Layer-Assisted Video Quality Adaptation for Improving QoE in Wireless Networks

JINHO PARK AND KWANGSUE CHUNG, (Senior Member, IEEE)
Department of Electronics and Communications Engineering, Kwangwoon University, Seoul 01897, South Korea
Corresponding author: Kwangsue Chung (kchung@kw.ac.kr)

This work was supported in part by the Institute for Information and Communications Technology Planning and Evaluation (IITP) funded by the Korea Government (MSIT) through the Development of the System to Analyze and Detect the Obscenity in Media Contents Utilizing Artificial Intelligence Technologies under Grant 2019-0-00287, and in part by the Research Grant of Kwangwoon University, in 2020.

ABSTRACT HTTP adaptive streaming (HAS) is a state-of-the-art technology for video streaming. Optimal adaptive streaming schemes should be designed to maximize the quality of experience (QoE) in network environments with unstable bandwidths such as wireless networks. QoE is degraded for Advanced Video Coding (AVC)-based HAS due to frequent quality switching and rebuffering in wireless networks. Meanwhile, Scalable Video Coding (SVC)-based HAS can maximize the QoE by flexibly requesting the base layer (BL) and enhancement layer (EL) based on variable network conditions. However, in SVC-based HAS, the required bandwidth is significantly high due to the encoding overhead. In this paper, we propose a layer-assisted video quality adaptation for improving QoE in wireless networks. The proposed scheme employs a video encoding method, composed of multiple BLs and ELs, using both AVC and SVC. The quality of the BL is determined using the buffer occupancy level and measured bandwidth to minimize rebuffering. The quality of the EL is determined using the buffer region and segment quality differences to minimize instability. Enhancement dummy (ED) consists of the selected ELs. The layer scheduler decides the segment that is to be requested based on the QoE model according to the BL and ED to maximize the QoE. Experimental results show that the proposed scheme achieves high QoE compared to the existing schemes because the instability is low and no rebuffering occurs.

INDEX TERMS HTTP adaptive streaming, wireless network, quality of experience, scalable video coding.

I. INTRODUCTION

Recently, video streaming services have gained massive popularity. Cisco predicts that video traffic will increase from 73% to 82% of the total traffic on the Internet [1]. With this form of increase, HTTP adaptive streaming (HAS) has been widely accepted for providing a smooth video streaming service [2]. Unlike video streaming using Real-time Transport Protocol (RTP) and User Datagram Protocol (UDP), HAS is not blocked at the network firewall using TCP-based HTTP. Moreover, HTTP faces no additional cost for installing a server for video streaming services. HAS has been involved in the development of several commercial services such as Microsoft’s HTTP Smooth Streaming [3], Apple’s HTTP Live Streaming [4], and Adobe’s HTTP Dynamic Streaming [5], but the basic behavior remains the same. The HAS server stores video content encoded with multiple qualities. Each quality is divided into multiple segments. High video quality has a high bitrate compared to that of a low video quality. The segment has a fixed length and contains data for video playback. The encoded video content information is described using a media presentation description (MPD) file and stored in the server. An HAS client requests and downloads an MPD file from the server when the video streaming service begins. The HAS client requests and downloads an MPD file from the server when the video streaming service begins. The client adaptively selects the video quality to improve the quality of experience (QoE) depending on the video content information described in the downloaded MPD file, network, and buffer status [2].

Video quality adaptation schemes should be designed to maximize QoE in dynamic network environments such as wireless networks. Advanced Video Coding (AVC)-based video quality adaptation scheme, in which video content is encoded in a single layer, selects the appropriate video quality using the network and buffer information.
In wireless networks, frequent and rapid video quality switches, due to bandwidth variation, and rebuffering occurs. Meanwhile, the Scalable Video Coding (SVC)-based video quality adaptation scheme in which video content is encoded in a base layer (BL) and enhancement layers (EL) can maximize QoE by flexibly requesting a layer according to the network conditions. However, SVC has an encoding overhead of at least a 10% higher bitrate of each layer compared to AVC [6], [7]. As the number of layers increases due to the video encoding overhead of the SVC, the required bandwidth also increases. Due to the increased bandwidth requirements, the SVC-based HAS has lower average quality.

