Transmit Power Aware Cross-Layer Optimization for LTE Uplink Video Streaming

Pinghua Zhao[1,2], Yanwei Liu[1,*], Jinxia Liu[3], Ruixiao Yao[1], Song Ci[2,4] and Antonios Argyriou[5]
[1]State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, 100093 Beijing, China
[2]Institute of Acoustics, Chinese Academy of Sciences, Beijing, China
[3]Zhejiang Wanli University, Ningbo, China
[4]University of Nebraska-Lincoln, Omaha, USA
[5]University of Thessaly, Volos, Greece

***Abstract*—The rapid developments of advanced wireless communication technologies and mobile devices are boosting the uplink multimedia applications. In this paper, a transmit power aware cross-layer optimization scheme is proposed to achieve a good trade-off between the transmit power and the perceived video quality for Long Term Evolution uplink video streaming. Specifically, the video coding quantization parameter and encoding mode at the application layer, and the uplink transmit power as well as modulation and coding scheme at the physical layer are jointly adjusted in the cross-layer optimization. To further improve the perceptual video experience for the end user with limited transmission resources, unequal quality control is performed by enhancing the video quality of region of interest. Additionally, the structural similarity is adopted as the video quality measurement metric to make the optimized video properly preserve the structural information during the cross-layer optimization process. Experimental results show that significant performance improvements in terms of the transmit power reduction and the perceptual video quality are achieved for the proposed transmit power aware cross-layer optimization scheme.**

***Index Terms*—Cross-layer design, transmit power aware, LTE uplink, perceptual video quality**

I. Introduction

Wireless mobile devices today like smartphones, tablets, and laptops are becoming even more powerful computationally. They are capable of receiving, rendering and also capturing and transmitting high quality videos. These recent advances have actually resulted in a reality that wants wireless video to be everywhere around us. An emerging particular class of uplink wireless video applications, such as video conferencing, video surveillance, video sharing for social web platforms, and so on, involve real-time capturing, encoding and transmission. Even though the transmission capacity of the Long Term Evolution (LTE) system has been improved significantly, for the aforementioned uplink applications it is still very challenging to not only get the optimal but even maintain a satisfactory visual experience for the end user. Why the above is the case can be mainly attributed to three reasons.

This work was supported in part by NSFC under Grant nos.61102077 and 61472388, National Key Technology R&D Program 2012BAH01B03, Zhejiang Provincial Natural Science Foundation of China under Contract LY13F010012, Public welfare project of Zhejiang Province under Contract 2014C31072.

*Corresponding author.

Firstly, the transmission resources allocated to the uplink physical layer are less than those of the downlink physical layer for both the Frequency Division Duplex (FDD) and Time Division Duplex (TDD) modes in LTE network. The throughput provided by the uplink physical layer can usually not afford the bandwidth demand of the high quality real-time uplink video streaming. Additionally, the battery-powered User Equipment (UE) is always energy-limited for long-term uplink video streaming. Too much transmit power not only results in the quick energy consumption but also makes people feel uncomfortable due to the radiation. Lastly, the wireless channel is usually characterized by the well-known impairments such as fading, path loss, and shadowing that contribute to bit errors and eventually lead to packet losses. Even worse, the loss of a single video packet may result in the drifting decoding distortion (error propagation) of the subsequent video data which refers to the previous incorrectly decoded video due to the high dependencies of the compressed video data.

Taking advantage of low Peak-to-Average Power Ratio (PAPR) property, Signal Carrier FDMA (SC-FDMA) rather than Orthogonal FDMA (OFDMA) is adopted in LTE uplink physical layer because it allows a lower power consumption for UE. Meanwhile, to balance the need for maintaining the link quality corresponding to the required Quality of Service (QoS) against the need to maximize UE's battery life, the uplink transmit Power Control (PC) is performed. The transmit power should be dynamically adapted to the characteristics of the radio propagation channel, such as path loss, penetration degradation, shadowing, fast fading, and so on. Specifically, the Fractional Power Control (FPC) [1] is adopted by 3GPP for LTE uplink transmission. The performance evaluation of FPC has been well investigated in the literatures [2]–[4]. However, these works only analyze the LTE networking performances including the cell edge and the overall throughput, but rarely consider the required user experience quality for the specific application during the power control process.

Actually, there have been a lot of works focusing on improving the end-to-end video quality for video streaming over LTE network. The cross-layer design [5]–[8] which jointly allocates the resources over each communication protocol layer has been recognized as an efficient way to achieve this purpose.

In [5], the system throughput, application quality of service constraint and scheduling fairness were jointly considered by the designed cross-layer optimization framework for the video delivery over LTE. In [6], a cross-layer optimization scheme which dynamically adjusted the overall throughput to meet the actual available bandwidth was proposed to improve the video quality for the m-Health SVC video streaming over LTE uplink. To further improve the perceptual video experience for the end user in LTE video streaming, the Quality of Experience (QoE) was introduced into the cross-layer optimization scheme [7], [8].

