Cross-Layer Resource Allocation Using Video Slice Header Information for Wireless Transmission Over LTE

Young-Ho Jung, *Member, IEEE*, Qing Song, *Student Member, IEEE*, Kyung-Ho Kim, Pamela Cosman, *Fellow, IEEE*, and Laurence B. Milstein, *Fellow, IEEE*

***Abstract*— In this paper, cross-layer resource allocation methods for wireless video transmission are proposed. We propose two practical metrics to measure the relative importance of each forward error correction (FEC) block to the user's quality of experience. They can be obtained from the header information, and require no overhead bits or amendment of the header format of the existing standard video codec. We discuss adaptive modulation and coding scheme allocation, and adaptive energy allocation for the FEC blocks of a group of pictures. We also provide low-complexity resource allocation solutions. In both approaches, we consider two types of slice packetization, which affect the priority of FEC blocks. The proposed cross-layer resource allocation methods have significant performance gain over equal error protection, and similar performance to more complicated algorithms.**

***Index Terms*— Unequal error protection, video communication, quality of experience, adaptive resource allocation, cross-layer optimization.**

I. Introduction

VIDEO packets can be corrupted in a wireless communication channel, which degrades the video quality of experience (QoE). Forward error correction (FEC) is usually used to preserve the video quality. For a given type of FEC, more FEC bits provide stronger protection. Therefore, when the total number of FEC bits for a video clip or a group of pictures (GOP) is fixed, unequal error protection (UEP) can usually be applied. More FEC bits are allocated to important video packets so that these packets can be transmitted more reliably.

There are many UEP schemes in the literature. For scalable video codecs [1], applying UEP is straightforward, since it is easy to measure the relative importance or priority of bits corresponding to different layers [2]–[9]. Hierarchical modulation and joint source-channel code rate optimization are examples. UEP approaches to maximize the QoE parameter, e.g., mean opinion score (MOS), are proposed in [10] and [11], however, only the relationship between the overall data rate and the MOS value was considered; the relative importance of each slice in a GOP was not considered. For pre-encoded non-scalable video streams, UEP is more challenging because it is difficult to obtain a utility metric to estimate the importance of each slice in a GOP. Yang *et al.* [12] considered frame loss in the low delay GOP structure (IPPP), and proposed to allocate more FEC bits to the frames at the beginning of each GOP, in order to reduce error propagation. For GOP structures including B frames, it is expected that I frames are more important than P frames which are more important than B frames. In order to obtain the quantitative value of importance of frames, more sophisticated methods have been proposed. For example, cumulative mean square error (CMSE) is used in [13]–[15]. CMSE is defined as the mean square error of all the frames that can be affected by a given frame or packet loss. It is measured using the distortion between the original video and the lossy video which is affected by that loss. Lin and Cosman [13] proposed an FEC code rate allocation by minimizing the overall CMSE with a constraint on the overall bit rate. A similar technique can be performed using video quality metric (VQM) [16]. Note that the computational complexity is greatly increased, because one has to tentatively discard one frame and decode at least one GOP to compute CMSE or VQM associated with that frame. The procedure has to be performed for each frame in the video. In other words, we have to decode the video as many times as the number of frames. In [17], a frame loss visibility model is built by analyzing the motion vectors and the residue after motion compensation. The computational complexity is reduced compared to the metrics using CMSE or VQM, because the motion and texture information can be extracted by decoding the video only once. Luo *et al.* [18], estimate the overall degradation due to a frame loss by multiplying a scale factor with the distortion of the lost frame. The scale factor depends on the position of the lost frame in the GOP. Video packets are divided into several classes according to the impact of their losses, and the FEC strength is determined for each class.

The priority of video packets becomes more complicated when each frame is packetized into slices, and each slice includes only a portion of a frame. For example, it is possible

Manuscript received October 28, 2016; revised March 26, 2017; accepted April 22, 2017. Date of publication April 18, 2017; date of current version August 3, 2018. This work was supported by the National Research Foundation of Korea through the Korean Government (MEST) under Grant 2016R1A2B4016207. This paper was recommended by Associate Editor L. Zhou. *(Corresponding author: Young-Ho Jung.)*

Y.-H. Jung is with the School of Electronics and Information Engineering, Korea Aerospace University, Goyang 10540, South Korea (e-mail: yhjung@kau.ac.kr).

Q. Song, P. Cosman, and L. Milstein are with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA.

K.-H. Kim is with Korea Aerospace Industries, Daejon 34068, South Korea.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2017.2701503

that a slice in a P frame is more important than a slice in an I frame [19]. The importance of slices depends not only on the frame type, but also on the number of macroblocks that are included, and how well they can be concealed at the decoder when they are lost. CMSE and VQM can still be used to measure the importance of slices. The computational complexity, however, would be even higher, since the video has to be decoded as many times as the number of slices. To reduce the complexity, slice loss visibility is modeled in [19] with the features extracted from the video. Using the visibility model, an FEC code rate optimization scheme is proposed in [20]. Similarly, Wu *et al.* [21] designed UEP scheme by classifying slices into four groups using CMSE which is predicted using a visibility model. The FEC code rates are allocated to the four groups by minimizing the overall distortion.

Heavy computation is one problem with the aforementioned approaches. The other problem is that the expected distortion values or the visibility scores cannot be easily obtained at the medium access control (MAC) / physical (PHY) layer without changing the header contents. These metrics are not practical, but they provide points of reference for what can be achieved by other practical utility metrics. In this paper, we propose new utility metrics which can be easily applied at the PHY or data link layers by using the existing header information.

We consider two types of UEP-based resource allocation problems: a) adaptive modulation and coding scheme (MCS) allocation with a constraint on the total number of PHY layer resource blocks (RBs), and b) adaptive energy allocation with a constraint on the total energy. The MCS or transmit energy of each FEC block is assigned according to the relative importance of FEC blocks. Here, an FEC block is a group of bits after channel encoding. We find feasible solutions which can be practically used with low complexity.

For the adaptive MCS allocation, the MCS of each FEC block is determined by the relative importance of each FEC block in a GOP, without changing the transmission energy per RB or the number of source bits. An integer programming based optimization problem is formulated and a practical low complexity solution is proposed. For the adaptive energy allocation, all FEC blocks have the same MCS level, but the energy per RB is assigned according to the importance of the FEC block. We solve the energy allocation problem using convex optimization.

In particular, our cross-layer optimizations are performed on the 3GPP Long Term Evolution (LTE) system. Since the end of 2015, video traffic has been taking up more than half of mobile traffic, and the ratio is expected to keep increasing [22]. Therefore, enhancing video transmission over LTE has become an important issue. There were several approaches to optimize video delivery efficiency over LTE [23]–[26]. Most consider resource allocation (or scheduling) for multiple video streams which share communication resources [23]–[25]. The optimization is performed with consideration of a delay constraint, average data rate and available data rate. Yang *et al.* [26] proposed to adaptively change the video source bit rate according to the variation of the expected data rate. Note that this requires re-encoding

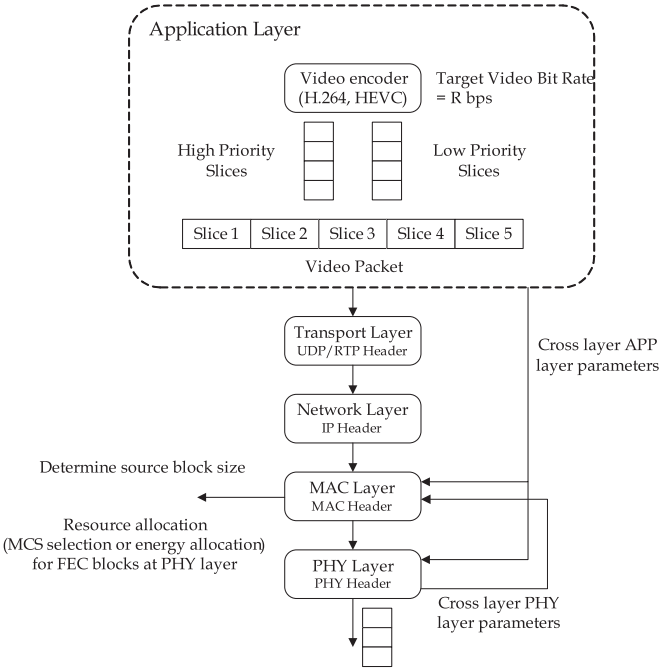


Fig. 1. Cross-layer optimization model between application layer and PHY/MAC layers for video QoE enhancement.

of a video stream. In this work, we consider pre-encoded single video stream transmission over LTE. The proposed algorithms can handle the granular RB structure and MCS options that fit with the current LTE system, and do not require any modification of the current PHY/MAC layer specification.

The rest of the paper is organized as follows: the system overview is described in Section II. In Section III, the two problems are formulated, and solutions are provided. We show the performance of the proposed methods in Section IV. Section V concludes the paper.

II. System Model

A. Cross-Layer Optimization Model

In this work, cross-layer optimized PHY-layer resource allocation methods are proposed. Fig. 1 shows a cross-layer optimization model between the application and PHY/MAC layers for video QoE enhancement. At the application layer, video streams are pre-encoded according to the target bit rates or pre-determined quality. Important application layer encoding parameters such as slice type, reference frame list, and frame number can be obtained at the MAC layer from the slice header. The application layer video packets are re-packetized at the transport layer and the network layer and some headers are attached. The transport and network layers are not within the scope of this paper. At the MAC layer, channel state information (CSI) is obtained from the PHY layer. Considering the CSI, an upper layer packet can be subdivided prior to channel encoding at the PHY layer. Then, UEP-based resource allocation is performed.

We consider an Orthogonal Frequency Division Multiple Access (OFDMA)-based PHY layer in frequency and time selective fading channels. We assume the coherence bandwidth

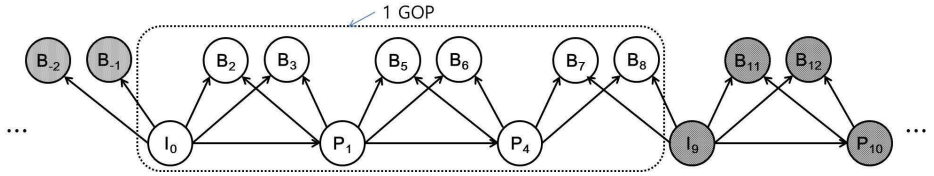


Fig. 2. An example of an open GOP structure of non-scalable video streams.

is larger than the subcarrier spacing and the coherence time is larger than the duration of an OFDM symbol.