QoE is affected by the video quality chosen by the client [8]. There are three main QoE factors which are (1) average video playback quality, (2) video quality switching during video streaming, and (3) the number and time of rebuffering (i.e., no data in the client buffer to represent a new screen) [9]. In particular, rebuffering impacts QoE the most [10]. Considering the factors affecting QoE, client-side video quality adaptation schemes for overall QoE improvement have been proposed. However, the existing video quality adaptation schemes find it difficult to maximize QoE because of the tradeoff problem between the QoE factors. For example, if the HAS client selects a low video quality to minimize rebuffering, the average video quality would become low. Average quality improves when a HAS client selects the highest video quality. However, rebuffering occurs when the highest video quality is not adequate for bandwidth. When an HAS client selects an appropriate video quality for the given bandwidth, the average quality is improved. Furthermore, rebuffering is minimized because the client reacts instantly when the transient bandwidth drops. However, QoE is degraded because video quality switches frequently when bandwidth variation occurs.

In this study, we propose a layer-assisted video quality adaptation for improving QoE in wireless networks. The proposed scheme utilizes the video encoding method with both AVC and SVC. The video encoding method consists of video content with multiple BLs and multiple ELs for each BL. The proposed quality adaptation scheme determines the layer quality based on the network and buffer information. To minimize rebuffering, the quality of BL is selected with the appropriate for network and buffer state. The quality of the EL is selected on the basis of user perceived quality and video quality difference of the segment to minimize instability. The enhancement dummy (ED) consists of the selected ELs. To maximize QoE, we determine the quality of segment requested next by using the predicted QoE according to BL and ED.

The rest of this paper is organized as follows: In section II we describe HTTP adaptive streaming. In section III, we present the layer-assisted video quality adaptation for minimizing rebuffering and instability. In section IV, we describe the experimental results for the proposed scheme. Finally, in section V, we conclude the paper.

II. BACKGROUND AND RELATED WORK

A. HTTP ADAPTIVE STREAMING

Traditional video streaming services use RTP and Real-time Transport Control Protocol (RTCP). However, RTP and RTCP based on UDP are blocked by network firewalls and require additional costs to provide video streaming services. HAS was proposed to solve the problem of firewall blocking and additional cost issues. This is made possible as HAS uses HTTP, which is a standard protocol based on TCP.

The HAS server consists of encoded video content, an MPD file that describes the video content information, and an HTTP module. The video content is encoded into multiple qualities, each of which is divided into multiple segments. The segments have a fixed length and contain data for video playback. In the MPD file, the video content information (bitrate, resolution, segment length, etc.) is described. The HAS client consists of a decoder/display, an adaptation module, and an HTTP module. The decoder/display module decodes the downloaded video segment and displays it on the screen. The adaptation module uses the context information to determine the segment of video quality that is to be requested. The context information includes the measured bandwidth and buffer occupancy information. The HAS client requests a video segment using an HTTP GET message.

Figure 1 shows the request flow for video segments in AVC-based and SVC-based HAS. The request flow of AVC based HAS is shown in Figure 1-(a). In the HAS server, the video content is encoded and stored on the basis of AVC [11], which constitutes a single layer of video quality. The client downloads an MPD file when video streaming begins. The video quality is determined according to the video content information and context information. The HAS client downloads a single segment corresponding to the determinedP video quality.

Figure 1. Request flow for video segments in AVC-based and SVC-based HAS (a) AVC-based HAS (b) SVC-based HAS.
The request flow of the SVC-based HAS is shown in Figure 1-(b). In the HAS server, the video content is encoded and stored on the basis of the BL for video playback and multiple ELs for enhancing the quality of BL. The EL enhances the quality of BL, resolution, and frame rate [12]. SVC [13]-based HAS, similar to AVC-based HAS, initially downloads an MPD file and requests a layer based on the video content information and context information. The client requests the BL for smooth video streaming. When enhancing the quality of the segment occupied in the buffer, the client requests an EL for the segment. In [14], the SVC-based HAS is integrated into the MPEG-DASH (Dynamic Adaptive Streaming over HTTP) standard and compared with AVC-based HAS in mobile networks. The use of SVC in HAS is more effective in improving QoE than that of AVC [15]. The results show that under stable networks, SVC-based HAS has a lower average quality caused by encoding overhead than that of non-SVC. However, in unstable networks, SVC-based HAS shows high QoE performance with a flexible request of BL and EL. It has lower playback interruption and instability than AVC-based HAS.

