

# JOINT ADAPTIVE INTRA REFRESHMENT AND UNEQUALLY ERROR PROTECTION ALGORITHMS FOR ROBUST TRANSMISSION OF H.264/AVC VIDEO

*Jin Xu, Zhimei Wu*

Multimedia Communication Research Center  
Institute of Software, Chinese Academy of Sciences (ISCAS), Beijing, China, 100080

## ABSTRACT

In this paper, we propose an efficient intra refreshment algorithm, which can achieve both the global optimization in determining the intra refreshing ratio and the local accuracy in selecting the proper macroblocks to be intra-coded before the current frame is actually encoded. Furthermore, incorporated with our proposed intra refreshment algorithm, an effective UEP scheme based on a dynamic FMO mapping mode is also proposed in this paper to better protect the intra macroblocks in a frame. The experimental results show that our joint intra refreshment and UEP algorithm can remarkably improve the reconstructed video quality in the packet lossy network.

## 1. INTRODUCTION

In the packet-loss environment (i.e., Internet), transmitting the hybrid-coded video may suffer from the error propagation. Intra update of macroblocks (MBs), also called intra refreshing, is a fairly efficient way to stop error propagation because the decoding of an intra MB does not need the information from previous frames. However, the increased error resilience provided by intra refreshing comes at the cost of reducing the coding efficiency. Therefore, there exists a tradeoff between coding efficiency and error resilience when adopting intra refreshing strategy to alleviate the effect of error propagation. To address this problem, many intra refreshing algorithms have been proposed in the literature [1, 2, 3, 5].

The state-of-the-art intra refreshing algorithms are generally based on rate-distortion optimization (RDO) framework [1, 2], in which the end-to-end distortion is taken in account to directly optimize the overall performance. However, the computation complexity of these classical RD-optimized intra refreshing algorithms is rather high. In the recent work [3], based on a new R-D model incorporating the intra refreshing rate, He et al proposed a global optimized adaptive mode selection algorithm which can determine the optimal intra refreshing rate for a frame even before the real quantization and entropy coding is applied to this frame. However, after the optimal intra refreshing rate for a frame is determined, a random selection is used to decide the individual coding mode for each MB,

which may unadvisedly choose the easy-concealed MBs as intra-mode. In this paper, based on a simple yet accurate R-D model incorporating the intra refreshing rate as well as a refined channel distortion estimation scheme, we devise a more efficient algorithm to properly intra-code the more important MBs.

If all the intra MBs of a coded frame can be better protected, we can expect that the effect of error propagation can be more quickly attenuated. In this paper, after the MBs in a frame needed to be intra-coded have been selected by adopting our proposed intra refreshment algorithm, we propose using a dynamic MBs mapping mode of FMO to better protect the intra MBs. The experiment results demonstrate that the proposed FMO scheme is well suited to transport prioritization.

The remaining of this paper is organized as follows. Section 2 gives a detailed description of our proposed integrated error resilience scheme. And then simulation and analysis are presented in the Section 3. At the end, conclusion is given in the Section 4.

## 2. OUR INTEGRATED ERROR RESILIENCE SCHEME

### 2.1. Optimal intra refreshment rate selection

In an end-to-end video coding and transmission system, the overall picture distortion for frame  $n$ , denoted as  $D(n)$ , consists of two parts: the source coding distortion  $D_s(n)$  and the channel distortion  $D_c(n)$ . Due to the sophisticated coding mechanism adopted by the most video standards and the time-varying characteristics of transmission channel, it is still a challenging problem to accurately estimate the overall picture distortion, especially when adaptive intra refreshing is adopted.

For source coding distortion, based on a so-called “ $\rho$ -domain” R-D analysis methodology, He et al introduced a new R-D model, which incorporates the intra refreshing rate  $\beta$  [3]. However, the calculation of its model parameters is not straightforward. Moreover, in order to conduct  $\rho$ -domain R-D analysis, the DCT coefficient distribution of

the intra-coded and inter-coded MBs is required to be known beforehand [4].

In this section, based on the R-D model presented in [4], we introduce a simple yet effective method to estimate the source coding distortion with intra refreshing rate. As reported in the literature [4], the coding bit rate  $R$  is a linear function of  $\rho$ , which is the percentage of zeros among the quantized DCT coefficients:

$$R(\rho) = \theta \cdot (1 - \rho)$$

where  $\theta$  is a constant. Meanwhile, within the  $\rho$ -domain, the distortion  $D$  of a video frame can be approximated by the follow exponential relationship:

$$D(\rho) = \sigma^2 e^{-\alpha(1-\rho)}$$

where  $\sigma$  is the standard deviation of the source data and  $\alpha$  is also an unknown constant. Both  $\theta$  and  $\alpha$  are related to the image content. Since here we only concern about the relation between  $R$  and  $D$ , we can acquire the following R-D function by combing  $R(\rho)$  and  $D(\rho)$ :