A physical resource block (RB), which is the smallest unit of the physical layer resources used by the scheduling algorithm, consists of a certain number of OFDM symbols in the time domain and subcarriers in the frequency domain. For example, in 3GPP LTE, a RB consists of 12 subcarriers on the frequency axis and 14 OFDM symbols on the time axis which span 1ms [27], [28]. The subcarriers and OFDM symbols in a RB do not need to be adjacent in time or frequency. They can be spread over all available bandwidth and time region, in order to obtain frequency and time diversity. We assume the following PHY layer characteristics in this paper.

- The transmitter does not have the instantaneous CSI for any specific RBs, but only has limited CSI, such as average signal-to-noise ratio (SNR) for the entire bandwidth.
- Different RBs experience similar amounts of time and frequency domain channel selectivity, and each RB has no expected performance difference.
- The total number of available RBs and energy of all FEC blocks in a GOP are determined by the scheduler.

The two PHY layer resource allocation problems that we consider are based on these assumptions. The first problem involves the optimization of RB assignment to FEC blocks without changing energy per RB, and the second problem involves energy assignment without changing MCS.

B. Video Encoding and Generation of Slices and FEC Blocks

We consider the transmission of pre-encoded non-scalable video streams. Fig. 2 shows an example of an open GOP composed of one I-frame, two P-frames and six B-frames. The subscripts denote the transmission order of frames. An open GOP allows the B-frames from one GOP to refer to an I- or P-frame in an adjacent GOP.

Encoded video bits corresponding to a variable number of macroblocks in a video frame are organized into slices. Each slice has a header which includes the slice type, the frame number and other important information that defines the slice. In order to transmit the encoded bits through a communication network, slices are encapsulated into network abstract layer (NAL) units by adding header bits to indicate the type of data in the NAL unit. By adding a starting code prefix at the beginning of a NAL unit or adding an ending code postfix at the end of a NAL unit, the boundaries of the NAL unit can be identified by searching the coded data for the start of the unique code prefix or postfix pattern.

We consider two types of packetization: 1) one NAL unit includes a whole frame, and 2) each NAL unit includes a fixed number of bits.

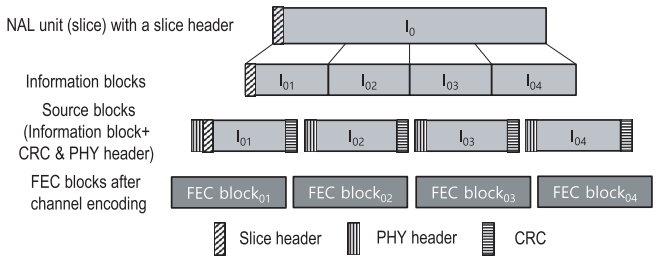


Fig. 3. Division of a NAL unit into multiple FEC blocks.

In the first scenario, since one NAL unit includes a frame, the number of bits of each NAL unit can vary considerably according to the type of frame (I, P or B) and video content. Because the number of information bits an FEC block can carry at the PHY layer is limited by the amount of PHY resources for a scheduling period, bits of a NAL unit are divided into multiple information blocks. Source blocks are generated by adding a cyclic redundancy check (CRC) and PHY header to each information block. Each source block is channel encoded to produce an FEC block. Each FEC block contains the same number of source block bits. Fig. 3 shows an example. A NAL unit corresponding to the I_0 frame is divided into four information blocks. Note that the first FEC block within these four information blocks is the most important, because only the first FEC block includes the slice header. If it is lost, the whole NAL unit cannot be decoded. Moreover, FEC blocks which are closer to the first block are more important than the blocks at the unit end. If FEC block_{02} is lost, then FEC block_{03} and FEC block_{04} are undecodable, since we cannot decode the syntax or locate the macroblock. We will call this the single slice per frame (sspf) scenario.

In the second scenario, each NAL includes a fixed number of information bits, so one frame can be divided into multiple NAL units. Each NAL unit can be fit into a single FEC block. Because each NAL unit is self-decodable, if one of the NAL units in a frame is missing, only the macroblocks in the missing NAL unit are lost and will be concealed. Note that the number of macroblocks in each NAL unit is not fixed. For example, a NAL unit might include a whole B-frame, while a NAL unit of an I-frame might include only one row of macroblocks. We will call this the multiple slices per frame (mspf) scenario.

III. Proposed Cross-Layer Resource Allocation

A. Proposed Utilities for Cross-Layer Optimized Resource Allocation

In this paper, we propose utility metrics which require no overhead for both types of NAL packetization.

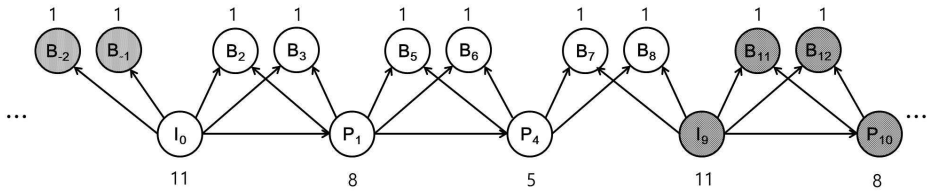


Fig. 4. An example of RNF_i.

1) Related Number of Frames (RNF): The RNF of the i-th FEC block is defined as

$$RNF_i = \text{the number of frames which can be affected by the loss of the } i\text{-th FEC block.} \quad (1)$$

Fig. 4 shows an example of RNF_i for the GOP structure in Fig. 2. Because FEC blocks of a B-frame are only related to the B-frame itself, the RNF is 1. In Fig. 2, the I-frame is related to all other frames in the same GOP and two B-frames at the end of the previous GOP, so the RNF is 11 for the FEC blocks of I-frames. The contents of FEC blocks of P_1 can be referred to by any other frames except I_0 in the GOP, so the RNF is 8. In a similar way, the RNF of each FEC block can be obtained requiring only the GOP structure and the frame number of the FEC block. Note that in both packetization scenarios, all FEC blocks in the same frame have the identical RNF values. We propose to use the RNF as the utility metric for resource allocation. We apply this as the slice importance to both packetization scenarios.

2) Multiple Slices Per Frame - Predicted VQM (P-VQM) Score: Unlike sspf, the importance of slices within a frame is unclear for mspf. When the number of bits of each NAL unit is fixed, the number of macroblocks in each NAL will be variable. A NAL unit can include only a portion of a frame. The importance of NAL units depends on many more factors than the previous scenario. In addition to the related number of frames, factors such as the number of macroblocks in the NAL unit, motion complexity of those macroblocks, and position of those macroblocks in the frame, have impacts on the importance. The work in [13] shows NAL units in P-frames can sometimes be even more important than NAL units in I-frames. A model using features extracted from NAL units to predict VQM scores is proposed in [19]. The problem with this method is still the complexity. Though the computational complexity is much lower than computing VQM scores, we still need to decode the video to extract motion vectors and residual energy of each macroblock. In our assumption, however, only the header information is available at MAC/PHY layer. To tackle that problem, we build a model to predict VQM scores using only the header information, and propose to use the P-VQM scores as the utility metric in this scenario. In the following, we describe the features used in this model and provide the coefficients of the model through training.

(i) Features
The header of a NAL unit includes the frame type, the frame number, and the starting macroblock number. From this information, we consider the following features:

a) *RNF* defined in the last section, which indicates the length of error propagation.
b) *IsIframe, IsPframe* and *IsBframe* are boolean factors according to the frame type of the NAL unit.
c) *NumMBs* is the number of macroblocks in the NAL unit.
d) *DevFromBdr* indicates how far the slice is from the border. It is computed as
DevFromBdr = *N/2* – *abs*(*Height* – *N/2*),
where *Height* is the spatial location of the NAL unit, *abs*(·) gives the absolute value, and *N* is the vertical resolution of the video divided by 16. For a CIF (352x288) video, *N* equals 18.
e) *NumBits* is the number of bits of the NAL unit. When we try to fix the number of bits of each NAL unit, we actually fix the maximum number of bits. Some NAL units can include fewer bits than the fixed number, because a NAL unit can include information of only one frame. For example, assume we set the maximum number of bits to 1000, and a frame is encoded with 1300 bits. The first NAL unit will include 1000 bits, while the second one will include 300 bits. The second NAL unit will probably have lower importance than the first one.
f) *IsFullUnit* and *NotFullUnit* denote whether the NAL unit includes the maximum number of bits or not, respectively. In the example above, we have *IsFullUnit* = 1 and *NotFullUnit* = 0 for the first NAL unit, and we have *IsFullUnit* = 0 and *NotFullUnit* = 1 for the second NAL unit.
g) *IsWholeFrame* and *NotWholeFrame* denote whether the NAL unit includes the whole frame or not. For example, assume that a B-frame is encoded with 600 bits, and an I-frame is encoded with 5000 bits. We set the maximum number of bits to 1000 again. The NAL unit of the B-frame will include the whole frame (*IsWholeFrame* = 1, *NotWholeFrame* = 0), and the NAL units of the I-frame will not include the whole frame (*IsWholeFrame* = 0, *NotWholeFrame* = 1).
h) *NumBitsPerMB* is the average number of bits per macroblock in the NAL unit. It is obtained by dividing *NumBits* by *NumMBs*.

TABLE I

CORRELATION COEFFICIENT AND MSE BETWEEN GROUND TRUTH AND P-VQM SCORES VERSUS FACTOR NUMBERS

Factor Number	Factors	Coefficient
α	1	-1.9874
1	*RNF* × *NumBits*	2.8983e-05
2	*IsPframe* × *IsFullUnit* × *NumMBs*	2.6309e-03
3	*IsIframe*	-1.9325e-01
4	*NumMBs* × *DevFromBdr* × *RNF* × *NotWholeFrame*	1.1235e-04
5	*IsPframe* × *NumBits*	4.9390e-05

(ii) Modeling

Like [19], we model VQM scores using logistic regression which is a type of generalized linear model (GLM). It can be represented as

$$\log\left(\frac{p}{1-p}\right) = \alpha + \sum_{j=1}^{K} x_j \beta_j, \tag{2}$$

where p is the normalized VQM score, α is a constant, x_j is a factor and β_j is its coefficient. The model coefficients, α and β_j, are trained using a training dataset.

