QoE Aware Resource Allocation for Video Communications over LTE Based Mobile Networks

(Invited Paper)

Ryan Perera, Anil Fernando, Thanuja Mallikarachchi,
Hemantha Kodikara Arachchi
I-Lab, Multimedia Communications Research, Centre for Vision Speech and Signal Processing
University of Surrey, Guildford GU2 7XH
United Kingdom
W.Fernando@surrey.ac.uk

Mahsa Pourazad
Institute for Computing, Information, and Cognitive Systems (ICICS)
University of British Columbia
289-2366 Main Mall
Vancouver, BC,
Canada, V6T 1Z4

***Abstract*— As the limits of video compression and usable wireless radio resources are exhausted, providing increased protection to critical data is regarded as a way forward to increase the effective capacity for delivering video data. This paper explores the provisioning of selective protection in the physical layer to critical video data and evaluates its effectiveness when transmitted through a wireless multipath fading channel. In this paper, the transmission of HEVC encoded video through an LTE-A wireless channel is considered. HEVC encoded video data is ranked based on how often each area of the picture is referenced by subsequent frames within a GOP in the sequence. The critical video data is allotted to the most robust OFDM resource blocks (RBs), which are the radio resources in the time-frequency domain of the LTE-A physical layer, to provide superior protection. The RBs are ranked based on a prediction for their robustness against noise. Simulation results show that the proposed content aware resource allocation scheme helps to improve the objective video quality up to 37dB at lower channel SNR levels when compared against the reference system, which treats video data uniformly. Alternatively, with the proposed technique the transmitted signal power can be lowered by 30% without sacrificing video quality at the receiver.**

Keywords: *HEVC, Video communications, LTE-A, Resource allocations, QoE.*

I. Introduction

The range of multimedia services used for everyday activities such as teleconferencing, mobile television and peer-to-peer video sharing are undergoing unprecedented growth recently. With this growth the quality of received video is of prime importance to users as well as service providers irrespective of where the users are connected from. A recent report by CISCO forecasts that in 2017 it would take an individual over 5 million years to watch the amount of video content crossing global IP networks within one month [1]. It further estimates that almost 90% of global data traffic will be video. This is mainly fuelled by the heap of consumer communication devices being introduced to the market and the user's higher expectation on video consumption. Since most users are mobile users, providing necessary capacity to handle these ever increasing video traffic possess a huge challenge on the wireless communications infrastructure since spectrum capacity is limited.

Even though the latest standards such as Long Term Evolution–Advanced (LTE-A), increases data rates in the downlink to as high as 1.5Gbps [2], this will not be sufficient when additional data is pumped together with actual content to support error resilience with more bandwidth-hungry video applications such as HD, UHD, SHD, 3D and HDR.

Meanwhile, new video coding standardisations, such as HEVC have been introduced to ease this situation by compressing the video data significantly compared to its predecessor and incorporating parallelism to a greater extent which helps real-time implementation [3]. Although HEVC introduces improved parallelism and compression efficiency, it disregards the transmission aspects of video data. When data is compressed more and more, recovering them at the receiver after it is transmitted through an error prone channel will be extremely challenging. A single error can cause the whole video sequence undecodable due to interdependencies between coding units, slices, frames, etc. In addition, bit errors caused by noisy channels and multipath propagation plays a crucial role in mobile wireless transmission environments. These errors create artefacts in the reconstructed video frames that propagate in both spatial and temporal domain due to the hierarchical prediction scheme employed in video compression stages. Therefore, it is essential to take the necessary precautions to mitigate these adverse effects when video is transmitted through error prone channels.

To address this problem, error resilience, error concealments, redundant data transmission are considered from the application layer. Some non-normative tools have been considered in the literature. However, the utilisation of error resilience tools and redundant data transmission is restricted by the channel bandwidth in transmission networks. The complexity of the codec and the loss of compression efficiency, also restrict the use of error resilient/concealment techniques in some application scenarios. On the other hand, to minimise the problem of video transmission in error prone channels, some approaches from the lower layers have also been considered.

This work was supported by EU FP7 project ACTION-TV (http://www.action-tv.net/), funded under European Commission's 7th Framework Program (Grant number: 611761).

However, these techniques have not exploited the properties of video coding and importance of the content of compressed video.