$$D(R) = \sigma^2 e^{-\gamma R}$$

where  $\gamma = \alpha / \theta$ . We can easily estimate this single model parameter by using the coding statistics of previous frame. Let  $D_{prev}$ ,  $R_{prev}$  and  $\sigma_{prev}$  be the actual coding distortion, bit rate and source standard deviation of the previous frame. Note that the bit rate does not include the bits for header information and motion vectors. Meanwhile, in the case of H.264 video, the source standard deviation of an inter-coded MB is the standard deviation of its motion-compensation residue while for an intra-coded MB it is the standard deviation of its intra prediction residue. Then, the estimate of  $\gamma$  is:

$$\gamma = \frac{1}{R_{prev}} \cdot \ln\left(\frac{\sigma_{prev}^2}{D_{prev}}\right)$$

Since the source standard deviation of each MB depends on its coding mode, if we change the coding mode of each MB or the intra refreshing rate  $\beta$  of a video frame, the overall source standard deviation of this video frame is also changed. If we denote the mean standard deviation of inter-coded MBs and intra-coded MBs of a video frame as  $\sigma^I$  and  $\sigma^P$ , respectively, the overall source standard deviation of a video frame with intra refreshing rate  $\beta$  can be approximately given by:

$$\sigma_\beta = \beta \cdot \sigma^I + (1 - \beta) \cdot \sigma^P$$

Although we cannot acquire the mean standard deviation of inter-coded MBs and intra-coded MBs for current frame before it is actually encoded, we can estimate these two mean standard deviations from the coding statistics of the previous frame because it is generally assumed that the frame to be coded is not much different from its previous frame. Then, our extended R-D model

incorporating the intra refreshing rate for frame  $n$  can be formulated as follow:

$$D_s(n) = D(R, \beta) = \sigma_\beta^2 e^{-\gamma R}$$

The experiment results presented in section 3 demonstrate that our proposed R-D model can give a rather accurate estimation of the source coding distortion with intra refreshing.

For channel distortion, a frame-level channel distortion model with intra refreshing rate is proposed in [3]. According to this literature, if the intra refreshing rate in frame  $n$  is  $\beta$ , the overall channel distortion for frame  $n$  is then given by:

$$D_c(n) = \beta D_c^I(n) + (1 - \beta) D_c^P(n)$$

where  $D_c^I(n)$  and  $D_c^P(n)$  represent the expected channel distortion of an intra MB and an inter MB, respectively.  $D_c^I(n)$  and  $D_c^P(n)$  can be estimated as follow:

$$D_c^I(n) = ap \cdot MSE_{org}(n, n-1) + p \cdot D_c(n-1)$$

$D_c^P(n) = [(1-p)b + p] \cdot D_c(n-1) + pa \cdot MSE_{org}(n, n-1)$  where  $p$  is the packet loss rate,  $a$  is a constant describing the energy loss ratio for the video encoder,  $b$  is a constant describing the motion randomness of the video scene,  $MSE_{org}(n, n-1)$  represents the mean square error (MSE) between the original frame  $n$  and  $n-1$ .

At last, for a given  $R_s$ , the value of  $\beta$  which can minimizes the overall distortion  $D(n) = D_s(n) + D_c(n)$  should be chosen as the optimal intra refreshing rate  $\beta^*$ .

## 2.2. Proposed intra refreshment strategy

At the end of coding each P-frame, we adopt the following procedure to update the channel distortion estimation for each MB on a block basis (i.e., 4x4).

Let  $D(n, (x, y))$  denotes the channel distortion estimation for a block at the spatial location  $(x, y)$  in frame  $n$ . If this block belongs to an inter-coded MB, assuming its motion-compensated block is in frame  $n-k$  and the corresponding motion vector is  $(vx, vy)$ , then if it is correctly received, it will inherit the channel distortion propagated from its motion-compensated block at the spatial location  $(x + vx, y + vy)$  in frame  $n-k$ , which is denoted as  $D(n-k, (x + vx, y + vy))$ . Since the motion-compensated block may locate on a sub-pixel position, we will compute  $D(n-k, (x + vx, y + vy))$  by weighting the channel distortions of the surrounding blocks in the reference frame that overlap with the motion-compensated block.

$$D_c(n-k, (x+vx, y+vy)) = \sum_{i=1}^4 \omega_i D_c(n-k, (x, y)_i)$$

where  $(x, y)_i$  represents the block  $i$  overlapping the considered block at position  $(x, y)$ , the weight  $\omega_i$  is the proportion ratio of the overlap area.

