of video in terms of the PSNR depends solely on the locations of the error bits as well as the error concealment method applied. In this experiment, simple non-motion compensated error concealment is used, therefore we expect that the higher PSNR improvement can be achieved if more sophisticated technique of error concealment is used in further study. Results show that the average PSNR is improved when channel prediction is applied to the proposed method (no IR + CP). This is because the number of lost important MBs is reduced when the locations of burst and guard sections are estimated. However, in the fast fading case, the improvement does not significantly increase in comparison with the case without channel prediction. The reason is that channel prediction algorithm is more precise in the slow fading case. In the fast fading case, there are more errors in locating burst and guard sections. Therefore, the number of lost important MBs in the fast fading case is higher than that of the slow fading case.

Figures 11 and 12 show the PSNR curve of the "Carphone" test sequence in the slow and fast fading case, respectively. From the curves, it can be observed that the average PSNR of the proposed method is higher than the others. This improvement is achieved by using an accurate method in stopping the effect of error propagation. Furthermore, by estimating locations of burst and guard sections, the important MBs are put in the error-free sections, thus the number of lost important MBs is reduced. Consequently, the PSNR of the new method is increased.

Tables 5 and 6 show the average PSNR of video sequences when comparing the proposed method using adaptive intra refresh rate with other methods using fix and random intra refresh rate. In the first case, the proposed method uses a fix intra refresh rate without channel prediction (FIR + no CP). The fix intra refresh rate is 11 MBs per frame. The results show that with considering effect of error propagation from the current frame to the next frame, the proposed
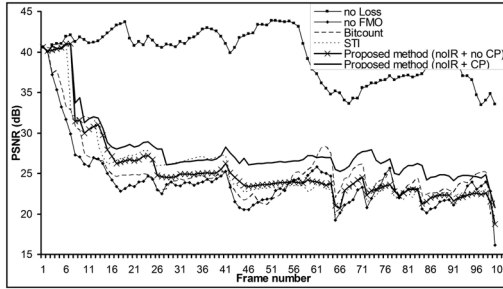
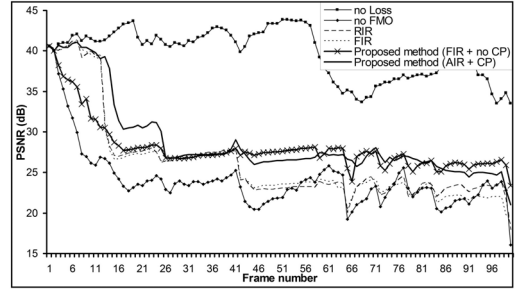**Fig. 11** PSNR comparison of "Carphone" in slow fading, 128 kbps, 20 fps.

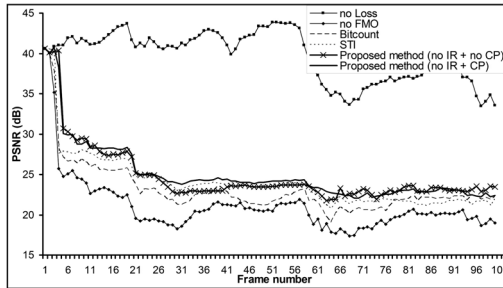**Fig. 13** PSNR comparison of "Carphone" in slow fading, 128 kbps, 20 fps.

Fig. 12 PSNR comparison of "Carphone" in fast fading, 128 kbps, 20 fps.
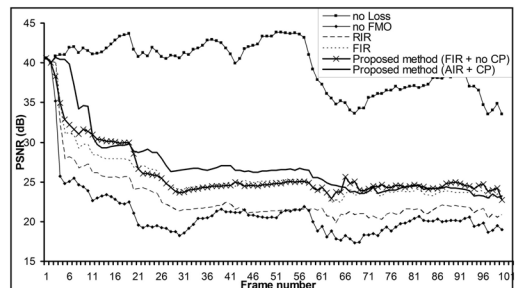

Fig. 14 PSNR comparison of "Carphone" in fast fading, 128 kbps, 20 fps.

Table 5 Comparison of the Average PSNR (dB) at 64 kbps, 10 fps.