In this paper, a QoE aware resource allocation for video transmissions over LTE-A based mobile networks is proposed. HEVC based compressed video is considered with their importance of the coded slices. Slices have been ranked based on their importance to the motion estimation/compensation and based on that physical layer resources are allocated at the resource allocations stage in LTE-A.

The rest of the paper is organised as follows. Section 2 provides related work on video communications over error prone channels. The proposed architecture is presented in section 3. Section 4 describes the simulation environment used to demonstrate the effectiveness of the proposed technologies. Section 5 presents the results. Finally, conclusions and future work is considered in section 6.

II. Related Work

Coordination between application and physical layer for optimum video communications with systems, such as H.264/AVC and UMTS, LTE/LTE-A, can be found in literature [4]-[12]. In [4], [5], the authors addressed the problem of adapting the source application layer based on feedback received from lower layers. In [4], channel conditions were evaluated using feedback obtained from the receiver and at the decoder frame losses were predicted using ARQ feedback. Accordingly, the algorithm adjusted the encoder rate control and adapted the reference picture selection. H.264/AVC-SVC, was exploited in [5], in which channel conditions were evaluated using a similar feedback mechanism proposed in [4]. It proposed a cross-layer (application and physical) solution with respect to adaptive transmission rates for SVC layers. However, both [4] and [5] failed to consider video content properties in their adaptation efforts. All IP packets were treated with equal priority at lower layers and subsequently, packets containing high priority information such as parameter sets and motion vectors were as equally vulnerable to channel distortions as low-priority information such as error signals. Moreover these techniques demanded feedback from lower layers to the application layer which resulted some practical issues for real-time video communications.

Apart from the above approaches, several attempts have been made in the Media Access Control (MAC) layer scheduling and physical layer operations, such as MIMO adaptation, error correction coding and modulation adaptation to harmonise them with video content to improve video transmission efficiency. Authors in [6] presented a criteria to switch between spatial multiplexing (SM) and transmit diversity (TD) in a MIMO channel based on a frame-by-frame motion intensity metric and the receiver SNR. Using the motion intensity and receiver SNR value, switching SNR value for each Group of Pictures (GOP) of the sequence was calculated and the best suited transmission mechanism (SM or TD) was selected for the transmission of the GOP. In [7], a priority based scheduling algorithm based on the importance of the picture types in the video sequence was proposed in MAC layer scheduling. Packets encompassing data from I-frames were regarded as high priority while packets from P-frames were considered as less important. However, the techniques proposed in [6] and [7] were based either on GOP or frame level rather than micro features of the video sequence. A more granular adaptation method which can work in slice/slice-segments is more suitable for optimising the network resources. A scheduling scheme in which user data rates were dynamically adjusted based on channel quality as well as the gradients of a utility function was proposed in [8]. The utility function was designed as a function of distortion of the received video. However the use of this algorithm for real time video transmission was not clearly discussed and several limitations need to be resolved for real-time video transmissions with state-of-the-art systems such as HEVC and LTE-A.

The potential error propagation distortion (PEPD) [9] is an effective mechanism for quantifying the importance of a slice to indicate the distortion introduced to the video stream if a particular slice was not received at the decoder. Authors in [10] proposed a mechanism to obtain a metric for the PEPD in the compressed domain without decompressing the video sequence. Using the above information in the compressed domain and the channel state information, a scheduling algorithm at the MAC layer for multiple mobile users was proposed to maximise the quality of video.

Algorithms have also been proposed with implementations in both application layer and lower layers in combination [11], [12]. In [11], the video quantisation parameter (QP), time-frequency RB selection and modulation and coding scheme are jointly determined to optimise the end-to-end QoS. Similarly, authors in [12] also proposed a similar approach to optimise the QoE. As mentioned before, these techniques also needed feedback from lower layers to the application layer which resulted some practical limitations in real-time video communications. Furthermore, in video coding, it was only considered QP for the optimisation which does not actually represent key information about the content of the compressed video.

One of the main problems in many of the algorithms in the literature was actual content information have not been considered at lower layers which can have a significant impact on video communications. In response, in this paper, a QoE aware resource allocation scheme based on the importance of slices in the compressed video is proposed for video communications in a lossy wireless channel. State-of-the-art video codecs (HEVC) and wireless systems (LTE-A) have been considered to demonstrate the effectiveness of the proposed scheme.

