

Parametric Planning Model for Video Quality Evaluation of IPTV Services Combining Channel and Video Characteristics

Jiarun Song, *Member, IEEE*, Fuzheng Yang, *Member, IEEE*, Yicong Zhou, *Senior Member, IEEE*, and Shan Gao

Abstract—Parametric planning models are designed for estimating the video quality, which can be applied to effective planning, implementation, and management of network video applications and communication networks. However, different from the bitstream-based evaluation models, the planning models are not allowed to exploit the video streams, with only limited information available for use, i.e., a few general parameters predetermined by the service providers and network operators. In this paper, a parametric planning model combining channel and video characteristics is proposed to estimate the video distortion caused by packet loss for Internet protocol television (IPTV) services. More specifically, the probability distribution of the channel states is determined by detailed analysis of the channel characteristics. Then, considering the influence of burst packet loss and the temporal dependence between frames, several sequence-level and frame-level parameters for video quality evaluation are derived from the perspective of the probability distribution of the channel states. Utilizing these parameters, the proposed model approximates the video quality considering the effects of direct packet loss and error propagation. Experimental results show that the proposed model has a superior performance for video quality estimation than the three commonly used parametric planning models.

Index Terms—Network planning, planning model, quality of experience (QoE) planning, video quality assessment, video streaming applications.

I. INTRODUCTION

RECENT years have witnessed an increasing proliferation of IPTV services. According to Cisco's report, the global

Manuscript received April 9, 2016; revised September 28, 2016 and November 19, 2016; accepted December 1, 2016. Date of publication December 12, 2016; date of current version April 15, 2017. This work was supported by the National Science Foundation of China under Grant 61601349, Grant 61371089, Grant 61571337, and Grant 61601348, and by the 111 Project B08038. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ivan Bajic.

J. Song is with the Collaborative Innovation Center of Information Sensing and Understanding and ISN, Xidian University, Xi'an 710071, China (e-mail: jrsong@xidian.edu.cn).

F. Yang is with the State Key Laboratory of ISN, Xidian University, Xi'an 710071, China, and also with the School of Electrical and Computer Engineering, Royal Melbourne Institute of Technology, Melbourne, VIC 3001, Australia (e-mail: fzhyang@mail.xidian.edu.cn).

Y. Zhou is with the Department of Computer and Information Science, University of Macau, Macau, China (e-mail: yicongzhou@umac.mo).

S. Gao is with the Media Technology Laboratory, Huawei Technologies Company Ltd., Shenzhen 518129, China (e-mail: simon.gaoshan@huawei.com).

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/TMM.2016.2638621

IP video traffic will be 82 percent of all consumer Internet traffic by 2020, up from 70 percent in 2015 [1]. In order to achieve a high level of user satisfaction for IPTV services, it is crucial to objectively predict the video quality for system design, network and Quality of Experience (QoE) planning, quality benchmarking and monitoring [1]–[4].

In terms of the employed information, objective quality assessment models for network video can be classified into five categories: parametric planning models, packet-layer models, bitstream-layer models, media-layer models, and hybrid models [5], as illustrated in Fig. 1. Different from the packet-layer model, the bitstream-layer model and the hybrid model which are usually used for quality monitoring [6], the parametric planning model is mainly applied to network and service planning. It is initially designed for service planners to identify beforehand how the video quality will be in a certain application and network parameters setting, to avoid over-engineering the applications, terminals, and networks while guaranteeing user's satisfaction [7]. More specifically, the video quality is estimated using a priori of parameters, and then the appropriate service and network parameters are chosen and deployed in practice according to the video quality. Thus, the parametric planning model is important and helpful for multimedia service providers and network operators.

However, unlike other kinds of assessment models that evaluate the video quality by exploiting the information of bitstream or media signals, as shown in Fig. 1, the parametric planning model estimates the video quality without resorting to the actual video streams during the planning phase, and uses only a few empirical parameters (e.g., coding bitrate, packet loss rate, and so on) supplied by the service providers and network operators. In such a case, the parametric planning model cannot obtain the detailed information about video coding (e.g., frame type, quantization parameter, motion vector, etc.), video content (e.g., temporal complexity, spatial complexity, pixel values, etc.), as well as packet loss (e.g., the actual position of packet loss, the number of lost packet, etc.). How to accurately estimate the video quality using parametric planning model challenging and still remains as an open issue.

Targeting service planning, a parametric planning model was standardized by ITU-T Recommendation G.1070 for video-phones, where the video quality was calculated by using application and network parameters, such as the bit rate, frame rate and packet loss rate [8]. Considering the influence of the

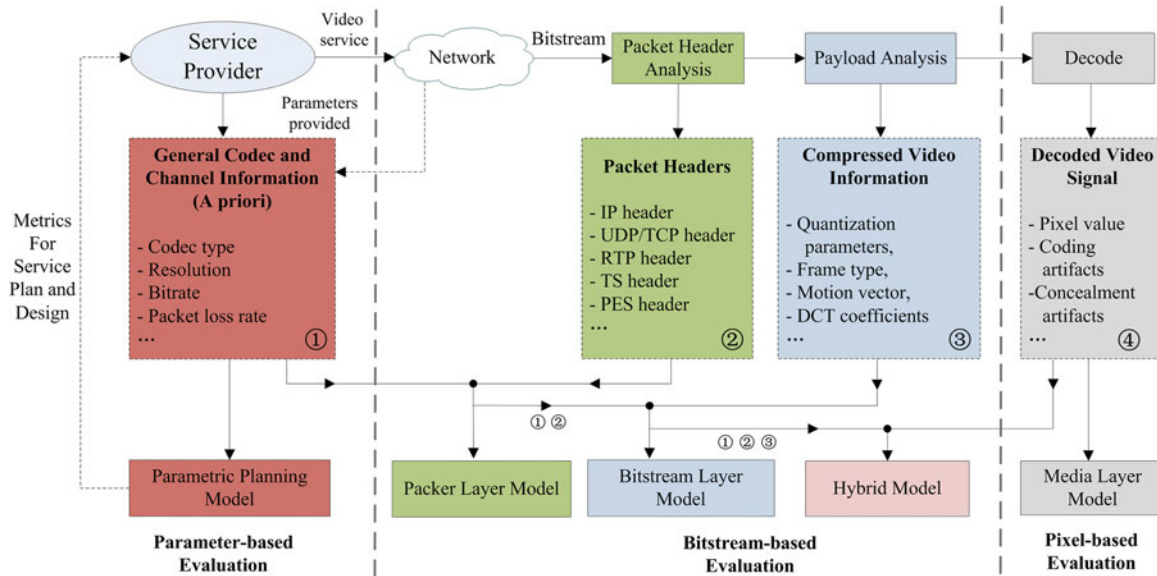


Fig. 1. Illustration of video quality evaluation model.

TABLE I
COMPARISON OF THE STATE-OF-THE-ART

| State-of-the-art models | Publication year | Typical input parameters | Burst packet loss included |
|-------------------------|------------------|---|----------------------------|
| G.1070 [8] | 2007 | Bit-rate, frame rate, packet loss rate | None |
| Yamagishi[9] | 2008 | Bit-rate, frame rate, packet loss frequency, and burst length | Yes |
| Garcia [10] | 2010 | Bit-rate, packet loss rate, the number of lost packets in a row | Yes |
| G.1071 [12] | 2015 | Bit-rate, frame rate, burst packet loss length, packets size, length of group of pictures (GOP) | Yes |

burst packet loss on the video quality, as shown in Table I, the burst length [9] and the number of lost packets in a row [10] were employed to estimate the video quality, respectively. The model in [9] was an update version of the ITU-T G.1070 model. However, all these parametric planning models simply use a few statistical parameters to estimate the video quality and fail to carefully analyze the influence of the packet loss on the coded frame and video streaming. In recent years, ITU-T Study Group 12 studied a new parametric planning model (G.OMVAS) for video streaming applications, which focuses on evaluating the impact of typical IP network impairments on the video quality [11]. The corresponding parametric planning model was presented in the ITU-T Recommendation G.1071 [12], whose formula and outputs are in accordance with those of packet layer model in ITU-T Recommendation P.1201 [13]. Particularly, the ITU-T G.1071 model provides a set of rules to convert the planning parameters into the forms of P.1201 inputs since these packet-layer inputs are not available in the planning phase. Though the influence of packet loss on the video streaming was studied in this model, the correlation between individual packet loss events is still not taken into account.

In practice, the packet loss process in the wired and wireless channels often exhibits finite temporal dependency and can be

well characterized via a finite-state Markov model [14]–[17]. In this paper, a parametric planning model is proposed to evaluate the video quality of IPTV services, which matches the G.OMVAS framework outlined by ITU-T SG12. The model covers the H.264/AVC coded video transmitted over channels modelled by a four-state Markov chain. Unlike most traditional methods which directly map the statistical coding and network parameters (e.g., coded bit-rate, packet loss rate, etc.) to video quality, the proposed model evaluates video quality combining channel and video characteristics. Specifically, the probability distribution of the channel states is calculated analyzing channel characteristics, which can better clarify different packet loss behaviors. Due to the fact that the burst packet loss and the temporal dependence between frames will lead to nonrandom frame distortion, several sequence- and frame-level parameters for quality evaluation of IPTV services are derived from the perspective of probability distribution of the channel states. Utilizing these parameters, the proposed model can better indicate the effects of direct packet loss and error propagation on the video streaming. It is noted that for IPTV services, the video streams are more sensitive to distortion caused by packet loss, while one-way delay and jitter are generally not problematic since the Set-Top-Box (STB) is able to provide proper de-jitter buffers [4]. Thus, this research is focused on how to investigate the influence of packet loss on the video quality for IPTV services.

The remainder of this paper is organized as follows. Section II describes the framework of the proposed parametric planning model. In Section III, the impacts of packet loss on video quality are evaluated. Performance evaluation and conclusion are given in Sections IV and V, respectively.

II. FRAMEWORK OF THE PROPOSED PARAMETRIC PLANNING MODEL

The framework of the proposed parametric planning model is illustrated in Fig. 2. The proposed model consists of five

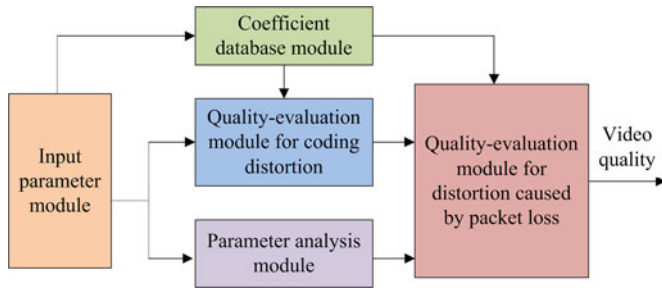


Fig. 2. Framework of the proposed parametric planning model.