Though the previous works can improve the LTE video streaming quality at some degree, there are still several problems to be studied for the LTE uplink video streaming. Firstly, the long-term video data transmission usually costs too much energy for the battery-powered UE. How to achieve a good trade-off between the required video experience for the video streaming and the battery life is rarely investigated in the previous works. Secondly, the previous distortion metrics, including Mean Squared Error (MSE) and Sum of Squared Error (SSE), have been shown to be not always suitable to quantify the visual distortion for both the image and video [9]. Even though the QoE can properly describe the perceived video quality, it is very difficult and time-consuming to be modeled.

In this paper, a transmit power aware cross-layer optimization scheme is proposed to improve the perceptual video quality for the LTE uplink video streaming. The LTE uplink adaption including the selection of Modulation and Coding Scheme (MCS) and the uplink transmit power control, and the video data adjustment at the application layer including the compressed video data rate and the encoding mode for each encoding unit (MacroBlock, MB), are jointly performed by the proposed cross-layer optimization scheme according to the time-varying channel states. The contributions of this paper are twofold. Firstly, the transmit energy consumption is well considered during the uplink video streaming process. By achieving the minimal perceptual video distortion under the constraint of transmit power budget, the proposed cross-layer optimization can achieve the purpose that balances the need for maintaining the required end-to-end video quality against the need to maximize the battery life. Secondly, the video quality of RoI is enhanced by allocating more bits to the RoI area. The adopted RoI-based distortion metric in which the Structural SIMilarity (SSIM) [10] is involved can make the transmitted video properly preserve structural information. Correspondingly, the perceptual experience of perceived video for the end user is improved.

II. Transmit power aware cross-layer optimization

A. The proposed cross-layer optimization framework

For the real-time LTE uplink video streaming, the data transmission from UE to Evolved Node B (eNodeB) has been known to be the communication bottleneck due to the limited capability of the mobile UE and the unequal resource assignment between the uplink and downlink channel. Thus, in this work we mainly focus on improving the quality of video streaming over LTE uplink channel from UE to eNodeB. As shown in Fig. 1, the transmission link adaption and the compressed video data adjustment are jointly considered to achieve a good trade-off between the decoding video quality and the uplink transmit power. The link adaption including the adjustment of MCS mode and transmit power at the uplink physical layer are performed to adapt to the time-varying channel states. Accordingly, the video bit rate is dynamically tuned to match the channel bandwidth by regulating the video coding quantization parameter. In addition, the optimal error-resilient encoding mode for each MB is selected to suppress the video quality degradation caused by the propagated error.

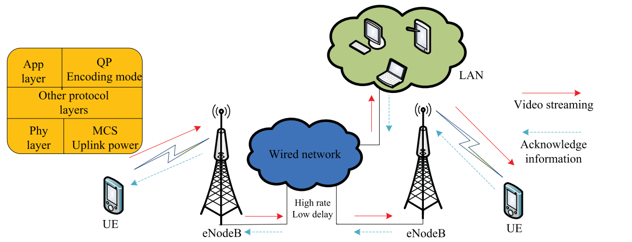


Fig. 1. Transmit power aware cross-layer optimization framework for LTE uplink video streaming.

1) Link adaption: The wireless transmission channel is characterized with time-varying signal degradations and noises. Different MCS modes can result in different levels of transmission payload and reliability. For the LTE uplink transmission, 15 MCS modes in Table I are configured to perform the link adaption process. In practice, the MCS for specific UE is selected by the eNodeB with the help of channel sounding, for example the Sounding Reference Signals (SRS). It has been known that the available spectrum is divided into some individual Resource Blocks (RBs) based on the time and frequency domains at the LTE physical layer. Each RB occupies the duration of one slot (0.5ms) in the time domain and 12 subcarriers (180kHz) in the frequency domain. Usually, the MCS mode is determined by maintaining the Block Error Rate (BLER) of each RB smaller than 10% for the traditional LTE uplink channel adaption.

It has been verified that the BLER $BLER(SINR(m))$ for the RB with the MCS mode m can be predicted as [11]

$$BLER(SINR(m) = \frac{1}{2}erfc(\frac{SINR(m) - b(m)}{\sqrt{2} \cdot c(m)}), \quad (1)$$

where $erfc(\cdot)$ is the complementary error function, $SINR(m)$ is the Signal Interference Noise Ratio (SINR) of the channel, $b(m)$ and $c(m)$ are the “transition center” and “transition width”, respectively. The values of $b(m)$ and $c(m)$ can be obtained by fitting (1) to the exact BLER in the specific communication system. In this work, the AWGN LTE uplink channel is simulated using the LTE link-level simulator [12]. Fig. 2 shows the BLER-SINR curves for the 15 MCS modes at the LTE uplink physical layer. The correspondingly modeled parameters $b(m)$ and $c(m)$ are shown in Table I.

