The Joint Effect of MMT AL-FEC and Error Concealment on Video Streaming QoE

Toshiro Nunome
Department of Computer Science and Engineering, Graduate School of Engineering,
Nagoya Institute of Technology, Nagoya 466–8555, Japan
nunome@nitech.ac.jp

***Abstract*—This paper deals with MMT (MPEG Media Transport), which is an application-level protocol for video transmission. We assess subjective QoE (Quality of Experience) and objective application-level QoS of video IP transmission with error concealment mechanisms of H.264/AVC and application-level forward error correction (AL-FEC) of MMT. In the experiment, we employ two contents, two types of the number of slices per picture frame, and two types of the total bitrate of video and its FEC code. We experiment with several load conditions. We then compare FEC schemes with three code rate values and no FEC scheme. From the assessment results, we show that the appropriate code rate for QoE enhancement depends on not only network conditions but also contents.**

***Index Terms*—audio and video IP transmission, MMT, AL-FEC, error concealment, subjective assessment**

I. Introduction

Video streaming services are widely used applications for any types of users over IP networks including cloud networking environments. In the services for consumers, QoE (Quality of Experience) [1] is the most important quality criterion.

As an application-level communication protocol for video transmission, MMT (MPEG Media Transport) [2] has been standardized. The protocol mainly aims to H.265/HEVC transmission and replaces MPEG2-TS, which has been widely used in broadcasting. One of the characteristics of MMT is consideration of IP transmission. In addition, MMT can be employed for various existing video formats such as H.264/AVC.

In MMT, we can employ AL-FEC (Application Level Forward Error Correction) for recovering from errors and packet losses over the networks. On the other hand, for mitigating the deterioration by the errors and packet losses, error concealment and frame skipping can be used. It has not been clarified the effect of the combination of AL-FEC and error concealment on QoE.

Thus, this paper assesses the joint effect of AL-FEC in MMT and the error concealment mechanism in H.264/AVC on subjective QoE and objective application-level QoS; the application-level QoS is closely related to QoE. Although the main target of MMT is H.265/HEVC transmission, H.265/HEVC omits error concealment mechanisms for encoding high definition video. And then FFmpeg [3], which is a decoder employed in this paper, also cut the mechanisms for H.265/HEVC. Thus, we use H.264/AVC, which supports various error concealment techniques.

II. MMT AL-FEC

MMT supports six AL-FEC schemes: Reed-Solomon, Structured LDPC, RaptorQ, RaptorQ LA, FireFort-LDGM, and Pro-MPEG. They have advantages and disadvantages in complexity, efficiency, ability, and so on. Among them, we employ Reed-Solomon in this paper.

ISO/IEC23008-1 Annex C specifies three SSBG (Source Symbol Block Generation) modes to deal with AL-FEC in MMT. We choose ssbg_mode1 from the three modes. The mode divides each source packet into fixed length blocks and then generates FEC codes. If a divided piece is shorter than the fixed length, a block is composed of the piece and padding bits.

In order to transmit H.264/AVC video with MMT, we consider a NAL unit, which has a slice, as a source packet for AL-FEC. The divided blocks from the NAL unit is called *source code blocks*. If the final piece of a divided NAL unit is shorter than the fixed length, padding is performed; it causes overhead.

We employ the one-stage FEC coding structure, in which FEC code generation is performed just one time for each source packet. From the source code blocks, the mechanism generates *repair blocks*. To compose an MMTP (MMT Protocol) packet, we attach an MMTP packet header and an MMTP payload header before the source code block and a source FEC payload ID after the source code block. For each repair block, we compose an MMTP packet with an MMTP packet header, a repair FEC payload ID, and the repair block.

III. Experimental method

Figure 1 shows the experimental system. All the links in the network are 100 Mbps full-duplex Ethernet. Media Server transmits video and audio streams to Media Client through MMTP/UDP. For audio, each MMTP/UDP packet includes an MU (Media Unit), which is an information unit for media synchronization control. As the interference traffic of audio and video, Web Server transmits Web traffic to Web Client according to requests generated by WebStone 2.5 [4], which is a web server benchmark tool. For the number of client processes, we employ 10, 15, and 20.