TABLE II
INPUT PARAMETERS OF THE PARAMETRIC PLANNING MODEL

| Category | Parameters |
|---|--|
| Input parameters for video quality estimation | Bit-rate, frame rate, packet loss rate, averaged burst packet loss length, packet size, length of GOP, number of slice, and characteristics of channel model (state transition probability). |
| Input parameters for coefficient database | Codec type (H.264, MPEG4), video resolution, packet loss concealment (freezing, slicing) |

modules: input parameter module, coefficient database module, parameter analysis module, quality-evaluation module for coding distortion, and quality-evaluation module for distortion caused by packet loss.

A. Input Parameter Module

The input parameter module includes the input parameters of the planning model. Different from the bitstream-based or pixel-based video quality assessment model that use the information extracted from an available video streaming, the inputs of the planning model are the statistical parameters supplied by the network operators and service providers during the planning phase. There are only a few parameters that can be used in the planning model. According to the proposal of ITU-T SG12 for the parametric planning models [18], the input parameters can be divided into two categories: the input parameters for quality estimation and those for the coefficient database. More specifically, the input parameters for quality estimation are used to evaluate the video quality, while the input parameters for the coefficient database are used to determine which group of model coefficients should be chosen to evaluate the video quality. The detailed input parameters for the parametric planning models are listed in Table II.

B. Coefficient Database Module

The coefficient database module stores a set of coefficients, which is designed to cope with the different video codec type, video resolutions or many others. The video streaming with different video codecs and resolutions may have different coefficients of the parametric planning model. Therefore, according to the input parameters (e.g., H.264 codec, 720P, slicing, zero motion error concealment), the coefficient database

module will provide the corresponding coefficients to evaluate the video quality.

C. Parameter Analysis Module

Combining input parameters with characteristics of the channel models, some other useful information for video quality evaluation can be deduced by the parameter analysis module. For instance, considering the influence of the packet loss on the video streaming, the average frame loss frequency, the expectation of the number of impaired frames and the average impaired ratio of each frame can be estimated combining the characteristics of the channel models. These parameters can be further used to estimate the video quality. The detailed procedure of the parameter analysis module is one of the contributions of the proposed planning model.

D. Quality-Evaluation Module for Coding Distortion

This module is used to evaluate the video coding distortion. According to [19], [20], the video quality affected by the coding distortion is closely related to the coded bitrate and frame rate. It has been effectively evaluated as follows:

$$Q_c = \begin{cases} 1 + v_1 \left(1 - \frac{1}{1 + (B_F / v_2)^{v_3}} \right), & F_R \geq 30 \\ \left(1 + v_1 \left(1 - \frac{1}{1 + (B_F / v_2)^{v_3}} \right) \right) \cdot \left(1 - v_4 \ln \left(\frac{30}{F_R} \right) \right), & F_R < 30 \end{cases} \quad (1)$$

where Q_c is the coding quality. F_R is the video frame rate. B_F is the average number of bits for coding a frame, which can be obtained by the ratio of the coded bit-rate and F_R . v_1 , v_2 , v_3 and v_4 are empirical parameters. Considering that all the mentioned parameters are available during the planning phase, the proposed parametric planning model utilizes (1) to estimate the video coding quality as well, where the parameters of v_1 , v_2 , v_3 and v_4 are retrained using the current database.

E. Quality-Evaluation Module for Distortion Caused by Packet Loss

This module evaluates the video quality when the video transmission in the presence of channel errors. According to the study in [7], when packets are lost during video transmission, significant errors may appear due to the corruption of related video data. Moreover, transmission errors in one frame may also propagate to the subsequent frames since the predictive coding structures are employed [20], [21]. This problem may be even worse in the video applications where the ‘‘IPPP’’ coding structure is employed. In this quality evaluation module, the video coding quality and the parameters calculated by the parameter analysis module are employed to evaluate the video quality affected by the packet loss. This module is also a focus of the proposed parametric planning model.

In the following section, the proposed parameter analysis method and the video quality evaluation method for the distortion caused by packet loss will be discussed in detail.

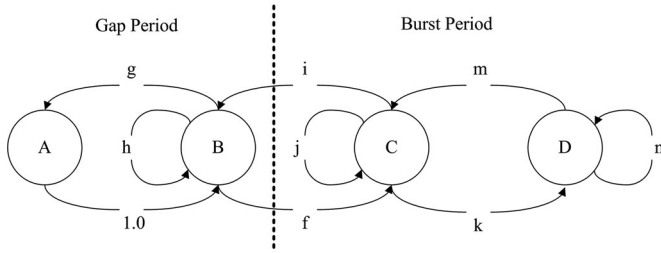


Fig. 3. Four-state Markov channel model.

III. EVALUATION OF DISTORTION CAUSED BY PACKET LOSS

For video streaming applications, the information of the impaired frame is very important to the video quality since the video sequence is constituted by frames. In order to accurately indicate the influence of packet loss on the video quality for IPTV services, this section first analyzes the channel characteristics, where the probability distribution of the channel states will be calculated. Considering the temporal dependence between frames, some frame- and sequence-level parameters will be presented to indicate the influence of the direct packet loss and error propagation on the video quality. A parametric planning model is finally proposed to evaluate the video quality.

A. Four-State Markov Channel Model Analysis

With respect to the parametric planning model, the characteristics of the channel (e.g., packet loss rate, burst packet loss rate, channel state transition probabilities) are known a priori in the planning phase. Generally, the finite-state Markov model is a good approximation of the actual packet loss processes for both wired and wireless channels [14]. However, the more elaborate division of channel states will increase the complexity of analyzing the finite-state Markov model and calculating the parameters of planning model. Therefore, it should make a trade-off between the accuracy and complexity of channel simulation [22].

According to the study in [15], the good and bad state run length distribution often tends to behave like a mixture of two geometric distributions. This characteristic just coincides with that of four-state Markov model (4SMM). Moreover, the transition probabilities in 4SMM can be established without having to run extensive physical layer simulations, thus it is comparatively easy to implement and analyze for 4SMM. Similar to the P.1201 and G.1071 models, this paper uses a 4SMM to emulate the sophisticated packet loss process [16], [23], as illustrated in Fig. 3. It is noted that in the 4SMM, two period types are involved, namely, the burst period and the gap period. The burst period, as discussed in [16], is defined as a longest sequence beginning and ending with a loss during which the number of consecutive received packets is less than a specified value (This value varies with the network and service scenarios, and an example of the suitable value for video service is 64 according to [16]). The other periods are classified as the gap periods. These two periods appear alternatively. In each period, there are two kinds of states corresponding to the fact that the packet is received and lost, respectively. Provided with the period to

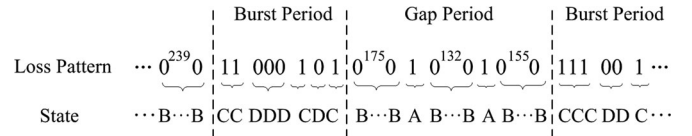


Fig. 4. Illustration of packet loss in the four-state Markov model by a binary sequence.

which the packet belongs, the four states of the 4SMM model are defined as:

1. State A: isolated packet lost in a gap period
2. State B: packet received successfully in a gap period
3. State C: packet lost in a burst period
4. State D: packet received in a burst period

It can be found from Fig. 3 that there are several rules for the transfer between different states. For example, State A can be transferred only to State B, while State B can be transferred to State A and C as well as to itself. State C can be transferred not only to State B and D, but also to itself. State D can be transferred to State C and itself, respectively.

Given a loss pattern of the 4SMM, the burst and gap periods are detected firstly according to their definitions. Then, the state for each packet is clarified by checking the packet is received or lost in a specified period. Fig. 4 illustrates the four states by a binary sequence where the value 1 indicates that current packet is lost and the value 0 denotes that current packet is correctly received. The values of 239, 175, 132 and 155 indicate the number of consecutive 0. It is obvious that each state is easily distinguished from other states and the state transitions obey the rules specified in Fig. 3.

For a specific 4SMM, as illustrated in Fig. 3, the matrix of transition probability Π can be expressed as

$$\Pi = \begin{bmatrix} p_{AA} & p_{AB} & p_{AC} & p_{AD} \\ p_{BA} & p_{BB} & p_{BC} & p_{BD} \\ p_{CA} & p_{CB} & p_{CC} & p_{CD} \\ p_{DA} & p_{DB} & p_{DC} & p_{DD} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ g & h & f & 0 \\ 0 & i & j & k \\ 0 & 0 & m & n \end{bmatrix} \quad (2)$$

where $p_{i,j}$ ($i, j \in \{A, B, C, D\}$) is the state transition probability between two states. It can be indicated as follows [24]:

$$p_{ij} = \Pr(X_1 = j | X_0 = i) \quad (3)$$

where \Pr is the probability, and X_0 and X_1 are random variables. This channel model can be characterized by the parameters g, h, f, i, j, k, m, n . The values of these parameters can be determined using the maximum likelihood estimators for a sample trace [25]. Taking the parameter g as an example, it indicates the state transition probability from State B to State A. The maximum likelihood estimator of g for a sample trace is

$$\hat{g} = \frac{n_{BA}}{n_B} \quad (4)$$

where n_{BA} is the number of times in the observed time series that state B follows state A and n_B is the number of times state B occurs in the trace. The values of other channel parameters (h, f, i, j, k, m, n) can be determined using the same method. Particularly, there are constant relationships among these parameters

[23], which can be expressed as follows:

$$h = 1 - f - g \quad (5)$$

$$k = 1 - i - j \quad (6)$$

$$n = 1 - m. \quad (7)$$

Therefore, the channel model can be determined when the values of g, f, i, j, m are provided. For the parametric planning model, the values of these parameters are assigned in advance and can be used as an input [18], [23].