PSNR (dB)	Akiyo		Foreman		Claire		Carphone	
	Slow	Fast	Slow	Fast	Slow	Fast	Slow	Fast
RIR	33.2	31.4	21.2	19.3	31.8	29.9	25.8	22.6
FIR	34.9	31.1	21.0	20.0	32.2	29.1	26.8	25.0
Proposed method (FIR + no CP)	34.5	31.9	22.6	21.2	32.4	29.7	27.3	24.2
Proposed method (AIR + CP)	**35.8**	**33.6**	**24.2**	**22.1**	**33.4**	**32.0**	**28.3**	**26.6**

Table 6 Comparison of the Average PSNR (dB) at 256 kbps, 30 fps.

PSNR (dB)	Akiyo		Foreman		Claire		Carphone	
	Slow	Fast	Slow	Fast	Slow	Fast	Slow	Fast
RIR	34.7	33.8	22.5	21.5	32.4	31.4	26.9	22.9
FIR	35.0	34.4	24.4	**23.5**	33.0	32.6	27	26.3
Proposed method (FIR + no CP)	33.5	33.4	24.4	22.0	32.3	**33.2**	29.5	26.5
Proposed method (AIR + CP)	**36.9**	**34.5**	**25.6**	22.7	**34.8**	33.0	**30.6**	**28.2**

Fig. 15 PSNR comparison of "Carphone" in slow fading, frame rate = 20 fps.

Fig. 16 PSNR comparison of "Carphone" in fast fading, frame rate = 20 fps.

method has higher the average PSNR than the FIR method used in [9]. In [9], only the errors propagated from the previous frame to the current frame are taken into account. Consequently, some MBs in the current frame are skipped in evaluating importance because these MBs may be not much affected by error propagation from the past frame. However, these MBs may cause a high distortion for the next frame. Therefore, it is necessary to consider both effects of error propagation from the previous frame to the current frame and from the current frame to the next frame. In the second case, the proposed method uses an adaptive intra refresh rate with consideration of channel prediction (AIR + CP). By locating the burst and guard sections, the number of lost intra
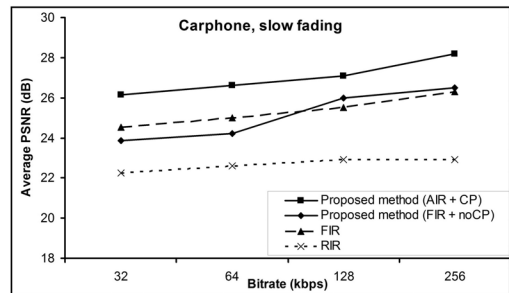
MBs is smaller than that of the other methods. Thus, average PSNR of the proposed method is higher.
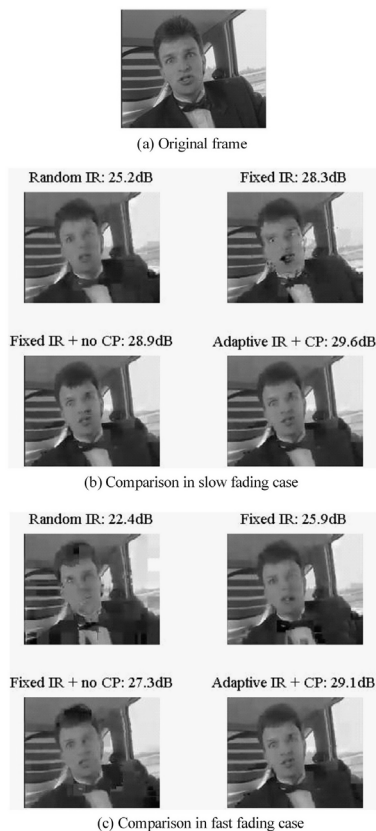
Figures 13 and 14 show the PSNR curve of methods

(a) Original frame

Random IR: 25.2dB   Fixed IR: 28.3dB

Fixed IR + no CP: 28.9dB   Adaptive IR + CP: 29.6dB

(b) Comparison in slow fading case

Random IR: 22.4dB   Fixed IR: 25.9dB

Fixed IR + no CP: 27.3dB   Adaptive IR + CP: 29.1dB

(c) Comparison in fast fading case

Fig. 17 Visual comparision of frame 49th of "Carphone" sequence between methods in slow and fast fading.

