Low Complexity Adaptive Intra-Refresh Rate for Real-Time Wireless Video Transmission

Angelo R. dela Cruz
Electronics Engineering Department
Faculty of Engineering
University of Santo Tomas, Manila, Philippines
ardelacruz@mnl.ust.edu.ph

Rhandley D. Cajote
Electrical and Electronics Engineering Institute
University of the Philippines, Diliman Campus
Quezon City, Philippines
rhandley@eee.upd.edu.ph

Abstract— **Real-time wireless video transmission systems must consider both error resiliency and low complexity. However, most error resilient features of recent video coding standards tend to increase computational complexity of the encoder. In this paper, we propose a low complexity error resilient joint source-channel adaptive intra-refresh rate scheme where the optimum number of intra-coded macroblocks is determined at frame level based on the minimum estimated end-to-end distortion. In this work, we propose source and transmission distortion models whose parameters are independent on sequence type which allows real-time video encoding. The source distortion model is based on residual information and quantization step using linear least square method. The residual information is estimated using the mean-absolute difference (MAD) prediction model based on the linear relationship between intra-refresh rate and MAD. The transmission distortion model is based on recursive model using reliable feedback channel. The proposed models are used to implement a joint source-channel video coding scheme using standard H.264/AVC encoder. Accurate estimate of the actual distortion at various refresh rates are achieved and able to estimate the distortion before encoding the frame. The proposed scheme is compared with random and periodic intra refresh schemes under wireless fading channel. Improvements in PSNR quality are measured which verifies the effectiveness of the proposed scheme especially in time varying channel conditions.**

Keywords—intra-refresh, source distortion, transmission distortion, error-control.

I. Introduction

Most of error resilient features of recent video coding standards tend to increase the computational complexity of video encoder [1-4]. However, the design of real-time wireless video transmission requires both error robustness and low complexity. Error control tools are only useful during unreliable channel condition and become unnecessary during good channel condition. Thus, in time varying channel, it is important to adjust the error control parameters based on the estimated channel condition. Among the various error control schemes of video encoder [5], intra-refresh (IR) schemes are still attractive because of its low complexity implementations and its efficiency in reducing transmission distortion. In this work, we propose a low complex adaptive intra-refresh rate scheme for wireless video transmission which optimizes the required number of intra-coded macroblocks (MBs) for a given frame based on the estimated channel condition. In line with this, we employ the source distortion model and intra-refresh rate prediction model in [6] and proposed a recursive transmission distortion model based on a reliable feedback channel. The estimated end-to-end distortion is used to select the best intra-refresh rate for a given frame.

In section II, the quantization distortion model based on residual information, quantization parameter and intra-refresh rate is presented and discussed. In section III, the proposed recursive transmission distortion model is presented. In section IV, the proposed framework and frame layer rate control will be discussed. Finally, section V discusses the experimental set-up and results of using the proposed framework under wireless fading channel. Objective quality using PSNR metric is measured and compared with intra-refresh schemes of JM16.0[7].

II. Source Distortion Analysis

Source distortion is the result of lossy compression caused by quantization which is mainly dependent on the available bitrate. Predicting source distortion allows the encoder to adjust the parameters that optimize either the quality and/or bit allocation for a given MB, slice or frame [8-10]. It is worth noticing that none of the existing source distortion models exploit the residual information to estimate quantization distortion. In H.264/AVC, the residual information measures the difference (in mean-absolute-difference MAD) between the original frame and motion-compensated predicted frame. It dictates the required quantization parameter (QP) for encoding the frame which greatly affects the source distortion. Note that higher MAD results to higher QP which results to an increase in source distortion. Thus, there exist a strong relationship among the source distortion, residual information in MAD and QP. Results in [6] show that a linear relationship between source distortion and the product of MAD and QP (MADxQP) exist as shown in Figure 1.

