The Effect of Viewpoint Change Strategies on Multi-View Video and Audio Transmission QoE over ICN/CCN*

Toshiro Nunome[1] and Yoshiyuki Takada[2]

Abstract— **In this paper, we exploit caches on intermediate nodes for QoE enhancement of multi-view video and audio transmission over ICN/CCN by controlling the content request start timing of consumers. We assume the selected single viewpoint transmission method; a consumer receives video and audio streams of a requested viewpoint. We perform a simple experiment with two consumers. When the consumers play video and audio with the time difference, we assess the effect of cached content by the former consumer's request on the output quality of the latter consumer. We deal with two types of viewpoint change strategies for the former consumer, which affect the efficiency of cache utilization. From the assessment results, we see that cache utilization has an important factor in enhancing QoE.**

I. Introduction

ICN (Information Centric Networking) [1] has been focused on as an information distribution network architecture. In ICN, the recipient of information content does not specify a source server to request. It just uses an identifier for retrieving the content from the network.

In ICN, intermediate router nodes can have a cache for storing received content. When a router receives a request and has the requested content in the cache, it can distribute the content without relaying the request to the content source. Although some architectures for ICN have been proposed, we employ CCN (Content Centric Networking) [2] as a realized architecture in this paper.

In traditional audiovisual streaming services, the user watches a precomposed view by the content producer. On the other hand, multi-view video [3] can transfer the right of viewpoint selection to the user. It can enhance the user's experience.

Even in ICN/CCN, QoE (Quality of Experience) [4] is an essential quality measure in network services. QoE is the perceived quality of the end user.

Owing to the familiarity of ICN/CCN for content distribution, there are a lot of studies of audiovisual transmission over ICN/CCN. Reference [5] evaluates the performance of video transmission over ICN/CCN with various cache replacements and decision policies. However, it does not consider multi-view video. In addition, QoE evaluation is not carried out.

*This work was supported by JSPS KAKENHI Grant Number 20K11788.

[1]Toshiro Nunome is with Computer Science Program, Department of Engineering, Graduate School of Engineering, Nagoya Institute of Technology, Nagoya 466–8555, Japan nunome@nitech.ac.jp

[2]Yoshiyuki Takada is with Computer Science Program, Department of Engineering, Graduate School of Engineering, Nagoya Institute of Technology, Nagoya 466–8555, Japan takada@inl.nitech.ac.jp

QoE-based study on ICN/CCN has been performed in [6]. It considers video streaming from multiple content stores and proposes an adaptive video streaming with distributed caching algorithm. Stohr et al. have analyzed the video streaming behavior of video quality adaptation algorithms in an emulated ICN/CCN environment [7]. Reference [8] compares cache decision policies in ICN/CCN on QoE of video and audio streaming. However, the papers also do not deal with multi-view video.

Reference [9] introduces multi-view video and audio transmission over IP networks and evaluates application-level QoS [10] and QoE. The application-level QoS is an objective quality metric at the application layer; it is closely related to QoE. Reference [11] deals with multi-view video over MPEG-DASH. However, they do not deal with ICN/CCN.

As for multi-view video over decentralized networks, Kurutepe et al. have considered multi-view video over a multi-tree peer-to-peer network [12]. However, the architecture is mainly different from ICN/CCN.

We face several challenging issues when we transmit multi-view video over ICN/CCN. In ICN/CCN, caches in the intermediate nodes have a significant role. Therefore, in this paper, we perform an experimental study of multi-view video and audio transmission over CCN. As a first step, we use two recipients with different content request start timing to exploit cached content on the intermediate nodes. We investigate the effect of viewpoint requesting behavior, playout buffering time, and interference traffic on application-level QoS and QoE. We then show the feasibility of QoE enhancement utilizing caches on intermediate nodes.

We organize the remainder of this paper as follows. Section II explains multi-view video and audio transmission over CCN. Section III introduces the experimental system. Section IV evaluates application-level QoS, and Section V assesses QoE. Finally, Section VI concludes this paper.

II. Multi-view video and audio over CCN

In this paper, we employ Cefore [13],[14] as a software platform for CCN. It is developed by NICT (National Institute of Information and Communications Technology), Japan. In this paper, the communication of Cefore is overlaid on UDP/IP.

A. CCN

CCN is an implemented architecture of ICN. In CCN, a client is called a consumer, and an information source is called a producer. CCN has two types of packet format: Interest and Data. Interest is a packet for information requests.