Considering that the states of 4SMM constitute an irreducible closed set, the stationary probabilities of State A, B, C and D (i.e., P_A, P_B, P_C, P_D) can be used to indicate the random packet loss rate, error free rate, burst packet loss rate, and error free rate in a burst period of the video stream, respectively. These stationary probabilities form the stationary probability distribution π . It is a (row) vector that keeps unchanged after being applied with the operation of transition matrix $\mathbf{\Pi}$ [14], [24], which can be expressed as

$$\pi \cdot \mathbf{\Pi} = \pi \quad (8)$$

$$\pi = [P_A, P_B, P_C, P_D] \quad (9)$$

where P_A, P_B, P_C, P_D are the stationary probabilities of State A, B, C and D, respectively. They are non-negative and sum to 1. According to (2), (8), and (9), the values of P_A, P_B, P_C and P_D can be achieved by

$$P_A = \frac{mgj}{(m+k) \cdot f + (1+g) \cdot mj} \quad (10)$$

$$P_B = \frac{mj}{(m+k) \cdot f + (1+g) \cdot mj} \quad (11)$$

$$P_C = \frac{mf}{(m+k) \cdot f + (1+g) \cdot mj} \quad (12)$$

$$P_D = \frac{fk}{(m+k) \cdot f + (1+g) \cdot mj}. \quad (13)$$

Therefore, when the values of channel parameters f, i, j, k , and m are provided, the packet loss rate and the burst packet loss rate will be determined as well. These stationary probabilities of channel states can be used to indicate the distribution of packet loss in the channel.

B. Packet Loss Metrics Affecting Video Quality

Based on the analysis above, several sequence- and frame-level parameters are then derived to evaluate the video quality affected by the packet loss. The sequence-level parameters include the average frame loss frequency (AFLF) and the expectation of the number of impaired frames (ENIF). The frame-level parameter contains the expectation of the impaired ratio of a frame (EIRF). More specifically, AFLF indicates the expected number of frames with packet loss per GOP. ENIF is the expected number of the impaired frames caused by the error propagation for each packet loss, as illustrated in Fig. 5. EIRF is the expectation of the impaired ratio of each frame suffering from the packet loss. These parameters can describe the impact of

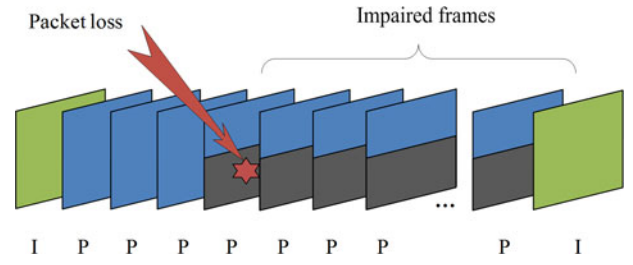


Fig. 5. Illustration of influence of packet loss on video streaming.

the direct packet loss and error propagation on video streaming from different prospects.

When the video sequences are coded at a low bit-rate, the number of bits for coding each frame is relatively small, and each frame is usually encapsulated into one packet for transmission. On the other hand, when the video sequences are coded at a high bit-rate, the number of bits for coding each frame is relatively large, and each frame is usually encapsulated into multiple packets for transmission. It is necessary to distinguish these two conditions because different patterns of packet encapsulation lead to different results of parameter calculation. Given a specific video service, the average number of packets in each frame can be determined by the ratio of B_F and packet size. In the rest of this subsection, these parameters will be calculated when each frame is encapsulated into one packet and into multiple packets, respectively.

1) *One Packet Per Frame*: When each frame is encapsulated into one packet for transmission, the packet loss will directly lead to the frame loss. In such a case, the frame loss process is the same as the packet loss process, which can be characterized via a four-state Markov model as well.

a) *AFLF*: For the video streaming suffering from packet loss, the value of AFLF V_{AFLF} can be calculated using the frame loss rate and the length of GOP, which is expressed as

$$V_{AFLF} = P_F \cdot L_G = P_P \cdot L_G = (P_A + P_C) \cdot L_G \quad (14)$$

where L_G is the length of GOP, P_F is the frame loss rate, P_P is the packet loss rate that is equal to the sum of the random packet loss rate P_A and burst packet loss rate P_C . Because each frame is encapsulated in one packet, $P_F = P_P$.

b) *ENIF*: When the i^{th} frame in a GOP is subject to errors caused by packet loss, its subsequent frames are usually error-prone because the i^{th} frame may be used as the reference. This contamination will not stop until the next synchronization point, typically an I-frame, is reached. Thus, the number of impaired frames is determined by the GOP length and the position of the lost frame.

For a video with the GOP length L_G , as shown in Fig. 6, there are a total of L_G loss patterns to be considered. More specifically, the value 1 indicates that the frame is affected by direct packet loss or error propagation, while the value 0 means that the frame is intact and not influenced by the packet loss. Taking the loss pattern of index 2 in Fig. 6 as an example, the second frame in a GOP is suffered from packet loss, and the total number of impaired frames is $L_G - 1$. In the Markov channel, the packet loss event exhibits dependencies over time [14]. Thus,

| Index | Loss Pattern | Probability | Impaired Frames |
|-----------|---|--|-----------------|
| 1 | 1 1 | $(P_A + P_C)/(1 - P_{Bh}^{L_G-1} - P_{Dn}^{L_G-1})$ | L_G |
| 2 | 0 1 1 | $(P_B(g+f) + P_{Dn})/(1 - P_{Bh}^{L_G-1} - P_{Dn}^{L_G-1})$ | $L_G - 1$ |
| 3 | 0 0 1 1 | $(P_{Bh}(g+f) + P_{Dnm})/(1 - P_{Bh}^{L_G-1} - P_{Dn}^{L_G-1})$ | $L_G - 2$ |
| 4 | 0 0 0 1 1 | $(P_{Bh}^2(g+f) + P_{Dn}^2 m)/(1 - P_{Bh}^{L_G-1} - P_{Dn}^{L_G-1})$ | $L_G - 3$ |
| ⋮ | ⋮ | ⋮ | ⋮ |
| $L_G - 1$ | 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 | $(P_{Bh}^{L_G-3}(g+f) + P_{Dn}^{L_G-3} m)/(1 - P_{Bh}^{L_G-1} - P_{Dn}^{L_G-1})$ | 2 |
| L_G | 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 | $(P_{Bh}^{L_G-2}(g+f) + P_{Dn}^{L_G-2} m)/(1 - P_{Bh}^{L_G-1} - P_{Dn}^{L_G-1})$ | 1 |

L_G

Fig. 6. Loss pattern in a GOP of a video sequence.

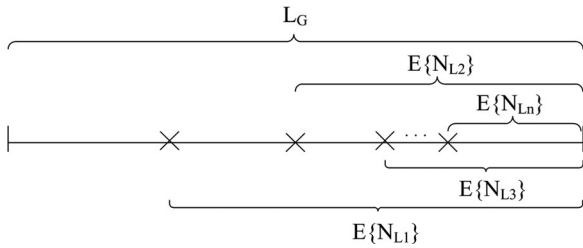


Fig. 7. Illustration of ENIF for each packet loss.

the probability of each loss pattern should be calculated using the stationary probabilities of different states and corresponding transition probabilities. Based on this analysis, the expectation of the impaired frame number V_{ENIF} for a single packet loss is calculated as

$$\begin{aligned}
 V_{ENIF} &= E\{N_L\} = \sum_{k=1}^{L_G} P_L(k) \cdot N_L(k) \\
 &= \left(L_G - \frac{P_B(1 - h^{L_G})}{1 - h} - \frac{P_D(1 - n^{L_G})}{1 - n} \right) \\
 &\quad \times (1 - P_B h^{L_G-1} - P_D n^{L_G-1})^{-1} \quad (15)
 \end{aligned}$$

where $N_L(k)$ and $P_L(k)$ are the number of the impaired frames and the probability of the loss pattern for the k^{th} index, respectively, as illustrated in Fig. 6.

When multiple packet losses occur in a GOP, the high frame loss frequency will lead to low expectation of the impaired frame number for each packet loss. As illustrated in Fig. 7, the values of EN_{L1} , EN_{L2} and EN_{L3} indicate the expectations of the impaired frame numbers for different packet loss events in a GOP. Obviously, $L_G > EN_{L1} > EN_{L2} > EN_{L3} > \dots > EN_{Ln}$. The relationship among these values can be expressed as follows:

$$\eta = \frac{E\{N_{L1}\}}{L_G} = \frac{E\{N_{L2}\}}{E\{N_{L1}\}} = \frac{E\{N_{L3}\}}{E\{N_{L2}\}} = \dots = \frac{E\{N_{Ln}\}}{E\{N_{L_{n-1}}\}} \quad (16)$$

where n is number of frame loss in a GOP and it is equal to V_{AFLF} , η is a constant determined by the ratio of the expectation of the impaired frame number for the first packet loss EN_{L1} and L_G , where EN_{L1} is equal to EN_L and can be calculated by (15). For the n^{th} loss packet loss in the GOP, the value of expectation number of the impaired frames EN_{Ln} can

be derived from (15) as

$$E\{N_{Ln}\} = \eta^{n-1} \cdot E\{N_{L1}\}. \quad (17)$$

In such a case, the value of V_{ENIF} is assumed as the mean value of all EN_{Ln} ($n = 1, 2, 3 \dots V_{AFLF}$) in a GOP, which can be expressed as

$$\begin{aligned}
 V_{ENIF} &= \frac{1}{V_{AFLF}} \sum_{n=1}^{V_{AFLF}} E\{N_{Ln}\} \\
 &= E\{N_{L1}\} \cdot \frac{1 - \eta^{V_{AFLF}}}{(1 - \eta) \cdot V_{AFLF}}. \quad (18)
 \end{aligned}$$

c) *EIRF*: If each frame is encapsulated into one packet, the packet loss will lead to the whole frame loss. In such a case, the expectation of the impaired ratio in a frame $V_{EIRF} = 1$.

