Iterative H.264 Video Decoding Using Mutual Information Exchangeable DSTS and SP Modulation

Nasru Minallah, Amaad Khalil and Nasir Ahmad
Department of Computer Systems Engineering, University of Engineering and Technology, Peshawar, Pakistan.
Email: {n.minallah}@uetpeshawar.edu.pk,
http://www.uetpeshawar.edu.pk

***Abstract*—In this paper we analyze the impact of unequal error protection (UEP) over Data-Partitioned (DP) H.264 video transmission over Correlated Rayleigh fading channel using Recursive Systematic Convolution Codes (RSC). Multi transmit and receive antenna system with Differential Space Time Spreading (DSTS) is employed in combination with Sphere Packing (SP) Modulation to demonstrate the overall BER reduction and PSNR improvement of the received bit-stream. Furthermore, Iterative Soft Bit Source Decoding (SBSD) along with Channel Decoding mechanism is employed, which results in mutual information exchange to correct possible errors and results in the enhanced quality of the final decoded video. Similarly the performance improvement due to additional Exit-Optimized Over-complete Mapping(OM) is also presented. It is important to note that the overall bitrate of all the consider coding schemes is kept constant, while conducting the performance comparison. In UEP important Partitions in H.264 video stream are more error protected compare to less important Partitions. It is observed that using DSTS along with SBSD and Over-complete Mapping results in BER reduction and PSNR improvement of the received H.264 Coded bit-stream.**

I. Introduction

The main goal of H.264 AVC standard is the design of high-compression efficient and Network friendly video transmission system. The H.264 Video Coding Layer (VCL) consists of different levels of elements like block, macro-block and slice and is responsible for significant video compression performance gain [1]. H.264 video coding techniques like multiple reference frames, deblocking filter, $1/4$ pixel motion compensation and integer transform resulted in an increase in coding performance [2]. The Network adaptation is achieved through Network Abstraction Layer (NAL). NAL provides header information and represent VCL in a manner appropriate for various wired and wireless networks. The H.264 results in significant bit-rate reduction, compare to all previous standards such as ISO/ IEC JTC1 MPEG4 and ITU-T Rec. H.263 at the same quality level [3]. However due to compression efficiency the generated bit-stream is more vulnerable to transmission errors. Different error resilience mechanisms are incorporated in H.264 to reduce this problem, but they are computationally complex and also negatively effect compression performance. Various robust video communication mechanisms such as layered video coding with unequal error protection are proved to be beneficial for H.264 video stream transmission.

In recent work [4], the error resilient coding schemes of H.264 are exploited and unequal error protection is applied to adaptive macro-block classified slices using Reed Solomon codes. Another approach is presented in [5], which uses data-partitioning mode of H.264/AVC. Unequal Error Protection (UEP) is performed using Adaptive Hierarchical QAM. Similarly data partitioning approach is also used in [6]. Data partitions were Unequal Error Protected based on their relative importance and was compared with non-scalable video coding under similar applications and transmission constraints over mobile channels. UEP to data partitioned H.264 video stream is also performed in [7], using turbo codes. Channel Adaptive Joint Source and Channel coding scheme is proposed in [8], where UEP is applied to different relative important partitions of H.264 using Rate Compatible Punctured Turbo Code (RCPT). In [9] a Joint Source and Channel Decoding method based on Maximum a posteriori probability (MAP) was proposed. The proposed method is applied to decode Motion Vectors in H.264 coded video stream. Similarly in [10] a joint source channel decoding system is proposed, where multidimensional intra-frame correlation is exploited by the receiver for error correction. An application of Iterative Source Channel Decoding for distributed video coding by modeling the video signal with Iterative Horizontal-Vertical Scanline Model (IHVSM) relying on a first-order Markov process is presented in [11].

In this paper we have performed Slice Error Sensitivity of data partitioned H.264 bit-stream. Using this knowledge UEP is performed on the video stream encoded using by employing Recursive Systematic Convolution codes. The resultant bit-stream is modulated using Sphere Packing (SP) modulation and is finally transmitted using Differential Space-Time Spreading (DSTS). The corresponding performance improvement is investigated. The performance improvement gained through SBSD and Over-complete Mapping is also demonstrated.