B. VIDEO QUALITY ADAPTATION SCHEMES

Video quality adaptation schemes are classified into throughput-based and buffer-based according to the context information used for video quality determination. Throughput-based schemes select video quality using only the measured bandwidth information. By choosing the video quality based on the measured bandwidth, the client can determine the highest quality among the video qualities appropriate for the bandwidth. Conventional [16] determines the appropriate video quality using the measured bandwidth according to the recently downloaded segment size and the segment download time. Bandwidth frequently changes due to TCP congestion control and interference of various signals. Due to bandwidth variation, the video quality switches unnecessarily. To minimize unnecessary switching of video quality, the bandwidth is calculated using smoothing techniques. FESTIVE [17] is a throughput-based algorithm that balances efficiency, stability, and fairness where multiple HAS clients stream video. In particular, FESTIVE selects the video quality according to the measured bandwidth using the harmonic mean for the last five segments. RAHS [18] uses the segment duration and the segment download time to determine the parameters representing the network state. To minimize video quality switching, the client selects the video quality by comparing the parameters with quality adaptation thresholds, which are calculated using the Lyapunov optimization technique. BOLA [25] formulates video quality adaptation as a utility maximization problem and devises an online control algorithm using the Lyapunov optimization technique. BOLA uses the ratio of the minimum buffer occupancy threshold and target buffer threshold to maintain the video quality. Le et al. [26] predicted the buffer variation using previously requested segments and selected the quality using the predicted buffer variation. They define the minimum and maximum buffer thresholds to minimize rebuffering. If the buffer occupancy level varies between two buffer regions, the video quality also decreases gradually. Bokani et al. [27] selected the video quality to be requested by predicting the buffer variation based on the Markov decision process.

SVC-DASH-M [28] is an SVC-based HAS system that uses multiple networks. This uses several parameters related to the QoE factor to determine the quality of the segment. WQUAD [29] determines the segment to be requested according to the buffer occupancy level using SVC. The only request to BL is to minimize buffering until the buffer is full. When the buffer is full, they request an EL to enhance the quality of the segment occupied by the buffer. Existing AVC-based HAS cannot change the quality of already fetched segments, and the SVC-based HAS has a problem of low average quality due to encoding overhead. We proposed a quality adaptation scheme to maximize QoE. The proposed scheme uses both AVC and SVC to request the BL and EL flexibly according to the context information.

III. PROPOSED SCHEME

A. VIDEO ENCODING METHOD

Figure 2 shows the video encoding method used in the proposed scheme. This method configures the video quality
using both AVC and SVC. The video quality is composed of a BL for video playback and an EL for enhancing the quality of BL. The BL is encoded into a single layer with multiple qualities, like AVC. The EL is encoded in multiple layers that enhance the quality of the BL, like SVC. The EL has a requirement of high bandwidth for the request because of the encoding overhead. The EL includes a different layer for each BL. However, it does not improve the quality of other BLs.

**B. PROPOSED SYSTEM ARCHITECTURE**

The overall structure of the proposed video quality adaptation scheme based on the video encoding method is shown in Figure 3. The proposed scheme measures the bandwidth using previously downloaded segments. The bandwidth is measured as expressed in (1).

\[ Th_k^e = \frac{D_k}{SDT_k} \]  

\( Th_k^e \) denotes the estimated bandwidth, \( D_k \) denotes the downloaded segment size, and \( SDT_k \) denotes the segment download time. The measured bandwidth frequently changes due to TCP congestion control and signal interference. The video quality faces unnecessary switching due to frequent changes of the measured bandwidth. To minimize frequent switching by measured bandwidth, we use the exponential weighted moving average (EWMA) to smooth the measured bandwidth, which is calculated as shown in (2).

\[ Th_k^s = \alpha Th_k^e + (1 - \alpha)Th_{k-1}^s \]  

\( Th_k^s \) denotes the smoothed bandwidth and \( \alpha \) denotes the smoothing weight. The quality of BL and EL is selected according to the measured bandwidth and the buffer occupancy level. The quality of BL is selected on the basis of the maximum available bandwidth to minimize rebuffering. The maximum available bandwidth is calculated using the target buffer occupancy level and smoothed bandwidth. The maximum available bandwidth is calculated as in (3).

\[ Th_k^{max} = \frac{B_{tar}}{B_{max}} \cdot Th_k^s \]  

\( B_{tar} \) denotes the target buffer occupancy level and \( B_{max} \) denotes the maximum buffer occupancy level. The maximum available bandwidth depends on the target buffer occupancy level. The target buffer occupancy level is calculated as shown in (4).

\[ B_{tar} = \frac{B_c + B_{max}}{2} \]  

\( B_c \) denotes the current buffer occupancy level. The quality of the BL is selected as shown in (5) by using the maximum available bandwidth.

\[ R_{BL_k+1}^{BL} = \max_{R_{BL_k+1}^{BL} \in \mathcal{R}_{BL_k+1}^{BL}} |R_{BL_k+1}^{BL} - Th_k^{max}| < Th_k^{max} \]  

\( R_{BL_k+1}^{BL} \) denotes the bitrate of the selected BL and \( \mathcal{R}_{BL_k+1}^{BL} \) denotes a set of BL bitrates. If the buffer occupancy level is low, the client selects a low video quality, even if the bandwidth is appropriate for selecting a high video quality. The determined quality minimizes the risk of rebuffering through increased buffer occupancy level due to short segment download times. The quality of the BL is determined by the measured bandwidth if the buffer occupancy level is sufficient. Therefore, the quality of BL has high instability in wireless networks. To minimize instability, the quality of the EL is selected.