2) *Multiple Packets Per Frame*: When each frame is encapsulated into multiple packets for transmission, the issue of burst packet loss may be actually less relevant between frames. This is because more packets per frame will make the burst packet loss concentrate on affecting individual frames, rather than spreading across multiple frames [14]. In such a case, the frame loss process can be simplified and characterized via a Bernoulli model [26], which is independent and identically distributed. Next, the proposed parameters will be calculated under the condition that each frame is encapsulated into multiple packets.

a) *AFLF*: If one of the packets of a frame is lost, this frame is defined as a lossy frame. For a frame with the number of V_{PpF} packets, the probability of the lossy frame can be calculated as

$$P_F = 1 - (P_B \cdot h^{V_{PpF}-1} + P_D \cdot n^{V_{PpF}-1}) \quad (19)$$

where the value of V_{PpF} can be calculated by the ratio of B_F and packet size. The value of $P_B \cdot h^{V_{PpF}-1}$ indicates the probability that all packets of current frame are in State B (error free), and the value of $P_D \cdot n^{V_{PpF}-1}$ indicates the probability that all packets of current frame are in State D (error free in a burst period). Thus, the value of $1 - P_B \cdot h^{V_{PpF}-1} - P_D \cdot n^{V_{PpF}-1}$ is the probability of a frame suffering from packet loss. Accordingly, the value of AFLF V_{AFLF} is achieved by

$$V_{AFLF} = P_F \cdot L_G. \quad (20)$$

b) *ENIF*: To calculate the value of ENIF, the probability of each loss pattern and the number of impaired frames under

| Index | Loss Pattern | Probability | Impaired Frames |
|----------|---|--|-----------------|
| 1 | 1 1 | $P_F/(1-(1-P_F)^{L_G})$ | L_G |
| 2 | 0 1 1 | $P_F(1-P_F)/(1-(1-P_F)^{L_G})$ | L_G-1 |
| 3 | 0 0 1 1 | $P_F(1-P_F)^2/(1-(1-P_F)^{L_G})$ | L_G-2 |
| 4 | 0 0 0 1 1 | $P_F(1-P_F)^3/(1-(1-P_F)^{L_G})$ | L_G-3 |
| \vdots | \vdots | \vdots | \vdots |
| L_G-1 | 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 | $P_F(1-P_F)^{L_G-2}/(1-(1-P_F)^{L_G})$ | 2 |
| L_G | 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 | $P_F(1-P_F)^{L_G-1}/(1-(1-P_F)^{L_G})$ | 1 |

L_G

Fig. 8. Loss pattern in a GOP of a video sequence.

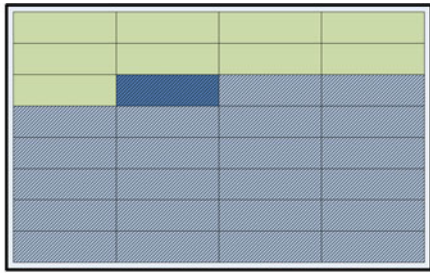


Fig. 9. Illustration of packet loss in one frame.

each loss pattern should be determined. Fig. 8 gives the probability of each loss pattern and the corresponding number of impaired frames for single packet loss. It can be found that the probability of each loss pattern is different from that in Fig. 6. This is because the probability distributions of the frame loss in these two conditions are different from each other. Based on the analysis above, the value of ENIF V_{ENIF} can be calculated by accumulating the product of the number of impaired frames for each pattern and the corresponding probability as follows:

$$\begin{aligned}
 V_{ENIF} &= E\{N_L\} = \sum_{k=1}^{L_G} P_L(k) \cdot N_L(k) \\
 &= \sum_{i=1}^{L_G} (L_G - i + 1) \frac{P_F(1 - P_F)^{i-1}}{1 - (1 - P_F)^{L_G}} \\
 &= \frac{L_G}{1 - (1 - P_F)^{L_G}} - \frac{1 - P_F}{P_F}. \quad (21)
 \end{aligned}$$

According to the L' Hospital's rule [27], the value of ENIF is equal to 0 when P_F is 0. This result indicates that if there is no packet loss occurs in the video streaming, the number of the impaired frames is equal to 0.

When multiple packet losses occur in a GOP, the high frame loss frequency will also lead to low expectation of the impaired frame number for each packet loss. In such a case, the average impaired frame number of each loss should be regularized by the frame loss frequency V_{AFLF} , which can refer to (18).

c) *EIRF*: When a packet of the frame is lost, as illustrated in Fig. 9, the successive packets of this frame have to be discarded because they usually cannot be decoded correctly.

Therefore, the impaired ratio of the frame is closely related to the number of the packets in a frame and the position of the packet loss occurs in the frame.

Considering the temporal dependency, the packet loss process in a frame can be characterized by a Markov model as well. Given a frame with V_{PpF} packets, there are V_{PpF} packet loss patterns to be considered. The expectation of the impaired ratio of the frame V_{EIRF} can be calculated by accumulating the product of the impaired ratio of the frame under each loss pattern and its probability as follows:

$$\begin{aligned}
 V_{EIRF} &= E\{R_E\} = \sum_{k=1}^{V_{PpF}} P_E(k) \cdot R_E(k) \\
 &= \frac{1}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}} \\
 &\quad \cdot \left(1 - P_B \frac{1 - h^{V_{PpF}}}{V_{PpF}(1 - h)} - P_D \frac{1 - n^{V_{PpF}}}{V_{PpF}(1 - n)} \right) \quad (22)
 \end{aligned}$$

where $R_E(k)$ is the impaired ratio of the frame under a k^{th} loss pattern, $P_E(k)$ is the corresponding probability of the packet loss pattern. Particularly, when V_{PpF} is equal to 1, the value of EIRF calculated by (22) is also equal to 1. It indicates that if each frame is encapsulated into one packet to transmit, the packet loss will lead to a whole frame loss.

C. Video Quality Assessment Model

Based on the analysis above, it can be found that apart from the basic statistical information (e.g., bit-rate and packet loss rate), other parameters (such as the V_{AFLF} , V_{ENIF} , and V_{EIRF}) are useful for reflecting the packet or frame loss behaviors and can be obtained by analyzing the characteristics of the channel and video coding. However, for the parametric planning model, how to find the relationship between these parameters and the video quality is also a key issue to be solved.

With respect to the video transmitted over the network, the video quality will be affected by the coding distortion and the distortion caused by packet loss. As mentioned in Section II, the video coding quality has been estimated using the parameters such as the coded bit-rate. Here, the focus is on the evaluation of the distortion caused by packet loss. For the video streaming with packet loss, the normalized distortion caused by packet

TABLE III
RESULTS OF ONE-WAY ANOVA TEST

| D_l and V_{AFLF} | | | | | |
|----------------------|----------------|-----|-------------|--------|------|
| | Sum of Squares | df | Mean Square | F | Sig. |
| Between Groups | 20.627 | 23 | .897 | 20.061 | .000 |
| Within Groups | 18.579 | 408 | .046 | | |
| Total | 39.206 | 431 | | | |
| D_l and V_{ENIF} | | | | | |
| | Sum of Squares | df | Mean Square | F | Sig. |
| Between Groups | 16.449 | 6 | 2.741 | 51.913 | .000 |
| Within Groups | 22.757 | 425 | .054 | | |
| Total | 39.206 | 431 | | | |
| D_l and V_{EIRF} | | | | | |
| | Sum of Squares | df | Mean Square | F | Sig. |
| Between Groups | 10.249 | 8 | 1.281 | 15.364 | .000 |
| Within Groups | 28.957 | 423 | .068 | | |
| Total | 39.206 | 431 | | | |

loss D_l can be expressed as

$$D_l = \frac{Q_c - Q_l}{Q_c - 1} \quad (23)$$

where Q_l is the video quality affected by packet loss. It can be found from (23) that the value of Q_l is closely related to D_l when Q_c is known.

In order to estimate the value of D_l for a given network scenario, the relationship between D_l and V_{AFLF} , V_{ENIF} , V_{EIRF} should be clarified, respectively. The one-way analysis of variance (ANOVA) tests are performed to check their relationship. The corresponding F-values are 20.06, 51.91, and 15.36. All p-values of the F-test results are smaller than 0.01 at 95% level, as illustrated in Table III. It indicates that the proposed parameters have a significant positive correlation with D_l . Here, a specific function D_l (V_{AFLF} , V_{ENIF} , V_{EIRF}) is defined. The monotonicity and the convergence of the function should meet the following requirements:

- The value of D_l should increase with the raise of V_{AFLF} , V_{ENIF} , and V_{EIRF} , respectively. This is because the larger values of V_{AFLF} , V_{ENIF} , and V_{EIRF} indicate more packet loss, more frames affected by error propagation, and severer error in each frame, respectively.
- If there is no packet loss in the transmission, the values of V_{AFLF} , V_{ENIF} , and V_{EIRF} are all equal to 0. In such a case, the value of D_l is also equal to 0.
- If the values of V_{AFLF} , V_{ENIF} , and V_{EIRF} are considerably large, the value of D_l should be approximated to the maximum value 1.

Correspondingly, an evaluation model for the video distortion caused by packet loss is proposed to satisfy the above constraints

$$D_l = 1 - \exp(-v_5 \cdot V_{AFLF}^{v_6} \cdot V_{ENIF}^{v_7} \cdot V_{EIRF}^{v_8}) \quad (24)$$

where v_5 , v_6 , v_7 and v_8 are empirical parameters obtained by regression. Submitting (23) into (22), the video quality affected

TABLE IV
NOTATIONS FOR THE PARAMETRIC PLANNING MODEL

| Notations | Parameters |
|--------------------------|--|
| Π | transition matrix |
| π | stationary probability distribution |
| L_G | length of GOP |
| V_{PPF} | average number of packets in a frame |
| V_{AFLF} | average frame loss frequency |
| V_{ENIF} | expectation of the number of impaired frames |
| V_{EIRF} | expectation of the impaired ratio of a frame |
| D_l | normalized distortion caused by packet loss |
| Q_l | video quality affected by packet loss |
| g, h, f, i, j, k, m, n | state transition probability |

by packet loss can be calculated as follows:

$$Q_l = 1 + (Q_c - 1) \cdot \exp(-v_5 \cdot V_{AFLF}^{v_6} \cdot V_{ENIF}^{v_7} \cdot V_{EIRF}^{v_8}). \quad (25)$$

Until now, the parametric planning model has been established, where the channel and video characteristics are combined together to evaluate the video quality. Table IV lists all the notations of this section. Considering the value of Q_c is estimated by B_F and F_R in (1), there are five parameters in the proposed parametric planning model, namely B_F , F_R , V_{AFLF} , V_{ENIF} , and V_{EIRF} . The parameters V_{AFLF} , V_{ENIF} , and V_{EIRF} can be further estimated from the probability distribution of the channel states (state transition probability). It considers the burst loss, the error propagation and the correlation between individual packet loss events.

IV. EXPERIMENTAL SETTINGS AND RESULTS

The experiments in this paper consist of two sessions: the training session and the validation session. For the training session, it mainly focuses on training the coefficients of the proposed model. For the validation session, it is used to check the performance of the proposed planning model.

A. Experiment Settings for Training

In the training session, four standard video sequences with different contents were employed. The involved sequences included: *CrowdRun*, *DucksTakeOff*, *Stockholm*, and *BasketballDrive*. Each video sequence with length of 10 seconds has three resolutions: QVGA, HVGA and 720P. All sequences were compressed by the FFmpeg 0.4.9 codec with x.264 library [28]. The coding structure was ‘‘IPPP’’. The number of slices per frame was 1. The length of GOP was 60 and the number of reference frames was 1. The detailed coding information such as the coded bit-rate and the frame rate are listed in Table V.