II. Data Partitioning (DP) employed by H.264

Apart from compression efficiency and network friendliness, H.264 video codec has also considered its support for advanced features resilience to error. In DP H.264, video stream is partitioned into 3 different types of Network Abstraction Layer Units (NALU's) [1].

A. *Type A NALU:*

It consists of header information of the slice, types of the MB, Quantization parameter and motion vectors. It's the most important type of partitions, because without it the information of the remaining two partitions can't be decoded.

B. Type B NALU:

It contains intra Coded Block Patterns (CBP's) and intra coefficients. The intra information is used to stop further drift and is more important then inter information. It's usually the smallest partition of the coded slice.

C. Type C NALU:

It carries inter CBP's and inter coefficients. The inter information is not helpful for re-synchronizing encoder and decoder and is therefore of least important. In most of the time it contain most of the data of the coded slice. Its usability is dependent on the availability of Type A partition but not on Type B partition.

The Block Diagram H.264 Slice with and without Data Partitioning is shown in Figure 1

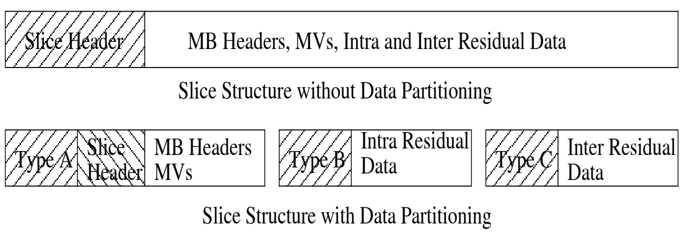

Fig. 1. Block Diagram H.264 Slice with and without Data Partitioning [12]

All these partitions need to be present for high quality video reconstruction. However acceptable reconstruction quality can be achieved even in the absence of intra or inter partitions, because of the availability of MB types and MV information, while only the texture information which is missing.

III. H.264 Slice Error Sensitivity

Before applying UEP to H.264 bit-stream, it was subjected to slice error sensitivity, in order to provide robust source matched UEP for its different types of slices. For this purpose we systematically corrupted bits in the first 4 "P" frames of 45 frame QCIF resolution akiyo video sequence. The PSNR degradations were recorded for each decoded frame under the following conditions:

1)While finding the sensitivity of specific slice A, B or C, the corresponding type of slice is corrupted sequentially one by one in the specified frame and the PSNR degradation resulted in the decoded sequence due to corresponding corrupt slice is recorded.

2)The above process is performed separately for each type of slice, only in the first 4 P frames of the sequence.

The details about number of Slices and their Size in bits for the first three P frames is shown in Table I

In H.264 video coded stream the most important type is slice A, which contain very important data such as Motion Vectors and header information. Slice B and C are always dependent on slice A and will not be decoded in the absence of A. If

TABLE I
Number of Slices and their Size in bits

P Frame #	A		B		C	
	Slices	Size	Slices	Size	Slices	Size
1	9	1152	3	856	9	48
2	9	1624	2	608	9	2128
3	9	1904	3	1152	9	3664
4	9	1336	3	952	9	720
Total	36	6016	11	3568	36	6560

A slice is dropped the whole slice including B and C will be dropped and the corresponding block will be considered as corrupted block. For corrupted blocks the decoder has enhanced capability to apply regular error concealment to the corrupted slice. Other wise if Partition A is correct, then header and motion vector information in Partition A is decoded and is used in the motion compensation process [6].

A. Slice Error Sensitivity of Partition A

Overview of the objective PSNR degradation caused by systematic slice A corruption events is given in Figure 2. It's observed from the figure that 2 out of 36 partitions have PSNR degradation/ frame about 10dB and 1partition has PSNR degradation above 1 dB range. All the remaining partitions PSNR degradation is below 1 dB range.