5. Conclusions

In this paper, the three-state Markov model is used to estimate the locations of burst and guard sections in channel. In addition, the importance of MB is measured based on the effect of error propagation. With the predicted information of channel, the important MBs are arranged into the guard sections and the unimportant MBs are arranged into the burst sections. Moreover, when considering the inter-frame error propagation, a suitable intra refresh rate is selected based on the channel state to reduce the EEP. Experimental results show that our proposed method gains some improvements in terms of the PSNR as compared with some conventional methods that have not taken the channel condition and the error propagation into consideration in generating an FMO map.

Acknowledgments

References

[1] ITU-T Recommendation H.264, "Advance video coding for generic audiovisual," March 2005.

[2] W. Hanatanong and S. Aramvith, "Analysis of macrobloc-to-slice group mapping for H.264 video transmission over packet-based wireless fading channel," 48th Midwest Symposium on Ctkt and Sys., vol.2, pp.1541–1544, Aug. 2005.

[3] J. Panyavaraporn, R. Cajote, and S. Aramvith, "Performance analysis of flexible macroblock ordering using bit-count and distortion measure for H.264 wireless video transmission," International Workshop on Smart Info-Media Systems in Bangkok (SISB 2007), Bangkok, Thailand, Nov. 2007.

[4] Y. Dhont, P. Lambert, and R. Van der Walle, "A Flexible macroblock scheme for unequal error protection," International Conference on Image Processing (ICIP'06), pp.829832, 2006.

[5] R.D. Cajote, S. Aramvith, R.C.L. Guevara, and Y. Miyanaga, "FMO slice group maps using spatial and temporal indicators for H.264 wireless video transmission," Circuits and Systems, ISCAS 2008, pp.3566–3569, May 2008.

[6] J.Y. Shih and W.J. Tsai, "A new unequal error protection scheme based on FMO," ICIP 2008, 15th IEEE Int'l Conference, pp.3068–3071, Oct. 2008.

[7] B. Katz, S. Greenberg, N. Yarkoni, N. Blaunstien, and R. Giladi, "New error-resilient scheme based on FMO and dynamic redundant slices allocation for wireless video transmission," IEEE Trans. Broadcast., vol.53, no.1, pp.308–319, March 2007.

[8] Institute of Computing Technology, "Proposed draft of adaptive rate control," Joint Video Team (JVT), doc. JVT-H012, 8th meeting Geneva, May 2003.

[9] B. Girod and N. Farber, "Error-resilient standard compliant video coding," available at link: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.51.6841

[10] Z.H. He, J.F. Cai, and C.W. Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," IEEE Trans. Circuits Syst. Video Technol., vol.12, pp.511–523, June 2002.

using "Carphone" video sequence in slow and fast fading. In the slow fading case, because of channel prediction and adaptive intra refresh rate, the PSNR curve of the proposed method is higher than that of "RIR" and "FIR". Since channel prediction is less precise in the fast fading, the average PSNR of "AIR + CP" is lower than "FIR + no CP" from the frame 71 to the frame 100 (Fig. 14). However, the results in Table 5 and Table 6 show that the average PSNR of "AIR + CP" is still higher than the other methods. Comparisons between the proposed method with "FIR" and "RIR" at different bit rates are shown in Figs. 15 and 16.

To further illustrate improvement of the proposed method, some frames from the "Carphone" test sequence are extracted for comparison. Figure 17 depicts qualities of original 49th frame of the "Carphone" sequence and the reconstructed frames from four different methods including "RIR", "FIR", "FIR + no CP" and "AIR + CP" in the slow and fast fading. It can be subjectively seen that the frame quality in "RIR" is severely affected by error. However, in the proposed method, this error can be substantially improved by using intra refresh and channel prediction. In our experimental condition with a normal PC system such as Pentium Dual-Core 1.86 GHz, Windows 7 Ultimate, 1 GB RAM, the proposed method requires a little longer processing time comparing with standard H.264. However, these additional processing time can be covered by using currently DSP and ASIC based systems.