The normalized VQM score is the VQM score of a slice divided by the maximum VQM score of the GOP. We try to estimate the normalized VQM instead of the actual VQM, because the VQM scores of different scenes can be very different from each other. The relative importance of slices is what we are interested in. The training dataset includes VQM scores of 15,000 NAL units from eight video sequences, which have different texture and motion complexity. Each video sequence includes 300 frames.

We apply an iterative feature selection technique in MATLAB to obtain the most useful features. We apply a 5-fold cross validation to obtain the model coefficients to prevent overfitting.

(iii) Model coefficients

We plot the correlation coefficient and mean squared error (MSE) between the ground truth and the P-VQM scores versus the number of features in Fig. 5. The correlation is 0.58 when 5 factors are included. Though the correlation is not high, it is higher than the model in [19] where more complex features are used. We show the five most important features and their coefficients in Table I, where × denotes interaction.

We will interpret the meaning of factors by the coefficients:

a) The number of affected frames still plays a crucial role. A larger value of *RNF* means more frames are affected by the loss of the packet, and thus the distortion will be shown longer on the screen. The two factors related to RNF have positive coefficients. *RNF* × *NumBits* means that when more frames are affected, and more bits are included in a NAL unit, the NAL unit is more important. More lost bits can be harder to conceal. Factor 4 indicates a NAL unit is more important if it affects more frames and larger areas, and is closer to the screen center. That is because distortion that covers a larger area is more likely to be detected by the viewer. The screen center usually draws the viewer's attention, and probably has more motion. Therefore, these factors contribute positively to the importance of packets.

b) There are two factors directly related to *IsPframe*, both with positive coefficients. In our settings, most P-frames are divided into several NAL units. *IsPframe* × *IsFullUnit* × *NumMBs* indicates, when the NAL unit is not the last one in that frame, the larger area it includes, the more important it is. This is consistent with the results in [19]. *IsPframe* × *NumBits* means the last NAL unit, if it has fewer bits than the other units in the same frame, has lower importance. The more bits the NAL unit includes, the more important it is.

c) *IsIframe* has a negative coefficient, but it does not mean I-frames have the lowest importance. I-frames have the largest *RNF*, and as mentioned before, the factor related to *RNF* has positive coefficients. Factor *IsIframe* just adjusts the importance of NAL units in I-frames. According to [16], NAL units in I-frames can be sometimes less important than NAL units in P-frames, because the former can include much fewer macroblocks than the latter.

We propose to use the model above to predict normalized VQM scores and use the scores as the utility for the mspf scenario.

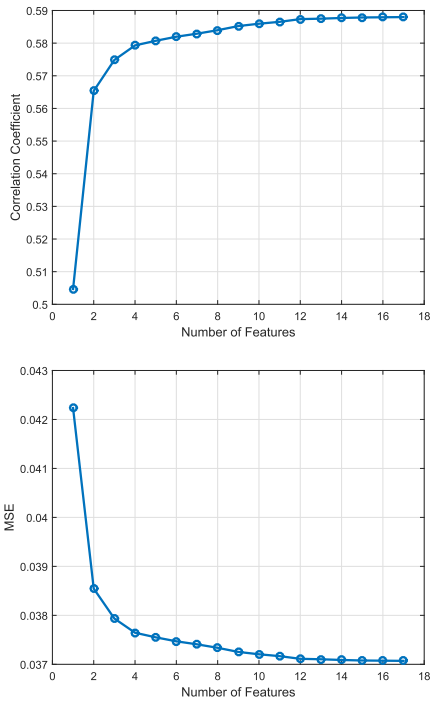


Fig. 5. Correlation coefficient and MSE between ground truth and P-VQM scores versus number of features.

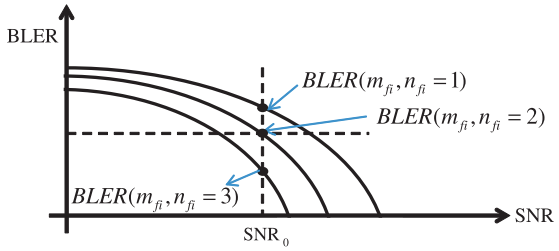


Fig. 6. Change of expected BLER according to the variation of the number of assigned RBs. This figure is a qualitative version of Fig. 9, which comes from [29].

B. Optimization Problem for Adaptive MCS Allocation

1) Problem Formulation: The optimization problem for cross-layer adaptive MCS allocation can be formulated as follows:

$$\arg\min_{\{n_{fi}\}} \sum_{f=1}^{F}\sum_{i=1}^{Q_f} u_{fi} BLER(m_{fi}, n_{fi}) \tag{3}$$

$$subject\ to \sum_{f=1}^{F}\sum_{i=1}^{Q_f} n_{fi} = N, \tag{4}$$

where F is the number of frames in a GOP, f is the index of the frame, Q_f is the number of FEC blocks in the f-th frame in a GOP, i is the index of the FEC block in a frame, n_{fi} is the number of RBs assigned to the i-th FEC block of the f-th frame, u_{fi} is the utility of the i-th FEC block of the f-th frame, and m_{fi} is the source block size of the i-th FEC block in the f-th frame. For the two scenarios described in Sec II: single slice per frame and multiple slices per frame, we use RNF and P-VQM as u_{fi}, respectively. For example, when the RNF value is used for the GOP structure in Fig. 4, $u_{fi} = 11$ for the FEC blocks corresponding to I-frames, and $u_{fi} = 1$ for the FEC blocks of B frames. $BLER(m_{fi}, n_{fi})$ is the expected block error rate (BLER) of the i-th FEC block of the f-th frame under a certain average SNR when the number of RBs assigned to the i-th FEC block is n_{fi}, and N is the total number of available RBs. In general, if we fix the source block size, almost all FEC blocks have the same m_{fi} value, but some of the last FEC blocks of each frame can have m_{fi} values which are less than the usual m_{fi} value. For the sspf packetization scenario, to give higher priority to the FEC block closer to the slice header, i.e., to those with lower index i in a frame, the following additional constraint is added:

$$n_{f1} \geq \cdots \geq n_{fQ_f}. \tag{5}$$

Q_f and N are given for a GOP but they can vary in different GOPs.

Since $BLER(m_{fi}, n_{fi})$ represents the loss probability of the i-th FEC block, the problem formulation of (3)-(4) aims to find a RB assignment combination for all FEC blocks of a GOP to maximize the total expected utility of the GOP at the receiver with the given N RBs. At a given average SNR in the assigned bandwidth, because the source coding is fixed, if more RBs are assigned to the i-th FEC block, the channel code rate can be reduced and more coding gain can be achieved. As shown in Fig. 6, when n_{fi} is increased, $BLER(m_{fi}, n_{fi})$ is decreased. We assume that CSI, such as average SNR and time/frequency selectivity of the channel, is fed back to the transmitter from the receiver, and $BLER(m_{fi}, n_{fi})$ is known at the transmitter corresponding to the channel conditions at the receiver. In addition, it is assumed that if m_{fi} is fixed, $BLER(m_{fi}, n_{fi})$ only depends on the number of assigned RBs, but does not depend on the specific RBs assigned to the i-th FEC block.

2) Proposed Heuristic Algorithm to Solve the Adaptive MCS Allocation Problem: The optimization in (3)-(5) is an integer programming problem, and is NP complete. Therefore, to obtain the optimum solution, exhaustive search is required. The computational complexity for exhaustive search is $O(Q^2N^3)$, where $Q = \sum_{f=1}^{F} Q_f$. A low complexity heuristic algorithm to solve (3)-(5) with good performance is desirable. Fig. 7 shows a heuristic resource allocation algorithm for the optimization problem. At the first step, $n_{\min}$ RBs are assigned to the FEC block that has the maximum utility, where $n_{\min}$ is the minimum number of RBs which produces $BLER(m_{fi}, n_{\min})$ less than a given threshold value. At the second step, we calculate the expected increment of the utility per additionally assigned RB for each FEC block. Then additional Δ_{RB} RBs or $n_{\min}$ RBs are assigned to the FEC block that has the maximum expected increment at Step 3. Steps 2 and 3 are repeated until all RBs are assigned. If there are multiple FEC blocks with the same utility, the FEC block with the lowest index has the priority.

For each iteration of Fig. 7, a comparison among Q candidates is required, and if $\Delta_{RB} = 1$, the maximum number of iterations is N. Therefore, the computational complexity of the proposed heuristic algorithm is O(QN), and is much lower than that of the exhaustive search.

C. Optimization for Adaptive Energy Allocation

1) Problem Formulation: The optimization problem for cross-layer adaptive energy allocation can be formulated as follows:

$$\arg\min_{\{E_{fi}\}} \sum_{f=1}^{F}\sum_{i=1}^{Q_f} u_{fi} BLER(E_{fi}) \tag{6}$$

$$subject\ to \sum_{f=1}^{F}\sum_{i=1}^{Q_f} n_{fi} E_{fi} = E_{tot}, \tag{7}$$

where E_{fi} is the energy of an RB assigned to the i-th FEC block of the f-th frame and E_{tot} is the total energy assigned to a GOP. The problem formulation of (6)-(7) aims to find an energy allocation for all FEC blocks of a GOP to maximize the total expected utility of the GOP at the receiver with the total energy E_{tot}. At a given SNR, if more energy is assigned to an FEC block, as shown in Fig. 8, the effective SNR is increased and the expected BLER is reduced. However, because there is a total energy constraint, we want to assign more energy to more important FEC blocks. For sspf packetization, to give higher priority to the FEC block

- Step 1: assign $n_{\min}$ RBs to the FEC block that has maximum utility
- Step 2: calculate $u_{fi}[BLER(m_{fi}, n_{fi\_old}) - BLER(m_{fi}, n_{fi\_old} + \Delta_{RB})]/\Delta_{RB}$ for FEC blocks with $n_{fi\_old} > 0$, and $u_{fi}\left(1 - BLER(m_{fi}, n_{\min})\right)/n_{\min}$ for FEC blocks with $n_{fi\_old} = 0$
- Step 3: allocate Δ_{RB} RBs or $n_{\min}$ RBs to the FEC block p^*, i^* that has maximum value of $u_{fi}[BLER(m_{fi}, n_{fi\_old}) - BLER(m_{fi}, n_{fi\_old} + \Delta_{RB})]/\Delta_{RB}$ or $u_{fi}\left(1 - BLER(m_{fi}, n_{\min})\right)/n_{\min}$
- Step 4: repeat steps 2 and 3 until all RBs are allocated

where Δ_{RB} is the step size of RB assignment and n_{min} is the minimum number of RBs which has $BLER(m_{fi}, n_{\min})$ less than given threshold value, e.g. 0.5.