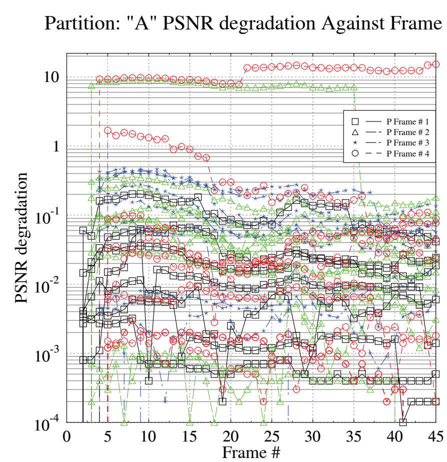

Fig. 2. H.264 Partition A Slice Error Sensitivity

As stated above, the whole slice including B and C is dropped as a result of corrupt Partition A. The complete slice is considered as corrupted and therefore the H.264 source decoder use to apply regular concealment by utilizing information from the previous frames. The error concealment techniques are normally very effective in case of low motion scene videos like akiyo and therefore the decoded video sequence has lower PSNR degradation.

B. Slice Error Sensitivity of Partition B

The objective PSNR degradation caused due to systematic slice B corruption event is shown in Figure 3 below. From the figure its observed that 5 out of 11 partitions have PSNR degradation in the range above 10dB, 2 are in 1dB range and the remaining have PSNR degradation lower then 1dB.

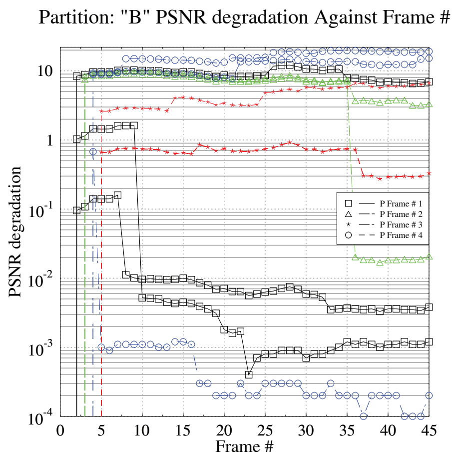

Fig. 3. H.264 Partition B Slice Error Sensitivity

Corruption of Partition B means that intra MB update will not be added to the reconstructed frame. The absence of intra MB update may results in producing a drift in the decoded video.

C. Slice Error Sensitivity of Partition C

From the Figure 4 its clear that only 1 out of 36 type C partitions have PSNR degradation in 10dB range while the remaining partitions have PSNR degradation lower than 1 dB.

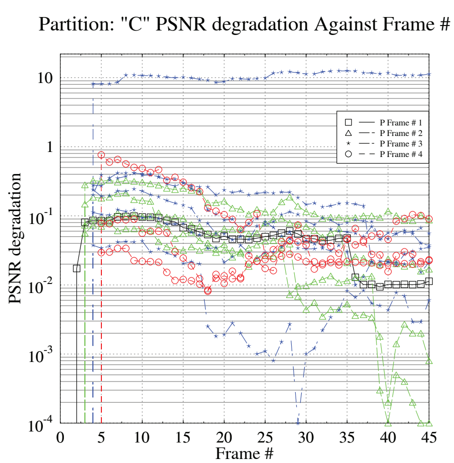

Fig. 4. H.264 Partition C Slice Error Sensitivity

Partition C contains intra MB update information. In case of low motion video sequence (video phoney) there is significant correlation among consecutive frames and the loss of partition C doesn't result in significant PSNR degradation.

IV. System Overview

UEP is provided through diverse rate RSC codes to the various H.264 video stream types. The different rate codes applied using Equal Error Protection (EEP) and Unequal Error Protection (UEP) are shown below in Table II,

TABLE II
Partition Type Vs Code-rate

Error Protection Type	Code Rate				
	A	B	C	3/4 OM	Overall
EEP1	1/3	1/3	1/3	No	1/3
EEP2	4/9	4/9	4/9	Yes	1/3
UEP1	1/3	1/4	5/12	No	1/3
UEP2	4/9	1/3	5/9	Yes	1/3

In order to evaluate the performance of the proposed system, we kept constant the overall bit rate for each of the error protection schemes. In UEP1 more protection to errors is considered for B type Partition to protect it more from channel errors compare to Partition A and C, because based on the Slice Error Sensitivity results, its observed that Partition B is one of the important partition and in case of corruption it results in significant reduction in PSNR. Also Partition B has the least size contribution, therefore it doesn't results in significant increase in over all bit-rate, as a result considerable error protection can still be applied for the Partition A and C. At the receiver iterative SBSD in combination with channel decoding is provided, which results in improved error correction ability of the system results in enhanced the subjective video quality.