By using linear least square method, we can obtain a linear model that estimates source distortion as functions of MAD and QP as shown in (1). The model parameters α_1 and α_2 are determined and updated for a given window length, L using information of previously encoded frames as shown in (2) and (3). x represents previous $MADxQP$ values of window length

978-616-361-823-8

L while y represents actual source distortion of previously encoded frames of window length L.

$$D_Q(n) = \alpha_1 MAD_n QP_n + \alpha_2 \quad (1)$$

$$\alpha_1 = \frac{L\sum xy - \sum x \sum y}{L\sum x^2 - (\sum x)^2} \quad (2)$$

$$\alpha_2 = \frac{\sum y \sum x^2 - \sum x \sum xy}{L\sum x^2 - (\sum x)^2} \quad (3)$$

Note that QP and MAD are computed without encoding the frame which allows us to estimate the source distortion before encoding the current frame. Figure 2 shows a sample accuracy of the proposed model using "mother-daughter" sequence at 96kbps.

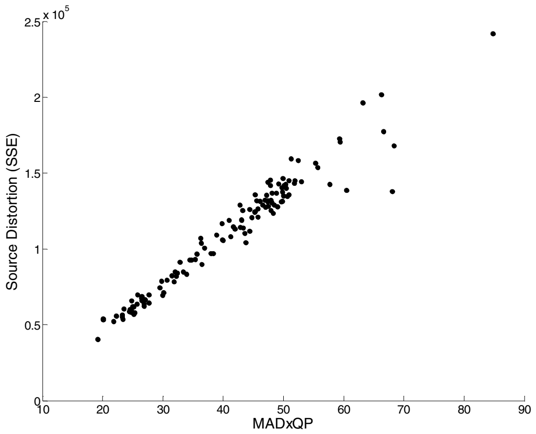

Figure.1. Relationship between Source Distortion and MADxQP for "mother-daughter" Sequence at 96kbps

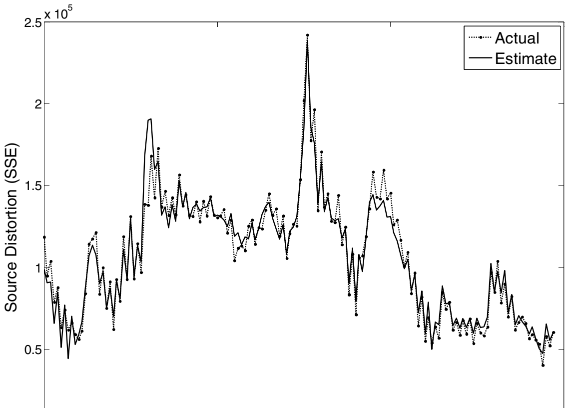

Figure.2. Estimated Quantization Distortion for "mother-daughter" Sequence at 96kbps

A. Intra-Refresh Rate

Given a constant bitrate, increasing the number of intra-coded MBs tends to increase source distortion because of higher QP required to encode the frame. Note that QP is estimated based on the residual information MAD and texture bits thus, it is intuitive that the actual MAD of a given frame should also increase. Results in [6] show that MAD is linearly increasing with intra-refresh rate as depicted in Figure 3. Thus, we can estimate MAD of current frame for any given refresh rate as shown in (4). $MAD_n(\Gamma_{n-1})$ is the predicted MAD from previous frame while $MAD_n(1)$ is the computed MAD with all MBs intra-coded. The proposed models in (1) and (4) are combined to estimate the source distortion as a function of intra-refresh rate.

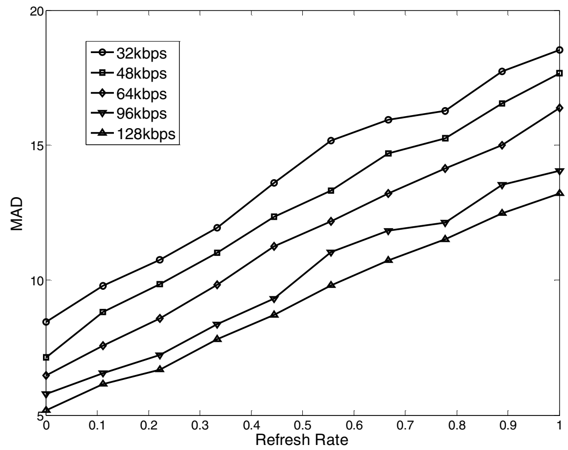

Figure 3. Relationship between MAD and Refresh Rate for "coastguard" Sequence at Frame 100

$$MAD_n(\Gamma) = \frac{MAD_n(1) - MAD_n(\Gamma_{n-1})}{1 - \Gamma_{n-1}}(\Gamma - \Gamma_{n-1}) + MAD_n(\Gamma_{n-1}) \quad (4)$$