Fig. 7. Proposed heuristic resource allocation method for (3)-(5).

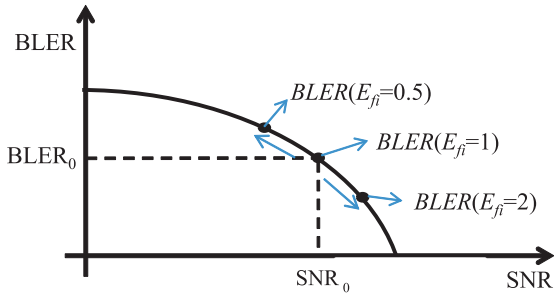


Fig. 8. Change of expected BLER according to the variation of assigned energy.

with lower index i, the following additional constraint is added:

$$E_{f1} \geq \cdots \geq E_{fQ_f}. \tag{8}$$

As in the case of adaptive MCS allocation, if we set u_{fi} to 1 for all FEC blocks, the optimization problem in (6)-(8) aims to minimize average BLER, and the solution is EEP, which has identical E_{fi} ($E_{fi} = E_{tot}/N$) for all RBs.

2) Convex Optimization-Based Solution for Adaptive Energy Allocation: When solving the optimization problem in (6)-(8), there is an obstacle to getting the solution efficiently, since the $BLER(E_{fi})$ in (6) is not a closed-form equation, but obtained from link level simulation results. Therefore, typical efficient optimization methods, such as convex optimization and linear programming, cannot be applied.

In order to apply a convex optimization method to (6)-(8), we need a closed-form convex equation for the reasonable approximation of each BLER curve within the SNR region of interest. As candidate convex functions for curve fitting, we consider the exponential form: $a\exp(-b\gamma)$, and power form: $a\gamma^b$, where a and b are curve fitting parameters, γ is the SNR, and $\exp(\cdot)$ is the natural exponential function. Fig. 10 shows an example of curve fitting for MCS 11 in Fig. 9. We can see that the exponential form approaches the original BLER curve quite well. Table II lists the approximated functions for MCS 7, 11 and 20, obtained by using the MATLAB curve fitting toolbox [30]. In Table III, the accuracy of curve fitting results for two candidate forms is compared. Two goodness-of-fit statistics are compared: R-square and root mean square error (RMSE) [31]. The range of R-square is from 0 to 1. When the approximation curve perfectly matches with the original curve, the R-square score is 1.

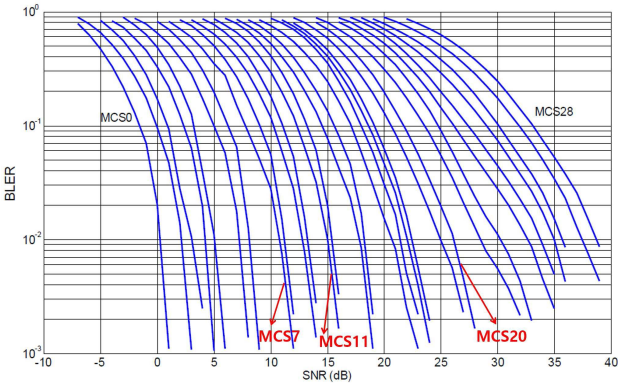


Fig. 9. An example of BLER performances for candidate MCSs [29] (LTE, two-layer spatial multiplexing, distributed virtual RB, ETU channel with Doppler frequency of 70Hz. All detailed parameters can be found in [29]).

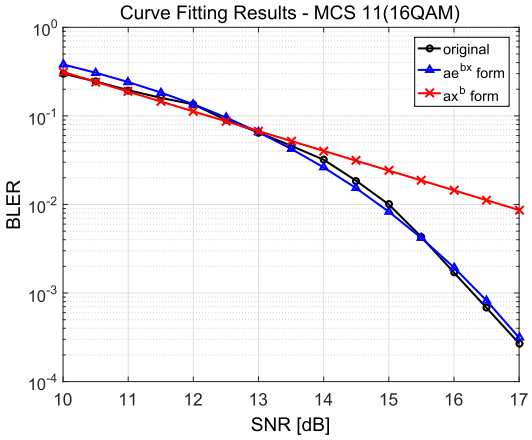


Fig. 10. An example of curve fitting of BLER simulation results to some convex functions.

For both goodness-of-fit statistics, the exponential-form approximation shows better fit, and we will substitute the exponential for the BLER function in (6). Then the problem can be re-expressed as

$$\arg\min_{\{e_{fi}\}} \sum_{f=1}^{F}\sum_{i=1}^{Q_f} u_{fi}a_m \exp(-b_m\gamma\, e_{fi}) \tag{9}$$

$$subject\ to\ -\mathbf{e} \leq 0$$

$$\mathbf{n}^T\mathbf{e} = N, \tag{10}$$

TABLE II
CURVE FITTING EXAMPLES FOR SOME MCSS IN FIG. 9

MCS Index	exponential form	power form
7	$1.3860e^{-0.4330\gamma}$	$10.8\gamma^{-2.637}$
11	$2.2410e^{-0.1771\gamma}$	$53.0\gamma^{-2.228}$
20	$0.3114e^{-0.0085\gamma}$	$515.4\gamma^{-1.726}$

TABLE III
CURVE FITTING ACCURACY COMPARISON FOR SOME MCSS IN FIG. 9

MCS Index	R-square		RMSE	
	exponential	power	exponential	power
7	0.9999	0.9905	6.4810^{-4}	0.6810^{-2}
11	1.0000	0.9851	1.7610^{-4}	1.2410^{-2}
20	0.9970	0.9857	1.7110^{-4}	0.6910^{-2}

where a_m and b_m are curve fitting parameters for the m-th MCS, e_{fi} is the ratio of E_{fi} and the energy of an RB for EEP at $SNR = \gamma$, $\mathbf{e} = [e_{11}\ e_{12} \cdots e_{FQ_F}]^T$, and $\mathbf{n} = [n_{11}\ n_{12} \cdots n_{FQ_F}]^T$. The second constraint in (10) comes from the total energy constraint. Because $\mathbf{n}^T\mathbf{e}$ is the sum of the energy ratios for N RBs, its value is N. The length of $\mathbf{e}$ and $\mathbf{n}$ is $Q = \sum_{f=1}^{F} Q_f$. For sspf, the following constraint is added:

$$e_{f1} \geq \cdots \geq e_{fQ_f}. \tag{11}$$

Applying Lagrangian multipliers $\lambda^* \in \mathbf{R}^Q$ for the inequality constraints and $v^* \in \mathbf{R}$ for the equality constraint, we obtain the following Karush-Kuhn-Tucker (KKT) conditions [32], [33] for (9)-(10) to solve the problem:

$$\begin{aligned}
&1)\ -\mathbf{e}^* \leq 0\\
&2)\ \mathbf{n}^T\mathbf{e}^* - N = 0\\
&3)\ \lambda^*_{fi} \geq 0, \quad f = 1, 2, \cdots, F,\ i = 1, 2, \cdots, Q_f\\
&4)\ \lambda^*_{fi} e^*_{fi} = 0, \quad f = 1, 2, \cdots, F,\ i = 1, 2, \cdots, Q_f\\
&5)\ \nabla f_0(e^*_{fi}) + \lambda^*_{fi}\nabla f_l(e^*_{fi}) + v^*\nabla h_1(e^*_{fi}) = 0,\\
&\quad f = 1, 2, \cdots, F, \quad i = 1, 2, \cdots, Q_f,
\end{aligned} \tag{12}$$

where f_0 is the objective function, $\{f_l\}$ are functions defining inequality constraints, and h_1 is the function defining the equality constraint. In addition, λ^*_l and v^* are optimum Lagrangian multipliers that meet the constraints, and e^*_l is the optimum value of e_l. Each term in 5) of (12) can be calculated as

$$\begin{aligned}
\nabla f_0(e^*_{fi}) &= \frac{\partial}{\partial e^*_{fi}}\left(u_{fi}a_m \exp(-b_m\gamma\, e^*_{fi})\right)\\
&= -u_{fi}a_m b_m\gamma\, \exp(-b_m\gamma\, e^*_{fi})\\
&= -\alpha_{fi}\exp(-b_m\gamma\, e^*_{fi}),
\end{aligned} \tag{13}$$

$$\begin{aligned}
\nabla f_{fi}(e^*_{fi}) &= \frac{\partial}{\partial e^*_{fi}}\left(-\mathbf{e}^*\right)\\
&= -1,
\end{aligned} \tag{14}$$

$$\begin{aligned}
\nabla h_1(e^*_{fi}) &= \frac{\partial}{\partial e^*_{fi}}\left(\mathbf{n}^T\mathbf{e}^* - N\right)\\
&= n_{fi},
\end{aligned} \tag{15}$$

- Step 1: select the MCS m based on EEP
- Step 2: find v^* that minimizes $\sum_{f=1}^{F}\sum_{i=1}^{Q_f} n_{fi}\max\left\{0, \frac{1}{b_m\gamma}\log\left(\frac{\alpha_{fi}}{v^* n_{fi}}\right)\right\} - N$ among candidates values of v^* within the range of $0 < v^* < \max\left\{\frac{\alpha_{fi}}{n_{fi}}\right\}$ by grid search
- Step 3: assign $e_{fi} = \max\left\{0, \frac{1}{b_m\gamma}\log\left(\frac{\alpha_{fi}}{v^* n_{fi}}\right)\right\}$

Fig. 11. The procedures of the proposed energy allocation based on convex optimization.

where $\alpha_{fi} = u_{fi}a_m b_m\gamma$. By applying (13)-(15) to (12)-5), we can obtain λ^*_i as

$$\lambda^*_{fi} = -\alpha_{fi}\exp(-b_m\gamma\, e^*_{fi}) + v^* n_{fi}. \tag{16}$$

Then by plugging (16) into (12)-4), it is re-written as

$$(-\alpha_{fi}\exp(-b_m\gamma\, e^*_{fi}) + v^* n_{fi})e^*_{fi} = 0. \tag{17}$$

There are two solutions of e^*_{fi}: 0 and $\frac{1}{b_m\gamma}\log\left(\frac{\alpha_{fi}}{v^* n_{fi}}\right)$. Because e^*_{fi} is non-negative, we can get the expression for optimum e^*_{fi} as

$$e^*_{fi} = \max\left\{0, \frac{1}{b_m\gamma}\log\left(\frac{\alpha_{fi}}{v^* n_{fi}}\right)\right\}. \tag{18}$$