III. Proposed Architecture

A. *Slice Ranking based on motion vector mapping*

Within a GOP, each frame is segmented into slices which are independently encoded. When the video is transmitted through an error prone channel, some slices are more important than others to stop error propagations. The importance of slices can be considered based on the actual contributions to the motion compensation. If a particular slice is contributing to a significant portion of the motion compensation in a given GOP, that particular slice is more important than others and this slice must be protected more to ensure better QoE at the video decoder. In this proposed architecture, the importance of slices are identified based on their actual contributions to the motion compensation process and ranked accordingly. For simulations, a GOP size of 4 is considered as in Figure 1.

The frames shown in Figure 1 are divided into slices such that each frame is transmitted using multiple NAL units rather than a single one. Different slice segmentation parameters are used depending on the frame resolution. For a CIF sequence, slice argument parameter 6 is used to indicate that a maximum of 6 CTUs will be used for a slice. This results in having 5 slices per frame for CIF resolution.

The importance of each slice within a particular GOP is measured using the total area that each slice corresponds to, within the GOP based on the motion vector information. In order to perform this prioritisation a slice database is maintained for each frame.

Let F be the set of frames (n) within the GOP such that,

$$F = \{f_1, f_2, f_3, \ldots, f_n\} \quad (1)$$

Let F_jS be the set of slices (k) within frame F_j,

$$F_jS = \{S_1, S_2, S_3, \ldots, S_k\} \quad (2)$$

$$j \in \{1,2,\ldots n\}$$

Then the S_kArea for the k^{th} slice is calculated as follows

$$S_k\text{Area} = \sum_{i=1}^{m} MV_{i,\,PUsize} \quad (3)$$

Where, m - total number of motion vectors pointed to the k^{th} slice (S_k) and S_kArea corresponds to the total area slice S_k of frame F_j that covers within the GOP based on the content captured from the slice S_k for motion compensation. All motion vectors belong to future frames within the GOP, which refer frame F_j and point to slice S_k are considered. $MV_{i,PUSize}$ is the area that is covered by the motion vector MV_i.

Finally, slice set F_jS is ranked based on the values of normalized S_kArea. Due to the access unit information that it carries, which is vital for the decoding process, slice S_1 is always ranked at 1st position in the ranking list. This ranking information is exchanged to the physical layer to allocate resources based on the channel state information.

B. Resource Allocation at the Physical Layer

The output of the HEVC encoder is a sequence of data that has no information about the nature of content. If a small part of this data is corrupted, its impact at the HEVC decoder is significant because HEVC uses a highly predictive algorithm (motion compensation and estimation). The proposed methodology passes an additional stream of data from the HEVC encoder which comprises details about the importance of each slice.

As described in the previous section, the slices of each frame will be ranked based on the number of times each slice is referred to by subsequent coding blocks during the motion estimation process. Based on this information, the sequence of NAL units is re-arranged such that high-priority NAL units of each frame are placed before low-priority NAL units in the transmitting order. Details of the prioritisation order is sent to the LTE-A receiver through the physical downlink control channel (PDCCH) which is the dedicated channel in the OFDM resource grid to transmit accompanying information such as scheduling information [13].

At the resource grid filling stage the Physical Downlink Shared Channel (PDSCH) RBs of each Transmission Time Interval (TTI, the time between transmitting one set of symbols to transmitting the next set of symbols) are sorted according to the channel distortion they experience which is predicted through Channel State Information (CSI) using a channel estimator as illustrated in the Figure 2. The re-arranged NAL units are allocated into the RBs such that the robustness of RBs match the prioritised NAL units, which effectively is a more efficient