The layer scheduler determines the segments to maximize QoE. The segment selection is driven by QoE predicted using the BL and ED. The predicted QoE shows average quality, instability, and rebuffering using the QoE model.

**C. QUALITY ADAPTATION**

The proposed scheme measures bandwidth based on the downloaded segments. The bandwidth is measured as expressed in (1).

\[ Th_k^e = \frac{D_k}{SDT_k} \]  

\( Th_k^e \) denotes the estimated bandwidth, \( D_k \) denotes the downloaded segment size, and \( SDT_k \) denotes the segment download time. The measured bandwidth frequently changes due to TCP congestion control and signal interference. The video quality faces unnecessary switching due to frequent changes of the measured bandwidth. To minimize frequent switching by measured bandwidth, we use the exponential weighted moving average (EWMA) to smooth the measured bandwidth, which is calculated as shown in (2).

\[ Th_k^s = \alpha Th_k^e + (1 - \alpha)Th_{k-1}^s \]  

\( Th_k^s \) denotes the smoothed bandwidth and \( \alpha \) denotes the smoothing weight. The quality of BL and EL is selected according to the measured bandwidth and the buffer occupancy level. The quality of BL is selected on the basis of the maximum available bandwidth to minimize rebuffering. The maximum available bandwidth is calculated using the target buffer occupancy level and smoothed bandwidth. The maximum available bandwidth is calculated as in (3).

\[ Th_k^{max} = \frac{B_{tar}}{B_{max}} \cdot Th_k^s \]  

\( B_{tar} \) denotes the target buffer occupancy level and \( B_{max} \) denotes the maximum buffer occupancy level. The maximum available bandwidth depends on the target buffer occupancy level. The target buffer occupancy level is calculated as shown in (4).

\[ B_{tar} = \frac{B_c + B_{max}}{2} \]  

\( B_c \) denotes the current buffer occupancy level. The quality of the BL is selected as shown in (5) by using the maximum available bandwidth.

\[ R_{BL_k+1}^{BL} = \max_{R_{BL_k+1}^{BL} \in \mathcal{R}_{BL_k+1}^{BL}} |R_{BL_k+1}^{BL} - Th_k^{max}| < Th_k^{max} \]  

\( R_{BL_k+1}^{BL} \) denotes the bitrate of the selected BL and \( \mathcal{R}_{BL_k+1}^{BL} \) denotes a set of BL bitrates. If the buffer occupancy level is low, the client selects a low video quality, even if the bandwidth is appropriate for selecting a high video quality. The determined quality minimizes the risk of rebuffering through increased buffer occupancy level due to short segment download times. The quality of the BL is determined by the measured bandwidth if the buffer occupancy level is sufficient. Therefore, the quality of BL has high instability in wireless networks. To minimize instability, the quality of the EL is selected.
on the basis of the buffer region and the instability score. Figure 4 shows a flowchart for the EL quality selection. We divide the buffer region for EL selection according to the user perceived quality. However, if the segment is decoded and video is played while downloading the EL, wastage of data occurs. To minimize data wastage, we calculate the buffer region considering the playback deadline. After determining the buffer region, we use the instability score to select the appropriate quality of EL among the selectable ELs. Instability scores are calculated using quality differences of adjacent segments and the bitrate of the EL. An example of buffer region calculation is shown in Figure 5. To minimize data waste, the buffer region excludes the playback deadline from the length of the buffer occupied by the segment. The playback deadline is determined on the basis of segment download time. The buffer region, excluding the playback deadline is divided into two types according to user perceived quality. A primary buffer region close to video playback indicates a region that has a high impact on user perceived quality. The EL of the segment belonging to this region is the first segment to be selected. The secondary buffer region is far from the video playback and represents a region where the impact of user perceived quality is relatively lower than that of the primary buffer region. The EL of the segment belonging to the secondary buffer region has a lower priority than that occupied by the primary buffer region. The primary buffer region is calculated as shown in (6).

\[
T_P = \begin{cases} 
T_{int}, & \text{if } T_k^{ED} > t_c - t_d \\
T_{c} - T_{d}, & \text{if } T_k^{ED} < t_c - t_d 
\end{cases} 
\]  

(6)

Next, the buffer region calculation is shown in (7).