Because the value of v^* is not fixed yet, we should find a feasible value of v^* that meets the other constraints of (12). Because a_m and b_m are positive, α_{fi} is also positive. In order to have positive e^*_{fi}, $\frac{\alpha_{fi}}{v^* n_{fi}}$ should be larger than 1. So, v^* should be within the range of

$$0 < v^* < \frac{\alpha_{fi}}{n_{fi}}. \tag{19}$$

Because v^* is the optimum Lagrangian multiplier used for all combinations of f and i, and at least one element of $\mathbf{e}^*$ should be a positive value, the range of v^* is the superset of (19) for all i:

$$0 < v^* < \max\left\{\frac{\alpha_{fi}}{n_{fi}}\right\}. \tag{20}$$

Then, finally, we can get v^* to meet the following condition in 5) of (12) from grid search in the range of (16):

$$\sum_{f=1}^{F}\sum_{i=1}^{Q_f} n_{fi}\max\left\{0, \frac{1}{b_m\gamma}\log\left(\frac{\alpha_{fi}}{v^* n_{fi}}\right)\right\} = N. \tag{21}$$

Once v^* is obtained, we can get the final value of e^*_{fi} by substituting v^* in (18) with the obtained value. Note that if the j-th FEC block of the f-th frame has a lower $\frac{\alpha_{fi}}{n_{fi}}$ than v^*, then $e^*_{fj} = 0$. Therefore, according to the results of the energy allocation, only the FEC blocks which have utility values larger than a threshold are assigned energy. Note that all FEC blocks in a frame have identical $\frac{\alpha_{fi}}{n_{fi}}$ values in sspf packetization, and all FEC blocks have identical e^*_{fi} value. Therefore, the constraint in (11) is automatically met.

Fig. 11 shows the procedures of the proposed energy allocation method based on convex optimization.

TABLE IV
CHARACTERISTICS OF VIDEO SEQUENCES AND VIDEO ENCODING OPTIONS USED FOR PERFORMANCE EVALUATION

NAL packetization	**single slice per frame (sspf)**	**multiple slices per frame (mspf)**
Test sequences	Bus(CIF), Foreman (CIF), News (CIF), Akiyo (CIF)	
# of frames	150	
QP	Bus, Foreman: 28/28/30, News, Akiyo: 22/22/24	
GOP structure	I B B P B B P B B (9 frames, open structure)	
Number of slices per frame	1	variable
Number of source bits per slice	variable	fixed (6000 bits)
# of FEC blocks per slice	variable	1
Source block size (max)	6000 bits	6000 bits
Resource allocation methods	adaptive MCS allocation (MCS 10 - 15, 18), adaptive energy allocation (MCS 11)	
Utility metrics used for optimization	- proposed: RNF - reference: VQM, CMSE, EEP	- proposed: RNF, P-VQM - reference: VQM, CMSE, EEP
RNF value	I: 11, P: 8 or 5, B:1	
P-VQM	-	predicted from slice header
Error concealment	Decode the decodable macroblocks and conceal the rest by copying the co-located macroblocks in the nearest available frame	Conceal the whole slice by copying the co-locate macroblocks in the nearest available frame

IV. PERFORMANCE EVALUATIONS AND DISCUSSION

A. Simulation Setup

This section evaluates the performance of the proposed cross-layer resource allocation methods. Table IV shows the characteristics of video sequences and video encoding options used for performance evaluation. Four CIF resolution (352×288) video sequences, *Bus*, *Foreman*, *News* and *Akiyo*, are used for testing. *Bus* and *Foreman* have higher motion activity than *News* and *Akiyo*. They are encoded using H.264/AVC JM 19.0 reference software [34] with an open GOP structure: I B B P B B P B B, as in Fig. 4. Each sequence includes 150 frames at 30 frames/sec. They are encoded with the two NAL packetization scenarios discussed in II-A: sspf and mspf. For sspf, all source bits corresponding to one frame are included in a single NAL unit (slice) and divided into multiple FEC blocks, each with 6000 source bits. For mspf, each slice has at most 6000 source bits, and each NAL unit is fit into a single FEC block. In other words, the source block size is 6000 for both encoding scenarios. The quantization parameters for *Bus* and *Foreman* are 28 for I/P-frames and 30 for B-frames. For *News* and *Akiyo*, they are 22 for I/P-frames and 24 for B-frames. The proposed heuristic adaptive MCS allocation and the adaptive energy allocation method based on convex optimization are applied to both NAL packetization scenarios. RNF and P-VQM are used as the utility metrics for mspf. Only RNF is used for sspf. In order to show the efficiency and performance of the proposed metrics, VQM, CMSE and EEP are used as reference. EEP means the same BLER is set for all the FEC blocks, i.e., the same MCS and the same energy are set for each FEC block. RNF values are determined only by the GOP structure, so they are the same for every GOP. In all the simulations, we consider the GOP structure in Fig. 4, and the corresponding RNF values are 11 for the I-frame, 8 for the first P-frame in the GOP, 5 for the second P-frame, and 1 for B-frames. P-VQM values are predicted using the information in the slice header, and they have different values for each GOP. Lost FEC blocks are concealed by copying the co-located macroblocks in the nearest available frame. Note that for sspf, if an FEC block is lost, all the source bits of this frame afterward cannot be decoded. We first decode the decodable FEC blocks in this frame, then conceal the rest by copying. For mspf, we conceal the whole lost slice.

TABLE V
EXPECTED BLER OF THE LTE SYSTEM FOR ETU70 CHANNEL ACCORDING TO THE VARIATION OF THE NUMBER OF RBS USED FOR SIMULATION OF ADAPTIVE MCS ALLOCATION

# of RBs	18	21	23	26	30	34	38
MCS	18	15	14	13	12	11	10
11dB	82%	65%	50%	39%	29%	19%	10%
12dB	72%	55%	40%	30%	20%	12%	5%
13dB	62%	45%	30%	20%	11%	7%	1.8%
14dB	57%	35%	20%	13%	6.5%	3.5%	0.3%
15dB	45%	25%	15%	7.5%	2%	1%	0.1%
16dB	35%	17%	8%	4%	0.4%	0.2%	0.02%

The performance of the proposed adaptive resource allocation is evaluated for the LTE system. The multipath fading channel is defined by a combination of a multipath delay profile and a maximum Doppler frequency. Annex B of the 3GPP technical specification [27] describes the multipath fading propagation conditions, including taps of the Extended Pedestrian A model (EPA), Extended Vehicular A model (EVA), and Extended Typical Urban model (ETU), as well as classical Doppler spectrum. In this work, we evaluate the performance for the ETU70 model (Doppler frequency = 70Hz).

In the case of adaptive MCS allocation, the expected BLERs according to the variation of n_i are known for the given channel conditions. Table V shows the BLER according to the variation of n_i for source block size of 6000 bits for the LTE system with two-layer spatial multiplexing in the ETU 70 channel. The MCSs and required number of RBs for the corresponding MCSs can be found from Table 7.1.7.1-1 and Table 7.1.7.2.1-1 of [28], and expected BLER values are obtained from the BLER curves corresponding to MCSs in the ETU 70 channel, which are shown in Fig. 9. For each GOP, we estimate the expected BLER for each FEC block. For example, if 1% BLER

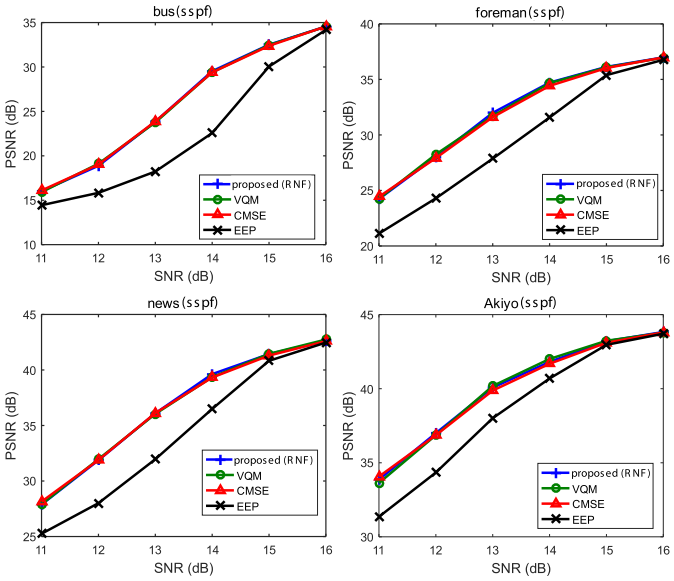


Fig. 12. PSNR comparison for adaptive MCS allocation methods with different utilities (sspf).

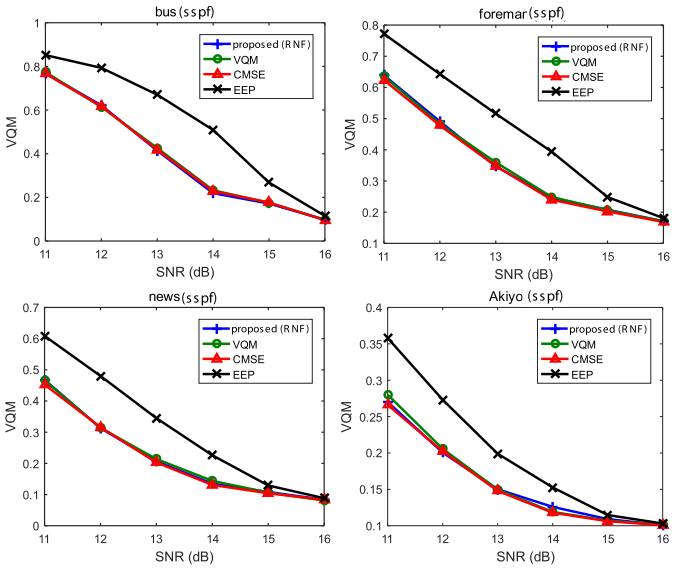


Fig. 13. VQM comparison for adaptive MCS allocation methods with different utilities (sspf).

