ADAPTIVE LAYERED MULTIPLE DESCRIPTION CODING FOR WIRELESS VIDEO WITH EXPANDING WINDOW RANDOM LINEAR CODES

Sajid Nazir, Vladimir Stanković

Department of Electronic and Electrical Engineering,
University of Strathclyde, Glasgow, UK
Email: {sajid.nazir, vladimir.stankovic}@eee.strath.ac.uk

Dejan Vukobratović

Department of Power, Electronics and Communication Engineering, University of Novi Sad, Serbia
Email: dejanv@uns.ac.rs

ABSTRACT

The error free communication of video data over multi-hop wireless networks is a challenging research problem. Multiple description coding has been proposed as a possible solution to leverage path diversity for error robustness. Forward error correction is an additional protection that can be provided to each description. Random linear codes have had renewed interest fostered by the multi-hop and multi-interface radio receivers. In this study, the descriptions are created using the encoding features of slicing and data partitioning for H.264/AVC video. The unequally protected video is protected with Expanding window-Random linear codes against channel errors. Fading channel error model is used to simulate real-world wireless channels. We also propose an adaptive scheme for video transmission over multiple paths. Such scheme may adapt to the varying channel conditions as is frequently the case in wireless transmission. The results show that the proposed scheme can be used for emerging wireless standards.

***Index Terms*—** Multiple Description Coding, Random Linear Codes, H.264/AVC, Data Partitioning, Multi-hop wireless networks

1. INTRODUCTION

H.264/AVC [1] is the latest video coding standard achieving significant compression efficiency and gaining widespread use in the emerging standards and applications. Video transmission over wireless networks poses many challenges. The problems of packet losses and bandwidth limitations are to be offset on a time-varying channel. The quality of video reconstruction can be improved by error correction and error resilience mechanisms. Another method is multipath video with Multiple Description Coding (MDC) [2] which can exploit the path diversity to improve reconstruction quality.

The source data is encoded in multiple descriptions, which are sent to the receiver over independent paths. The receiver can reconstruct the encoded data, at an acceptable quality from any subset of the descriptions received. The quality of reconstruction increases with more descriptions received. The obvious advantage is of having video rendering at the receiver despite a complete path failure.

A class of rateless codes which has become popular recently are Random Linear Codes (RLC) [3]. RLC applied over a source message produces encoded symbols as random linear combinations of source symbols with coefficients randomly selected from a given finite field. As a packet level Application-Layer Forward Error Correction (AL-FEC) solution, RLC is simple to implement and perform as near-optimal erasure codes. For short lengths of the source messages, the decoding complexity of Gaussian Elimination is acceptable (see [4] and references therein).

The multiple descriptions are normally created with duplication of the important data over both descriptions, thereby affording inherent unequal error protection (UEP) ensuring a better reconstruction quality. In [5], redundant slices are used for creating multiple descriptions. In [6], the base layer and the enhancement layer packets are sent over different paths and Selective ARQ is used to notify base layer losses and re-transmissions. It is difficult to cope with long error bursts for any scheme which is non-adaptive to the channel.

In the proposed scheme we create independent descriptions with least duplication and additionally propose expanding window – random linear codes (EW-RLC) [7] for FEC. The advantage being that the degree of protection can be adapted to suit a particular wireless channel.

Given a source and destination connected via two paths, the focus of this study is to find the best rate distribution for the descriptions that maximize Peak signal to noise ratio PSNR. The contributions of this work are (1) an adaptive scheme based on slicing and data partitioning for multiple- path layered video communication. (2) multi-path expanding window – RLC as a robust solution to unequally protect the data of each description.

The proposed scheme is fully compatible with the H.264/AVC standard. The possible applications are in multi-interface networks and Video on Demand (VOD) applications.

The rest of the paper is structured as follows. Section 2 covers the necessary background. The proposed system is described in Section 3. Section 4 and 5 cover the simulation results for uniform and burst loss channel models respectively. The conclusion and future research directions are highlighted in Section 6.