979-8-3503-0219-6/23/$31.00 ©2023 IEEE

Data is a packet for the content itself. The content consists of chunks. A chunk is an information unit for content request and retrieval.

The face is a logical interface for connecting adjacent nodes in CCN. It is an enhanced interface from the conventional one for connecting nodes to transfer data to applications.

A node in CCN has three tables for routing and caching: FIB (Forwarding Information Base), PIT (Pending Interest Table), and CS (Content Store). FIB is a table storing the relationship between content name and face for routing Interests. Each node refers FIB to relay Interests to the following routers. PIT memorizes the relationship between the content name and the face for sending Data on the reverse route through Interest. CS is a cache for content. The cache is managed by a unit of chunk. CS stores the content and its name.

B. Multi-view video and audio

Multi-view video can provide video and audio from the viewpoint requested by the user. In this paper, we employ the selected single viewpoint transmission method [11]. It transmits only the requested viewpoint. Hence, when the viewpoint changes, the server changes the transmitted stream to the requested viewpoint.

In traditional IP networks, the client communicates with the server by the IP address. When the user changes the viewpoint, the client requests the viewpoint change to the server.

In CCN, the consumer, the recipient of content, requests content via Interest. When the user changes the viewpoint, the consumer changes the requested viewpoint by changing the content identifier in Interest. If neighbor routers cache the content, the consumer can achieve the content from the routers. Therefore, in CCN, the consumer does not always have to communicate with the producer.

III. Experimental system

A. Network configuration

Figure 1 shows the experimental network with five nodes in this study. The five nodes are desktop PCs working with Ubuntu Linux. A Producer, two Consumers, and two CCN Routers are connected with full-duplex Ethernet links. The communication speed between Producer and CCN Router 2 is 1000 Mbps, and between CCN Router 1 and each Consumer is 100 Mbps. We consider the link between two CCN Routers as a bottleneck link; the communication speed is restricted to 10 Mbps by means of tc (traffic control) command in Linux.

Producer caches the content in its CS. CCN Router 1 also has CS, while CCN Router 2 does not cache content. Producer has enough capacity to cache the whole of the content in this study. The cache size of CCN Router 1 is 5000 chunks. The cache principle is FIFO.

As the interference traffic of audiovisual stream from Producer to Consumers, CCN Router 2 generates 1500 bytes of IP datagrams with exponentially distributed intervals and sends them to CCN Router 1 through UDP/IP. We consider three amounts of interference traffic: 5, 6, or 7 Mbps.

B. Specification of video and audio

Table I shows the specification of video and audio in this study. An MU (Media Unit) is an application-level information unit for media synchronization. The content is a toy train running on plastic rails, as shown in Fig. 2. A video picture frame is an MU. It is divided into ten slices. Each video chunk has a video slice. In audio, an audio chunk consists of an audio MU.

Consumer employs playout buffering as an intra-stream media synchronization algorithm. It stores an MU in a receive buffer until the target output time determined by the MU birth time and the buffering time, and a delayed MU by the target output time is discarded. In this paper, to assess the effect of the playout buffering time on audiovisual quality and viewpoint change response, we employ nine values of the playout buffering time from 100 ms to 500 ms with a 50 ms interval. The playout buffering time causes a tradeoff between output quality and viewpoint change response. We utilize error concealment when Consumer does not receive video slices. For the I slice, we conceal the slice from slices in the same video frame. We use the slice in the reference video frame for concealment of the P slice.

C. Methodology

In the experiment, we assume a situation in which Consumer can exploit caches by controlling the content request's start timing. Consumer 1 first starts the content request. 500 ms after starting the content request of Consumer 1, Consumer 2 starts to request content. We consider the case in which Consumer 2 can exploit the cache in CCN Router 1 stored according to Consumer 1's request.

When Consumer transmits Interest, it is relayed by router nodes and finally arrives at nodes with the content, i.e., CCN Router 1 or Producer. The node with the content sends the Data packet, including the requested chunk, to Consumer, which sends Interest. The intermediate node can cache the content when the Data packet is relayed via the node.

The viewpoint change behavior of two Consumers can affect the application-level QoS and QoE. If Consumers 1 and 2 request the same viewpoint, Consumer 2 can use the cached content. Hence, in this case, the video and audio stream through the bottleneck link is for a viewpoint. Otherwise, the streams through the bottleneck link are for two viewpoints requested by two Consumers.

