MOTION BASED ERROR RESILIENCE TECHNIQUE FOR H.264 CODED, MEDIUM AND SLOW MOTION VIDEO SEQUENCES

Maheshi B. Dissanayake[1*]

[1*] Corresponding Author, Dept. of Electrical and Electronic Engineering, Faculty of Engineering, University of Peradeniya, Peradeniya, Sri Lanka, Email: maheshid@ee.pdn.ac.lk

ABSTRACT

The proposed error resilience algorithm incorporates motion information of predictive based video coding to improve the robustness of the transmitted data. The motion information of the video data plays a vital role in frame reconstruction. Therefore, the technique proposed in this research highlights the effect of providing extra protection for the motion information. The proposed system is implemented based on the existing redundant coding algorithm of the scalable extension of H.264/AVC and it encodes redundant data only for the motion information. The paper analyzes the effect of proposed error resilience technique, redundant motion information coding, on both base layer and enhancement layer picture quality. The performance of the algorithm is evaluated in both error free and error prone Internet Protocol(IP) packet network environments. The proposed algorithm increases the bandwidth utilization with slight degradation in the primary picture quality for error free conditions, compared to the existing redundant coding method of H.264/AVC standard. Furthermore, the simulation results under packet loss environments show that the proposed algorithm outperforms the existing redundant picture coding technique of H.264/AVC standard.

***Key words*:** H.264, scalable video coding, error resilience, motion vectors

1. INTRODUCTION

In video transmission environments, due to dynamic nature of the transmission channels errors such as packet losses, bit errors, burst errors are almost inevitable. These errors degrade the quality of the reconstructed video frames and results in a mismatch between the transmitted and reconstructed frames which may propagate spatially as well as temporally[1]. Hence, error resilience mechanisms are employed in video communication system to mitigate these adverse effects caused by packet losses.

In boarder sense, error resilience tools are of four types: Localizations techniques, Data partitioning techniques, Redundant coding techniques and Concealment driven coding techniques[1]. However, the introduction of error resilient tools to the encoded stream increases the amount of redundancy present and decreases the compression efficiency of the encoded data[2]. Therefore, it is essential to maintain the right balance between the error resilience information and primary video quality. The research presented in this paper, investigate the error resilience provided by means of redundant coding technique.

The key error resilient technique supported in H.264/AVC standard is redundant picture coding[2]. However, this scheme is a potentially inefficient error resilient technique, specifically in low bandwidth and low packet loss rate environments due to heavy redundancy associated. Since the transmission of the exact duplicates of the primary requires the same amount of bandwidth as the primary stream, the total bandwidth available for the transmission will be shared equally by the two streams. Yet, in error free environments the duplicates of the primary picture, which is transmitted for the sole purpose of error resilience will be discarded completely, resulting in waste of bandwidth.

For this reason, the main focus of current error resilience research is on reducing the amount of redundant information transmitted while achieving better error resilience. For instance, references [3-6] present some redundant coding based proposals, which concentrate on achieving high error recovery with minimum redundancy. However, some of these techniques introduce higher complexity to the encoder, while others introduce noticeable delay to the system. The new redundant coding technique presented in this paper achieves higher output video quality while maintaining the redundancy at an acceptable level. Also, the proposed technique is a simple system without an additional delay.

The rest of the paper is organized as follows. Section 2 present the background theory related to the research presented. Section 3 elaborates on the proposed system architecture and the system design stages, while Section 4 describes the test conditions. The simulation results are presented in Section 5. Finally the Section 6 concludes the paper.

2. BACKGROUND THEORY

Video coding uses motion compensated prediction (MCP) as a spatial and temporal redundancy removal technique. It exploits the correlation between the successive frames and takes the necessary measures to remove these redundancies. The MCP consists of two components called Motion Estimation (ME) and Motion Compensation (MC).

Fig. 1 exhibits the basic operation of MCP. First of all, the current frame to be encoded is sub partitioned into blocks of NxM. Then, Motion estimation and motion compensation is applied on to each and every NxM macroblocks (MB). Next, a closely matching reference MB to the current MB is searched in previously encoded frames in reference list 0 and 1. The process of finding the best match MB is called Motion Estimation. The distance between the reference MB position and the current MB position is given by the motion vector. The selected predictor MB is subtracted from the current MB to get the residual. In other word, the motion vectors (MV) indicates the direction of the movement of the objects in the frame, and residual represents the difference between the motion compensated reference picture and the current picture. This complete process is known as the Motion compensation.

Next, the coded residual and the motion vector of the current MB are transmitted to the decoder to facilitate efficient decoding. The decoder uses the MV to recreate the predictor MB and after decoding the residual data, it adds the two components together to obtain the decoded MB. Although, motion compensation is the most computationally intensive operation in video coding, it improves the overall coding efficiency of the system.