For the LTE uplink video streaming, one video packet usually occupies several RBs that share the common video packet synchronization mark. Thus, the loss for any RB belonging to the video packet may lead to the loss of the whole video packet. The video Packet Loss Probability (PLP) $\rho_{n,i}(m)$ for the slice

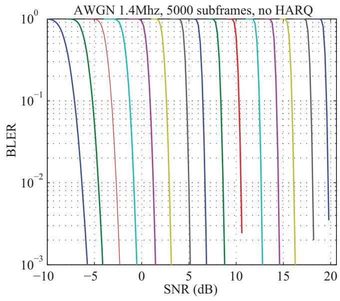


Fig. 2. BLER-SINR curves for all 15 MCS modes (MCS 1 (leftmost) to MCS 15 (rightmost)).

TABLE I
THE CANDIDATE LTE UPLINK MCS MODES

MCS mode (m)	Modulation order	Rate (bit-s/symbol)	$b(m)$	$c(m)$
1	QPSK	0.1523	-8.074	0.7534
2	QPSK	0.2344	-5.981	0.6018
3	QPSK	0.3770	-3.812	0.4832
4	QPSK	0.6010	-1.823	0.4258
5	QPSK	0.8770	0.4875	0.3163
6	QPSK	1.1758	2.265	0.2826
7	16QAM	1.4766	4.503	0.2082
8	16QAM	1.9141	6.207	0.2134
9	16QAM	2.4063	8.103	0.2211
10	64QAM	2.7305	10.05	0.1941
11	64QAM	3.3223	12.22	0.1865
12	64QAM	3.9023	14.03	0.1833
13	64QAM	4.5234	15.66	0.1907
14	64QAM	5.1152	17.70	0.1751
15	64QAM	5.5547	19.35	0.1676

$s_{n,i}$ carried by the MCS mode m is related to the BLERs for all the RBs that the video packet contains as

$$\rho_{n,i}(m) = 1 - \prod_{k=1}^{B_{num}} (1 - BLER_k(SINR(m))), \quad (2)$$

where B_{num} is the RB number that the video packet contains, and $BLER_k(SINR(m))$ is the BLER for the k^{th} RB in the video packet corresponding to the slice $s_{n,i}$. Thus, the MCS mode should be selected by well considering the effect of BLER on the video PLP for the wireless video streaming.

At the condition of the same level of transmission signal degradation and noise, higher transmit power can result in higher SINR. Thus, the uplink transmit power control serves as an important role in the mobile communication system because it can balance the need for sufficient transmit energy to provide the required SINR against the need to maximize the battery life. In LTE uplink, the adopted power control scheme is based on a combination of open-loop and closed-loop control. Specifically, the FPC is adopted as the open-loop control scheme, and it can properly adjust the transmit power to compensate the long-term signal degradations including the path loss and shadowing. Faster link adaption for short-term inference and noise is performed with the help of closed-loop power control.

The overall UE transmit power P_m (measured in dBm) for the MCS mode m is set as

$$P_m = \min\{P_{\max}, P_0 + \alpha \cdot L + \Delta_m + f(\Delta_i) + 10 \cdot \lg M\}, \quad (3)$$

where P_{max} is the maximum transmit power; $P_0+\alpha \cdot L$ indicates the open-loop power control, and P_0 is a semi-static power level comprising a nominal power level that is common for all UEs in the cell and a UE-specific offset, α is a fractional path-loss compensation factor, L is the estimated downlink path-loss; $\Delta_m + f(\Delta_i)$ is the dynamic offset for the closed-loop power control, Δ_m is the MCS-dependent power offset, $f(\Delta_i)$ is the UE-specific Transmitter Power Control (TPC) command with relative or absolute increase depending on $f(\cdot)$ function; $10 \cdot \lg M$ is the bandwidth factor, and M is the number of allocated RBs for the UE. Knowing the transmit power P_m, the experienced $SINR(m)$ (measured in dB) is defined as

$$\begin{aligned} SINR(m) =& P_m - L - IoT - TN \\ =& P_0 + 10 \cdot \lg M + (\alpha - 1) \cdot L \\ & + \Delta_m + f(\Delta_i) - IoT - TN, \end{aligned} \quad (4)$$

where IoT refers to the Interference over Thermal (IoT), and TN is the thermal noise. To maintain the required link SINR, the transmit power should be adaptively regulated according to the time-varying inferences and noises.