III. TRANSMISSION DISTORTION ANALYSIS

To estimate the transmission distortion before encoding the frame, the model parameters must be independent on the sequence type. The transmission distortion model used in [8] as depicted in (6), assumes the availability of the model parameters α and β before encoding which makes the model sequence dependent. It is also assumed that the model parameters for each type of video sequence are constant. These limitations do not allow the transmitter to encode the sequence of real-time video information because it requires *a priori* information about the test sequence that will be encoded.

In this work, the transmission status of previously transmitted video packets is transmitted from the receiver side using a reliable feedback channel. Thus, the actual end-to-end distortion of previously encoded frames are easily computed at the encoder side. This previous information allows us to compute the model parameters recursively as shown in (8) where p is the packet loss probability and FD is difference of two consecutive frames (in SSE). In this work, we assume that the source distortion D_Q and transmission distortion D_C are uncorrelated as depicted in (5). D_Q is estimated using (1) while

D_C is estimated using (6). Note that the actual distortion of previously transmitted frames are available at the encoder side. Figure 4 shows the validity and accuracy of the transmission distortion model.

$$\widehat{D}_n = \widehat{D}_{Q_n} + \widehat{D}_{C_n} \quad (5)$$

$$\widehat{D}_{C_n} = D_{C_{n-1}}[p + (1-p)\beta(1-\Gamma)] + p\alpha FD_n \quad (6)$$

$$FD_n = E[(f_n - f_{n-1})^2], \quad (7)$$

$$\begin{bmatrix}\beta \\ \alpha\end{bmatrix} = \begin{bmatrix}(1-p)(1-\Gamma_{n-1})D_{C_{n-2}} & pFD_{n-1} \\ (1-p)(1-\Gamma_{n-2})D_{C_{n-3}} & pFD_{n-2}\end{bmatrix}^{-1} \begin{bmatrix}D_{C_{n-1}} - pD_{C_{n-2}} \\ D_{C_{n-2}} - pD_{C_{n-3}}\end{bmatrix} \quad (8)$$

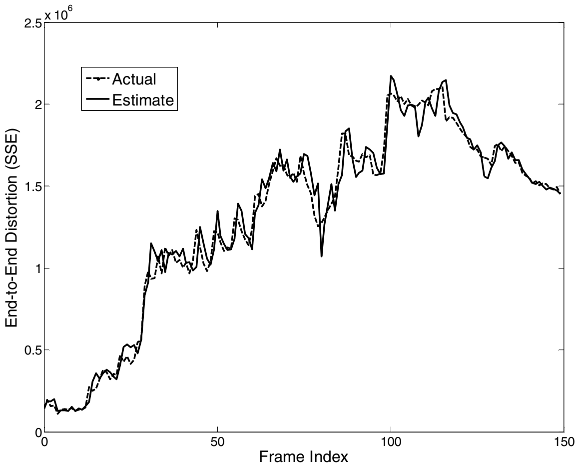

Fig.4. Estimated End-to-End Distortion for "mother-daughter" Sequence at SNR = 15dB

IV. Low Complexity Adaptive Intra-Refresh Rate Framework

Since both source distortion and transmission distortion are easily estimated at the encoder side before encoding the frame, we can select the optimum intra-refresh rate Γ, that will minimize the estimated end-to-end distortion for a given constant bitrate R_S as shown in (9).

$$\Gamma^* = \arg \min_{R_S \le R_T} \{D_Q(\Gamma) + D_C(\Gamma)\}, \quad (9)$$

After determining the optimum refresh rate Γ^*, the encoder selects random locations of MBs to be intra-coded. The proposed frame layer rate control is described as follows. Assume P-frame is to be encoded:

1. Compute target, header and texture bits.
2. Compute $MAD_n(\Gamma_{n-1})$ and $MAD_n(1)$ by encoding the entire frame as intra-mode.
3. Compute actual end-to-end distortion of previous frames.
4. Loop for all refresh rates Γ
   - 3.1. Predict $MAD_n(\Gamma)$ using linear model in (4) and its corresponding QP.
   - 3.2. Estimate source distortion, D_Q at frame level using (1).
   - 3.3. Estimate transmission distortion and end-to-end distortion, $D(n)$ at frame level using (5,6).
5. Choose the optimum refresh rate, Γ^* and its corresponding QP that minimize $D(n)$.
6. Intra-code Γ^* random locations of MBs.
7. Compute actual source distortion after encoding entire frame n.
8. Adjust model parameters $\{\alpha_1, \alpha_2, \alpha, \beta\}$ for next frame encoding.

V. Results and Analysis

The proposed scheme is implemented using baseline profile of JM16.0 standard reference encoder for H.264/AVC. Transmission of video packets are simulated under wireless fading channel generating error patterns that mimics the behavior of typical wireless transmission. We consider bitrates from 32kbps to 128kbps to represent wide range of low bitrate channel. To combat the effects of random packet errors, we simulate 30 trials for each bitrate and SNR condition. Lost packets are concealed using non-motion compensated error concealment method. Various QCIF test video sequences are encoded to represent different levels of motion intensities. The performance of the proposed scheme is measured using PSNR metric and compared with two intra-refresh schemes namely; (a) random intra-refresh (RIR) with 10%, 25%, 50% and 75% refresh rates and (b) periodic intra-refresh (PIR) with fixed refresh rate of 10%.

A. PSNR Quality

In Figure 5, RIR methods with higher refresh rates outperform PIR scheme during low SNR state because higher refresh rates reduces transmission distortion. But during high SNR condition, PIR scheme provides better PSNR quality than RIR scheme with higher refresh rates because source distortion is more dominant. It is observed that for all SNR conditions, the proposed scheme outperforms the abovementioned schemes due to its ability to adjust its refresh rate based on the condition of the channel. Table 1 summarizes the average PSNR gain for all SNR conditions at various bitrates.

Table 1. Average PSNR Gain of the Proposed Scheme for "mother-daughter" Sequence

Error Control Method	**Bandwidth (kbps)**				
	32	**48**	**64**	**96**	**128**
RIR 10%	0.8985	1.5131	0.7002	1.9216	2.4556
RIR 25%	2.1350	2.1355	1.2306	2.0866	1.6724
RIR 50%	4.5876	3.9648	2.8935	2.8041	2.4399
RIR 75%	5.4485	6.3549	4.6915	4.3361	3.8458
PIR 10%	0.7104	0.7706	0.2510	1.2240	1.2512

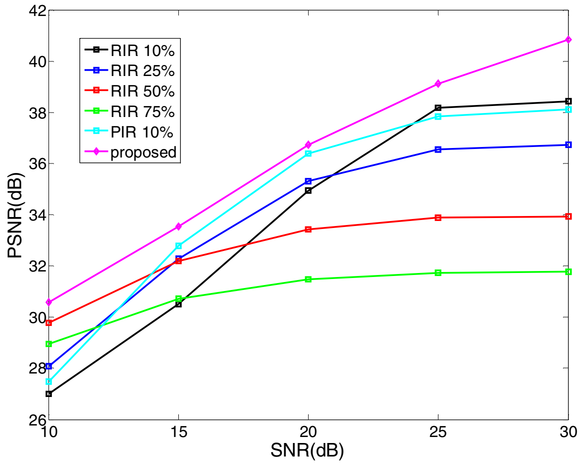

Figure 5. PSNR vs. SNR for "mother-daughter" at 96kbps

Various test sequences with different levels of motion intensity and spatial texture are used to validate the effectiveness of the proposed scheme. Table 2 shows the summary of PSNR improvements as compared with the PIR and RIR schemes. The results in Table 2 verifies the effectiveness of the proposed joint source-channel coding scheme as compared with the inherent intra-refresh tool of recent video coding standard.