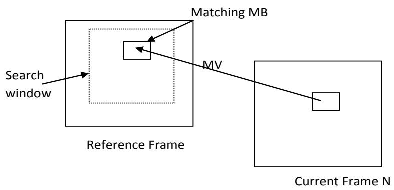

Figure 1: Motion Compensated Prediction

As a result of the MCP based coding approach adopted in H.264 coding standard, the network abstraction layer unit (NALU) of the encoded frames consists of coded residual data and motion information. The data which aids the decoder to recreate the MC frame at the decoding end is called Motion information (MI). It plays a more important role compared to the residual information, when it comes to displaying the image properly [7]. Therefore, the technique proposed in this paper highlights the effect of providing extra protection for the motion information.

3. PROPOSED ARCHITECTURE

Fig. 2 illustrates the basic block diagrams of H.264/AVC codec, Joint Scalable Video Model (JSVM). It consists of the motion compenstaion prediction process and residual data coding process. In JSVM, the exisiting redundant pictures coding process retrace the whole primary picture coding process. In the proposed method, only the motion information is transmitted as redundant data. Neither motion compensation nor residual coding is performed at redundnat data genreation cycle. Hence, redundant picture Network Abstraction Layer (NAL) units contain the MVs of the primary picture. Accordingly the proposed method reduces the amount of redundant information transmitted and in turn saves the bandwidth for high quality primary data transmission.

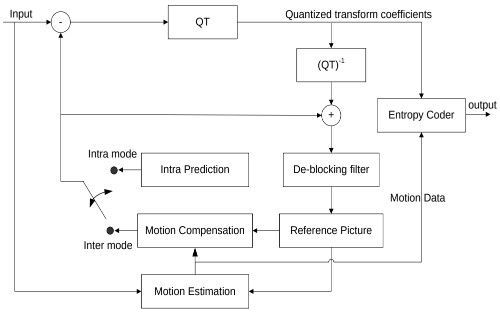

Figure 2: Structure of the JSVM encoder

During the JSVM decoding process, if the primary data is lost then redundant data is used to recover the error. In the proposed method if the NAL unit of the primary picture is lost then the motion information is recovered using the uncorrupted redundant NAL unit. Further, if both primary and redundnat NALUs are lost, then the frame copy error concelment is used to recover from the error.

4. TEST CONDITIONS

The JSVM 8.9 codec [8] is used to generate the H.264/AVC data stream. The applicability of the proposed system for heterogeneous environment was tested with the higher layers of the scalable extension of H.264/AVC standard. For this purpose the experiments were carried out both at QCIF base layer level and CIF enhancement layer level. The encoded bit-streams with and without the existing redundant coding method of JSVM 8.9 are used as the benchmark for the analysis of the results.

The simulations are carried out using three sequences, for 4000 frames at QCIF/CIF resolutions. The sequences were selected depending on the inheriting motion of it. For example, “News” features slow motion, and “Forman” contains medium motion, while “Paris” contains considerable amount of motion in a static background. Also, the channel bandwidth and frame rates are varied according to the specifications given in[9]. An IP packet loss simulator is used to simulate packet loss scenario[10], where packet loss rates (PLRs) of 0%, 3%, 5%, 10 % and 20% were considered. Further, the performance of the proposed and the two reference schemes is compared in terms of Peak Signal to Noise Ratio (PSNR).

5. RESULTS

According to Fig 3 to 6, the proposed system exhibits a significant improvement, compared to the standard scheme with redundant data, for all of the packet loss rates (PLR) tested. Also, at higher loss rates the proposed performs better than the scheme without any redundant information's, except on Fig. 4, which is of slow motion sequence, "News". The improvement in Fig.3 and 6 are due to the fact that frame copy error concealment fails to provide sufficient error recovery at high packet loss environments for medium motion sequences. Further, the residual of the medium motion sequences requires considerable amount of bandwidth, since the correlation between adjacent frames is low. Hence, a considerable amount of bandwidth gain can be achieved in the primary data, which in turn improves the reconstructed video quality, when the residual is dropped from the redundant data stream.

In slow motion sequence "News" at low bandwidth and low resolution, frame copying provides good reconstructed picture quality due to its low scene motion. Although, scheme without any redundant coding methods shows a sharp degradation in quality with the increase in packet loss rates, the PSNR improvement achieved at 0% PLR is 3 dB ahead of the proposed method according to Fig. 4. This helps to provide an output quality of 36.5 dB at 20% PLR which is equal to the PSNR of the proposed scheme at that PLR. Moreover, because of the high correlation present between adjacent frames in slow moving sequences, it is impractical to get any gain in PSNR by means of transmitting the whole redundant pictures. However, by employing the proposed method, the output quality fluctuation between frames in a variable PLR environment can be kept at a minimum.