If this block belongs to an intra-coded MB and is correctly received, there will be no propagation error inherited from the previous frames. But if this block is lost and concealed with the collocated block in the previous frame, regardless of the coding mode of the MB, the error-concealed distortion of this block is composed of two parts: the channel distortion of the error-concealed block, which is denoted as  $D(n-1, (x, y))$  and the newly introduced distortion between error-concealed block and the reconstructed block. The latter distortion can be calculated as the block MSE between the reconstructed frame  $n$  and frame  $n-1$ , which is denoted as  $MSE_{rec}(n, n-1, (x, y))$ .

The aforementioned two cases can be formulated as follow:

$$D_c(n, (x, y)) = \begin{cases} p \cdot D_c(n-1, (x, y)) + p \cdot MSE_{rec}(n, n-1, (x, y)) & \text{intra-coded} \\ (1-p) \cdot D_c(n-k, (x+vx, y+vy)) & \text{inter-coded} \\ + p \cdot D_c(n-1, (x, y)) + p \cdot MSE_{rec}(n, n-1, (x, y)) & \end{cases}$$

Thus, the channel distortion estimation of a MB can be acquired by summing up all its constituted blocks' channel distortion estimation.

Then, at the beginning of coding a P-frame, once the optimal intra refreshing rate  $\beta^*$  is determined according section 2.1, with the updated channel distortion estimation for each MB in the previous frame, we use the follow formula to estimate the channel distortion for each MB of current frame and forcedly intra-code  $\beta^* \cdot N$  MBs with higher channel distortion estimation ( $N$  is the total number of MBs in a frame):

$$D_c^{MB}(n) = [(1-\beta^*)(1-p) + p] \cdot D_c^{MB}(n-1) + p \cdot MSE_{org}^{MB}(n, n-1)$$

### 2.3. Unequal error protection with dynamic FMO

Since intra MBs can effectively stop error propagation, if the intra MBs in a frame can be better protected, we should expect that the effect of error propagation could be more quickly attenuated. In the previous video coding standards, the MBs in a frame can only be handled in a raster scan, which make it difficult to deliberately protect the intra MBs in a frame unless each MB is packetized in a packet. However, in H.264/AVC, by using FMO, the MBs in a frame can be assigned to slices in an order other than the normal raster scan order.

The fact that MBs can be arbitrarily grouped in slices offers new possibilities in unequally protecting different

portions of a video frame. In this work, after  $\beta^* \cdot N$  MBs with higher channel distortion estimation in a frame have been selected to be forcedly intra-coded, all the MBs in current frame will be mapped to two slice groups: slice group 0 and slice group 1. The intra MBs will be assigned into slice group 0 and the inter MBs into slice group 1. If there is few MBs chosen to be intra-coded in a frame, in order to reduce the overhead introduced by slice header, some inter MBs with higher channel distortion estimation can be assigned into slice group 0. In this situation, the number of inter MBs mapped into slice group 0 can be adjusted based on channel status. In this paper, under such circumstance, we simply prescribe that the size of slice group 0 should not exceed 1/3 of the total MBs of a frame. Since the mapping pattern is changing from frame to frame, we can use a new picture parameter set (PPS) to convey an individual macroblock allocation map (MBAmapping) for each frame [6]. Obviously, slice group 0 has a higher importance than slice group 1. In this paper, we use a stronger RS code to protect slice group 0.

### 3. SIMULATION RESULTS

A series of experiment have been carried out to verify the performance of our proposed integrated error resilience scheme. All simulations are based on the modified H.264/AVC reference software JM 8.6. The main coding parameters are as follows: the first frame of the sequence is intra-coded and the rest frames are all inter-coded; no B frame is used; two reference frames are used for motion-compensated prediction.

Firstly, we test the accuracy of the proposed R-D model for source coding with intra refreshing. Fig. 1 shows the PSNR match for QCIF Foreman coded at 64kbps with  $\beta=0.1$  and 0.5. The coded frame rate is 10 fps.

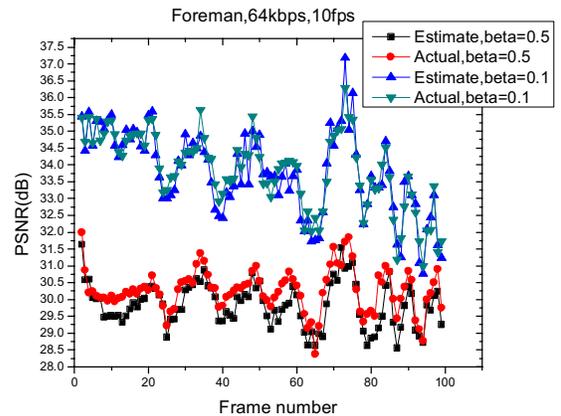


Fig 1. The comparison between actual and estimated source coding distortion

Secondly, we compare the proposed intra refreshing scheme with the Random Intra Update algorithm (RIU) [5]

and the Adaptive Intra Refreshing algorithm (AIR) [3]. From Fig. 2, it can be observed that both the AIR scheme and our proposed scheme acquire a better average PSNR than the RIU scheme. This is because these two schemes can adjust the intra refreshing rate according to the channel conditions and the characteristics of the input video data. Moreover, since our proposed scheme can appropriately choose the MBs with higher channel distortion estimation to be intra-coded, we can see that its performance is steadily better than that of AIR.