2) Video data adjustment: The throughput of LTE uplink channel changes with the selected MCS mode, which is determined in terms of the SINR. Thus, the video bit rate should be accordingly adjusted to fully utilize the bandwidth for LTE uplink channel. As we all know, the video data compression is a kind of lossy compression. Different Quantization Parameters (QPs) can result in the encoded video streams with different bit rates and qualities. Larger QP can produce the video stream with larger quantization-induced distortion and lower bit rate than the video stream that smaller QP produces. In the proposed scheme, to decrease the quantization-induced encoding distortion and the effect of transmission delay on the perceived quality of the real-time video application, the QP for each video packet is adaptively regulated.

In addition, the compressed video data is always with high dependencies in both the spacial and temporal domains. The loss of one video packet may result in incorrectly decoding of the subsequent video data which refers to the previous lost video packet. To restrain the effect of error propagation on the video quality degradation, the insertion of intra-coded MBs has been shown to be an effective technique since the decoding of the intra-coded MB does not need any reference information from the previous frames. However, the increasing numbers of intra-coded MB may lead to large amounts of encoding bits and decrease the coding efficiency. In the proposed cross-layer optimization scheme, the SSIM-based error-resilient Rate-Distortion Optimization (RDO) [13] is introduced into the encoding process for each MB to select the optimal encoding mode which can suppress the effect of error propagation on the video quality degradation.

B. Cross-layer optimization problem formulation

In the LTE uplink video streaming, each video slice is packed to be a video packet for transmitting. To proceed with a concrete formulation of the cross-layer problem, let us denote $s_{n,i}$ the i^{th} slice in the n^{th} frame f_n. The expected end-to-end video distortion $E\{D_{n,i}\}$ for the slice $s_{n,i}$ is minimized at the condition that the transmit power $P_{n,i}$ is not larger than the transmit power budget $P^b_{n,i}$. Mathematically, the cross-layer optimization problem for slice $s_{n,i}$ is modeled as

$$\min\{E\{D_{n,i}\}\} \quad s.t. \quad P_{n,i} \le P^b_{n,i}. \tag{5}$$

The transmit power $P_{n,i}$, the link MCS mode $MCS_{n,i}$ at the physical layer, and the encoding quantization parameter $QP_{n,i}$ at the application layer for the slice $s_{n,i}$ are jointly regulated to achieve the optimal optimization result. Inherently, to restrain the effect of error propagation on the video quality degradation, the encoding mode for each MB inside the slice $s_{n,i}$ is selected by the SSIM-based error-resilient RDO. Let us denote $mb_{n,i,j}$ the j^{th} MB in the slice $s_{n,i}$. For the MB $mb_{n,i,j}$, the optimal error-resilient encoding mode $EM_{n,i,j}$ is selected by achieving a good trade-off between the expected SSIM-based MB decoding distortion $E\{DSSIM_{n,i,j}\}$ and the number of encoding bits $R_{n,i,j}$ as

$$\min_{\{\mathrm{EM}_{n,i,j}\}}\{J_{n,i,j}\}, J_{n,i,j}=E\{DSSIM_{n,i,j}\}+\lambda^{ep}_{ssim(n,i,j)} \cdot R_{n,i,j}, \tag{6}$$

where $J_{n,i,j}$ and $\lambda^{ep}_{ssim(n,i,j)}$ are the Lagrange cost for each candidate encoding mode and the SSIM-based Lagrange multiplier in the error-prone transmission environment, respectively. The SSIM-based Lagrange multiplier $\lambda^{ep}_{ssim(n,i,j)}$ is used to balance the expected decoding distortion $E\{DSSIM_{n,i,j}\}$ and the number of encoded bits $R_{n,i,j}$ for the MB $mb_{n,i,j}$, and it is obtained in terms of the Lagrange parameter derivation [13].

III. The solution: Lagrange relaxation method

For the slice $s_{n,i}$, the expected decoding distortion $E\{D_{n,i}\}$ is minimized under the constraint of the transmit power budget as (5). In practice, the constraint optimization problem defined in (5) can be solved via the Lagrangian relaxation as

$$\min\{J_{n,i}\}, J_{n,i} = E\{D_{n,i}\} + \lambda_{n,i} \cdot P_{n,i}, \tag{7}$$

where $J_{n,i}$ is the Lagrange cost, $\lambda_{n,i}$ is the Lagrange multiplier to balance the transmit power $P_{n,i}$ and expected end-to-end video distortion $E\{D_{n,i}\}$. In the following, the expected end-to-end video distortion $E\{D_{n,i}\}$ estimation and the dynamic programming solution for (7) are presented.