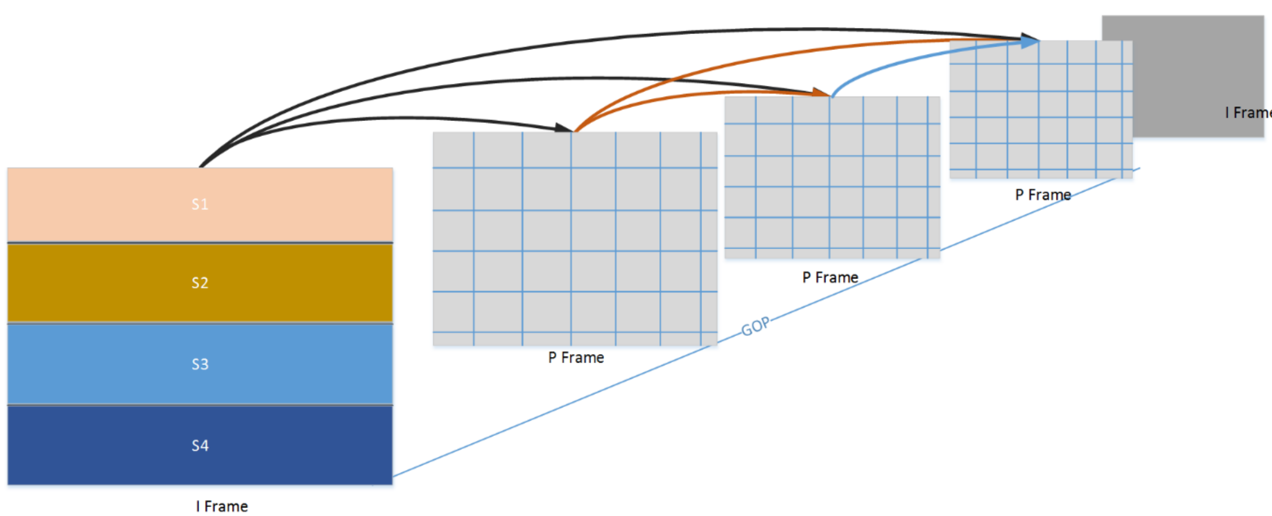


Figure 1. GOP structure and prediction references

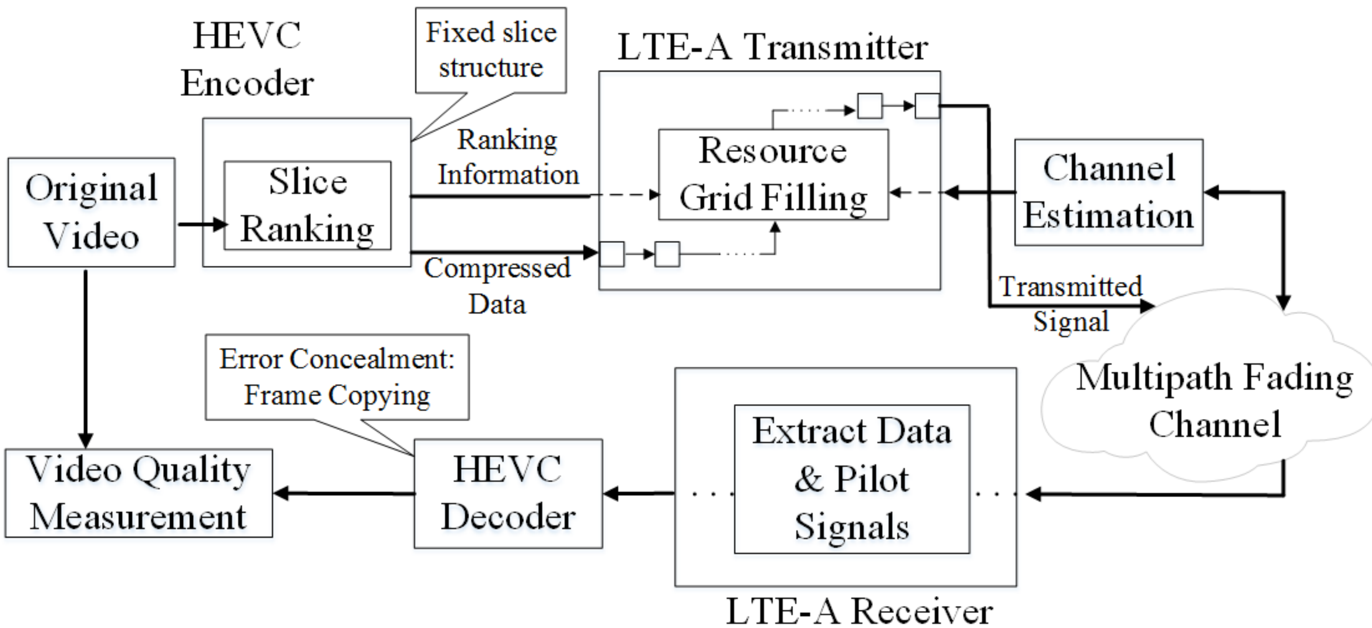


Figure 2. The proposed architecture

resource allocation mechanism to maximise video quality. Due to the proposed resource allocation technique described above, it is highly unlikely that significant NAL units will be corrupted by noise or multipath fading. In the event of a NAL unit being affected, it will be most likely a low priority NAL unit that will pose marginal impact on the received video quality because the impact of this lost NAL unit(s) on motion compensation is minimal.