2. BACKGROUND

2.1. H.264/AVC Slicing and DP

H.264/AVC provides many error-resilience features to mitigate the effect of lost packets during transmission. The feature of slicing effectively creates many resynchronization points. The partitioning of a frame into slices can be used to create multiple descriptions

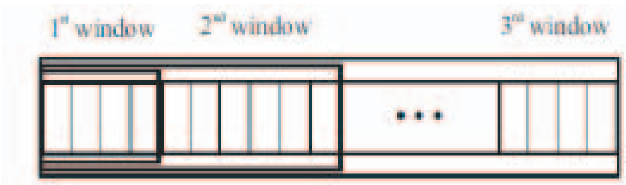

Fig. 1. Expanding window structure.

with fine granularity. Another scheme available in the extended profile is data partitioning (DP) [8] which supports the partitioning of a frame/slice in up to three partitions. Partition A contains the most important data comprising slice header, quantization parameters, and motion vectors. Partition B contains the intra-coded macroblocks (MB) residual data, and partition C contains inter-coded MB residual data.

It is thereby possible to assign different protection levels to different partitions based on importance. The decoding of DP A is always independent of DP B and C. However, if DP A is lost the remaining partitions cannot be utilized. The decoding of DP B can be made independent of DP C, using Constrained Intra Prediction (CIP) parameter in H.264/AVC encoder. The effect of error propagation can be limited by inserting periodic macroblock intra updates (MBIU).

2.2 Expanding Window-Random Linear Codes

Expanding window fountain (EWF) codes [7] are a class of UEP fountain codes based on the idea of creating a set of "nested windows" over the source block. The rateless encoding process is then adapted to use this windowing information while producing encoded packets. We use the EWF concept to create EW RLC over two windows (EW2) from consecutive source blocks containing k symbols (data packets).

To obtain source blocks amenable to UEP, we define the set of windows over the groups of source symbols of unequal importance, which are generated as a result of DPH.264/AVC video stream. Note that in order to define a window over a subset of source symbols, a particular priority DP has to be aggregated together. The general layout of a window structure with three importance layers is shown data is the first window (W1) and the importance of data additionally included in windows progressively decreases as we proceed to the third window (W3). The subset data of W1 is contained in all the subsequent windows and is hence the best protected. Apart from W1, each window in addition to some of its own data also encloses the data of the higher importance windows. The size and structure of a window depends upon the elements meeting particular set criteria from a specific subset window. The number of windows is governed by the aggregation scheme employed to group encoded elements.

The decoding of a window is same as RLC decoding, in that, a window is recoverable if the receiver collects at least the same amount of linearly independent encoded symbols obtained from the window (or the windows contained in it) as there were in the window [9].

The encoding process for EW RLC has one important initial step that is to first select a window from which the RLC encoded symbol is to be generated. We term probability of selection of W1 as PS_1 and that of W2 as PS_2. After a window is selected, the encoding is the standard RLC encoding performed over the source packets contained in that particular window only [9].

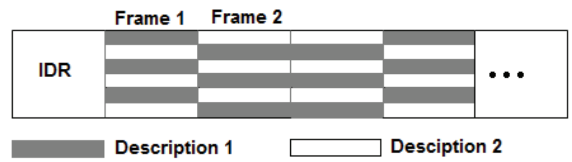

Fig. 2. Slice layout for creating descriptions. Each rectangle represents a slice.

Table I. System configuration.

Cat.	GOP		MDC1		MDC2	
	Size (bytes)	Cum. PSNR	Size (bytes)	Cum. PSNR	Size (bytes)	Cum. PSNR
IDR	11515	-	11515	-	11515	-
DP A	9922	32.45	4892	28.37	5030	27.92
DP B	4352	33.91	1744	28.76	2608	28.49
DP C	14828	40.1	7538	29.55	7290	28.92
Total	40617	40.1	25689	29.55	26443	28.92

3. PROPOSED SYSTEM

In this study we assume two independent wireless channels connecting a source S to the destination D, which could exist in a multi-interface receiver in a multi-hop wireless network. The system parameters relating to the channels are the data rate and packet erasure rates.