Table 2. Summary of Average PSNR Gain of the Proposed Scheme as Compared with PIR and RIR Methods

Test Sequence	**Bandwidth (kbps)**									
	32		**48**		**64**		**96**		**128**	
	RIR	PIR	RIR	PIR	RIR	PIR	RIR	PIR	RIR	PIR
akiyo	2.64	2.90	1.95	1.86	1.53	1.44	2.00	1.80	2.46	2.42
carphone	1.56	0.84	0.74	0.22	1.37	0.75	0.79	-0.3	1.47	0.52
coastguard	2.15	1.04	2.10	0.72	1.68	0.38	3.42	1.43	2.71	0.87
foreman	1.17	0.68	1.19	0.67	0.90	0.40	1.73	0.13	0.72	-0.5

B. *Subjective Quality*

We also compare the subjective quality performance of decoded frames using RIR, PIR and the proposed scheme. Note in Figure 6 that during low SNR condition, the transmission distortion is more severe in PIR scheme than RIR method due to higher refresh rate used by RIR scheme. But it can also be observed that the RIR method heavily quantized the encoded frame due to large number of intra-coded MBs. In Figure 7, during high SNR state, source distortion is more dominant than transmission distortion especially at low bitrate condition. Note that PIR and RIR schemes have the same refresh rate which heavily quantized the encoded frame resulting to a higher source distortion compared with the proposed scheme. Note that the proposed scheme provides better decoded quality for both SNR conditions by selecting the appropriate refresh rate based on the estimated end-to-end distortion.


(a) RIR 75% (b) PIR (c) Proposed

Figure 6. Comparison of Subjective Quality for "mother-daughter" Sequence at Frame 50 at 128kbps, SNR = 10dB


(a) RIR 10% (b) PIR (c) Proposed

Figure 7. Comparison of Subjective Quality for "mother-daughter" Sequence Frame 50 at 32kbps, SNR = 30dB

C. *Complexity*

The average complexity is computed by obtaining the average encoding time for the entire test sequences. Test sequences are encoded individually using a laptop computer with Intel Core i7-3630QM 2.40 processor with 24GB DDR3 RAM running in Windows 8. The average encoding time is computed by determining the encoding time (in msec) for each frame for the entire 150 frames/sequence. It is intuitive that the original JM16.0 encoder setting is the most complex among the test methods considered in this work because it uses the least number of intra-coded MBs. It is important to note that intra-coding provides less computation because of the absence of temporal predictions. Thus, it is intuitive that RIR with 75% refresh rate has the least complex method as depicted in Figure 8 and Table 3. However, even during good channel condition where intra-coded MBs are not needed, RIR 75% scheme still provides fixed number of intra-coded MBs. Note that our proposed scheme tends to decrease its complexity during low SNR conditions because it intelligently increases the required number of intra-coded MBs. During high SNR condition however, the proposed scheme increase its complexity by adding inter-coded MBs to minimize quantization distortion. Table 3 summarizes the reduction of complexity with respect to original JM16.0 settings. Note that the time consumed by distortion estimation unit is negligible because the estimation is done before encoding process. This can be verified in Table 3 at 30dB SNR where the distortion estimation unit consumes extra 0.37% of JM encoding time which is almost negligible. This result verifies that the proposed scheme can be adopted to low delay wireless video transmission applications.