To implement AL-FEC, we employ OpenFEC [5], which is an open source library of FEC. x264 and FFmpeg are used as an encoder and a decoder, respectively. The video image size is 1920×1080 pixels, and the frame rate is 29.97 frame/second (MU/s). A video frame is divided into several slices, and a NAL unit is composed of a slice. In this paper, we divide a video frame into 4 or 17 slices.

As the values of the code rate of FEC, we use 1/2, 2/3, and 5/6. We compare them with a method which does not perform FEC code generation (i.e., the code rate is 1). In the method without FEC, we divide a slice into fixed length source code

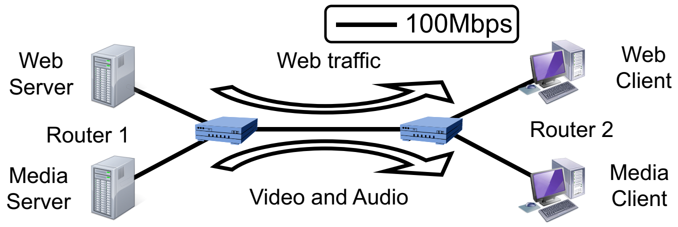

Fig. 1. Experimental system

blocks with padding if needed. In addition, in any code rate except for 1, we generate at least a repair block for a NAL unit. Thus, note that the actual code rate can have a difference from the setting code rate. In this paper, we set the block size is 512 bytes. The assessment with various block length is a future study.

In addition, we do not employ FEC for audio transmission owing to the simplicity of implementation. This is because we focus on the effect of FEC on video quality. The audio codec is aac (stereo, 128 kbps, 46.875 MU/s). We consider only the media transfer phase. We have not implemented signaling mechanisms before/after the media transfer phase.

As the video total bitrate of the source video stream and FEC blocks, we consider two values: 3 Mbps and 6 Mbps. When we do not employ FEC, we use 3 Mbps or 6 Mbps as the video encoding bitrate. For employing FEC, we set the video encoding bitrate with consideration of the code rate: 1.5 Mbps and 3 Mbps for the code rate 1/2, 2 Mbps and 4 Mbps for the code rate 2/3, and 2.5 Mbps and 5 Mbps for the code rate 5/6. Note that the network-level transmission rate becomes larger as the code rate decreases.

We employ the error concealment scheme as the video output. We utilize *Frame Copy* and the interpolation from neighboring macroblocks for video error concealment. We use the simple playout buffering control with the playout buffering time 500 ms.

IV. Experimental Results

In this paper, we treat the audio MU loss ratio, the video MU loss ratio, the slice loss ratio for video, and PSNR (Peak Signal-to-Noise Ratio) as the application-level QoS parameters. The slice loss ratio represents the percentage of lost slices in an output frame; it shows the image quality of video stream. The MU loss ratio is the ratio of the number of MUs not output at the recipient to the number of MUs transmitted by the sender. PSNR is widely used as a measure of image quality.

Fig. 2 shows the video MU loss ratio. The video slice loss ratio is depicted in Fig. 3. Fig. 4 represents the audio MU loss ratio. Figs. 2 through 4 are the results for the video bitrate 6 Mbps. Figs. 5 and 6 presents the PSNR of video luminance for the video bitrate 3 Mbps and that for the video bitrate 6 Mbps, respectively.

We find in Fig. 2 that the video MU loss occurs for the number of slices 4 without FEC. The loss ratio increases as the interference traffic (i.e., the number of Web clients) increases. When the number of slices is 4, the size of a slice is larger than that for the number of slices 17. Thus, the effect of a lost MMTP packet becomes large in the 4 slices. On the other hand, when we employ FEC, it can recover lost MMTP packets. Thus, the MU loss ratio can be reduced.

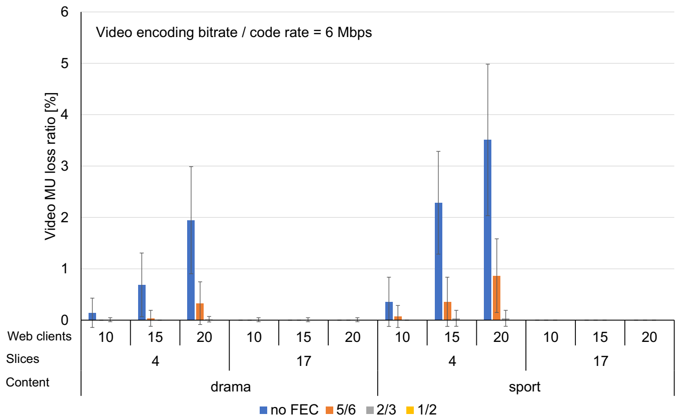

Fig. 2. Video MU loss ratio (video bit rate 6 Mbps)