Therefore, we model two types of viewpoint change behavior for the objective application-level QoS assessment. One is when the user follows the train's movement; we call it "sequential." The other is when the user changes the viewpoint randomly, called "random." The interval of viewpoint change is five seconds in both cases.

IV. Application-level QoS assessment

In the application-level QoS assessment, we deal with two cases of viewpoint change strategies of two Consumers.

979-8-3503-0219-6/23/$31.00 ©2023 IEEE

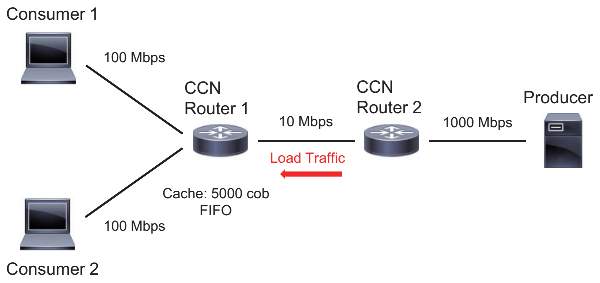

Fig. 1. Experimental network

TABLE I
SPECIFICATION OF VIDEO AND AUDIO

item	video	audio
codec	H.264 (JM19.0)	MPEG-4 AAC-LC
picture pattern	IPPPP	-
image size [pixels]	640 × 480	-
slices	10	-
average MU size [bytes]	I:23615 P:4532	348
MU rate [MU/s]	30.0	46.875
average bitrate [kb/s]	2000	115

- Case 1: Both Consumer 1 and 2 are sequential.
- Case 2: Consumer 1 is random, and Consumer 2 is sequential.

A. Application-level QoS parameter

In this study, we evaluate the audio MU loss ratio, the video MU loss ratio, and the viewpoint change delay. The MU loss ratio is the ratio of the MUs not output to the total number of MUs. The viewpoint change delay is the interval from the viewpoint change request to the output of the requested video.

B. Assessment results

Figures 3 through 11 show the application-level QoS assessment results for each viewpoint change strategy and the amount of interference traffic. The tendency of audio MU loss ratio is almost the same as that of video MU loss ratio. Therefore, this section mainly discusses the video MU loss ratio and the viewpoint change delay.

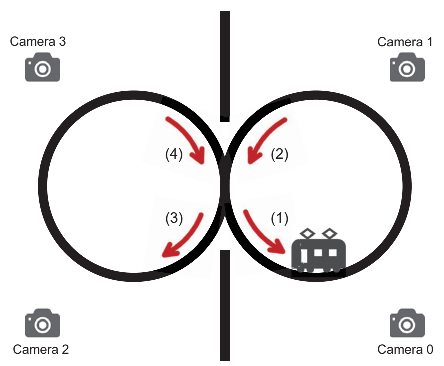

Fig. 2. Content

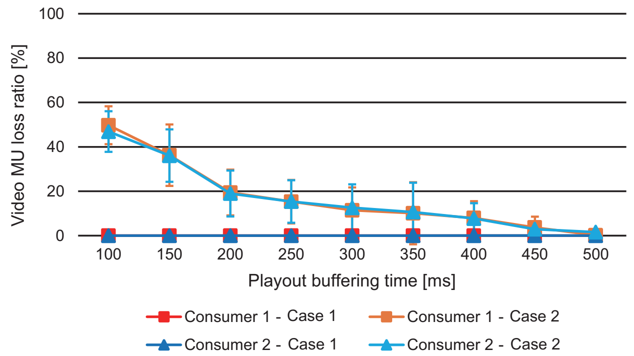

Fig. 3. Video MU loss ratio, average load 5 Mbps

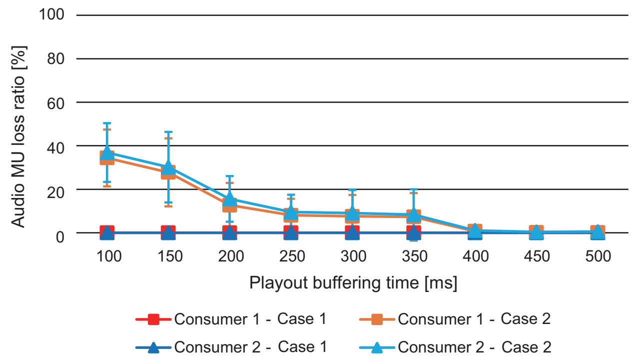

Fig. 4. Audio MU loss ratio, average load 5 Mbps

The results in the figures are the average of ten experimental runs. The figures also show 95 % confidence intervals. The abscissa means the playout buffering time, and the legend shows the Consumer and the case of the viewpoint change strategy.