A. Soft Bit Source Decoding

The SBSD scheme utilizes the redundant information contained in the H.264 coded bit-stream after source coding is used to determine extrinsic information. The presence of residual redundancy is evident from the existence of the non-uniform probability distribution M array symbols $P(s_k), s_k = (b^k(1), b^k(2), ...b^k(M))$ as shown in the table III.

TABLE III
Symbol Vs Probability

Symbol	Probability
0	0.22791640
1	0.11727422
2	0.11945398
3	0.10481402
4	0.11365674
5	0.10745756
6	0.10268064
7	0.10674644

This distribution is obtained by chopping the H.264 coded video-stream into 3 bits per symbol [13].

For our simulation the 300 frame "Akiyo" video sequence, 150 frame "Miss America" video sequence and 300-frame "Mother and Daughter" video sequences were encoded using H.264/AVC and they were used as training video sequences. These video sequences were selected on the basis of having relatively small head and shoulders motion without any fast object motion and screen changes, which justify our video-phoney video scene scenario.

B. Over-Complete Mapping

The performance of the SBSD is depended on presence of residual redundancy or inherent inter-symbol correlation within the source coded stream after source coding to reduce the adverse effects of transmission errors on video quality. However in the presence of efficient source codec such as H.264 limited redundancy is left in the source encoded bit-stream which results in slight extrinsic information provision of SBSD and therefore it provides in less performance improvement beyond 2 decoding iterations [13]. Therefore in this scheme we intentionally increased the redundancy in the encoded bit-stream with the aid of over-complete mapping. The mapping scheme of rate $3/4$ was used and is shown below in Table IV,

TABLE IV
SYMBOL VS MAPPING

Symbol	Mapping
000	0000
001	1001
010	1010
011	0011
100	1100
101	0101
110	0110
111	1111

Therefore an addition Exit Optimized Over-complete mapping [13] is introduced in the source coded stream. This results in introducing intentional redundancy in the source coded bit-stream for and performance improvement.

The system diagram of the proposed model is shown below in the Figure 5 and 6.

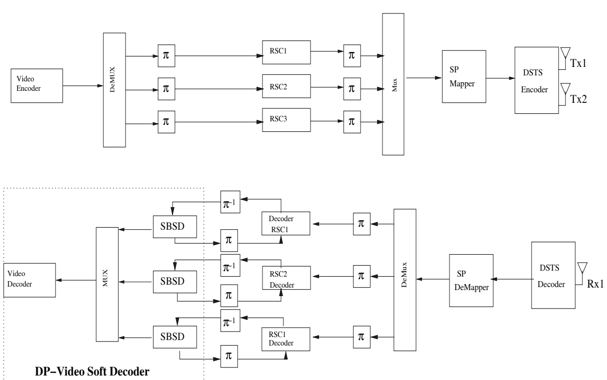

Fig. 5. Iterative source and channel decoding model

V. PERFORMANCE RESULTS

Both EEP and UEP is applied to partition A, B and C of the H.264 encoded video stream using various rate RSC codes. For performance comparison the corresponding final data rate of both EEP and UEP type systems with and without OM is fixed to 1/3.

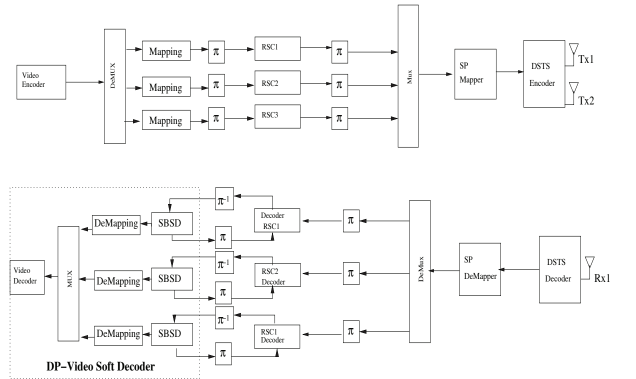

Fig. 6. Iterative source and channel decoding model with Soft-bit Source Decoding