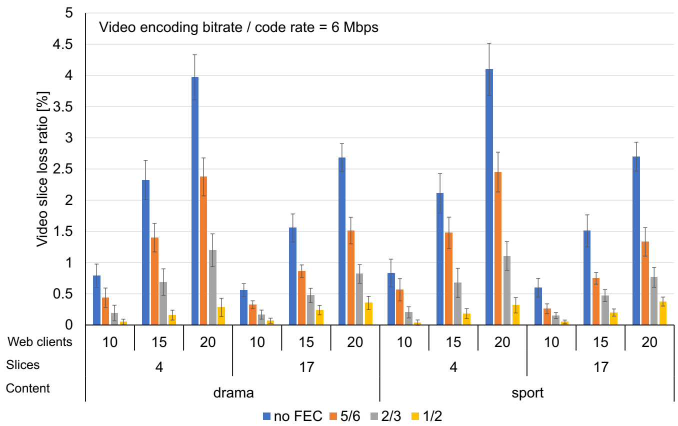

Fig. 3. Slice loss ratio of output video (video bit rate 6 Mbps)

We also see in Fig. 3 that the slice loss ratio increases as the number of slices per frame decreases, and FEC can reduce the slice loss ratio. Besides, the small coding ratio is more efficient for decreasing the slice loss ratio. This is because not only the enhancement of error correction performance by the large redundancy but also the decrement of video source encoding bitrate affect the loss ratio.

Fig. 4 shows that the audio MU loss ratio is smaller than about 1 %; there is no large degradation of audio quality. The audio MU loss ratio does not have a clear tendency against the number of slices, the contents, and the code rate of FEC.

We notice in Figs. 5 and 6 that the scheme without FEC has the larger PSNR than the schemes with FEC when the number of Web clients is 10. This is because under the lightly loaded condition, the large source encoding bitrate can provide good image quality. Besides, as the interference traffic increases, the schemes with FEC have the larger PSNR than the scheme without FEC. When the video bitrate is 3 Mbps with sport under the 20 Web clients condition, the moderate code rate obtains the best PSNR. Thus, the appropriate code rate is important to enhance QoS.

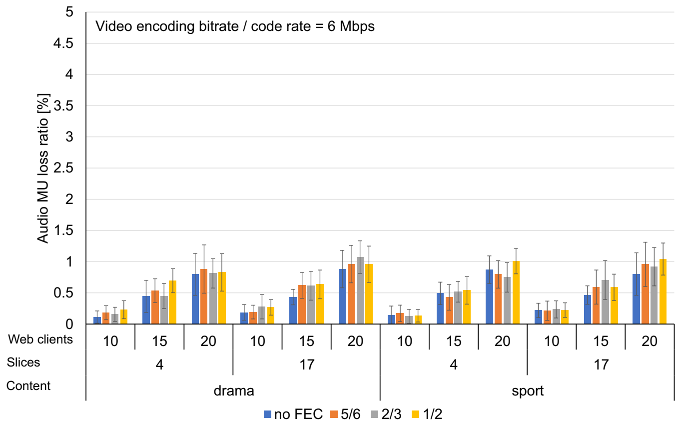

Fig. 4. Audio MU loss ratio (video bit rate 6 Mbps)

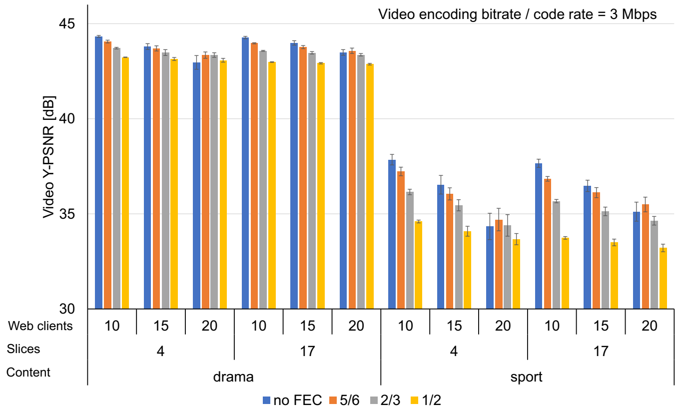

Fig. 5. Video PSNR (video bit rate 3 Mbps)