\[
t_d = t_p + SDT_k 
\]  

(7)

The length of the primary buffer region is calculated as shown in (8).

\[
\begin{align*}
T_{int} &= t_d - t_k^{ED} \\
\text{subject to } &\tau < t_c - t_k^{ED} 
\end{align*} 
\]  

(8)

The length of the primary buffer region decreases due to frequent EL requests as the rebuffering risk is low when the network is stable. Because of the primary buffer region, a segment close to playback is selected. However, due to frequent requests of the EL, the primary buffer region becomes shorter than the segment duration, thus, leading to the quality being selected by the secondary buffer region. As the secondary buffer region is similar to the current buffer occupancy level, the quality of the EL is determined using the instability score. Thus, selected EL does not rely on user perceived quality. For EL requests of segments close to the playback, the primary buffer region must be longer than the segment duration. The secondary buffer region is calculated on the basis of the primary buffer region as shown in (9).

\[
T_S = \begin{cases} 
S_c - (t_d + T_{int}) \cdot \left( \frac{t_d + T_{int}}{\tau} < S_j < N \right) 
\sum \left( \frac{t_d + T_{int}}{\tau} \right) \text{if } T_{int} > t_c - t_d \\
0 & \text{else} 
\end{cases} 
\]  

(9)

\(N\) denotes the number of downloaded segments. The secondary buffer region is calculated according to the primary buffer region. If the network is unstable, the length of the primary buffer region is increased by requesting frequent BLs due to the high rebuffering risk. As the length of the primary buffer region increases, the length of the secondary buffer region decreases. After determining the buffer region, the EL is selected on the basis of the instability score. The instability score is calculated as shown in (10).

\[
IS_{i,j} = \frac{R_{i,j}^{EL}}{R_{i,j}^{EL}} \cdot Q_{diff} 
\]  

(10)

\(R_{i,j}^{EL}\) denotes the EL bitrate of the ith segment and \(Q_{diff}\) denotes the quality difference from adjacent segments. The instability score is calculated as shown in (11).

\[
Q_{diff} = \begin{cases} 
\min \left( IS_{i+1,j} - IS_{i-1,j} \right) & \text{if } IS_{i+1,j} - IS_{i-1,j} \neq 0 \\
\min \left( |q(i) - q(i+1)|, |q(i-1) - q(i)| \right) & \text{else} 
\end{cases} 
\]  

(11)

\(q(i)\) denotes the video quality level of the segment.

Instability scores are calculated using the bitrate of the EL and video quality difference between adjacent segments. Even if segments in the buffer have the same video quality, the bitrates of ELs for quality improvement are different depending on the quality of their BL. The EL quality is selected as shown in (12) using an instability score.

\[
R_{k+1}^{ED} = \max \left( \sum R_{i,j}^{EL} < TH_k \right) 
\]  

(12)
D. LAYER SCHEDULER

The layer scheduler determines the segments using the BL and ED to maximize QoE. Figure 6 shows the overall operation of the layer scheduler. It predicts QoE according to BL and ED. The predicted QoE represents the average quality, instability, and rebuffering for each segment. This value is calculated by using the values of the BL and ED requests. The next request segment is determined on the basis of the predicted QoE value. If the difference between the two predicted QoE values is more than 0, it indicates that requesting the BL represents a high QoE improvement. Therefore, the layer scheduler selects the BL. On the other hand, if the difference between the two is less than 0, it indicates that the requesting ED improves QoE over the BL requests. Therefore, the layer scheduler selects ED. The predicted QoE value is calculated as shown in (13).

$$Q_{oE}^k = \frac{\sum_{i=1}^{N} q(i) - \delta \sum_{i=1}^{N-1} |q(i+1) - q(i)|}{N-1} - \mu Q_{rebuff,k}$$  \hspace{1cm} (13)

$Q_{rebuff,k}$ denotes rebuffering due to the $k$th layer request, $\delta$ denotes the video quality switching weight, and $\mu$ denotes the rebuffering weight. The predicted QoE value is calculated on the basis of average quality, instability, and rebuffering model through a segment request, which is calculated as shown in (14).

$$Q_{rebuff,k} = \begin{cases} \sum_{i=1}^{N} \left( \frac{D_i}{T_{hi}} - B_i \right) + \tau & \text{if } k = BL \\ \sum_{i=1}^{N} \left( \frac{D_i}{T_{hi}} - B_i \right) & \text{if } k = ED \end{cases}$$  \hspace{1cm} (14)

IV. PERFORMANCE EVALUATION

A. SIMULATION SETUP

To evaluate the performance of the proposed scheme, NS-3 [30] was used to construct the environment as shown in Figure 7. The implementation included the HAS server and [Further content continues...].
HAS client with a buffer length set to 30 s and bandwidth set to 5 Mbps. Table 1 lists the set of bitrates for video quality.