A. End-to-end distortion estimation

The SSIM-based distortion (DSSIM) $DSSIM(x, y)$ for two MBs x and y is defined as

$$\begin{aligned} DSSIM(x,y) &= 1-SSIM(x,y) \\ &= 1 - \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \end{aligned} \tag{8}$$

where $SSIM(x, y)$ is the SSIM index between the MBs x and y, μ_x, σ_x and σ_{xy} are the mean, standard deviation and cross correlation between the two image MBs, respectively. C_1 and C_2 are used to ensure stability when the means and variances are close to zero [10].

For the wireless video streaming, the decoding distortion is induced not only by the irreversible quantization, but also by the packet loss and the error propagation. When the slice $s_{n,i}$ is discarded in the transmission process, all the MBs in $s_{n,i}$ are decoded using the error concealment strategy. For the MB $mb_{n,i,j}$ in the slice $s_{n,i}$, denote $b_{n,i,j}$ and $b^{e\_c}_{n,i,j}$ the original MB pixels and the concealed MB pixels, respectively. If the slice $s_{n,i}$ is lost in the transmission process, the SSIM-based decoding distortion for the MB $mb_{n,i,j}$ is $1-SSIM(b_{n,i,j}, b^{e\_c}_{n,i,j})$. When the slice $s_{n,i}$ is correctly received, the MBs that belong to the slice $s_{n,i}$ are decoded with the intra or inter decoding tools. Denote $b^{n\_l}_{n,i,j}$ the decoded MB pixels without packet loss for the MB $mb_{n,i,j}$. At this condition, the SSIM-based decoding distortion for the MB $mb_{n,i,j}$ is $1-SSIM(b_{n,i,j}, b^{n\_l}_{n,i,j})$. Thus, the expected SSIM-based decoding distortion $E\{DSSIM_{n,i,j}\}$ for the MB $mb_{n,i,j}$ can be correspondingly calculated as

$$\begin{aligned} E\{DSSIM_{n,i,j}\} &= \rho_{n,i} \cdot (1 - SSIM(b_{n,i,j}, b^{e\_c}_{n,i,j})) \\ &\quad + (1 - \rho_{n,i}) \cdot (1 - SSIM(b_{n,i,j}, b^{n\_l}_{n,i,j})) \\ &= 1 - \rho_{n,i} \cdot SSIM(b_{n,i,j}, b^{e\_c}_{n,i,j}) \\ &\quad - (1 - \rho_{n,i}) \cdot SSIM(b_{n,i,j}, b^{n\_l}_{n,i,j}), \end{aligned} \tag{9}$$

where $\rho_{n,i}$ is the PLP for the video packet corresponding to the slice $s_{n,i}$, and it can be estimated by (2).

After getting the expected decoding distortion for each MB that belongs to the slice $s_{n,i}$, a RoI-based decoding distortion expectation $E\{D_{n,i}\}\_RoI$ is used to estimate the decoding distortion for the slice $s_{n,i}$ as

$$\begin{aligned} E\{D_{n,i}\}\_RoI =& \frac{1}{N^{non}_{n,i}} \cdot \alpha_{non} \sum_{mb_{n,i,j} \in A_{non}} E\{DSSIM_{n,i,j}\} \\ &+ \frac{1}{N^{r}_{n,i}} \cdot \alpha_{r} \sum_{mb_{n,i,j} \in A_{r}} E\{DSSIM_{n,i,j}\}, \end{aligned} \tag{10}$$

where $N^{non}_{n,i}$ and $N^{r}_{n,i}$ are the non-RoI MB number and the RoI MB number that slice $s_{n,i}$ contains, A_{non} and A_r indicate the non-RoI and RoI MB sets respectively, α_{non} and α_r are the corresponding weights with $\alpha_{non} + \alpha_r = 1$ [14]. In this work, the interactive RoI decision method which enables the end user to select the RoI while watching the received video is adopted [15]. Furthermore, to enhance the video quality of RoI, the quantization parameters for the MBs belonging to the RoI are adjusted to be smaller than those of the other MBs. When performing the RoI-based cross-layer optimization, the quantization parameter $QP_{n,i}$ for the slice $s_{n,i}$ refers to the quantization parameter pairs $\{QP^{RoI}_{n,i}, QP^{RoI}_{n,i} + QP_{scale}\}$, in which $QP^{RoI}_{n,i}$ and QP_{scale} indicate the quantization parameter for the MB belonging to the RoI and the quantization parameter difference between the MBs belonging to RoI and the other MBs.