Akiyo video consisting of 45 frames in QCIF format is encoded with H.264 video codec at 15 fps is used as a test sequence. The resultant video sequence has is of IPPPPP... form, consisting of error free first I frame and remaining 44 P frames i.e. the video sequence will follow the insertion of 'I' anchor frame after every 3 seconds which will stop the propagation of errors. The video sequence based on these specific parameters were selected by keeping in view that the drift and average quality in data-partition video stream is dependent on P frames in each group of pictures (GOP). The bit-rate of the coded video is 96 kbps. Each frame is partioned into 9 slices. Furthermore, each slice is made of 11 MB's.

Video sequences were encoded using JM 12.1 of the H.264/AVC standard. Error resilient tools like Flexible Macro block ordering (FMO) [14] and multiple frames used as reference for inter motion search result in a very limited performance improvement in low motion scene video like "akiyo". Therefore in our simulation FMO is turned off. Similarly only one previous frame is considered for inter frame motion vector search, therefore it corresponds to the reduced number of computations as compare to considering 5 previous frames for motion search. For the sake of confidence we ran each experiment 60 times and recorded the average results. The PSNR and BER performance of these different schemes using Correlated Rayleigh fading channel is graphically represented in Figure 7 and 8. It is clear from the presented results that the UEP scheme with SBSD result in the best EbN0db Vs BER performance. Furthermore, the employed transmission mechanism with UEP resulted in $5db$ performance gain at PSNR value of 40, with reference to Equal Error Protected (EEP) scenario.

VI. CONCLUSION

In this paper we investigated the performance of Unequal Error Protection (UEP) on Data-Partitioned (DP) H.264 video transmission over Correlated Rayleigh fading channel using Recursive Systematic Convolution Codes (RSC). Multi transmit and receive antenna system with Differential Space Time