The proposed technique has removed early termination at the presence of a codeblock[1] (CB) Cyclic Redundancy Check (CRC) detection. Early termination is useful where ARQ is conceivable and saves decoding computational power because an erroneous intermediate CB would render the entire transport block (TB) useless at the TB CRC detection stage. However, the proposed technique makes use of all the decoded data and by turning early termination off, retains the unaffected NAL units intact. When early termination is turned off, the receiver does not discard the following CBs of a TB when it detects an error, but continues decoding until the end of the TB.

At the receiver's end, the prioritisation details received from the encoder are turbo decoded and used to arrange the NAL units back to their respective locations before they are passed to the HEVC decoder.

The proposed architecture is illustrated in Figure 2. As described above, at the HEVC encoder slices in each frame are ranked and this ranking information is passed to the LTE transmitter. The ranking information is only used in the resource grid filling stage. The compressed video data received from the application layer is input to the TB CRC generation stage which is the initial stage of the LTE-A downlink. Subsequently, this data goes through CB segmentation, turbo coding, rate matching and bit selection, CB concatenation, scrambling, QAM modulation, layer mapping and precoding stages before being input to the resource grid filling stage.

After the data is transmitted over the multipath fading channel, the LTE-A receiver will regenerate the data using the side information which is transmitted through the PDCCH.

Finally, the output of the HEVC decoder is compared against the original video sequence and PSNR is calculated. The benchmark for the comparison is an exact duplicate of the proposed system except for the slice ranking and RB sorting.

IV. SIMULATION

For simulation purposes the HEVC encoder is configured such that each slice would contain a maximum of 6 Large Coding Units (LCUs) and hence each slice is a horizontal bar of 352×64 pixels. This results in 5 slices per frame (the resolution used is 352×288). During the simulation, a two-codeword spatial multiplexed transmission using closed-loop codebook-based precoding is used at the LTE-A transceiver. The simulation encompasses two antennas both at the receiver and the transmitter and a wireless channel with seven multipath elements is assumed. A 0.7 channel coding rate and a modulation of 64QAM is used. For each frame, 5 NAL units are inserted into each codeword. Since two codewords are transmitted simultaneously using spatial multiplexed transmission, two frames will be transmitted within each TTI. At the receiver, the turbo decoder executes the full eight iterations on the received data. The following table illustrates the main system parameters that are used during the simulation.

[1] Transport blocks are further segmented to code blocks to accommodate the interleaver at the turbo coder

TABLE I. SYSTEM PARAMETERS

Video Coding Parameters	
Frame size	352×288
Frame rate	30fps
Number of frames	21
Number of slices per frame	5
Reference frame	4
GOP size	4
Frame structure	IPPPIPPP…
Configuration	Low delay P main
Intra period	4
Slice argument[2]	6
Decoding refresh type	IDR
Slice mode	1
LTE-A parameters	
Channel bandwidth	10MHz
Number of RBs	50
Number of subcarriers per RB	12
Subcarrier bandwidth	15kHz
Channel coding rate	0.7
Modulation	64QAM
MIMO mode	Spatial multiplexing
Antenna configuration	2×2
Transmission time interval	1ms
Number of turbo decoder iterations	8
Early termination	Off

Since early termination is turned off at the LTE-A decoder, in the context of the simulation, a detected error will discard only one NAL unit instead of discarding the entire frame.

V. PERFORMANCE EVALUATION

To demonstrate the effectiveness of the proposed technique, three different video sequences have been considered (*news*, *highway* and *coastguard*). These sequences represent different motion characteristics (low, medium and high). To match the channel bandwidth, different QPs have been selected based on video characteristics.

As described above, since 5 NAL units are inserted within a transport block, the number of bits in 5 successive NAL units must not exceed the maximum allowable number of bits permitted for a TB. Using the LTE-A configuration explained in Table I, the maximum number of bits that can be inserted into a TB is approximately 25000. Therefore, NAL units of Intra coded frame (the type of frame that has the highest number of bits) should satisfy this requirement such that the 5 NAL units of the intra frame have less than 25000 accumulated bits.