1) MU loss ratio: In Fig. 9, we see that the video MU loss occurs in Case 1 for the average load of 7 Mbps with a playout buffering time of 100 ms, although the MU loss ratio with a large buffering time is almost 0. This is because the MU cannot arrive by the scheduled playout timing owing to the congestion and short playout buffering time.

In Case 2, we find in Figs. 3, 6, and 9 that the video MU loss ratio decreases as the playout buffering time increases, irrespective of the amount of load traffic. This is because

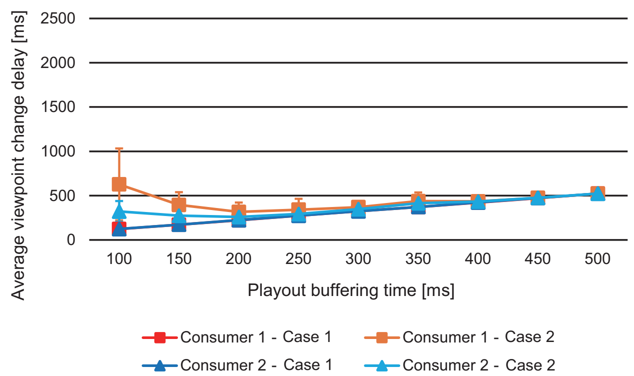

Fig. 5. Viewpoint change delay, average load 5 Mbps

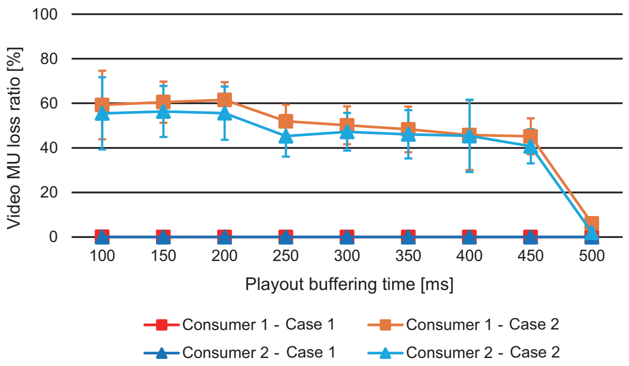

Fig. 6. Video MU loss ratio, average load 6 Mbps

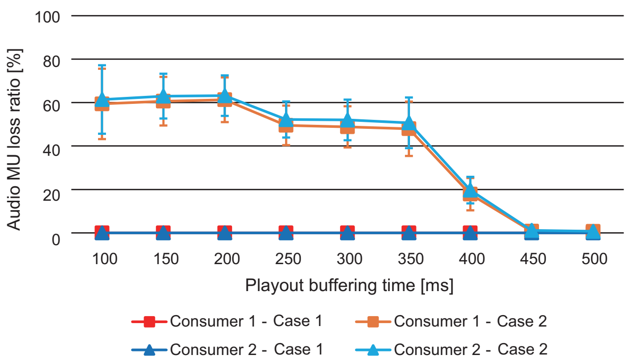

Fig. 7. Audio MU loss ratio, average load 6 Mbps

sufficient playout buffering time can wait for many delayed arrivals.

We also notice in Fig. 9 that Consumer 2 has a smaller MU loss ratio than Consumer 1 for the average load of 7 Mbps in Case 2. This is because Consumer 2 can exploit the cached content even in Case 2.

2) Average viewpoint change delay: In Figs. 5, 8, and 11, we find that the viewpoint change delay increases as the playout buffering time increases in Case 1. This is because the playout buffering time increases the waiting time before output. We also notice in Fig. 5 that the viewpoint change delay in Case 2 with the average load of 5 Mbps also has the same tendency as in Case 1.