[11] R. Zhang, S.L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," IEEE J. Sel. Areas Commun., vol.18, no.6, pp.966–976, June 2000.

[12] P. Nunes, L.D. Soares, and F. Pereira, "Error resilience macroblock rate control for H.264/AVC video coding," ICIP 2008, 15th IEEE Int'l Conference, pp.2132–2135, 2008.

[13] J. Xu and Z. Wu, "Joint adaptive intra refreshment and unequally error protection algorithms for robust transmission of H.264/AVC video," Multimedia and Expo, 2006, IEEE Int'l Conference, pp.693–696, July 2006.

[14] T. Sato, K. Tokuda, M. Kawabe, and T. Kato, "Simulation of burst error models and an adaptive error control scheme for high speed data transmission over analog cellular system," IEEE Trans. Veh. Technol., vol.40, no.2, pp.443–453, May 1991.

[15] T. Sata, M. Kawabe, T. Kato, and A. Fukasawa, "Throughput analysis method for hybrid ARQ schemes over burst error channels," IEEE Trans. Veh. Technol., vol.40, no.1, pp.110–118, Feb. 1993.

[16] S. Aramvith, I.-M. Pao, and M.-T. Sun, "A rate-control scheme for video transport over wireless channels," IEEE Trans. Circuits Syst. Video Technol., vol.11, no.5, pp.569–580, May 2001.

[17] B.D. Fritchman, "A binary channel characterization using partitioned Markov chains," IEEE Trans. Inf. Theoy, vol.IT-13, no.2, pp.221–227, April 1967.

[18] T.-C. Chen, L.-F. Chang, A.H. Wong, M.-T Sun, and T.R. Hsing, "A real-time software based end-to-end wireless visual communications simulation platform," Proc. SPIE Visual Comms. and Image Processing, vol.3, pp.1068–1074, May 1995.

[19] W.C. Jakes, Jr., ed., Microwave Mobile Communications, Chapter 1.7, pp.68–73, John Wiley & Sons, 1974.

[20] H.264 JM Ref. Software, available at link: http://iphome.hhi.de/suehring/tml/download/

Supavadee Aramvith received a B.S. (first class honors) degree in computer science from Mahidol University, Bangkok, Thailand, in 1993. She received M.S. and Ph.D. degrees in electrical engineering from the University of Washington, Seattle, Washington, in 1996 and 2001, respectively. She joined Chulalongkorn University in June 2001. She is currently an assistant professor and vice chair in Information and International Affairs and head of Communication Engineering Division in the Department of Electrical Engineering, Chulalongkorn University, Bangkok, Thailand. She was IEEE Thailand section secretary (2007 to 2010) and currently serves as IEEE Region 10 educational activities coordinator (2011 to 2012). She also serves as conference manager and executive secretary of IEICE Bangkok area representative. Her research interests include computer vision techniques for surveillance applications, rate-control for video coding, error resilient video coding for wireless video transmissions, and image/video retrieval techniques.

Yoshikazu Miyanaga received his B.Eng., ME., and D.Eng. degrees from Hokkaido University, Sapporo Japan, in 1979, 1981, and 1986, respectively. He is currently a Professor in the Laboratory for Information Communication Networks, Division of Media and Network Technology at the Graduate School of Information Science and Technology, Hokkaido University. His research interests are in the areas of adaptive signal processing, nonlinear signal processing and parallel/pipelined VLSI systems. Dr. Miyanaga is a member of the Institute of Electrical and Electronics Engineers (U.S.A), and the Acoustical Society of Japan.

Tien Huu Vu was born in Vietnam, in 1979. He received the B. Eng. in electrical engineering from Hanoi University of Technology, Hanoi, Vietnam in 2002. He is currently pursuing the Ph.D. degree under the AUN/SEED-Net Scholarship from JICA at the Digital Signal Processing Research Laboratory, Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University.