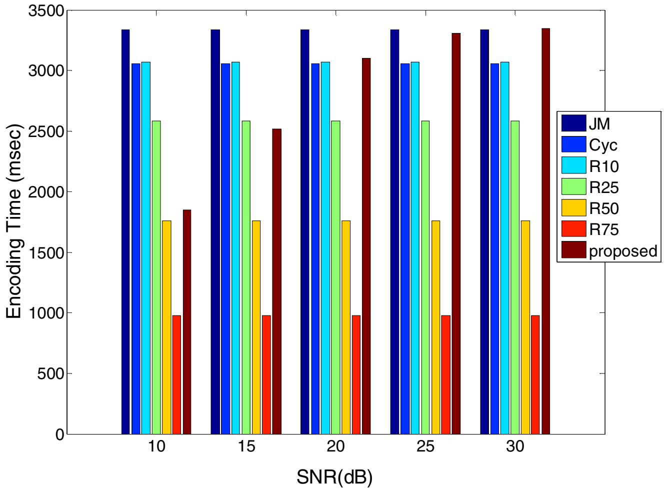

Figure 8. Average Complexity of Various Error Control Methods

Table 3. Average Complexity Reduction with respect to JM16.0

Error Control Method	**SNR**				
	10dB	**15dB**	**20dB**	**25dB**	**30dB**
Cyclic or PIR	8.37%	8.37%	8.37%	8.37%	8.37%
Random 10%	7.94%	7.94%	7.94%	7.94%	7.94%
Random 25%	22.57%	22.57%	22.57%	22.57%	22.57%
Random 50%	47.17%	47.17%	47.17%	47.17%	47.17%
Random 75%	70.74%	70.74%	70.74%	70.74%	70.74%
proposed	44.50%	24.46%	6.97%	0.80%	-0.37%

VI. CONCLUSION

This paper has presented a low complexity error resilient video transmission under wireless fading channel by selecting the best intra-refresh rate at any given frame. Source distortion model and transmission distortion model which allows the encoder to estimate the end-to-end distortion before encoding are presented. A novel linear model that estimates the residual information, MAD as a function of intra-refresh rate is proposed. Accurate estimate of MAD and source distortion at various intra-refresh rates, bitrates and test sequences are achieved. The proposed models are employed and implemented in a standard H.264/AVC encoder frame layer rate control. Experimental results have shown significant improvements for both objective (PSNR) and subjective quality. The proposed scheme provides low complexity compared to original baseline profile encoder settings especially at severe channel conditions.

REFERENCES

[1] W.-J. Tsai, Y.-C. Sun, "Error-Resilient Video Coding using Multiple Reference Frames", *IEEE International Conference on Image Processing 2013*, pp. 1875-1879, Melbourne, Sept. 2013.

[2] J. Zheng, L.P. Chau, "Multiple Description Coding using Multiple Reference Frame for Robust Video Transmission", *IEEE International Symposium on Circuits and Systems 2005*, Vol. 4, pp. 4006-4009, May 2005.

[3] B. Katz, S. Greenberg, N. Yarkoni, N. Blaunstien, R. Giladi, "New Error_Resilient Scheme Based on FMO and Dynamic Redundant Slices Allocation for Wireless Video Transmission", *IEEE Trans. on Broadcasting*, Vol. 53, No.1, pp. 308-319, March 2007.

[4] J.C. Schmidt, K. Rose, "Jointly Optimized Mode Decisions in Redundant Video Streaming", *IEEE Trans. on Circuits and Systems for Video Technology*, Vol.21, No.4, pp. 513-518, April 2011.

[5] T. Stockhammer, M.M. Hannuksela, T. Wiegand, "H.264/AVC in Wireless Environments", *IEEE Trans. on Circuit and Systems for Video Technology*, Vol. 13, No. 7, pp. 657-673, 2003.

[6] A.R. dela Cruz, R.D. Cajote, "Joint Source-Channel Rate Allocation for Wireless Video Transmission Based on Distortion Estimation", *International Conference on Information Science and Applications (ICISA)*, Seoul South Korea, pp. 164-167, May 2014.

[7] Joint Video Team, JVT Reference Software Encoder, version 16.0., http://www.hhi.de/seuhring/tml/

[8] Z. He, J. Cai, C.W. Chen, "Joint Source Channel Rate-Distortion Analysis for Adaptive Mode Selection and Rate Control in Wireless Video Coding", *IEEE Trans. on Circuits and Sytems for Video Technology*, vol. 12, No.6 June 2002.

[9] Z. Chen, D. Wu, "Rate Distortion Optimized Cross-Layer Rate Control in Wireless Video Communication", *IEEE Trans. on Circuit and Systems for Video Technology* , vol. 22, No. 3, March 2012.

[10] W.N. Lie, Z.W. Gao, T.L. Liu, P.C. Jui, "Joint Source-Channel Video Coding Based on the Optimization of End-to-End Distortions", *Springer-Verlag Berlin Heidelberg*, PSIVT 2006, LNCS 4319, pp. 842-851, 2006.