To simulate the packet loss process in IP networks, the compressed bit streams were packetized following the format defined by the RTP payload format for H.264 [29], where the maximum size of the packet was set as 1500 bytes. A four-state Markov model was employed to simulate the packet loss distribution. The channel state transition probabilities i, j, k, m, n were set at 0.3, 0.65, 0.05, 0.25, 0.75, respectively, which were recommended by ITU Study Group 12 for network video

TABLE V
EXPERIMENTAL SETTINGS FOR THE TRAINING SET

| Index | Bitrate (kbps) | Frame rate | Packet loss rate | Burst packet loss rate |
|-------|-----------------------|------------|-------------------------|----------------------------------|
| QVGA | 128, 192 | 15 | 0, 0.5%, 1%, 2%, 3%, 5% | 0, 0.15%, 0.3%, 0.6%, 0.9%, 1.5% |
| HVGA | 384, 768 | 30 | | |
| | 256 | 15 | | |
| 720P | 512, 768, 1536 | 30 | | |
| | 512, 1024, 2048, 4096 | 30 | | |

services [23]. The value of f was equal to g and the changes of f and g led to different packet loss rates. Here, the values of f and g were set at 0.0012, 0.0023, 0.0047, 0.0072 and 0.0122, which correspond to the packet loss rates 0.5%, 1%, 2%, 3% and 5%, and the burst packet loss 0.15%, 3%, 0.6%, 0.9% and 1.5%, respectively, as illustrated in Table V. All sequences were decoded using H.264 decoder FFmpeg 0.4.9 with the zero motion error concealment techniques. When a packet in a frame was lost, the successive packets of this frame were all discarded as well. To enable decoding, the first packet (in the instantaneous data refresh frame) and the last packet (the loss of this packet cannot be detected by the decoder) were not dropped in the experiment.

In order to obtain the video quality, a subjective test was carried out following the guide-lines specified by ITU-T recommendation P.913 [30]. The test environment was controlled, where the room luminosity was between 100 and 200 Lux and the noise of the laboratory was 40-50 dB. The monitor used for display was 22-inch LCD Flat Panel, with a resolution of 1920×1080 pixels. The videos were displayed at their native resolution to prevent any distortions caused by scaling operations. The viewing distance was set between 4H-5H (H is the picture height).

A total of 25 non-expert viewers participated in this subjective test, including 12 females and 13 males. All participants were university students aged between 23 and 28 years, and they were screened for visual acuity and color blindness. Subjective video quality was assessed using a single-stimulus presentation method and a 5-point absolute category rating (ACR) scale [30]. The duration of the subjective test was limited to 30 minutes to prevent eye strain and fatigue. Before the formal test, the views were asked to watch four examples to get familiar with the rating process. During the formal test, the users were instructed to watch each video clip once, and to rate the video quality immediately after watching it. All videos in the experiment were viewed by each subject. After the test, the users were screened with regard to the accuracy of their rating values. The standard exclusion procedures were followed as specified in [31]. After this screening process, the rating samples of 2 out of 25 subjects (8%) were discarded. The video quality of each video sequence was finally obtained by averaging all the 23 subject's rating results, also known as the mean opinion score (MOS) [30]. There were 288 video quality scores for 288 video clips and all these test data constituted the training data set TR.

Operational coefficients given in Table VI were trained using the data in TR. The values of v_1 , v_2 , v_3 and v_4 were obtained by

TABLE VI
COEFFICIENTS IN THE PROPOSED METHOD

| Index | v_1 | v_2 | v_3 | v_4 | R^2 | RMSE |
|-------|-------|-------|-------|-------|-------|-------|
| QVGA | 3.75 | 1.07 | 2.36 | 0.20 | 0.723 | 0.363 |
| HVGA | 3.79 | 1.11 | 2.17 | 0.21 | 0.748 | 0.352 |
| 720P | 3.82 | 1.16 | 2.04 | 0.25 | 0.763 | 0.348 |
| Index | v_5 | v_6 | v_7 | v_8 | R^2 | RMSE |
| QVGA | 0.32 | 1.21 | 0.04 | 1.69 | 0.630 | 0.561 |
| HVGA | 0.26 | 0.96 | 0.04 | 1.52 | 0.664 | 0.542 |
| 720P | 0.72 | 1.23 | 0.03 | 2.21 | 0.688 | 0.538 |

TABLE VII
EXPERIMENTAL SETTINGS FOR VALIDATION SET

| Index | Bitrate (kbps) | Frame rate | Packet loss rate | Burst packet loss rate |
|-------|----------------------|------------|-------------------------|----------------------------------|
| QVGA | 96, 128 | 15 | 0, 0.5%, 1%, 2%, 3%, 5% | 0, 0.15%, 0.3%, 0.6%, 0.9%, 1.5% |
| HVGA | 256, 512 | 30 | | |
| | 384, 512, 832, 1600 | 30 | | |
| 720P | 512, 768, 1536, 3840 | 30 | | |

training the coded data without packet loss in TR, in accordance with the method proposed in [19]. More specifically, the values of v_1 , v_2 , v_3 were fitted using the coded data in TR with the frame rate of 30fps by the least square error fitting. Then, the value of v_4 was fitted using the coded data with frame rate of 15 fps. Considering the effect of packet loss, the value of D_l for the data with packet loss in TR was calculated according to (23). Then, the coefficients v_5 to v_8 were trained using the values of D_l , V_{AFLF} , V_{ENIF} , and V_{EIRF} by the least square error fitting. It should be noted that the coefficients under different video resolutions should be trained separately.

B. Experimental Settings for Validation

In the validation session, the performance of the proposed method was checked using six sequences including: *ParkRun*, *Riverbed*, *IntoTree*, *Sunflower*, *ParkJoy*, and *Touchdown*. The detail experiment settings are listed in Table VII. The procedures of the subjective test were the same as those in the training test. All the subjective rating values constituted the validation data set VL.

Three metrics suggested by VQEG [32] were employed in this work. They were Pearson correlation coefficient (PCC) for linearity, root-mean-squared error (RMSE) for accuracy, and Spearman rank order correlation coefficient (SROCC) for monotonicity. The maximum value of PCC and SROCC is 1. Generally, the smaller of the RMSE and the larger of the PCC and SROCC indicate the better performance of the model.

The proposed model was compared with other three parametric planning models [9], [10], [12]. All these models consider the effect of burst packet loss for video quality evaluation. Because the video resolutions are not included in these models, the corresponding model coefficients were re-trained on the

TABLE VIII
PERFORMANCE OF EACH MODEL

| Index | | Proposed Model | Yamagishi [9](VQA1) | Garcia [10] (VQA2) | G.1071 [12] (VQA3) |
|-------|-------|----------------|---------------------|--------------------|--------------------|
| QVGA | PCC | 0.809 | 0.714 | 0.733 | 0.760 |
| | RMSE | 0.521 | 0.592 | 0.584 | 0.557 |
| | SROCC | 0.755 | 0.643 | 0.665 | 0.714 |
| HVGA | PCC | 0.824 | 0.814 | 0.787 | 0.822 |
| | RMSE | 0.511 | 0.556 | 0.552 | 0.534 |
| | SROCC | 0.816 | 0.805 | 0.771 | 0.807 |
| 720P | PCC | 0.850 | 0.821 | 0.834 | 0.835 |
| | RMSE | 0.546 | 0.615 | 0.605 | 0.579 |
| | SROCC | 0.861 | 0.785 | 0.786 | 0.797 |
| ALL | PCC | 0.828 | 0.787 | 0.786 | 0.807 |
| | RMSE | 0.519 | 0.588 | 0.581 | 0.568 |
| | SROCC | 0.821 | 0.789 | 0.771 | 0.804 |

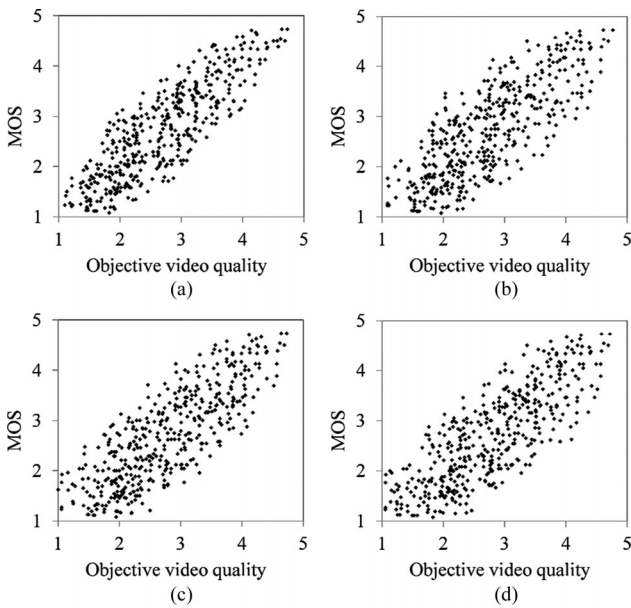


Fig. 10. Scatter plots of the objective and subjective scores evaluated by the different models. (a) Proposed model, (b) VQA1, (c) VQA2, and (d) VQA3.

proposed TR dataset, according to their recommended methods in [9], [10] and [12], respectively. Moreover, for simplicity, the models in [9], [10], and [12] are denoted as VQA1, VQA2 and VQA3, respectively.

Table VIII compares the performance of each model. It is observed that the proposed model has larger PCC and SROCC values and smaller RMSE values than other models under various resolutions. Fig. 10 visualizes the performance of each model by using the scatter plots of the objective and subjective scores in VL dataset. There is a better linear relationship between the predicted video quality by the proposed model and MOS, which indicates that the video quality predicted by the proposed model is more close to the MOS.