On the other hand, in Case 2 with heavy load traffic in

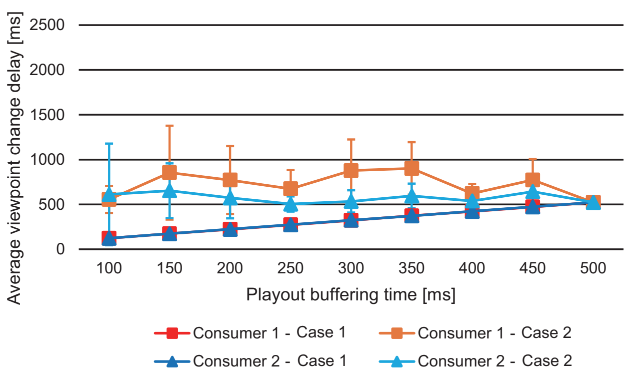

Fig. 8. Viewpoint change delay, average load 6 Mbps

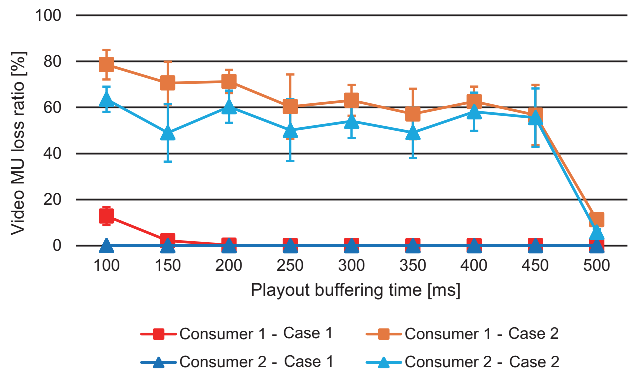

Fig. 9. Video MU loss ratio, average load 7 Mbps

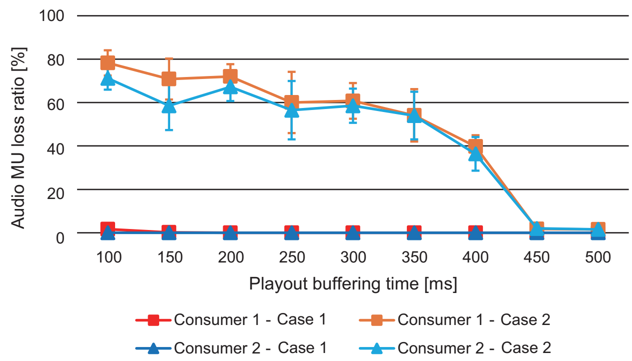

Fig. 10. Audio MU loss ratio, average load 7 Mbps

Figs. 8 and 11, the viewpoint change delay does not have a strong relationship with the increase in the playout buffering time. This is because of unstable MU output timing due to congestion.

V. QoE ASSESSMENT

In the subjective QoE assessment, we employ the same conditions of average load traffic and the playout buffering time as in the application-level QoS assessment. The assessors watch stimuli at Consumer 2. They can change viewpoints as they like.

The viewpoint change strategies in this section are as follows.