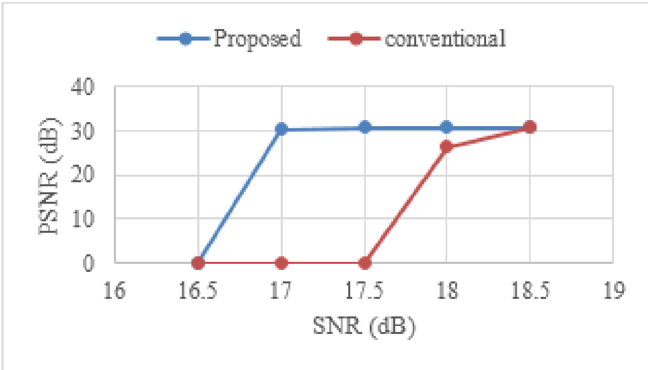
In order to satisfy the upper bit budget for the number of bits in each TB, the QP for each sequence is selected accordingly. Depending on the spatial characteristics of these video sequences, QP values have been selected as 32 for *highway*, 40 for *news* and 45 for *coastguard*.

The benchmark system used for the comparisons are the conventional HEVC encoder/decoder and the basic LTE-A transceivers. The output of the HEVC encoder is a sequence of NAL units which is input to the LTE-A transmitter in the same order as it is output from the HEVC encoder. Upon reaching the resource grid filling stage of the LTE-A transmitter, the data is sequentially allotted to the RBs. In this allocation process the RBs dedicated for synchronisation signals, broadcast channel, control channel and cell specific reference signals (CSRs) are skipped and only the RBs pertaining to the PDSCH is used. At the LTE receiver this data is detected in the same order it was allocated and is input to the decoding process. Subsequently, the sequence of received NAL units is input to the HEVC decoder and the raw video sequence is obtained.

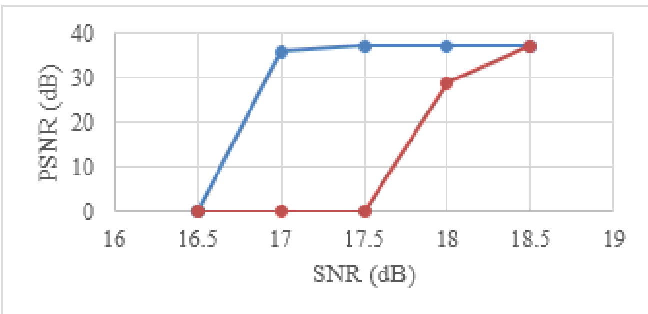
During the simulations, simple frame coping is used as the error concealment strategy employed at the HEVC decoder. Finally PSNR of the distorted video sequence is calculated with reference to the original video sequence under different channel SNR for the proposed technique and the benchmark system which was explained before. Figure 3 illustrates the average PSNR for the above three video sequences for different channel SNR values. Irrespective of the motion characteristic, all results show a similar performance. The proposed technique outperforms the conventional method of resource allocations up to 37dB at some channel SNR. Alternatively, the proposed technique can lower the transmitted signal power by 30% without sacrificing video quality at the receiver as illustrated in Figure 3. This can be considered as a significant step towards reducing the signal power in LTE-A base stations which results in higher capacity at the same end user QoE. The reason for the improvements achieved in the proposed methodology is the unequal treatment for the NAL units. Even though the channel BERs are similar in both cases the proposed methodology shows significant improvements in PSNR. The bit errors are occurred in the most vulnerable RBs. Since in the proposed technique, these vulnerable RBs carry least significant NAL units, their impact on HEVC decoder is minimal. This is the main reason for the performance improvement of the proposed technique.

All three figures show a steep fall in PSNR when the channel SNR falls below a threshold. The reason for this can be attributed to the characteristics of LTE-A physical downlink where the BER vs. channel SNR is designed with a steep gradient at the SNR threshold. Factors such as turbo decoding with 8 iterations and CRC insertion both at TB and CB levels contribute to this steepness. The reason for this characteristic is to enable adaptive modulation and coding at the base station. Accordingly, after a certain SNR threshold, the BER will rise exponentially distorting the majority of NAL units.