V. QoE Assessment

In the experiment, we ask the assessors to evaluate video and audio output at Media Client. To reproduce the experimental conditions easily, we employ trace files which record the receive timing of video slices and audio MUs. The

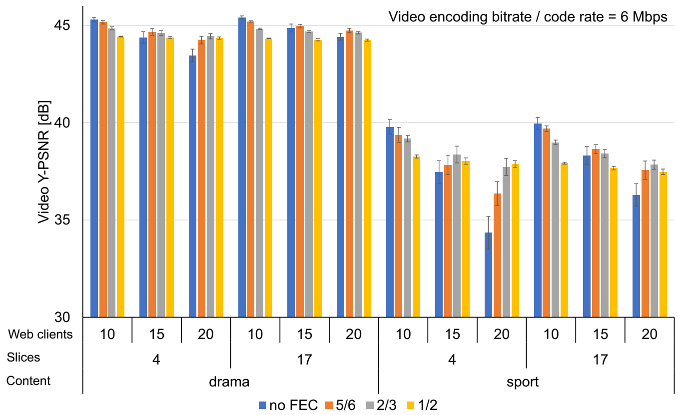

Fig. 6. Video PSNR (video bit rate 6 Mbps)

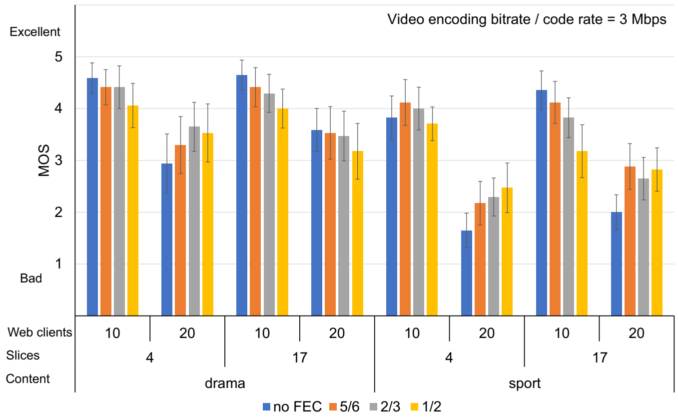

Fig. 7. MOS ("overall quality of video and audio is bad – excellent", video bitrate 3 Mbps)

assessors evaluate the first 10 seconds of the video and audio transmission in the experiment.

For the sake of mitigating assessors' burden, we exclude 15 Web clients from the experimental patterns. Thus, the number of stimuli is 32 for each content, i.e., the combination of the three FEC code rates and without FEC, two types of the video transmission bitrate, two types of the number of slices per video frame, and two values of the number of Web clients. Furthermore, we add five dummy stimuli.

In this paper, we treat a polar term "overall quality of video and audio is bad – excellent". The criterion is evaluated to be one of five grades (score 5 is the best, and score 1 is the worst). We then calculate the mean opinion score (MOS), which is an average of the rating scale scores for all the users. The assessors are 17 male students of our university who major in computer science.

Figure 7 is the assessment result with the video transmission rate 3 Mbps. We see in the figure that the MOS value without FEC is larger than that with FEC when the number of Web clients is 10. Besides, the appropriate code rate differs as the number of slices per frame and contents for the number of Web clients 20.

VI. Conclusions

In this study, we assessed QoE of video IP transmission with error concealment mechanisms of H.264/AVC and AL-FEC of MMT. We then found that the appropriate code rate can enhance QoE according to not only network conditions but also contents.

In the future study, the assessment under various conditions including wireless networks is needed. We also need to evaluate with an FEC mechanism for audio.

References

[1] ITU-T Rec. P.10/G.100, "Vocabulary for performance, quality of service and quality of experience," Nov. 2017.

[2] ISO/IEC 23008-1, "Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 1: MPEG media transport (MMT)," Second edition, Aug. 2017.

[3] "FFmpeg, " *http://www.ffmpeg.org/*.

[4] Mindcraft Inc, "WebStone benchmark information," *http://www.mindcraft.com/webstone/*.

[5] "OpenFEC.org," *http://openfec.org/*.