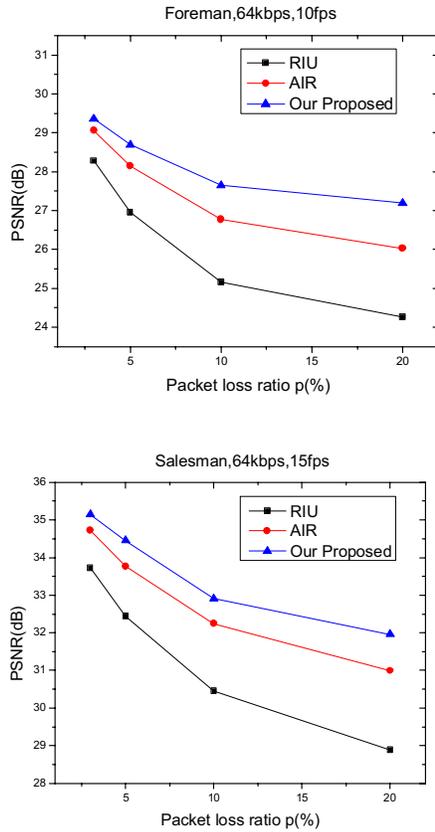


Fig. 2 Average PSNR performance comparison among different intra refreshing schemes

Thirdly, we evaluate the performance of our proposed scheme when the UEP technique is incorporated. For the slices in slice group 0, a stronger RS(6,3) code is used while for the slices in slice group 1, a weaker RS(4,3) code is used. For comparison, we also implement our proposed scheme with equal error protection (EEP), in which all MBs in a frame are assigned to one single slice group in normal raster scan order and all slices are protected with RS(5,3) code. From Fig. 3, we can see that the proposed UEP scheme steadily outperforms the EEP scheme, especially for high packet loss rate.

#### 4. CONCLUSIONS

In this paper, an efficient intra refreshment algorithm is proposed. Since our intra refreshment algorithm can achieve both global optimization and local accuracy in intra refreshing for a frame, its error-resilient performance is significantly improved. Furthermore, because our intra refreshment algorithm can determine the MBs needed to be intra-coded before the current frame is actually encoded, an effective UEP scheme based on a dynamic FMO mapping mode is also proposed to better protect the intra MBs in a frame. The experimental results show that the proposed UEP scheme can markedly improve the reconstructed video quality.

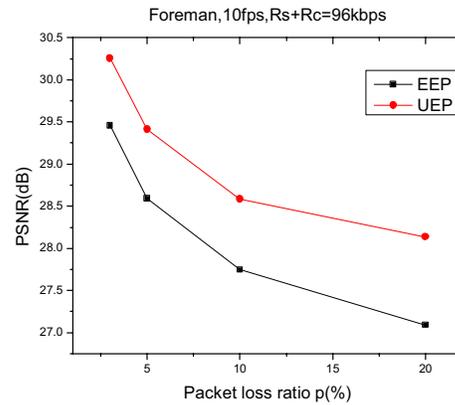


Fig. 3 The performance comparison between UEP and EEP

#### 5. REFERENCES

- [1] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE J. Select. Areas Commun.*, vol. 18, no. 6, pp. 966–976, Jun. 2000.
- [2] G. Cote, S. Shirani, and F. Kossentini, "Optimal mode selection and synchronization for robust video communications over error-prone networks," *IEEE J. Select. Areas Commun.*, vol. 18, no. 6, pp. 952–965, Jun. 2000.
- [3] Z. He, J. 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, no. 6, pp. 511–523, Jun. 2002.
- [4] Z. He, " $\rho$ -domain rate-distortion analysis and rate control for visual coding and communication," PhD Dissertation, University of California, Santa Barbara, 2001.
- [5] G. Cote and F. Kossentini, "Optimal intra coding of blocks for robust video communication over the Internet," in *Image Commun.*, Sept. 1999, pp. 25–34.
- [6] "Text of ISO/IEC FDIS 14496-10: Information Technology - Coding of audio-visual objects - Part 10: Advanced Video Coding", ISO/IEC JTC 1/SC 29/WG 11 N5555, Mar. 31, 2003.