3.1. MDC Configurations

Two independent descriptions with equal size and importance can be created by dividing the slices of a frame into two sets. The decoder is thus able to create the missing information in case of losing any description.

In order to make use of EW-RLC codes over each description, the DP within each slice can be assigned to different layers for unequal protection.

3.2. System Parameters

The simulations have been performed using Foreman sequence in CIF format. The CIP flag has been set and MBIU are used to contain the error propagation. The GOP size is 8, and we use 6 slices per frame. The video data rate is 1000 kbits/sec for 25 frames per sec. Each simulation is run 1000 times for the results to converge. Two descriptions using slice distribution are created for use in all the simulations. Each of the description contains 3 slices from each frame as shown in Figure 2. Note that each of the descriptions has the Instantaneous Decoder Refresh (IDR) frame to make possible the decoding of each description.

In order to enable unequal protection to each of the description's data, we identify two importance layers. The data within each of the slice is already partitioned into DP A, B and C. The most important data for video re-construction, consisting of IDR, DP A and DP B is termed as high- priority layer (HPL). The remaining data i.e. DP C constitutes low-priority layer (LPL).

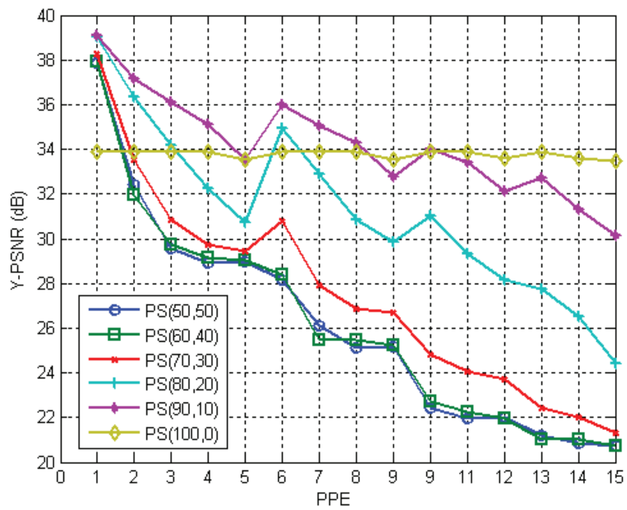

Fig. 3. PSNR at different PPE values of Section 4.1.

Table II. PSNR Contribution.

Categories Received		PSNR (dB)
MDC1	MDC2	40.1
MDC1	HPL2	36.27
MDC1	Fail2	29.55
HPL1	Fail2	28.76
HPL1	MDC2	36.6
Fail1	MDC2	28.92
Fail1	HPL2	28.49
HPL1	HPL2	33.91
Fail1	Fail2	20.71

The sizes of the different partitions for the first GOP of the Foreman sequence together with the assignment of the partitions to MDC1 and MDC2 are shown in Table I.

The PSNR contribution for each of the possible combinations has been calculated and is listed in Table II. As can be seen from the Table, the best video quality is achievable when both the descriptions are received correctly. However, the novel way that the descriptions have been designed, the video can be decoded with just the HPL of any one description only. This layering of video data within a description makes this scheme adaptive and practical for varying channel conditions. In case where both descriptions are lost entirely and nothing is decodable for the GOP then the decoder applies simple error concealment technique by replacing the lost GOP by the last frame of the previously successfully decoded GOP, which comes to 20.71 dB.

The video data belonging to each description is divided into packets of 512 bytes. The details of the packetization are as shown

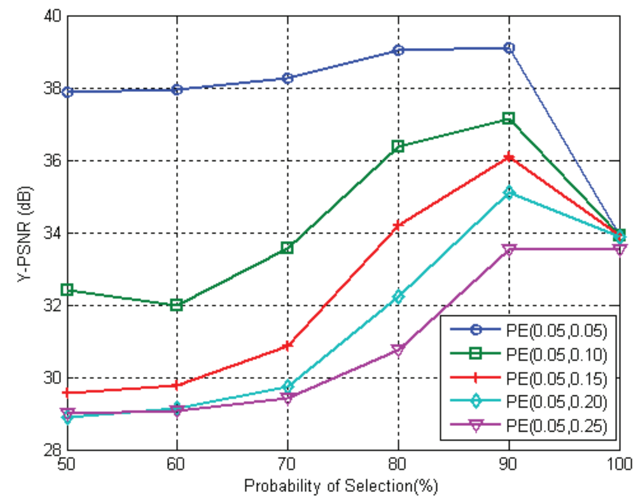

Fig. 4. PSNR at different PE.