Six video qualities were used in the experiment. Each video quality constituted BL with a different quality, like AVC. Each configured BL had multiple ELs that could enhance the quality, like SVC. We refer to the performance evaluation scenarios in Figure 8. Scenario 1 produces an abrupt bandwidth change. We evaluate how the algorithms adapt the video quality as the bandwidth changes drastically. Scenario 2 produces bandwidth fluctuations of variable amplitude. We evaluate how the algorithms adapt the video quality as the amplitude of the throughput varies. To evaluate the performance of the proposed scheme, we compared it with four quality adaptation schemes. The selected adaptation algorithms that Conventional and BOLA using AVC, WQUAD using SVC, and LAAVS [31] using both AVC and SVC. To evaluate the QoE of each quality adaptation scheme, the QoE is calculated as shown in (17) based on the QoE model proposed in [32].

\[
\text{QoE} = \frac{1}{N} \sum_{i=1}^{N} q(i) - \frac{1}{N-1} \sum_{i=1}^{N-1} |q(i+1) - q(i)| - \sigma \frac{1}{N} \sum_{i=1}^{N} \text{rebuff}(i)
\]

### TABLE 1. Set of bitrates for video quality (Mbps).

| Quality Level | Base Layer Bitrate | Enhancement Layer Bitrate |
|---------------|---------------------|--------------------------|
| 1             | 0.30                | 0.57                     |
| 2             | 0.75                | 0.63                     |
| 3             | 1.20                | 0.93                     |
| 4             | 1.85                | 1.43                     |
| 5             | 2.85                | 2.1                      |
| 6             | 4.30                |                           |

**FIGURE 8. Performance evaluation scenarios.**

**FIGURE 9. Average quality and instability of buffer region.**

C. QOE PERFORMANCE IN SCENARIO 1

Figure 10 shows QoE of the quality adaptation schemes in Scenario 1. Conventional selects the video quality based on the measured bandwidth. It determines the highest quality for the bandwidth, so the average quality is high and there is no rebuffering due to bandwidth variation. Conventional, however, reacts quickly to rapid and frequent bandwidth changes in wireless networks. Consequently, Conventional has high instability. BOLA determines the video quality using the Lyapunov optimization technique. It has low because it uses buffer information only. However, although the bandwidth is appropriate for segment of high quality requests, the average quality is low due to the selection of low video quality. BOLA causes rebuffering due to the selection of video quality that is not appropriate for bandwidth fluctuations, resulting in low QoE.

WQUAD is an SVC-based quality adaptation scheme that requests the BL and EL using the buffer information. WQUAD requests the EL of the next quality level when the buffer is full. WQUAD has low instability due to conservative video quality improvement. However, WQUAD manages to...
achieve a low average quality due to SVC encoding overhead. QoE is low as rebuffering occurs when an unsuitable layer is requested during the drop of transient bandwidth.

LAAVS uses the network and buffer information to determine the video quality based on the video encoding method by utilizing both AVC and SVC. To minimize rebuffering, the quality of the BL is selected on the basis of the measured bandwidth using the harmonic mean up to a fixed buffer threshold. Because the quality of the segment is determined using the measured bandwidth, the average quality is high and there is no rebuffering. The EL of the segment with the lowest quality among the segments occupied in the buffer is selected. The EL is requested if the bandwidth is higher than the sum of the BL and EL bit rates.

The proposed scheme requests the BL according to the buffer occupancy level and measured bandwidth so that rebuffering does not occur. The EL is selected on the basis of the buffer region and instability score. An EL of a segment close to the video playback is requested, and an EL for a segment far from the video playback is also requested depending on the network and buffer conditions. The proposed scheme has an average quality similar to that of LAAVS, but its instability is lower than that of the existing schemes. Consequently, the proposed scheme achieves high QoE.

D. QoE PERFORMANCE IN SCENARIO 2

Figure 11 shows QoE of the quality adaptation schemes in Scenario 2. In Scenario 2, rebuffering does not occur in all the algorithms. As in Scenario 1, Conventional has high average quality by quickly reacting to bandwidth fluctuation of amplitude variation. Conventional, however, adjusts video quality unnecessarily due to erroneous measurement according to drastic bandwidth changes. As a result, instability is high. BOLA, like Scenario 1, determines video quality based on the buffer status, so the average quality is low. BOLA shows high instability, unlike the results in Scenario 1. The reason behind instability is that video quality is inappropriately selected by unstable buffer status caused by bandwidth fluctuation of variable amplitude.