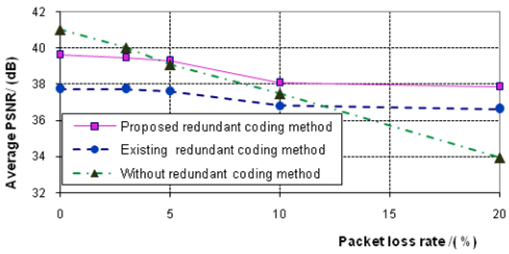

Figure 3: Foreman with QCIF @ 64 kbps & 7.5fps

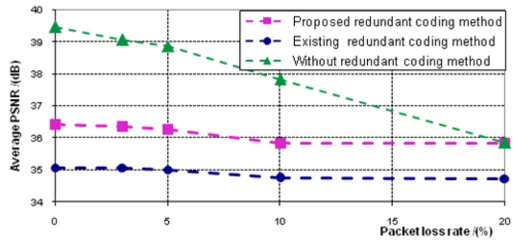

Figure 4: News with QCIF @ 48 kbps & 10 fps

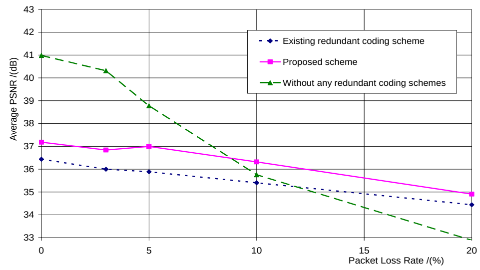

Figure 5: News with CIF @ 144kbps & 10fps

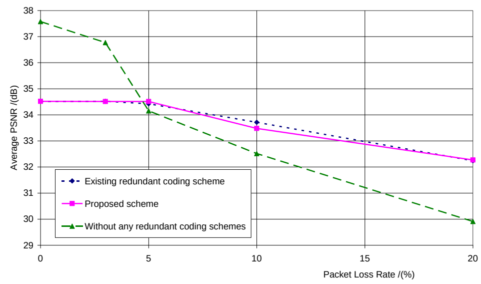

Figure 6: Paris with CIF @ 384kbps &

6. CONCLUSION

An error resilient scheme which improves the bandwidth utilization with limited primary picture quality degradation is proposed for H264/AVC. This proposal uses motion information as redundant data to achieve the above objectives. The experimental results show that the proposed technique outperforms the existing redundant picture coding scheme of JSVM for all the tested PLRs.

According to the results, in a fluctuating transmission network scenario, an adaptive error resilience scheme which has the capacity to change the redundant data allocation ratio dynamically will perform better than having a fixed amount of redundant data embedded into the system. Moreover, in the future this technique can be extended and tested in wireless environments such as WiMAX to establish the findings for scalable video in heterogeneous wireless transmission systems.

7. REFERENCES

[1]. A. Vetro, J. Xin, and H. F. Sun, "Error resilience video transcoding for wireless communications," *IEEE Wireless Communications*, vol. 12, no. 4, pp. 14-21, Aug.2005.

[2]. T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 657-673, July2003.

[3]. P. Baccichet and A. Chimienti, "Error resilience by means of coarsely quantized redundant descriptions," JVT Document, JVT-S046, Geneva, Switzerland, Apr.2006.

[4]. B. S. Choi and T. Sung, "Cross XORed Redundant picture," JVT document, JVT-X045, Geneva, 2009.

[5]. Hendry and M. Kim, "SVC/SVC Sliding window XOR-based redundant picture for error resilience," JVT document, JVT-X056, Geneva, June2007.

[6]. D. Y. Suh and G. H. Park, "Proposal for extension of redundant picture (XRP)against loss of key pictures," JVT Document, JVT-S015, Geneva, Apr.2006.

[7]. M.Ghanbari, *Standard Codecs: Image Compression to Advanced Video Coding* Institution of Electrical Engineers, London, 2003, pp. 1-407.

[8]. "JSVM 8.9 software," downloaded from CVS server, garcon.ient.rwth-aachen.de/cvs/jvt.: 2009.

[9]. Wien M., H. Schwarz, and T. Oelbaum, "Performance analysis of SVC," JVT Document, JVT-U141, Hangzhou, Oct.2006.

[10]. Guo Y., Wang Y, and Li H., "SVC/AVC Loss simulator donation," JVT document, JVT-Q069, Nice, France, JVT-Q069, Oct.2005.