Table III. Packet allocation to different layers.

Layer	MDC1			MDC2		
	Size (bytes)	PSNR	Pkts	Size (bytes)	PSNR	Pkts
HPL	18151	28.76	36	19153	28.49	38
LPL	7538	29.55	15	7290	28.92	15
Total	25689	29.55	51	26443	28.92	53

in Table III. The simulations are carried out with both uniform loss and fading channel model to represent wireless channels.

For the video transmission, depending on the bandwidth and packet erasure rate, PS1 could be varied to maximize the PSNR. An interesting scenario is with PS1 as 100%, thus effectively sending HPL only.

In the absence of any errors the PSNR achieved with both the descriptions is the same as that of standard coding, considering slicing and data partitioning.

4. UNIFORM LOSS MODEL

4.1. Simulation Setup

In this section we restrict the analysis to uniform loss. The probabilities PS1 and PS2 are kept the same for both paths. The data rate for each of the paths is kept as 10% over the source rate. The probability of error (PE) on paths are varied in increments of 0.05 from 0.05 to 0.25, to yield 15 combinations (0.05, 0.05), (0.05, 0.10), (0.05, 0.15), (0.05, 0.20), (0.05, 0.25), (0.10, 0.10), (0.10, 0.15), (0.10, 0.20), (0.10, 0.25), (0.15, 0.15), (0.15, 0.20), (0.15, 0.25), (0.20, 0.20), (0.20, 0.25), and (0.25, 0.25). We term this pair of PE as PPE1 through PPE15.

4.2 Results and Analysis

The results for each set of PE with different PS1 are shown in Fig. 3. Each curve (from left to right) depicts a PPE from (0.05,

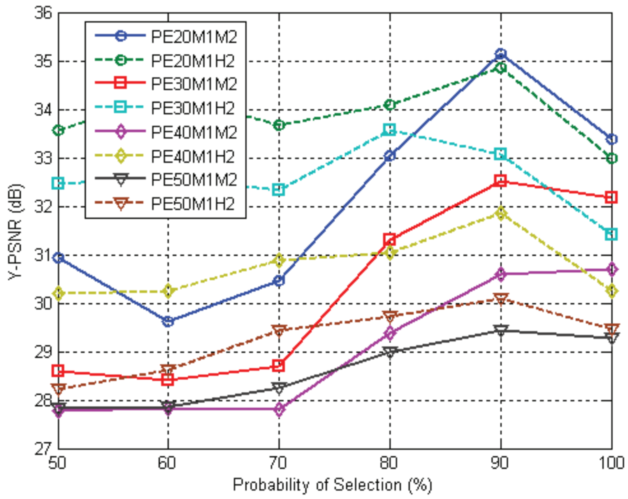

Fig. 5. Comparison of sending MDC1, MDC2 with MDC1, HPC2.

0.05) to (0.25, 0.25) and shows a downward trend subject to cumulative PE on both channels. The scheme with PS (100, 0) is seen to be virtually independent of the increasing packet erasures. Thus, by increasing the protection of HPL for a description it is possible to offset the effect of higher packet loss rates, and hence, a consistent reconstruction quality can be obtained.

The results for some PE are shown in Fig. 4. It can be seen that as the PS is increased, all the schemes converge at around 33.91 dB which is the PSNR achieved with receiving only the HPL for both MDC1 and MDC2.

5. FADING CHANNEL MODEL

5.1. Simulation Setup

In this section we assume one of the channels C1 to be better than the other C2. The PE1 is kept as 0.05, whereas PE2 is varied as 0.2, 0.3, 0.4 and 0.5. Gilbert model is used to simulate the burst losses with an average burst length of 5 packets for PE1 and PE2. The PS is varied from 0.5 to 1.0 as in the case of uniform loss. The data rate for each of the paths is kept as 10% over the source rate.