To further check the prediction accuracy of the model, the statistical significance analysis and hypothesis test were performed. Firstly, the assumption of Gaussianity of the scores estimated by each objective VQA model was checked using the Kolmogorov-Smirnov test (K-S test). According to the test results by SPSS

TABLE IX
STATISTICAL SIGNIFICANCE MATRIX BASED ON RESIDUAL BETWEEN MOS AND THE PREDICTED VIDEO QUALITY

| Index | | Proposed Model | Yamagishi [9](VQA1) | Garcia [10](VQA2) | G.1071 [12](VQA3) |
|----------------------|--|----------------|---------------------|-------------------|-------------------|
| Proposed Model | | 1.000 | 0.027 | 0.033 | 0.041 |
| Yamagishi [9] (VQA1) | | 0.027 | 1.000 | 0.119 | 0.087 |
| Garcia [10](VQA2) | | 0.033 | 0.119 | 1.000 | 0.052 |
| G.1071 [12] (VQA3) | | 0.041 | 0.087 | 0.052 | 1.000 |

TABLE X
PERFORMANCE PROPOSED MODEL UNDER DIFFERENT DEVIATIONS OF CHANNEL PARAMETERS

| Dev | | -10% | -5% | 5% | 10% | ME | MSD |
|------|-------|-------|-------|-------|-------|------|-------|
| QVGA | PCC | 0.805 | 0.807 | 0.812 | 0.814 | 0.6% | 0.004 |
| | RMSE | 0.513 | 0.517 | 0.523 | 0.525 | 1.5% | 0.005 |
| | SROCC | 0.751 | 0.752 | 0.756 | 0.758 | 0.5% | 0.003 |
| HVGA | PCC | 0.828 | 0.827 | 0.823 | 0.823 | 0.5% | 0.002 |
| | RMSE | 0.524 | 0.516 | 0.506 | 0.507 | 2.5% | 0.007 |
| | SROCC | 0.822 | 0.820 | 0.815 | 0.815 | 0.7% | 0.003 |
| 720P | PCC | 0.846 | 0.849 | 0.852 | 0.853 | 0.7% | 0.003 |
| | RMSE | 0.561 | 0.551 | 0.544 | 0.549 | 2.7% | 0.007 |
| | SROCC | 0.864 | 0.862 | 0.858 | 0.858 | 0.4% | 0.003 |

17.0 [33] (scores for each resolution, $p > 0.05$), the null hypothesis (the scores have a standard normal distribution) could not be reject at the 5% level for any models. Therefore, the assumption of Gaussianity was valid for the scores estimated by all models.

Then, an F-test based on the residuals between MOS and the video quality estimated by different objective VQA models was performed to statistically compare performance of these models. The null hypothesis was that the variance of residuals from one objective video quality evaluation model was statistically indistinguishable with 95% confidence from that of another objective model. It could be expressed as follows:

$$\text{Null Hypothesis} \Leftrightarrow \sigma^2_{MOS-MOSp1} = \sigma^2_{MOS-MOSp2} \quad (26)$$

where $MOSp1$ and $MOSp2$ were the video quality scores estimated by two different VQA models. $\sigma^2_{MOS-MOSp1}$ indicates that the variance of residuals between MOS and $MOSp1$. The results of the statistical significance test are presented in Table IX, where the values are the probabilities when the null hypothesis of equal variances is not rejected. According to the results in Tables VIII and IX, the proposed model demonstrates a significantly superior prediction performance to other models. It can efficiently estimate the video quality for IPTV services which suffered from packet loss.

Considering that there may be a deviation between the actual network parameters and their statistical values, the sensitivity of the proposed parametric planning model to the accuracy of the channel parameters was also checked. Here, the deviations of the channel parameters f , g , i , j , and m were set at -10%, -5%, 0%, 5% and 10%, respectively, as illustrated in Table X. It can be found that, when the deviation of channel

model parameters is in a range from -10% to $+10\%$, the maximum errors (ME) of PCC, RMSE and SROCC between MOS and the predicted video quality are 0.7%, 2.7% and 0.7% (within 95% level), respectively. The maximum standard deviations (MSD) of PCC, RMSE and SROCC are 0.004, 0.007 and 0.003, respectively. However, the average standard deviations of PCC, RMSE and SROCC between the predicted video quality by different methods and MOS are 0.02, 0.032 and 0.021, respectively. The standard deviations of PCC, SROCC and RMSE are comparatively small when the channel model parameters slightly changes. Thus, the performance of the proposed model has certain robustness to the accuracy of the channel model parameters.

V. CONCLUSION

Because the actual video streams cannot be obtained during the planning phase in practice, only little information can be used in the parametric planning model. Consequently, how to accurately estimate the video quality is rather challenging. In this paper, a parametric planning model was proposed for IPTV services. The principal contributions of this work are to check the necessity of combining the channel and video characteristics to evaluate the video quality in the planning phase. Considering the fact that the burst packet loss and the temporal dependence between frames will lead to nonrandom frame distortion, several sequence-level and frame-level parameters for video quality evaluation have been derived from the channel models. Using these parameters, the proposed planning model can better indicate the effects of direct packet loss and error propagation. Experimental results have demonstrated that the proposed model has excellent performance in video quality evaluation and it can be used as an effective tool for network or QoE planning.

It should be noted that the video quality may also be affected by other factors such as different encoding structure, frame or slice type, error control and concealment techniques. Our future work will consider the influence of these factors on the video quality assessment. Moreover, the proposed model can be further extended to interactive video services by taking other factors into consideration, such as the one-way delay and jitter.

APPENDIX

PROOF OF (10)~(13), (15), (21), AND (22)

Proposition 1:

$$P_A = \frac{mgj}{(m+k) \cdot f + (1+g) \cdot mj}$$

$$P_B = \frac{mj}{(m+k) \cdot f + (1+g) \cdot mj}$$

$$P_C = \frac{mf}{(m+k) \cdot f + (1+g) \cdot mj}$$

$$P_D = \frac{fk}{(m+k) \cdot f + (1+g) \cdot mj}$$

Proof:

According to (2) and (5)–(9) of this paper, The numbers of the appendix have been revised following the numbers of text. there are

$$\pi \cdot \mathbf{\Pi} = \pi \quad (27)$$

$$\mathbf{\Pi} = \begin{bmatrix} p_{AA} & p_{AB} & p_{AC} & p_{AD} \\ p_{BA} & p_{BB} & p_{BC} & p_{BD} \\ p_{CA} & p_{CB} & p_{CC} & p_{CD} \\ p_{DA} & p_{DB} & p_{DC} & p_{DD} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ g & h & f & 0 \\ 0 & i & j & k \\ 0 & 0 & m & n \end{bmatrix} \quad (28)$$

$$\pi = [P_A, P_B, P_C, P_D]. \quad (29)$$

Take (28) and (29) into (27), we can obtain that

$$[P_A, P_B, P_C, P_D] \begin{bmatrix} p_{AA} & p_{AB} & p_{AC} & p_{AD} \\ p_{BA} & p_{BB} & p_{BC} & p_{BD} \\ p_{CA} & p_{CB} & p_{CC} & p_{CD} \\ p_{DA} & p_{DB} & p_{DC} & p_{DD} \end{bmatrix} = [P_A, P_B, P_C, P_D].$$

The equation can be expressed as follows:

$$\begin{cases} P_A \cdot p_{AA} + P_B \cdot p_{BA} + P_C \cdot p_{CA} + P_D \cdot p_{DA} = P_A \\ P_A \cdot p_{AB} + P_B \cdot p_{BB} + P_C \cdot p_{CB} + P_D \cdot p_{DB} = P_B \\ P_A \cdot p_{AC} + P_B \cdot p_{BC} + P_C \cdot p_{CC} + P_D \cdot p_{DC} = P_C \\ P_A \cdot p_{AD} + P_B \cdot p_{BD} + P_C \cdot p_{CD} + P_D \cdot p_{DD} = P_D. \end{cases} \quad (30)$$

According to (28), there are

$$\begin{bmatrix} p_{AA} & p_{AB} & p_{AC} & p_{AD} \\ p_{BA} & p_{BB} & p_{BC} & p_{BD} \\ p_{CA} & p_{CB} & p_{CC} & p_{CD} \\ p_{DA} & p_{DB} & p_{DC} & p_{DD} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ g & h & f & 0 \\ 0 & i & j & k \\ 0 & 0 & m & n \end{bmatrix}.$$

Take g, i, j, k, m and n into (30), we can obtain that

$$\begin{cases} P_B g = P_A \\ P_A + P_B h + P_C i = P_B \\ P_B f + P_C j + P_D m = P_C \\ P_C k + P_D n = P_D. \end{cases} \quad (31)$$

Considering that

$$\begin{cases} h = 1 - f - g \\ k = 1 - i - j \\ n = 1 - m \\ P_A + P_B + P_C + P_D = 1. \end{cases} \quad (32)$$

Take (32) into (31) and solve equations, we can obtain that

$$\begin{cases} P_A = \frac{mgj}{(m+k) \cdot f + (1+g) \cdot mj} \\ P_B = \frac{mj}{(m+k) \cdot f + (1+g) \cdot mj} \\ P_C = \frac{mf}{(m+k) \cdot f + (1+g) \cdot mj} \\ P_D = \frac{fk}{(m+k) \cdot f + (1+g) \cdot mj} \end{cases}$$

To calculate the values of V_{ENIF} and V_{EIRF} , there are two cases to be considered.

Case 1: When each frame is encapsulated into one packet for transmission, the packet loss will directly lead to the frame loss. In such a case, the value of V_{ENIF} can be derived as follows:

Proposition 2:

$$\begin{aligned} V_{ENIF} &= E\{N_L\} = \sum_{k=1}^{L_G} P_L(k) \cdot N_L(k) \\ &= \left(L_G - \frac{P_B(1-h^{L_G})}{1-h} - \frac{P_D(1-n^{L_G})}{1-n} \right) \\ &\quad \times (1 - P_B h^{L_G-1} - P_D n^{L_G-1})^{-1}. \end{aligned}$$

Proof: Equation (33) as shown at the bottom of this page. Let

$$S_{n1} = \sum_{t=1}^{L_G-1} h^{t-1} (L_G - t), \quad S_{n2} = \sum_{r=1}^{L_G-1} n^{r-1} (L_G - r).$$

There is

$$\begin{aligned} S_{n1} &= \sum_{t=1}^{L_G-1} h^{t-1} (L_G - t) \\ &= (L_G - 1) + h(L_G - 2) + h^2(L_G - 3) + \dots + h^{L_G-2} \end{aligned} \quad (34)$$

$$\begin{aligned} h \times S_{n1} &= \sum_{t=1}^{L_G-1} h^t (L_G - t) \\ &= h(L_G - 1) + h^2(L_G - 2) + h^3(L_G - 3) \\ &\quad + \dots + h^{L_G-1}. \end{aligned} \quad (35)$$

Subtract (34) from (35), it can obtain that

$$\begin{aligned} (h-1) \times S_{n1} &= -(L_G - 1) + h + h^2 + \dots \\ &\quad + h^{L_G-2} + h^{L_G-1} = -L_G + \frac{1-h^{L_G}}{1-h} \\ S_{n1} &= \frac{L_G}{1-h} - \frac{1-h^{L_G}}{(1-h)^2}. \end{aligned} \quad (36)$$