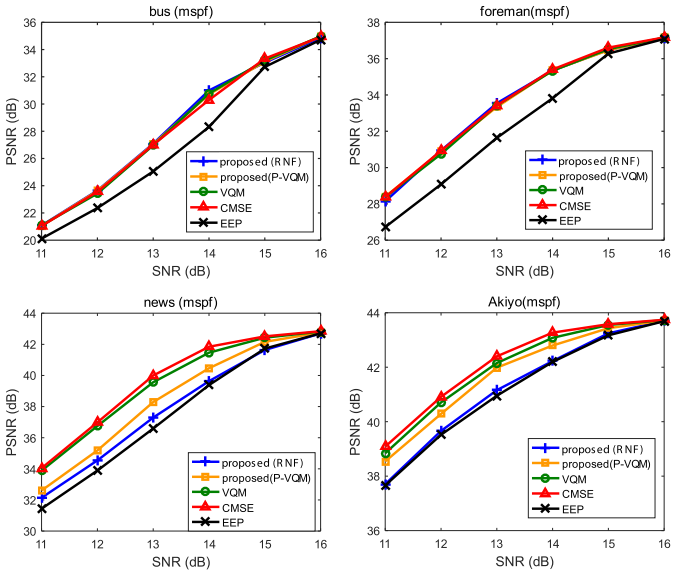


Fig. 14. PSNR comparison for adaptive MCS allocation methods with different utilities (mspf).

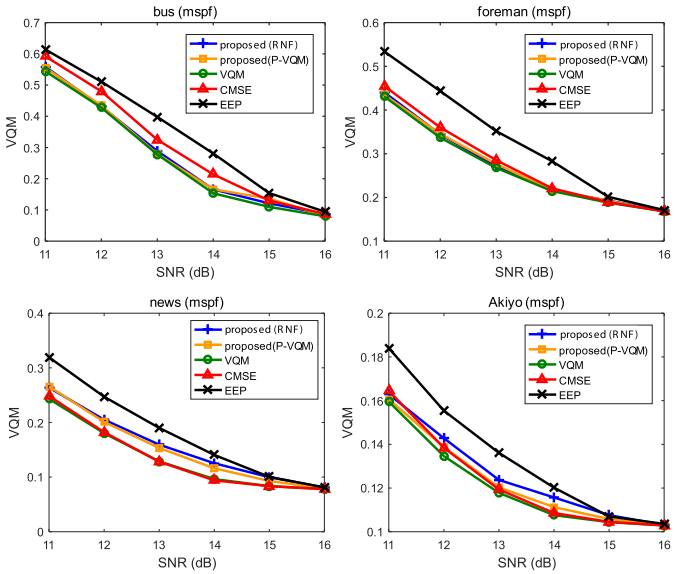


Fig. 15. VQM comparison for adaptive MCS allocation methods with different utilities (mspf).

is targeted at 15dB SNR for EEP, MCS 11 is selected and 34 RBs should be assigned to each FEC block. According to the result of the UEP based adaptive MCS allocation, if 38 RBs are assigned to an FEC block with higher utility value, 0.1% BLER is expected for it.

In the case of adaptive energy allocation, we have an approximate BLER function of the MCS 11 as described in III-*C-2)*. We obtain the expected BLER with different allocated energy. In the simulation, we assume MCS 11 was selected for all FEC blocks and 34 RBs are assigned to an FEC block with 6000 bits. If a GOP has 50 FEC blocks, the total number of available RBs is 1700.

B. Simulation Results and Discussion

Figs. 12 - 15 show the performance of the adaptive MCS allocation for different utility metrics. Figs. 12 and 13 are results for sspf packetization. For all video sequences, the proposed RNF utility shows almost the same PSNR and VQM as utilizing VQM and CMSE as utility. As summarized in Tables VI and VII, the average PSNR gain over EEP at an SNR of 13dB is 3.98dB, and the average VQM gain over EEP is 0.153.

Figs. 14 and 15 are adaptive MCS allocation results for mspf packetization. As summarized in Tables VI and VII, P-VQM has an average PSNR gain 1.45 dB and an average VQM gain of 0.061 over EEP at an SNR of 13 dB. The average PSNR and VQM gains are 1.21dB and 0.058 for RNF. P-VQM and RNF perform almost the same as CMSE and VQM for the high motion sequences Bus and Foreman. They have some loss to CMSE and VQM for News and Akiyo, but still outperform EEP.

TABLE VI

PSNR GAIN OVER EEP (DB) (a), AND PSNR GAIN ACHIEVEMENT RATIO (b) AT SNR = 13dB

sequence	sspf (MCS)	sspf (energy)	mspf (MCS)		mspf (energy)	
			RNF	P-VQM	RNF	P-VQM
Bus	**5.61**	**2.63**	**2.04**	**2.04**	**1.15**	**1.23**
Foreman	**4.13**	**1.94**	**1.91**	**1.70**	**1.12**	**1.32**
News	**4.13**	**1.64**	0.70	**1.02**	0.13	**1.02**
Akiyo	**2.06**	**1.55**	0.20	**1.03**	0.86	**1.18**
average	**3.98**	**1.94**	**1.21**	**1.45**	0.81	**1.19**

(a)

sequence	sspf (MCS)	sspf (energy)	mspf (MCS)		mspf (energy)	
			RNF	P-VQM	RNF	P-VQM
Bus	**99%**	**99%**	**100%**	**100%**	**84%**	**91%**
Foreman	**100%**	**100%**	**100%**	**95%**	**87%**	**100%**
News	**99%**	**93%**	21%	37%	6%	47%
Akiyo	**95%**	**100%**	14%	**71%**	50%	69%
average	**98%**	**98%**	59%	**76%**	57%	**77%**

(b)

TABLE VII

VQM GAIN OVER EEP (a), AND VQM GAIN ACHIEVEMENT RATIO (b) AT SNR = 13dB

sequence	sspf (MCS)	sspf (energy)	mspf (MCS)		mspf (energy)	
			RNF	P-VQM	RNF	P-VQM
Bus	**0.256**	**0.121**	**0.110**	**0.118**	**0.065**	**0.057**
Foreman	**0.169**	**0.059**	**0.081**	**0.075**	0.024	0.025
News	**0.138**	0.043	0.031	0.037	0.007	0.015
Akiyo	0.049	0.018	0.012	0.016	0.004	0.004
average	**0.153**	**0.060**	**0.058**	**0.061**	0.025	0.025

(a)

sequence	sspf (MCS)	sspf (energy)	mspf (MCS)		mspf (energy)	
			RNF	P-VQM	RNF	P-VQM
Bus	**100%**	**100%**	**91%**	**98%**	**80%**	**70%**
Foreman	**93%**	68%	**96%**	**89%**	**90%**	**93%**
News	**97%**	**76%**	56%	59%	17%	36%
Akiyo	**95%**	**73%**	68%	**86%**	37%	41%
average	**96%**	**79%**	**78%**	**83%**	56%	60%

(b)

The different performances of the proposed utilities for different videos are due to motion in the video. For *Bus* and *Foreman*, almost all the macroblocks include motion information. Therefore, slices in I-frames have similar VQM scores, i.e., similar impacts on the overall quality, if they are lost. We plot the VQM scores of two GOPs of sequence *Bus* in Fig. 16. Using RNF, these slices are given the same priority that matches the VQM scores. That accounts for the similar performance of VQM, CMSE and RNF. For *News* and *Akiyo*, however, slices in the same frame actually have very different impacts on the overall quality. Static slices which contain mostly background can be concealed very well by copying the co-located blocks in the closest available frame, while slices with moving objects are hard to conceal. Using RNF, these slices in the same frame are given the same priority, which is not consistent with the VQM scores. We plot the VQM scores and RNF of *Akiyo* in Fig. 17. Another reason is that UEP allocates fewer RBs to the FEC blocks of B-frames. The saved RBs are assigned to FEC blocks of the I-frame.

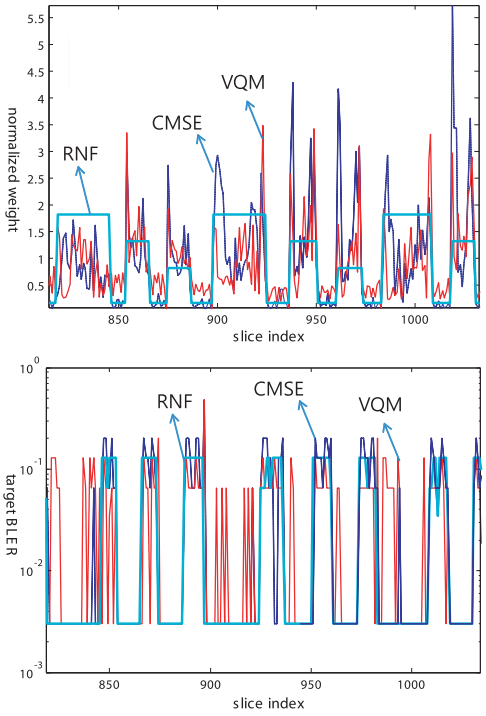


Fig. 16. Variations of utility metrics and target BLER according to adaptive MCS allocation for *Bus*.

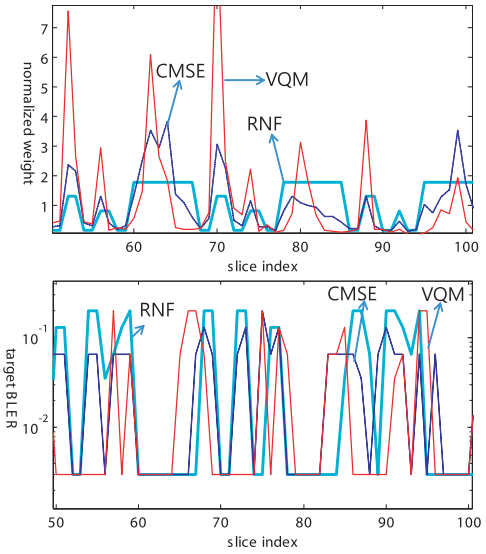


Fig. 17. Variations of utility metrics and target BLER according to adaptive MCS allocation for *Akiyo*.

For video sequences with low motion, the size of slices of the B-frame is small, so there are not many RBs that can be saved for the I-frame. These two reasons account for the performance loss of RNF over VQM and CMSE based utilities.

P-VQM performs as well as RNF for *Bus* and *Foreman*, and performs better than RNF for *News* and *Akiyo*. Unlike RNF, P-VQM treats slices in the same frame differently. Some slices are given higher priority. For sequence Akiyo, P-VQM performs almost the same as VQM and CMSE. The performance loss is due to inaccurate prediction. Note that P-VQM is computationally much simpler than VQM or CMSE.