B. Dynamic programming solution for the formulated cross-layer problem

For the slice $s_{n,i}$, the proposed cross-layer optimization is performed by selecting the optimal parameters from the possible values of the parameter set $\{P_{n,i}, MCS_{n,i}, QP_{n,i}, \{EM_{n,i,j}|mb_{n,i,j} \in s_{n,i}\}\}$. As discussed in Section II.B, the encoding modes $\{EM_{n,i,j}|mb_{n,i,j} \in s_{n,i}\}$ for the MBs inside the slice $s_{n,i}$ are selected by (6). When performing the cross-layer optimization for the slice $s_{n,i}$ as (7), the optimal Lagrange multiplier $\lambda_{n,i}$ has to be determined. It has been shown [16] that if there is a $\lambda^*_{n,i}$ such that

$$\{P^*_{n,i}, MCS^*_{n,i}, QP^*_{n,i}\} = arg \ \ \min J_{n,i}(P_{n,i}, MCS_{n,i}, QP_{n,i}) \quad (11)$$

makes $P_{n,i} = P^b_{n,i}$, then $\{P^*_{n,i}, MCS^*_{n,i}, QP^*_{n,i}\}$ is also an optimal parameter set for (7). It is known that the transmit power-distortion curve is non-increasing for the transmitted video packet. $\lambda^*_{n,i}$ can be obtained by the bisection algorithm as shown in Algorithm 1. Then, given the transmit power budget, the optimal cross-layer optimization can be performed as (6) and (7).

Algorithm 1 Bisection searching algorithm for $\lambda^*_{n,i}$

1: Choose two values of $\lambda_{n,i}$, $\lambda^l_{n,i}$ and $\lambda^u_{n,i}$ with $\lambda^l_{n,i} \leq \lambda^u_{n,i}$ which satisfy $P^*_{n,i}(\lambda^u_{n,i}) \leq P^b_{n,i} \leq P^*_{n,i}(\lambda^l_{n,i})$
2: Set $\lambda^{next}_{n,i} \leftarrow \frac{\lambda^l_{n,i}+\lambda^u_{n,i}}{2}$
3: Perform the optimization as (7) with $\lambda^{next}_{n,i}$
4: **if** $P^*_{n,i}(\lambda^{next}_{n,i}) = P^*_{n,i}(\lambda^u_{n,i})$ **then**
5: $\quad \lambda^*_{n,i} = \lambda^u_{n,i}$, algorithm stop
6: **else if** $P^*_{n,i}(\lambda^{next}_{n,i}) > P^b_{n,i}$ **then**
7: $\quad \lambda^l_{n,i} \leftarrow \lambda^{next}_{n,i}$, go to step 2
8: **else**
9: $\quad \lambda^u_{n,i} \leftarrow \lambda^{next}_{n,i}$, go to step 2
10: **end if**

IV. Experimental Results

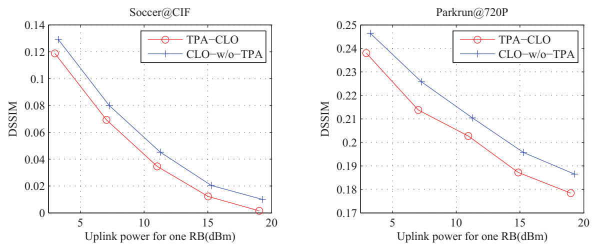
With the help of H.264/AVC reference software JM16.1 and the LTE uplink channel model [12], the proposed cross-layer optimized LTE uplink video streaming is simulated. The main experimental parameters are shown in Table II.

TABLE II
EXPERIMENTAL PARAMETERS

P_0	(-82, -78, -74, -70, -66)dBm
Distance attenuation	$L = 35.3 + 37.6 \cdot \log(d)$ dB d: distance in meter
UE to eNodeB distance d	200m
Fractional power factor α	0.7
Penetration loss	20dB
Thermal noise per RB	-116dBm
Maximum transmit power	24dBm
IoT	Rayleigh distribution
Candidate MCS modes	Table I
Transmission rate	300k symbol/s

A. Transmit power efficiency

The RoI quality enhancement is not adopted in this evaluation, and we set α_{non} and α_r to equally be 0.5. The Cross-Layer Optimization scheme without Transmit Power Awareness (CLO-w/o-TPA) which adopts the constant transmit power (open-loop power control) is used as the anchor scheme. At the condition of the average IoT $\overline{IoT} = -27dBm$, the Distortion-Transmit Power (DSSIM-TP) curves of the proposed Transmit Power Aware Cross-Layer Optimization (TPA-CLO) scheme and the CLO-w/o-TPA scheme are presented in Fig. 3. It can be seen from Fig. 3 that the proposed TPA-CLO scheme can significantly decrease the decoding distortion at the same transmit power cost compared to the CLO-w/o-TPA scheme.



(a) (b)

Fig. 3. The DSSIM-TP curves of the TPA-CLO scheme and the CLO-w/o-TPA scheme for the sequences (a) Soccer (CIF) and (b) Parkrun (720P).