Considering that S_{n1} and S_{n2} have the same form, therefore S_{n2} can be derived following the above steps as well

$$S_{n2} = \frac{L_G}{1-n} - \frac{1-n^{L_G}}{(1-n)^2}. \quad (37)$$

Take (36) and (37) into (33), it can obtain that

$$\begin{aligned} V_{ENIF} &= \frac{(P_A + P_C)L_G + P_B(g+f)S_{n1} + P_D m S_{n2}}{1 - P_B h^{L_G-1} - P_D n^{L_G-1}} \\ &= \left((P_A + P_C)L_G + P_B(g+f) \left(\frac{L_G}{1-h} - \frac{1-h^{L_G}}{(1-h)^2} \right) \right. \\ &\quad \left. + P_D m \left(\frac{L_G}{1-n} - \frac{1-n^{L_G}}{(1-n)^2} \right) \right) \\ &\quad \times \frac{1}{1 - P_B h^{L_G-1} - P_D n^{L_G-1}}. \end{aligned}$$

Because $h = 1 - f - g$, and $n = 1 - m$

$$\begin{aligned} V_{ENIF} &= \left((P_A + P_C)L_G + P_B(1-h) \left(\frac{L_G}{1-h} - \frac{1-h^{L_G}}{(1-h)^2} \right) \right. \\ &\quad \left. + P_D(1-n) \left(\frac{L_G}{1-n} - \frac{1-n^{L_G}}{(1-n)^2} \right) \right) \end{aligned}$$

$$\begin{aligned} V_{ENIF} &= E\{N_L\} = \sum_{k=1}^{L_G} P_L(k) \cdot N_L(k) \\ &= P_L(1) \cdot N_L(1) + P_L(2) \cdot N_L(2) + P_L(3) \cdot N_L(3) + \dots + P_L(L_G) \cdot N_L(L_G) \\ &= \frac{(P_A + P_C)}{1 - P_B h^{L_G-1} - P_D n^{L_G-1}} \cdot L_G + \frac{P_B(g+f) + P_D m}{1 - P_B h^{L_G-1} - P_D n^{L_G-1}} \cdot (L_G - 1) \\ &\quad + \frac{P_B h(g+f) + P_D n m}{1 - P_B h^{L_G-1} - P_D n^{L_G-1}} \cdot (L_G - 2) + \frac{P_B h^2(g+f) + P_D n^2 m}{1 - P_B h^{L_G-1} - P_D n^{L_G-1}} \cdot (L_G - 3) \\ &\quad + \dots + \frac{P_B h^{L_G-2}(g+f) + P_D n^{L_G-2} m}{1 - P_B h^{L_G-1} - P_D n^{L_G-1}} \cdot 1 \\ &= \frac{(P_A + P_C)L_G + P_B(g+f) \sum_{t=1}^{L_G-1} h^{t-1} (L_G - t) + P_D m \sum_{r=1}^{L_G-1} n^{r-1} (L_G - r)}{1 - P_B h^{L_G-1} - P_D n^{L_G-1}} \end{aligned} \quad (33)$$

$$\begin{aligned}
 & \times \frac{1}{1 - P_B h^{L_G-1} - P_D n^{L_G-1}} \\
 = & \left((P_A + P_C) L_G + P_B \left(L_G - \frac{1 - h^{L_G}}{1 - h} \right) \right. \\
 & \left. + P_D \left(L_G - \frac{(1 - n^{L_G})}{1 - n} \right) \right) \\
 & \times \frac{1}{1 - P_B h^{L_G-1} - P_D n^{L_G-1}} \\
 = & \left((P_A + P_B + P_C + P_D) L_G - \frac{P_B (1 - h^{L_G})}{1 - h} \right. \\
 & \left. - \frac{P_D (1 - n^{L_G})}{1 - n} \right) \\
 & \times \frac{1}{1 - P_B h^{L_G-1} - P_D n^{L_G-1}} \\
 = & \left(L_G - \frac{P_B (1 - h^{L_G})}{1 - h} - \frac{P_D (1 - n^{L_G})}{1 - n} \right) \\
 & (1 - P_B h^{L_G-1} - P_D n^{L_G-1})^{-1}.
 \end{aligned}$$

Case 2: When each frame is encapsulated into multiple packets for transmission, the values of V_{ENIF} and V_{EIRF} can be derived as follows:

Proposition 3:

$$\begin{aligned}
 V_{ENIF} &= E \{N_L\} = \sum_{k=1}^{L_G} P_L(k) \cdot N_L(k) \\
 &= \sum_{i=1}^{L_G} (L_G - i + 1) \frac{P_F (1 - P_F)^{i-1}}{1 - (1 - P_F)^{L_G}} \\
 &= \frac{L_G}{1 - (1 - P_F)^{L_G}} - \frac{1 - P_F}{P_F}.
 \end{aligned}$$

Proof:

$$\begin{aligned}
 V_{ENIF} &= E \{N_L\} = \sum_{k=1}^{L_G} P_L(k) \cdot N_L(k) \\
 &= P_L(1) \cdot N_L(1) + P_L(2) \cdot N_L(2) + P_L(3) \cdot N_L(3) \\
 & \quad + \dots + P_L(L_G) a \\
 &= \frac{P_F}{1 - (1 - P_F)^{L_G}} \cdot L_G + \frac{P_F (1 - P_F)}{1 - (1 - P_F)^{L_G}} \cdot (L_G - 1) \\
 & \quad + \frac{P_F (1 - P_F)^2}{1 - (1 - P_F)^{L_G}} \cdot (L_G - 2) \\
 & \quad + \dots + \frac{P_F (1 - P_F)^{L_G-2}}{1 - (1 - P_F)^{L_G}} \cdot 2 + \frac{P_F (1 - P_F)^{L_G-1}}{1 - (1 - P_F)^{L_G}} \cdot 1 \\
 &= \sum_{r=1}^{L_G} (L_G - i + 1) \frac{P_F (1 - P_F)^{i-1}}{1 - (1 - P_F)^{L_G}}. \tag{38}
 \end{aligned}$$

Let

$$S_n = \sum_{i=1}^{L_G} (L_G - i + 1) \cdot (1 - P_F)^{i-1}.$$

There is

$$\begin{aligned}
 S_n &= L_G + (L_G - 1)(1 - P_F) + (L_G - 2) \\
 & \quad \times (1 - P_F)^2 + \dots + (1 - P_F)^{L_G-1} \tag{39} \\
 (1 - P_F) \times S_n &= L_G (1 - P_F) + (L_G - 1)(1 - P_F)^2 \\
 & \quad + (L_G - 2)(1 - P_F)^3 \\
 & \quad + 2 \times (1 - P_F)^{L_G-1} + (1 - P_F)^{L_G}. \tag{40}
 \end{aligned}$$

Subtract (39) from (40), it can obtain that

$$\begin{aligned}
 -P_F \times S_n &= -L_G + (1 - P_F) + (1 - P_F)^2 + \dots \\
 & \quad + (1 - P_F)^{L_G-1} + (1 - P_F)^{L_G} \\
 &= -L_G + \frac{(1 - P_F)(1 - (1 - P_F)^{L_G})}{1 - (1 - P_F)} \\
 S_n &= \frac{L_G}{P_F} - \frac{(1 - P_F)(1 - (1 - P_F)^{L_G})}{P_F^2}. \tag{41}
 \end{aligned}$$

Take (41) into (38), it can obtain that

$$\begin{aligned}
 V_{ENIF} &= \frac{P_F}{1 - (1 - P_F)^{L_G}} \\
 & \quad \times \left(\frac{L_G}{P_F} - \frac{(1 - P_F)(1 - (1 - P_F)^{L_G})}{P_F^2} \right) \\
 &= \frac{L_G}{1 - (1 - P_F)^{L_G}} - \frac{1 - P_F}{P_F}.
 \end{aligned}$$

Proposition 4:

$$\begin{aligned}
 V_{EIRF} &= E \{R_E\} = \sum_{k=1}^{V_{PpF}} P_E(k) \cdot R_E(k) \\
 &= \frac{1}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}} \\
 & \quad \cdot \left(1 - P_B \frac{1 - h^{V_{PpF}}}{V_{PpF}(1 - h)} - P_D \frac{1 - n^{V_{PpF}}}{V_{PpF}(1 - n)} \right).
 \end{aligned}$$

Proof: Equation (42) as shown at the top of the next page.

Let

$$S_{n1} = \sum_{t=1}^{V_{PpF}-1} h^{t-1} (V_{PpF} - t), S_{n2} = \sum_{r=1}^{V_{PpF}-1} n^{r-1} (V_{PpF} - r).$$

There is

$$\begin{aligned}
 S_{n1} &= \sum_{t=1}^{V_{PpF}-1} h^{t-1} (V_{PpF} - t) \\
 &= (V_{PpF} - 1) + h(V_{PpF} - 2) + h^2(V_{PpF} - 3) \\
 & \quad + \dots + h^{V_{PpF}-2} \tag{43}
 \end{aligned}$$

$$\begin{aligned}
V_{EIRF} &= E\{R_E\} = \sum_{k=1}^{V_{PpF}} P_E(k) \cdot R_E(k) \\
&= P_E(1) \cdot R_E(1) + P_E(2) \cdot R_E(2) + P_E(3) \cdot R_E(3) + \dots + P_E(V_{PpF}) \cdot R_E(V_{PpF}) \\
&= \frac{(P_A + P_C)}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}} \times 1 + \frac{P_B(g+f) + P_D m}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}} \times \frac{V_{PpF} - 1}{V_{PpF}} \\
&\quad + \frac{P_B h(g+f) + P_D n m}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}} \times \frac{V_{PpF} - 2}{V_{PpF}} + \frac{P_B h^2(g+f) + P_D n^2 m}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}} \times \frac{V_{PpF} - 3}{V_{PpF}} \\
&\quad + \dots + \frac{P_B h^{V_{PpF}-2}(g+f) + P_D n^{V_{PpF}-2} m}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}} \times \frac{1}{V_{PpF}} \\
&= \frac{P_A + P_C + P_B(g+f) \sum_{t=1}^{V_{PpF}-1} h^{t-1} \frac{(V_{PpF}-t)}{V_{PpF}} + P_D m \sum_{r=1}^{V_{PpF}-1} n^{r-1} \frac{(V_{PpF}-r)}{V_{PpF}}}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}} \quad (42)
\end{aligned}$$

$$\begin{aligned}
h \times S_{n1} &= \sum_{t=1}^{V_{PpF}-1} h^t (V_{PpF} - t) \\
&= h(V_{PpF} - 1) + h^2(V_{PpF} - 2) + h^3(V_{PpF} - 3) \\
&\quad + \dots + h^{V_{PpF}-1}. \quad (44)
\end{aligned}$$