Also, to establish the importance of adapting the description to the channel bandwidth and error characteristics, we assume for C1 same configuration as described above in this section. However, for the channel C2 we transmit HPL only. The reason for this is that for a low bandwidth channel like C2 here, the protection of whole of the description MDC2 is not possible. We present results for the adaptive scheme to arrive at an optimal rate-allocation by varying the coding parameters, i.e., PS, and the information (complete description or HPL).

5.2 Results and Analysis

The results for transmission comparing the PSNR with burst loss, for transmitting MDC1 and MDC2 in full, with transmitting MDC1 with only HPC2 at different PS are shown in Fig. 5. The configuration PE20M1H2 is a scheme with 20% burst errors wherein MDC1 and HPC2 are being transmitted. It can be seen that the schemes with only HPC2 being transmitted have better results for the entire range of PS. Thus, in any such scenario, where the channel is incapable of successfully communicating the whole of source data, it is advantageous to reduce the source data itself. This reduction brings a drop in PSNR which is offset by gain in successful reception (from reduced information), hence providing an overall improved result.

6. CONCLUSION

The design of two path MDC scheme has been evaluated under various channel conditions. The scheme is fully compatible with the standard H.264/AVC encoder. The proposed scheme of UEP with slicing and data partitioning of H.264/AVC constructs layered video data amenable to FEC with EW-RLC. The transmission of only the HPL layer has been shown to offset the worst channel conditions. This makes such adaptation both beneficial and necessary for low bandwidth channels. This optimal scheme for rate adaptation over multiple paths for wireless video transmission can adaptively operate at GOP level.

This work will be extended to encompass transmission scenarios with more than two paths. The application of proposed scheme over the emerging LTE-A standard will also be explored.

REFERENCES

[1] T. Wiegand, G. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Cir. Sys. Video Techn.*, vol. 13, no. 7, pp. 560-576, Jul. 2003.

[2] Vivek K. Goyal, "Multiple Description Coding: Compression Meets The Network," *IEEE Signal Processing Mag.*, vol. 8, no. 5, pp. 74–93, Sept. 2001.

[3] M. Wang and B. Li, "How practical is network coding?," *IEEE IWQoS 2006, Intl. Workshop on Quality of Service*, pp. 274-278.

[4] S. Nazir, D. Vukobratović, and V. Stanković , "Performance evaluation of Raptor and Random Linear Codes for H.264/AVC video transmission over DVB-H networks," in *Proc. IEEE ICASSP-2011, Intl. Conf. on Acoustics, Speech, and Signal Processing, Prague, Czech Republic*, May 2011.

[5] T. Tillo, M. Grangetto, and G. Olmo, "Redundant slice optimal allocation for H.264 multiple description coding," *IEEE Trans. On Circuits and Systems for Video Technology,* vol. 18, no. 1, Jan 2008.

[6] S. Mao, S. Lin, S. Panwar, Y. Wang, and E. Celebi, "Video transport over ad hoc networks: Multistream coding with multipath transport," IEEE J. on Selected Areas in Communs., vol. 21, no. 10, pp. 1721–1737, 2003.

[7] K. Nybom, S. Gronroos, and J. Bjorkqvist, "Expanding window fountain coded scalable video in broadcasting," in *Proc. IEEE ICME-2010, Intl. Conf. on Multimedia and Expo., Singapore*, July 2010.

[8] ITU-T Recommendation H.264 (03/2010) Advanced video coding for generic audiovisual services.

[9] D. Vukobratović, V. Stanković, D. Sejdinović, L. Stanković, and Z. Xiong, "Scalable video multicast using Expanding window fountain codes," *IEEE Trans. Multimedia*, vol. 11, pp. 1094-1104, Oct. 2009.

[10] V. Stankovic, R. Hamzaoui, and Z. Xiong. "Packet loss protection of embedded data with fast local search," in *Proc. Int'l Conf. Image Processing*, Rochester, NY, Sep. 2002.