To get the accurate amplitude of the transmit power saving ratio for the proposed scheme over the CLO-w/o-TPA scheme, the transmit power saving ratios for nine video sequences at different channel conditions are summarized in Table III. At the conditions of $\overline{IoT} = -38dBm$, $\overline{IoT} = -27dBm$ and $\overline{IoT} = -20dBm$, the proposed scheme can achieve 24.28%, 22.64%, and 21.18% reductions in transmit power on average, respectively. It can be seen that the proposed TPA-CLO scheme can achieve significant transmit power reduction compared to the CLO-w/o-TPA scheme. This is because that the transmit power can be dynamically adjusted to meet the SINR demand for the required video quality by compensating the signal power at unsatisfactory channel condition and decreasing the transmit power at good channel condition.

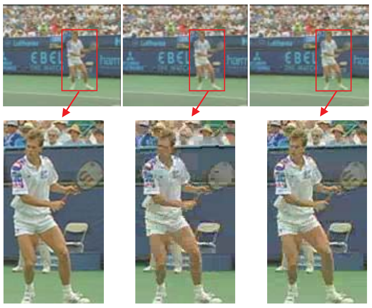
TABLE III
TRANSMIT POWER SAVING RATIO

Sequences	$\Delta P(\%)$		
	$\overline{IoT} = -38dBm$	$\overline{IoT} = -27dBm$	$\overline{IoT} = -20dBm$
Bus(CIF)	-21.44	-19.79	-17.97
Football(CIF)	-19.53	-17.36	-16.55
Foreman(CIF)	-17.32	-16.54	-15.75
Stefan(CIF)	-23.08	-21.64	-19.65
Paris(CIF)	-24.33	-21.75	-20.09
Soccer(CIF)	-18.43	-16.78	-15.69
Parkrun(720P)	-33.49	-32.28	-30.26
Mobacal(720P)	-29.43	-27.98	-26.34
Shield(720P)	-31.47	-29.68	-28.22
AVG	-24.28	-22.64	-21.18

B. Video experience quality improvement

In this subsection, the video experience quality improvement due to the RoI quality enhancement is evaluated. α_{non} and α_r are set to 0.1 and 0.9, respectively. The quantization parameter difference QP_{scale} between the MBs belonging to the RoI and the other MBs is set to 5. As an example, the region that the tennis player moves is recognized as the RoI for the sequence Stefan(CIF). At the condition of $P_0 = -78dBm$ and $\overline{IoT} = -27dBm$, one sample picture for the original sequence and the decoded sequences for the TPA-CLO without RoI enhancement (TPA-CLO-w/o-RoI) scheme and the TPA-CLO with RoI quality enhancement (TPA-CLO-w-RoI) scheme are shown in Fig. 4 (a), (b), and (c), respectively.

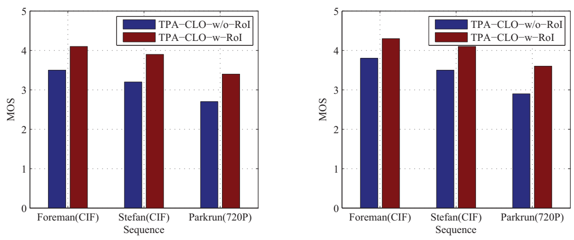
Subjectively, compared to the TPA-CLO-w/o-RoI scheme, we can find that the TPA-CLO-w-RoI scheme can achieve better visual experience for the user who is interested in the tennis player, especially for the face of the player and the tennis packet. As shown in the caption of Fig. 4, the average SSIM-based distortions DSSIM (α_{non} = α_r = 0.5) for the decoded frames of both the TPA-CLO-w/o-RoI scheme and TPA-CLO-w-RoI are similar. However, the TPA-CLO-w-RoI scheme can achieve significant less RoI-based SSIM distortion DSSIM_RoI ($\alpha_{non} = 0.1$, $\alpha_r = 0.9$) than the TPA-CLO-w/o-RoI scheme. It can be seen that the adopted RoI-based distortion metric can properly measure the perceptual video quality degradation for RoI quality enhancement optimization.



(a) (b) (c)

Fig. 4. The 3^{rd} frame for (a) original sequence, the decoded sequences optimized by (b) TPA-CLO-w/o-RoI scheme, $DSSIM = 0.0737$, $DSSIM\_RoI = 0.0744$, and (c) TPA-CLO-w-RoI scheme, $DSSIM = 0.0782$, $DSSIM\_RoI = 0.0491$.

In addition, the Degradation Category Rating (DCR) [17] is used as the evaluation method to verify the overall perceptual quality improvement due to the RoI-based unequal quality control. The opinion score is continuously mapped to the value of 0 (bad perceptual quality) to 5 (excellent perceptual quality), and twenty viewers participate in the subjective evaluation process. As shown in Fig. 5, the Mean Opinion Score (MOS) values for three decoded sequences at different channel conditions are shown. It can be seen from Fig. 5 that the video streams optimized by the proposed TPA-CLO-w-RoI scheme can achieve higher MOS values than those of the TPA-CLO-w/o-RoI scheme.