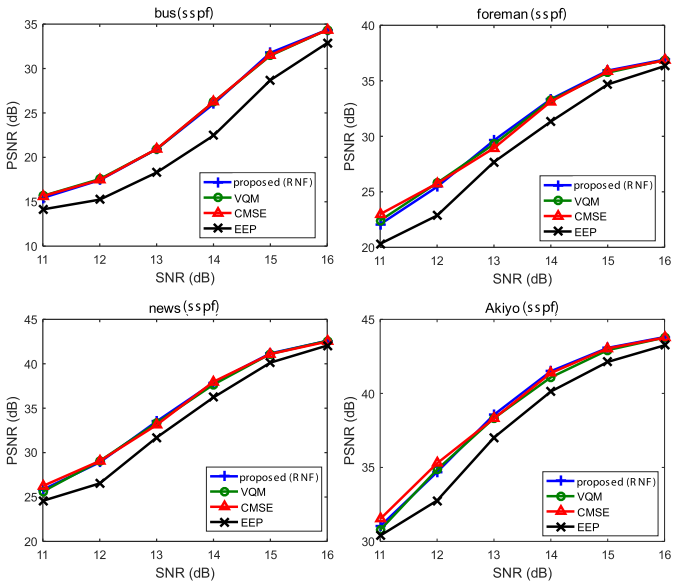


Fig. 18. PSNR comparison for adaptive energy allocation methods with different utilities (sspf).

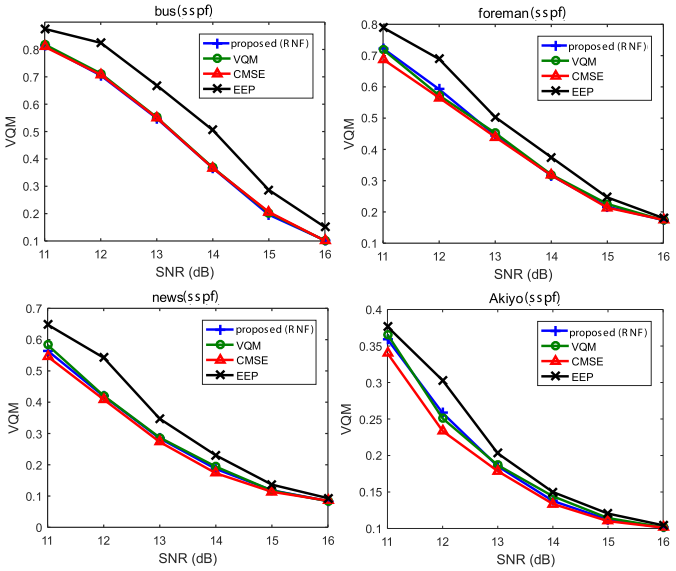


Fig. 19. VQM comparison for adaptive energy allocation methods with different utilities (sspf).

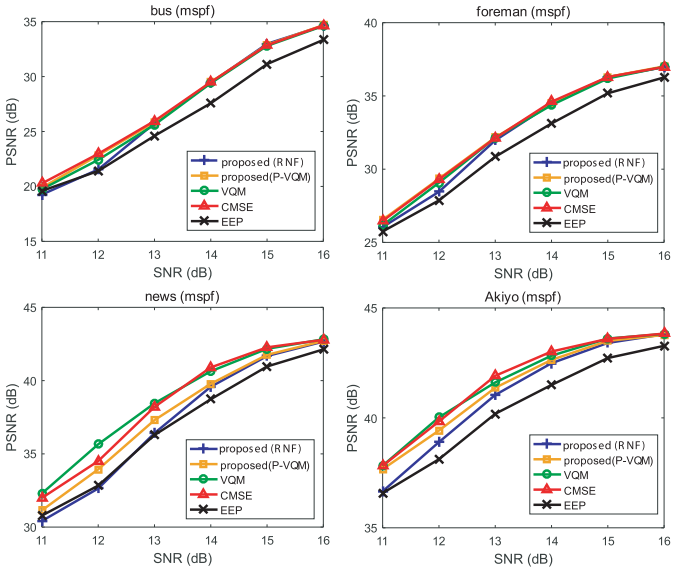


Fig. 20. PSNR comparison for adaptive energy allocation methods with different utilities (mspf).

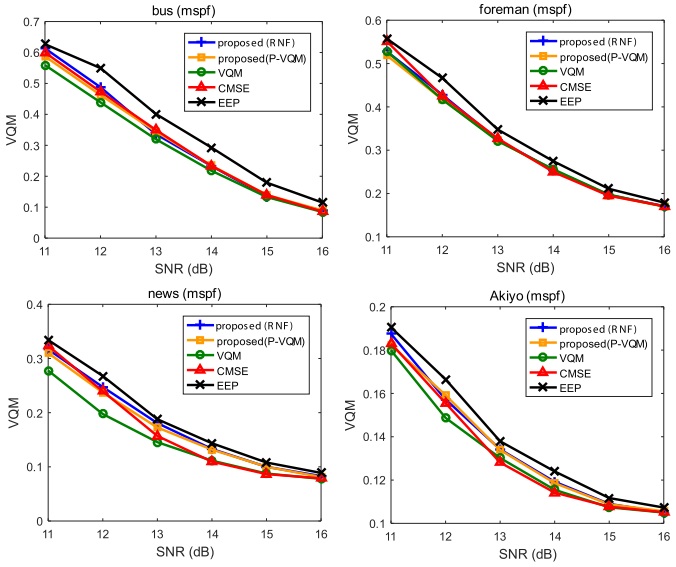


Fig. 21. VQM comparison for adaptive energy allocation methods with different utilities (mspf).

Figs. 18 - 21 show the performance of the adaptive energy allocation for different utility metrics. Figs. 18 and 19 are the results of sspf packetization. For all video sequences, the proposed RNF utility shows meaningful PSNR and VQM gains which are very similar to the performance of VQM and CMSE. As summarized in Tables VI and VII, the average PSNR gain over EEP at 13dB SNR is 1.94dB, and the average VQM gain over EEP is 0.06. The PSNR gain ratio, which is the ratio between the PSNR gain of the proposed utility metric and the maximum PSNR gain, is compared in Table VI-(b). We define the maximum PSNR gain as the larger of the PSNR gain of the ideal VQM utility metric over EEP and the PSNR gain of the CMSE utility metric over EEP for the given test video sequence and encoding parameters. An identical value of PSNR gain achievement ratio which is 98% can be achieved for both adaptive MCS allocation and adaptive energy allocation. In Table VII-(b), the VQM gain achievement ratio which is the ratio between the VQM gain of the proposed utility metric and the maximum VQM gain is compared. The maximum VQM gain is the larger value of VQM gain between the VQM gains of the ideal VQM utility metric and CMSE utility metric over EEP for the given test video sequence and encoding parameters. In contrast to the PSNR gain achievement ratio, the VQM gain achievement ratio of adaptive energy allocation is 17% less than that of adaptive MCS allocation.

Figs. 20 and 21 compare the PSNR and VQM of adaptive energy allocation for different utility metrics in the mspf

scenario. The average PSNR gain of RNF is 0.81 dB and the average PSNR gain of P-VQM is 1.19dB which correspond to 57% and 77% of VQM gain achievement ratio. The average VQM gain of both RNF and P-VQM is 0.025, but P-VQM can achieve a meaningful improvement in VQM gain for *News* and *Akiyo* compared to RNF.

According to the results of the simulations, we can see the effectiveness of the proposed resource allocation methods and new utility metrics. The utility metric-based resource allocation methods show meaningful PSNR and VQM gains over EEP for several video encoding options and video sequences with different motion activity. The most important merit in applying the proposed utility metrics is they can be obtained from the current header information and there is no additional feedback overhead or heavy computation.

V. Conclusion

We proposed cross-layer optimized adaptive MCS allocation and adaptive energy allocation for wireless video transmission. We considered two packetization scenarios in both problems. The proposed metrics, RNF and P-VQM, require no additional feedback overhead or heavy computation, and utilize only the information in the slice headers. They can be easily applied to various wireless communication systems. The UEP using the proposed metrics has much better performance than EEP, and similar performance to UEP using CMSE and VQM. However, the computational complexity is greatly reduced compared to CMSE and VQM. Another contribution of this paper is that we proposed practical cross layer resource allocation methods which can handle the granular RB structure that fits with real wireless communication systems, i.e., 3GPP LTE. Future research includes UEP resource allocation for hybrid automatic repeat request transmission, and multicast video packets.

References

[1] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[2] F. M. S. Cho and W. A. Pearlman, "Robust image transmission using a new joint source channel coding algorithm and dual adaptive OFDM," in *Proc. GlobeCom*, Jan. 2004, p. 636.

[3] S. S. Arslan, P. C. Cosman, and L. B. Milstein, "Progressive source transmissions using joint source-channel coding and hierarchical modulation in packetized networks," in *Proc. GlobeCom*, Nov. 2009, pp. 1–6.

[4] L. Toni, Y. S. Chan, P. C. Cosman, and L. B. Milstein, "Channel coding for progressive images in a 2-D time-frequency OFDM block with channel estimation errors," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2476–2490, Nov. 2009.

[5] S. S. Arslan, P. C. Cosman, and L. B. Milstein, "Concatenated block codes for unequal error protection of embedded bit streams," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1111–1122, Mar. 2012.

[6] H. Ha and C. Yim, "Layer-weighted unequal error protection for scalable video coding extension of H.264/AVC," *IEEE Trans. Consum. Electron.*, vol. 54, no. 2, pp. 736–744, May 2008.

[7] Y. Huo, M. El-Hajjar, R. G. Maunder, and L. Hanzo, "Layered wireless video relying on minimum-distortion inter-layer FEC coding," *IEEE Trans. Multimedia*, vol. 16, no. 3, pp. 697–710, Apr. 2014.

[8] M. Zhao, X. Gong, J. Liang, W. Wang, X. Que, and S. Cheng, "QoE-driven cross-layer optimization for wireless dynamic adaptive streaming of scalable videos over HTTP," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 3, pp. 451–465, Mar. 2015.

[9] Y. Huo, M. El-Hajjar, and L. Hanzo, "Inter-layer FEC aided unequal error protection for multilayer video transmission in mobile TV," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 9, pp. 1622–1634, Sep. 2013.