WQUAD achieves a low average quality due to its SVC encoding overhead. Since WQUAD determines the layer to be downloaded on the basis of the buffer state, if the buffer occupancy level decreases, it downloads the BL of the new segment first to quickly fill the buffer. Therefore, in the case of bandwidth fluctuations such as Scenario 2, the instability of WQUAD is high because it cannot efficiently improve the quality of segments close to video playback.

LAAVS determines the quality of the BL according to the measured bandwidth. LAAVS has low instability by selecting an appropriate video quality for bandwidth without unnecessary quality switching. However, LAAVS determines EL when...
the bandwidth is higher than the bitrate sum of BL and EL, so the average quality is similar to Conventional. This scheme shows high QoE compared to other existing schemes.

The proposed scheme increases the quality of BL step by step according to the buffer occupancy level even if the bandwidth is sufficient. In addition, the proposed scheme predicts QoE and selects the next layer to be requested. Even if there are changes in video quality due to bandwidth variation, instability is low because an appropriate EL is selected to reduce the quality changes. As a result, the proposed scheme has higher QoE because of higher average quality and lower instability than the existing video quality adaptation schemes.

V. CONCLUSION

In this paper, we proposed a layer-assisted video quality adaptation for improving QoE in wireless networks. The proposed scheme utilizes a video encoding method that uses both AVC and SVC to flexibly request segments. To minimize rebuffering, the quality of the BL is determined according to the network and buffer conditions. To minimize instability, the quality of the EL is selected using an instability score determined using the segment quality difference and buffer region by user perceived quality. The EL has a quality that is appropriate for the given bandwidth. ED is composed of the determined ELs. The layer scheduler selects segments on the basis of QoE prediction values according to BL and ED to maximize QoE. Experimental results show that the proposed scheme has a low instability and high average quality when the buffer region is divided by considering of user perceived quality. Consequently, the proposed scheme outperforms the existing solutions for various scenarios.

REFERENCES

[1] Cisco, “Cisco visual networking index: Forecast and trends, 2017–2022,” Cisco, San Jose, CA, USA, White Paper, Feb. 2019.

[2] A. Bentaleb, B. Taami, A. C. Begen, C. Timmerer, and R. Zimmermann, “A survey on bitrate adaptation schemes for streaming media over HTTP,” IEEE Commun. Surveys Tuts., vol. 21, no. 1, pp. 562–585, 1st Quart., 2019.

[3] Microsoft. Smooth Streaming. Accessed: Nov. 15, 2018. [Online]. Available: http://www.iis.net/downloads/microsoft/smooth-streaming

[4] Apple. HTTP Live Streaming. Accessed: Nov. 15, 2018. [Online]. Available: https://developer.apple.com/streaming

[5] Adobe. HTTP Dynamic Streaming. Accessed: Nov. 15, 2018. https://www.adobe.com/products/hds-dynamic-streaming.html

[6] T. Oelbaum, H. Schwarz, M. Wien, and T. Wiegand, “Subjective performance evaluation of the SVC extension of H.264/AVC,” in Proc. 15th IEEE Int. Conf. Image Process., Oct. 2008, pp. 2772–2775.

[7] J. M. Boyce, Y. Ye, J. Chen, and A. K. Ramasubramonian, “Overview of SHVC: Scalable extensions of the high efficiency video coding standard,” IEEE Trans. Circuits Syst. Video Technol., vol. 26, no. 1, pp. 20–34, Jan. 2016.

[8] R. K. P. Mok, W. Li, and R. K. C. Chang, “IRate: Initial video bitrate selection system for HTTP streaming,” IEEE J. Sel. Areas Commun., vol. 34, no. 6, pp. 1914–1928, Jun. 2016.

[9] M. Seufert, S. Egger, M. Slanina, T. Zinner, T. Hoßeld, and P. TranGia, “Survey on quality of experience of HTTP adaptive streaming,” IEEE J. Sel. Areas Commun., vol. 32, no. 4, pp. 719–733, Apr. 2014.
[17] J. Jiang, V. Sekar, and H. Zhang, “Improving fairness, efficiency, and stability in HTTP-based adaptive video streaming with festive,” IEEE/ACM Trans. Netw., vol. 22, no. 1, pp. 326–340, Feb. 2014.

[18] C. Liu, L. Bouazizi, and M. Gabbou, “Rate adaptation for adaptive HTTP streaming,” in Proc. 2nd Annu. ACM Conf. Multimedia Syst., Feb. 2011, pp. 169–174.