(a) (b)

Fig. 5. The MOS values for different sequences at the conditions of (a) $\overline{IoT} = -27dBm$ and (b) $\overline{IoT} = -32dBm$.

V. Conclusions

In this paper, a transmit power aware cross-layer optimization scheme is proposed to optimize the perceptual quality of LTE uplink video streaming. The video coding quantization parameter and encoding mode for each MB at the application layer, and the transmit power and MCS mode at the physical layer are jointly adjusted to achieve the optimal trade-off between the acquired video quality and the transmit power. In addition, the video quality for RoI is enhanced at the limited transmit resource condition. Extensive experimental results show that the proposed scheme can achieve about 23% transmit power cost reduction at the condition of the same perceived video quality compared to the traditional CLO-w/o-TPA scheme, and that significant perceptual quality improvement due to RoI quality enhancement for LTE uplink video streaming is achieved.

References

[1] R1-073224. Way Forward on Power Control of PUSCH. 3GPP TSGRAN WG1 49-bits, Jun. 2007.

[2] C. Ú. Castellanos, D. L. Villa, C. Rosa, K. I. Pedersen, F. D. Calabrese, P. Michaelsen, and J. Michel, "Performance of Uplink Fractional Power Control in UTRAN LTE," IEEE Vehicular Technology Conference, pp. 2517-2521, May 2008.

[3] A. Simonsson and A. Furuskär, "Uplink Power Control in LTE - Overview and Performance," IEEE Vehicular Technology Conference, pp. 1-5, Sept. 2008.

[4] K. Sagjan, S. Strzyż, K. I. Pedersen, J. Steiner, and C. Rosa, "Open Loop Power Control Parameter Settings Impact on LTE HetNet Uplink Performance," IEEE International Conference on Communications, pp. 1134-1138, June 2013.

[5] H. Luo, S. Ci, D. Wu, J. Wu, and H. Tang, "Quality-driven cross-layer optimized video delivery over LTE," IEEE Communications Magazine, vol. 48, no. 2, pp. 102-109, Feb. 2010.

[6] S. Cicalò, Mazzotti, S. Moretti, V. Tralli, and M. Chiani, "Cross-layer optimization for m-health SVC multiple video transmission over LTE uplink," IEEE International Conference on e-Health Networking, Application & Services, pp. 212-217, Oct. 2013.

[7] M. Shehada, B. Fu, S. Thakolsri, and W. Kellerer, "QoE-based resource reservation for unperceivable video quality fluctuation during Handover in LTE," IEEE Consumer Communications and Networking Conference, pp. 171-177, Jan. 2013.

[8] R. Vishwanath and O. Ozgur, "Video-QoE aware radio resource allocation for HTTP adaptive streaming," IEEE Conference on Communications, pp. 1076-1081, June 2014.

[9] L. Toni, P. Frossard, "MSE cross-layer optimization criteria: What else?" IEEE COMSOC MMTC E-Letter,Vol.6, No.2, pp. 13-16, Feb. 2011.

[10] Z. Wang, L. Lu, and A.C. Bovik, "Video quality assessment based on structural distortion measurement," Signal Processing: Image Communication, vol. 19, no. 2, pp. 121-132, Feb. 2004.

[11] K. Sayana, J. Zhuang, and K. Stewart, "Link performance abstraction based on mean mutual information per bit (MMIB) of the LLR channel," IEEE Std. 802.16 BWA WG, C802.16m-07/0972007.

[12] [Online]. Available: http://www.nt.tuwien.ac.at/ltesimulator/.

[13] P. Zhao, Y. Liu, J. Liu, S. Ci, and R. Yao, "SSIM-based error-resilient rate-distortion optimization of H.264/AVC video coding for wireless streaming," Signal Processing: Image Communication, vol. 29, no. 3, pp. 303-315, Mar. 2014.

[14] W. Yang, L. Wu, Y. Fan, and Z. Wang, "A method of iamge quality assessment based on region of interest," World Congress on Intelligent COntrol and Automation, pp. 6840-6843, June 2008.

[15] M. Makar, A. Mavlankar, P. Agrawal, and B. Girod, "Real-time video streaming with interactive region-of-interest," International Conference on Image Processing, pp. 4437-4440, Sept. 2010.

[16] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," IEEE Transaction on Image Processing, vol. 2, no. 2, pp. 160-175, Apr. 1993.

[17] ITU-T RECOMMENDATION, P.910, "Subjective video quality assessment methods for multimedia applications," 1999.