[10] S. Thakolsri, W. Kellerer, and E. Steinbach, "QoE-based cross-layer optimization of wireless video with unperceivable temporal video quality fluctuation," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2011, pp. 1–6.

[11] L. Zhou, Z. Yang, Y. Wen, H. Wang, and M. Guizani, "Resource allocation with incomplete information for QoE-driven multimedia communications," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 3733–3745, Aug. 2013.

[12] X. Yang *et al.*, "Unequal loss protection for robust transmission of motion compensated video over the Internet," *Signal Process. Image Commun.*, vol. 18, no. 3, pp. 157–167, Mar. 2003.

[13] T.-L. Lin and P. C. Cosman, "Efficient optimal RCPC code rate allocation with packet discarding for pre-encoded compressed video," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 505–508, May 2010.

[14] K. K. R. Kambhatla, S. Kumar, S. Paluri, and P. C. Cosman, "Wireless H.264 video quality enhancement through optimal prioritized packet fragmentation," *IEEE Trans. Multimedia*, vol. 14, no. 5, pp. 1480–1495, Oct. 2012.

[15] L. He and G. Liu, "Quality-driven cross-layer design for H.264/AVC video transmission over OFDMA system," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 6768–6782, Dec. 2014.

[16] E. Danish, A. Fernando, O. Abdul-Hameed, and M. Pourazad, "Perceptual quality-driven resource allocation in energy-aware wireless video multicasting," in *Proc. 10th Int. Conf. Heterogeneous Netw. Quality, Rel., Secur. Robustness (QShine)*, Aug. 2014, pp. 101–107.

[17] Y.-L. Chang, T.-L. Lin, and P. C. Cosman, "Network-based H.264/AVC whole-frame loss visibility model and frame dropping methods," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3353–3363, Aug. 2012.

[18] Z. Luo, L. Song, S. Zheng, and N. Ling, "Raptor codes based unequal protection for compressed video according to packet priority," *IEEE Trans. Multimedia*, vol. 15, no. 8, pp. 2208–2213, Dec. 2013.

[19] Y.-L. Chang, T.-L. Lin, and P. C. Cosman, "Network-based IP packet loss importance model for H.264 SD videos," in *Proc. IEEE 18th Int. Packet Video Workshop*, Dec. 2010, pp. 178–185.

[20] L. Toni, P. C. Cosman, and L. B. Milstein, "Channel coding optimization based on slice visibility for transmission of compressed video over OFDM channels," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 7, pp. 1172–1183, Aug. 2012.

[21] Y. Wu, S. Kumar, F. Hu, Y. Zhu, and J. D. Matyjas, "Cross-layer forward error correction scheme using raptor and RCPC codes for prioritized video transmission over wireless channels," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 6, pp. 1047–1060, Jun. 2014.

[22] *Cisco Visual Networking Index: Forecast and Methodology 2015–2020*, CISCO Inc. San Jose, CA, USA, 2016.

[23] A. E. Essaili, D. Schroeder, E. Steinbach, D. Staehle, and M. Shehada, "Quality-driven cross-layer optimized video delivery over LTE," *IEEE Commun. Mag.*, vol. 48, no. 2, pp. 102–109, Feb. 2010.

[24] S. Karachontzitis, T. Dagiuklas, and L. Dounis, "Novel cross-layer scheme for video transmission over LTE-based wireless systems," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2011, pp. 1–6.

[25] A. E. Essaili, D. Schroeder, E. Steinbach, D. Staehle, and M. Shehada, "QoE-based traffic and resource management for adaptive HTTP video delivery in LTE," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 6, pp. 988–1001, Jun. 2014.

[26] J. Yang, Y. Ran, S. Chen, W. Li, and L. Hanzo, "Online source rate control for adaptive video streaming over HSPA and LTE-style variable bit rate downlink channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 2, pp. 643–657, Feb. 2016.

[27] *User Equipment (UE) Radio Transmission and Reception, V11.0.0*, document TS 36.101, 3GPP, Mar. 2012.

[28] *Physical Layer Procedures, V10.5.0*, document TS 36.213, 3GPP, Mar. 2012.

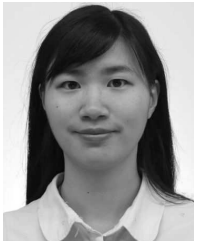
[29] (May 2014). W. B. Yang and M. Souryal. *NISTIR 7986: LTE Physical Layer Performance Analysis*. [Online]. Available: http://dx.doi.org/10.6028/NIST.IR.7986

[30] (2016). MathWorks Inc. *Curve Fitting Toolbox*. [Online]. Available: http://www.mathworks.com/products/curvefitting/

[31] (2016). MathWorks Inc. *Evaluating Goodness of Fit*. [Online]. Available: http://www.mathworks.com/help/curvefit/evaluating-goodness-of-fit.html

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
[33] D. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 2008.
[34] (2015). *H.264/AVC Reference Software JM19.0*. [Online]. Available: http://iphome.hhi.de/suehring/tml/download/

Young-Ho Jung (S'98–M'04) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 1998, 2000, and 2004, respectively. From 2004 to 2007, he was with the Samsung Advanced Institute of Technology and Samsung Electronics, South Korea. He was a Visiting Scholar with the Department of Electrical Engineering, Stanford University, in 2005. From 2015 to 2016, he was a Visiting Scholar with the Department of Electrical Engineering, University of California at San Diego. Since 2007, he has been with the School of Electronics and Information Engineering, Korea Aerospace University, Goyang, South Korea, where he is currently an Associate Professor. His primary research interests include cross-layer optimization, signal processing for communication systems and radar systems, and communication systems for unmanned aerial vehicles.



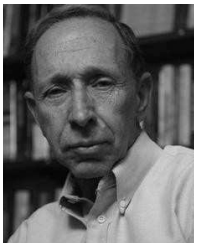
Qing Song (S'15) received the B.Eng. degree in automation from Tongji University, Shanghai, China, in 2011, and the M.S. degree in electrical engineering from the University of California at San Diego, La Jolla, CA, USA, in 2013, where she is currently pursuing the Ph.D. degree. Her industry experience includes internships at Amazon Lab126, Cupertino, CA, USA, in 2013, and Dolby Labs, Inc., Sunnyvale, CA, USA, in 2014 and 2016. Her research interests include image/video processing, compression, and transmission.

Kyung-Ho Kim received the B.S. and M.S. degrees in electronics and information engineering from Korea Aerospace University, Goyang, South Korea, in 2014 and 2017, respectively. Since 2017, he has been with Korea Aerospace Industries, Ltd., Daejeon, South Korea. His research interests include 5G communication systems, communication systems for unmanned aerial vehicles, and satellite communication systems.

Pamela Cosman (S'88–M'93–SM'00–F'08) received the B.S. degree (Hons.) in electrical engineering from the California Institute of Technology in 1987 and the Ph.D. degree in electrical engineering from Stanford University in 1993. She was an NSF Post-Doctoral Fellow with Stanford University and the University of Minnesota from 1993 to 1995. She joined the Department of Electrical and Computer Engineering, University of California at San Diego, as a Faculty Member, where she is currently a Professor.

Her research interests are in the areas of image and video compression and processing, and wireless communications. She has authored over 250 technical papers in these fields, and one children's fiction book that includes error correction coding. She received numerous awards include the ECE Departmental Graduate Teaching Award, the Career Award from the National Science Foundation, the Powell Faculty Fellowship, the Globecom 2008 Best Paper Award, the HISB 2012 Best Poster Award, and the 2016 UC San Diego Affirmative Action and Diversity Award. Her administrative positions include serving as the Director of the Center for Wireless Communications from 2006 to 2008, the ECE Department Vice-Chair from 2011 to 2014, and an Associate Dean for Students from 2013 to 2016.

Dr. Cosman is a member of Tau Beta Pi and Sigma Xi. She was a Guest Editor of the June 2000 Special Issue of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS on Error-Resilient Image and Video Coding and the Technical Program Chair of the 1998 Information Theory Workshop in San Diego. She has been a member of the Technical Program Committee or the Organizing Committee for numerous conferences, including ICIP 2008-2011, QOMEX 2010-2012, ICME 2011-2013, VCIP 2010, PacketVideo 2007–2013, WPMC 2006, ICISP 2003, ACIVS 2002-2012, ICC 2012, Asilomar Conference on Signals, Systems and Computers 2003, and EUSIPCO 1998. She was an Associate Editor of the IEEE COMMUNICATIONS LETTERS from 1998 to 2001 and an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS from 2001 to 2005. She was the Editor-in-Chief from 2006 to 2009, and a Senior Editor from 2003 to 2005 and 2010 to 2013 of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.



Laurence B. Milstein (S'66–M'68–SM'77–F'85) received the B.E.E. degree from the City College of New York, New York City, NY, USA, in 1964, and the M.S. and Ph.D. degrees in electrical engineering from the Polytechnic Institute of Brooklyn, Brooklyn, NY, USA, in 1966 and 1968, respectively.

From 1968 to 1974, he was with the Space and Communications Group, Hughes Aircraft Company, and from 1974 to 1976, he was a member of the Department of Electrical and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA. Since 1976, he has been with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA, USA, where he is currently the Ericsson Professor of Wireless Communications Access Techniques and also the former Department Chairman, involved in the area of digital communication theory with special emphasis on spread-spectrum communication systems. He has also been a consultant to government and industry in the areas of radar and communications.

Dr. Milstein was a recipient of the 1998 Military Communications Conference Long Term Technical Achievement Award, the Academic Senate 1999 UCSD Distinguished Teaching Award, the IEEE Third Millennium Medal in 2000, the 2000 IEEE Communications Society Armstrong Technical Achievement Award, and various prize paper awards. He was also a recipient of the IEEE Communications Theory Technical Committee (CTTC) Service Award in 2009 and the CTTC Achievement Award in 2012. In 2015, he received the UCSD Chancellors Associates Award for Excellence in Graduate Teaching. He was an Associate Editor of the Communication Theory for the IEEE TRANSACTIONS ON COMMUNICATIONS, an Associate Editor of Book Reviews for the IEEE TRANSACTIONS ON INFORMATION THEORY, an Associate Technical Editor of the *IEEE Communications Magazine*, and the Editor-in-Chief of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He has been a member of the Board of Governors of the IEEE Communications Society and the IEEE Information Theory Society. He was the Vice President for Technical Affairs in 1990 and 1991 of the IEEE Communications Society. He was a former Chair of the IEEE Fellows Selection Committee.