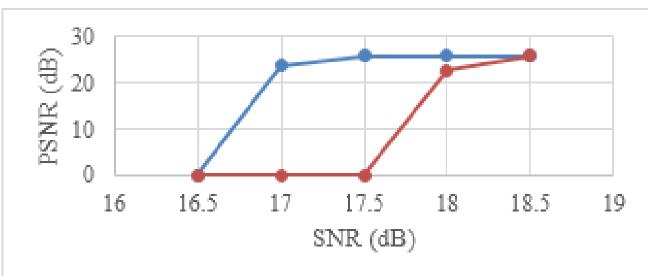
[2] Indicates maximum number of CTUs that a slice contains



a. News Sequence



b. Highway Sequence



c. Coastguard Sequence

Figure 3. System performance

VI. CONCLUSIONS

In this paper, a QoE aware resource allocation mechanism is proposed for LTE-A based wireless system to improve the system capacity. The proposed algorithm prioritises slices based on their importance in preventing error propagation in the decoded video. Based on CSI and slice ranking information the NAL units are allocated to the RBs. Simulation results indicate that the proposed technique outperforms the conventional method of resource allocations up to 37dB at some channel SNR. Instead, the base station can lower the transmitted signal power by 30% without sacrificing QoE at the receiver. However, the proposed technique needs to pass slice ranking information to the LTE-A transmitter which needs some additional overhead. However, unlike in other cross layer approaches, the proposed technique does not require any feedback channels between the physical layer and the application layer which will hinder real-time video communication. Furthermore, the additional overhead is negligible compared to the gains obtained in the video transmission. Future work will focus on investigating a more effective technique of ranking the slices and efficiently utilise adaptive modulation and coding together with intelligent RB allocation. Thereby, the base station will be able to further improve the utilization of system resources in video transmission to maximise the system capacity.

REFERENCES

[1] http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.html .

[2] E. Dahlman, S. Parkvall and J. Skold, 4G LTE / LTE Advanced for Mobile Broadband, Elsevier, 2011.

[3] G. Sullivan, J. Ohm, W. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," in *IEEE Transactions on Circuits and Systems for Video Technology,* pp. 1649 - 1668, 2012.

[4] C. Zhang, Y. Wang, S. Zheng and H. Zhang, "Multilevel Cross-Layer Feedback based Adaptive Video Coding for Wireless Transmission," in *8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, pp. 1-4, 2012.

[5] I. Ahmed, L. Badia, D. Munaretto and M. Zorzi, "Analysis of PHY/Application Cross-layer Optimization for Scalable Video Transmission in Cellular Networks," in *14th International Symposium and Workshops on a World of Wireless, Mobile and Multimedia Networks (WoWMoM) IEEE*, pp. 4-7, 2013.

[6] M. Hafeez, S. Jangsher and S. A. Khayam, "A Cross-Layer Architecture for Motion-Adaptive Video Transmission over MIMO Channels", in *IEEE International Conference on Communications (ICC)*, pp. 1-5, 2011.

[7] S. Karachontzitis, T. Dagiuklas and L. Dounis, "Novel Cross-Layer Scheme for Video Transmission over LTE-Based Wireless Systems," in *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1-6, 2011.

[8] P. Pahalawatta, R. Berry and T. Pappas, "Content-Aware Resource Allocation and Packet Scheduling for Video Transmission over Wireless Networks," in *IEEE Journal on Selected Areas in Communications,* pp. 749-759, 2007.

[9] Y. Zhang, W. Gao, H. Sun, Q. Huang and Y. Lu, "Error resilience video coding in H.264 encoder with potential distortion tracking," in *International Conference on Image Processing, (ICIP)*, pp. 163-166, 2004.

[10] L. Fan, L. Guizhong and H. Lijun, "A cross-layer scheduling algorithm for H.264 video transmission over wireless networks," in *Second International Workshop on Cross Layer Design (IWCLD)*, pp. 1-6, 2009.

[11] Y. Hu, G. Lv, S. Ci and H. Tang, "A Cross-Layer Video Transmission Scheme with Guaranteed End-to-End QoS over MIMO OFDM Systems," in *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 206-211, 2012.

[12] G. Lv, Y. Hu, S. Ci and H. Tang, "QoE-driven Cross-layer Optimized Video Transmission Scheme over MIMO-OFDM systems," in *18th IEEE International Conference on Networks (ICON)*, pp. 364-369, 2012.

[13] 3GPP Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (Release 10), 3GPP TS 36.213 v10.0.0, 2010-12.