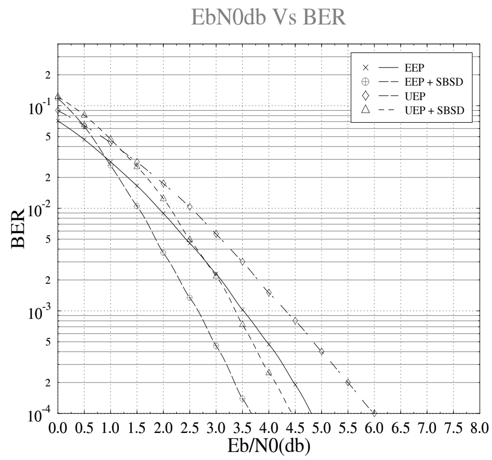

Fig. 7. EbN0db Vs BER curves of the consider system setup

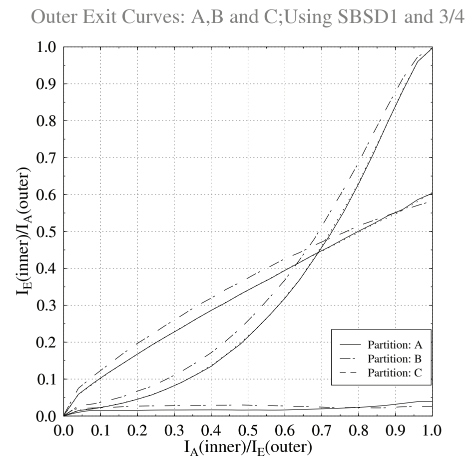

Fig. 9. Outer Exit characteristics of of the consider system setup

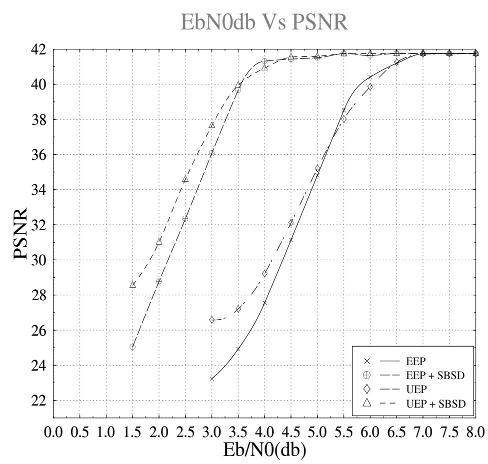

Fig. 8. EbN0db Vs PSNR curves of the consider system setup

Spreading (DSTS) is employed in combination with Sphere Packing (SP) Modulation to demonstrate the overall BER reduction and PSNR improvement of the received bit-stream. Furthermore, Iterative Soft Bit Source Decoding (SBSD) along with Channel Decoding mechanism is employed, which results in mutual information exchange to correct possible errors and results in the enhanced quality of the final decoded video. Similarly the performance improvement due to additional Exit-Optimized Over-complete Mapping(OM) is also presented. It is observed that using DSTS along with SBSD and Over-complete Mapping results in BER reduction and PSNR improvement of the received H.264 Coded bit-stream.

REFERENCES

[1] S. Wenger, "H.264/AVC over IP," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, pp. 645–656, July 2003.

[2] T. Wiegand, G. J. Sullivan, G. Bjntegaard, and A. Luthra, "Overview of the h.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, pp. 560–576, July 2003.

[3] 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, pp. 657–673, July 2003.

[4] N. Thomos, S. Argyropoulos, N. V. Boulgouris, and M. G. Strintzis, "Robust transmission of h.264/AVC video using adaptive slice grouping and unequal error protection," in *Multimedia and Expo, 2006 IEEE International Conference on*, (Toronto, ON, Canada), pp. 593–596, July 2006.

[5] Y. C. Chang, S. W. Lee, and R. Komiya, "A low-complexity unequal error protection of h.264/AVC video using adaptive hierarchical QAM," *IEEE Transactions on Consumer Electronics*, vol. 52, pp. 1153–1158, Nov. 2006.

[6] T. Stockhammer and M. Bystrom, "H.264/AVC data partitioning for mobile video communication," *2004. ICIP '04. 2004 International Conference on Image Processing*, vol. 1, pp. 545–548, Oct. 2004.

[7] M. El-Hajjar, O. Alamri, and L. Hanzo, "Differential space-time spreading using four transmit antennas and iteratively detected sphere packing modulation," in *Spread Spectrum Techniques and Applications, 2006 IEEE Ninth International Symposium on*, (Manaus-Amazon), pp. 322–326, Aug. 2006.

[8] N. D. Dao and W. A. C. Fernando, "Channel coding for h.264 video in constant bit rate transmission context over 3g mobile systems," in *Circuits and Systems, 2003. ISCAS '03. Proceedings of the 2003 International Symposium on*, vol. 2, May 2003.

[9] Y. Wang and S. Yu, "Joint source-channel decoding for h.264 coded video stream," *IEEE Transactions on Consumer Electronics*, vol. 51, pp. 1273–1276, Nov. 2005.

[10] Y. Huo, C. Zhu, and L. Hanzo, "Spatio-temporal iterative source channel decoding aided video transmission," *Vehicular Technology, IEEE Transactions on*, vol. 62, pp. 1597–1609, May 2013.

[11] Y. Huo, T. Wang, R. Maunder, and L. Hanzo, "Two-dimensional iterative source-channel decoding for distributed video coding," *Communications Letters, IEEE*, vol. 18, pp. 90–93, January 2014.

[12] P. Raibroycharoen, M. M. Ghandi, E. V. Jones, and M. Ghanbari, "Performance analysis of h.264/AVC video transmission with unequal error protected turbo codes," *2005. VTC 2005-Spring. 2005 IEEE 61st Vehicular Technology Conference*, vol. 3, pp. 1580–1584, May/June 2005.

[13] A. Q. Pham, L. L. Yang, and L. Hanzo, "Joint optimization of iterative source and channel decoding using over-complete source mapping.," in *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*, (Dublin), pp. 2404–2408, Apr. 2007.

[14] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with h.264/AVC: tools, performance, and complexity," *IEEE Circuits and Systems Magazine*, vol. 4, pp. 7–28, / 2004.