Subtract (43) from (44), it can obtain that

$$\begin{aligned}
(h-1) \times S_{n1} &= -(V_{PpF} - 1) + h + h^2 + \dots + h^{V_{PpF}-1} \\
&= -V_{PpF} + \frac{1 - h^{V_{PpF}}}{1 - h} \\
S_{n1} &= \frac{V_{PpF}}{1 - h} - \frac{1 - h^{V_{PpF}}}{(1 - h)^2}. \quad (45)
\end{aligned}$$

Considering that S_{n1} and S_{n2} have the same form, therefore S_{n2} can be derived following the above steps as well

$$S_{n2} = \frac{V_{PpF}}{1 - n} - \frac{1 - n^{V_{PpF}}}{(1 - n)^2}. \quad (46)$$

Take (45) and (46) into (42), it can obtain that

$$\begin{aligned}
V_{EIRF} &= \frac{P_A + P_C + \frac{P_B(g+f)S_{n1}}{V_{PpF}} + \frac{P_D m S_{n2}}{V_{PpF}}}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}} \\
&= \left(P_A + P_C + \frac{P_B(g+f)}{V_{PpF}} \cdot \left(\frac{V_{PpF}}{1 - h} - \frac{1 - h^{V_{PpF}}}{(1 - h)^2} \right) \right. \\
&\quad \left. + \frac{P_D m}{V_{PpF}} \cdot \left(\frac{V_{PpF}}{1 - n} - \frac{1 - n^{V_{PpF}}}{(1 - n)^2} \right) \right) \\
&\quad \times \frac{1}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}}.
\end{aligned}$$

Because $h = 1 - f - g$, and $n = 1 - m$

$$\begin{aligned}
V_{EIRF} &= \frac{P_A + P_C + P_B + P_D - \frac{P_B(1-h^{V_{PpF}})}{V_{PpF}(1-h)} - \frac{P_D(1-n^{V_{PpF}})}{V_{PpF}(1-n)}}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}} \\
&= \frac{1}{1 - P_B h^{V_{PpF}-1} - P_D n^{V_{PpF}-1}} \\
&\quad \cdot \left(1 - P_B \frac{1 - h^{V_{PpF}}}{V_{PpF}(1 - h)} - P_D \frac{1 - n^{V_{PpF}}}{V_{PpF}(1 - n)} \right).
\end{aligned}$$

REFERENCES

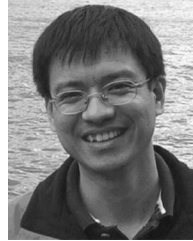
- [1] "Cisco Visual Networking Index: Forecast and methodology, 2015-2020," White Paper, Cisco Systems, Inc., San Jose, CA, USA, 2016. [Online]. Available: <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html>
- [2] D. Rui, G. Liu, and J. Yang, "Utility-based optimized cross-layer scheme for real-time video transmission over HSDPA," *IEEE Trans. Multimedia*, vol. 17, no. 9, pp. 1495-1507, Sep. 2015.
- [3] K. Miller, D. Bethanabhotla, G. Caire, and A. Wolisz, "A control-theoretic approach to adaptive video streaming in dense wireless networks," *IEEE Trans. Multimedia*, vol. 17, no. 8, pp. 1309-1322, Aug. 2015.
- [4] *Quality of Experience Requirements for IPTV Services*, ITU-T Recommendation, G.1080, Dec. 2008.
- [5] A. Raake *et al.*, "IP-based mobile and fixed network audiovisual media services," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 68-79, Nov. 2011.
- [6] A. Takahashi, D. Hands, and O. V. Barriac, "Standardization activities in the ITU for a QoE assessment of IPTV," *IEEE Commun. Mag.*, vol. 46, no. 2, pp. 78-84, Feb. 2008.
- [7] F. Yang and S. Wan, "Bitstream-based quality assessment for networked video: A review," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 203-209, Nov. 2012.
- [8] *Opinion Model for Video-Telephony Applications*, ITU-T Recommendation, G.1070, Apr. 2007.
- [9] K. Yamagishi and T. Hayashi, "Parametric packet-layer model for monitoring video quality of IPTV services," in *Proc. IEEE Int. Conf. Commun.*, May 2008, pp. 110-114.
- [10] M. N. Garcia and A. Raake, "Parametric packet-layer video quality model for IPTV," in *Proc. Int. Symp. Signal Process. Appl.*, 2010, pp. 349-352.
- [11] P. Coverdale, S. Moller, A. Raake, and A. Takahashi, "Multimedia quality assessment standards in ITU-T SG12," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 91-97, Nov. 2011.
- [12] *Opinion Model for Network Planning of Video and Audio Streaming Applications*, ITU-T Recommendation, G.1071, Jun. 2015.
- [13] *Parametric Non-Intrusive Assessment of Audiovisual Media Streaming Quality*, ITU-T Recommendation, P.1201, Oct. 2012.

- [14] Z. Li, J. Chakareski, X. Niu, Y. Zhang, and W. Gu, "Modeling and analysis of distortion caused by Markov-model burst packet losses in video transmission," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 7, pp. 917–931, Jul. 2009.
- [15] Y. Yu and S. L. Miller, "A four-state Markov frame error model for the wireless physical layer," in *Proc. IEEE Int. Conf. Wireless. Commun. Netw.*, Mar. 2007, pp. 2053–2057.
- [16] *Packet Loss Distributions and Packet Loss Models*, ITU Report COM12-D97-E, Jan. 2003.
- [17] S. W. Yuk, M. G. Kang, B. C. Shin, and D. H. Cho, "An adaptive redundancy control method for erasure-code-based real-time data transmission over the internet," *IEEE Trans. Multimedia*, vol. 3, no. 3, pp. 366–374, Sep. 2001.
- [18] *Study Group 12 Call for Proposal on Opinion Model for Audio and Video Streaming Applications (G.O.MVAS)*, ITU-T TSB Circular 32 COM 12/JKK, Apr. 2009.
- [19] F. Yang, J. Song, S. Wan, and H. R. Wu, "Content-adaptive packet-layer model for quality assessment of networked video services," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 6, pp. 672–683, Oct. 2012.
- [20] F. Yang, S. Wan, Q. Xie, and H. R. Wu, "No-reference quality assessment for networked video via primary analysis of bit-stream," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 11, pp. 1544–1554, Nov. 2010.
- [21] Z. Yuan, W. Gao, Y. Lu, Q. Huang, and D. Zhao, "Joint source-channel rate-distortion optimization for H.264 video coding over error-prone networks," *IEEE Trans. Multimedia*, vol. 9, no. 3, pp. 445–454, Apr. 2007.
- [22] A. J. Goldsmith and P. P. Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 868–886, May 1996.
- [23] *Processing Chain for P.NAMS/P.NBAMS Hypothetic Reference Circuits*, ITU Report TD 472rev1 (GEN/12), Jan. 2011.
- [24] H. Wang and N. Moayeri, "Finite-state Markov channel—A useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163–171, Feb. 1995.
- [25] M. Yajnik, S. Moon, J. Kurose, and D. Towsley, "Measurement and modelling of the temporal dependence in packet loss," in *Proc. IEEE Conf. Comput. Commun.*, Mar. 1999, pp. 345–352.
- [26] B. V. Gnedenko, *The Theory of Probability*. New York, NY, USA: Chelsea, 1962.
- [27] A. E. Taylor, "L'Hospital's rule," *Amer. Math. Monthly*, vol. 59, no. 1, pp. 20–24, 1952.
- [28] F. Bellard and M. Niedermayer, "FFmpeg," 2012. [Online]. Available: <http://ffmpeg.org>
- [29] S. Wenger and T. Stockhammer, *RTP Payload Format for H.264 Video*, RFC 3984, 2005.
- [30] *Subjective Video Quality Assessment Methods for Multimedia Applications*, ITU-T Recommendation P.913, 2014.
- [31] *Methodology for the Subjective Assessment of the Quality of Television Pictures*, ITU-R Recommendation BT.500-11, 2006.
- [32] *Objective Perceptual Multimedia Video Quality Measurement in the Presence of a Full Reference*, ITU-T Recommendation J.247, 2008.
- [33] B. S. Kang and K. S. Kim, *SPSS 17.0: Statistical Analysis of the Social Sciences*, Hannarae Academy, Seoul, South Korea, 2009.



Jiarun Song (M'16) received the B.S. degree in telecommunication engineering and the Ph.D. degree in communication and information system from Xidian University, Xi'an, China, in 2009 and 2015, respectively.

He is currently a Postdoctoral Researcher in the State Key Laboratory of Integrated Services Networks, Xidian University. His research interests include QoE, video quality assessment, and multimedia communication.



Fuzheng Yang (M'10) received the B.E. degree in telecommunication engineering and the M.E. and Ph.D. degrees in communication and information system from Xidian University, Xi'an, China, in 2000, 2003, and 2005, respectively.

In 2005 and 2006, he became a Lecturer and an Associate Professor with Xidian University, respectively. Since 2012, he has been a Professor of communications engineering at Xidian University. He is also an Adjunct Professor with the School of Engineering, Royal Melbourne Institute of Technology, Melbourne, VIC, Australia. From 2006 to 2007, he served as a Visiting Scholar and Postdoctoral Researcher in the Department of Electronic Engineering, Queen Mary University of London, London, U.K. His research interests include video quality assessment, video coding, and multimedia communication.



Yicong Zhou (S'08–M'10–SM'14) received the B.S. degree from Hunan University, Changsha, China, and the M.S. and Ph.D. degrees from Tufts University, Medford, MA, USA, in 1992, 2008, and 2010, respectively, all in electrical engineering.

He is currently an Associate Professor and the Director in the Vision and Image Processing Laboratory, Department of Computer and Information Science, University of Macau, Macau, China. His research interests include chaotic systems, multimedia security, image processing and understanding, and

machine learning.

Dr. Zhou is an Associate Editor of the *Journal of Visual Communication and Image Representation*, an Editorial Board Member of *Neurocomputing*, and a Leading Co-Chair of the Technical Committee on Cognitive Computing of the IEEE Systems, Man, and Cybernetics Society. He was the recipient of the third place prize of Macau Natural Science Award in 2014.



Shan Gao received the M.Sc. degree from the University of Surrey, Guildford, U.K., in 2006.

In 2006, he joined the Huawei Technologies Co., Ltd., Shenzhen, China, where he is a Senior Research Engineer with the Media Technology Laboratory. Since 2007, he has been actively involved in the ITU-T standardization. He has more than 50 patent applications and several adopted standard proposals. His research interests include QoE, end-user quality perception, video communication, and video coding.