[19] R. K. P. Mok, X. Luo, E. W. W. Chan, and R. K. C. Chang, “QDASH: A QoE-aware DASH system,” in Proc. 3rd Multimedia Syst. Conf. (MMSys), Feb. 2012, pp. 11–22.

[20] T. Thang, Q.-D. Ho, J. Kang, and A. Pham, “Adaptive streaming of audiovisual content using MPEG DASH,” IEEE Trans. Consum. Electron., vol. 58, no. 1, pp. 78–85, Feb. 2012.

[21] I. Ironid, Q. Wang, C. Grecos, J. Calero, and P. Casaseca, “Efficient QoE-aware scheme for video quality switching operations in dynamic adaptive streaming,” ACM Trans. Multimedia Comput. Commun. Appl., vol. 15, no. 1, pp. 1–23, Feb. 2019.

[22] T.-Y. Huang, C. Ekanadham, A. J. Berglund, and Z. Li, “Hindsight: Evaluate video bitrate adaptation at scale,” in Proc. 10th ACM Multimedia Syst. Conf., Jun. 2019, pp. 140–149.

[23] H. T. Le, D. V. Nguyen, N. P. Ngoc, A. T. Phan, and T. C. Thang, “Buffer-based bitrate adaptation for adaptive HTTP streaming,” in Proc. Int. Conf. Adv. Technol. Commun., Oct. 2013, pp. 33–38.

[24] T.-Y. Huang, R. Johari, N. McKeown, M. Trunnell, and M. Watson, “A buffer-based approach to rate adaptation: Evidence from a large video streaming service,” in Proc. ACM Conf. SIGCOMM (SIGCOMM), Aug. 2014, pp. 187–198.

[25] K. Spiteri, R. Urgaonkar, and R. K. Sitaraman, “BOLA: Near-optimal bitrate adaptation for online videos,” in Proc. 35th Annu. IEEE Int. Conf. Comput. Commun. (IEEE INFOCOM), Apr. 2016, pp. 1–9.

[26] H. T. Le, N. Pham Ngoc, and C.-T. Truong, “Bitrate adaptation for seamless on-demand video streaming over mobile networks,” Signal Process., Image Commun., vol. 65, pp. 154–164, Jul. 2018.

[27] A. Bokani, M. Hassan, and S. Kanhere, “HTTP-based adaptive streaming for mobile clients using Markov decision process,” in Proc. 20th Int. Packet Video Workshop, Dec. 2013, pp. 1–8.

[28] S. Ibrahim, A. H. Zahran, and M. H. Ismail, “SVC-DASH-M: Scalable video coding dynamic adaptive streaming over HTTP using multiple connections,” in Proc. 21st Int. Conf. Telecommun. (ICT), May 2014, pp. 400–404.

[29] J. Hwang, J. Lee, and C. Yoo, “Eliminating bandwidth estimation from adaptive video streaming in wireless networks,” Signal Process., Image Commun., vol. 47, pp. 242–251, Sep. 2016.

[30] The Network Simulation NS-3. Accessed: Dec. 4, 2018. [Online]. Available: https://www.nsnam.org

[31] A. T. Nasrabadi and R. Prakash, “Layer-assisted adaptive video streaming,” in Proc. 28th ACM SIGMM Workshop Netw. Operating Syst. Support Digit. Audio Video (NOSSDAV), Jun. 2018, pp. 31–36.

[32] X. Yin, A. Jindal, V. Sekar, and B. Sinopoli, “A control-theoretic approach for dynamic adaptive video streaming over HTTP,” in Proc. ACM Conf. Special Interest Group Data Commun. (SIGCOMM), Aug. 2013, pp. 325–338.

JINHO PARK received the B.S. degree from the Electronics and Communications Engineering Department, Kwangwoon University, Seoul, South Korea, in 2018, where he is currently pursuing the Ph.D. degree. His research interests include QoS, quality of experience support, multimedia systems, and streaming protocols.

KWANGSUE CHUNG (Senior Member, IEEE) received the B.S. degree from the Electrical Engineering Department, Hanyang University, Seoul, South Korea, the M.S. degree from the Electrical Engineering Department, Korea Advanced Institute of Science and Technology (KAIST), Seoul, and the Ph.D. degree from the Electrical Engineering Department, University of Florida, Gainesville, FL, USA. Before joining the Kwangwoon University, in 1993, he spent ten years with the Electronics and Telecommunications Research Institute (ETRI), as a member of a Research Staff. He was also an Adjunct Professor with KAIST, from 1991 to 1992. He was a Visiting Scholar with the University of California, Irvine, from 2003 to 2004. His research interests include communication protocols and networks, QoS mechanisms, and video streaming.

* * *