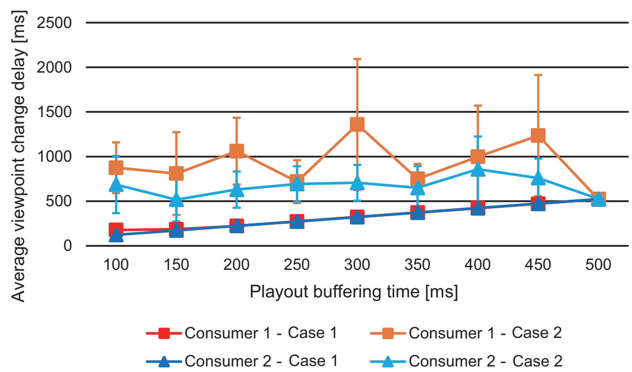

Fig. 11. Viewpoint change delay, average load 7 Mbps

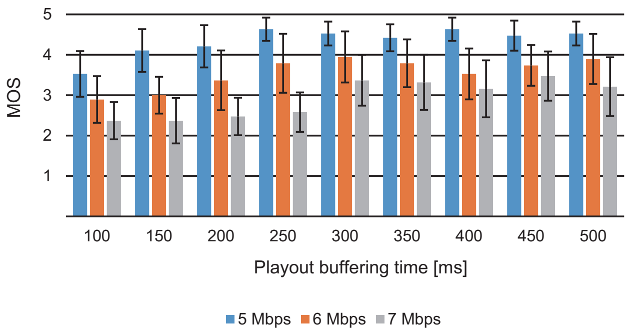

Fig. 12. MOS of audiovisual quality in Case 1

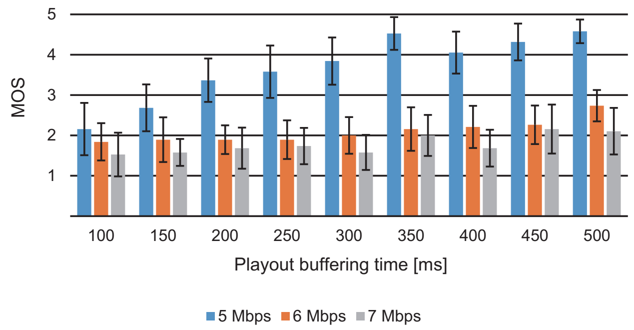

Fig. 13. MOS of audiovisual quality in Case 2

- Case 1: Consumer 1 is sequential, and Consumer 2 is manipulated by the assessor.
- Case 2: Consumer 1 is random, and Consumer 2 is manipulated by the assessor.

A. Methodology

The assessment criteria are audiovisual quality, viewpoint change response, and overall quality. The five-grade impairment scale evaluates the audiovisual quality: 5 imperceptible, 4 perceptible but not annoying, 3 slightly annoying, 2 annoying, and 1 very annoying. The other two criteria, the viewpoint change response and the overall quality, are scored within the five-grade absolute quality scale: 5 excellent, 4 good, 3 neutral, 2 poor, and 1 bad. The integer value is regarded as a subjective score. We then calculate MOS (Mean Opinion Score) as the quantitative measure of perceptual quality.

In the subjective evaluation, at first, we present the multi-view video and audio without degradation as the basis of evaluation. We then show the stimuli with various load and playout buffering conditions. The assessors see the video and audio with viewpoint change through 20 seconds in each experimental run and then score the three criteria.

The assessors are 19 Japanese male students in their twenties. Our future study will investigate the impact of the assessors' attributes.

B. Assessment result

Figures 12 and 13 show the MOS of audiovisual quality in Cases 1 and 2, respectively. In the same way, Figs. 14 and 15 represent the MOS of viewpoint change response. The MOS of overall quality is depicted in Figs. 16 and 17. The abscissa of the figures shows the playout buffering time. The legend shows the amount of interference traffic.

1) Video and audio quality: We find in Figs. 12 and 13 that the MOS of audiovisual quality increases as the playout buffering time increases for the average load of 5 Mbps. This is because MUs that can arrive at the receiver by the target output time increase as the playout buffering time increases.

On the other hand, when the average load is 6 or 7 Mbps, the MOS degrades against that with the average load of 5 Mbps owing to MU loss, as we find in the application-level QoS assessment. We find that the MOS in Case 1 is larger than in Case 2. This is because Consumer 2 can use the cached chunks according to Consumer 1's request in Case 1. In Case 2, the viewpoint of Consumer 1 is mainly different from that of Consumer 2, and then Consumer 2 cannot utilize cached chunks efficiently.

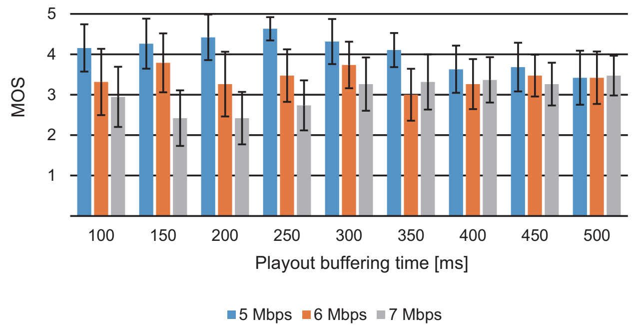

Fig. 14. MOS of viewpoint change response in Case 1

2) Viewpoint change response: In Figs. 14 and 15, we see that the MOS for viewpoint change response decreases as the playout buffering time increases from 300 ms to 500 ms when the average load is 5 Mbps. This is because the playout buffering time increases the waiting time for the output.

When the average load is 6 or 7 Mbps, the variance of MOS value against the playout buffering time is smaller than that for the average load of 5 Mbps. This is because the effect of MU loss becomes dominant under heavily loaded conditions.

3) Overall quality: In Fig. 16, we notice that the MOS of overall quality for Case 1 with a 5 Mbps load has a peak against the playout buffering time at around 250 ms.

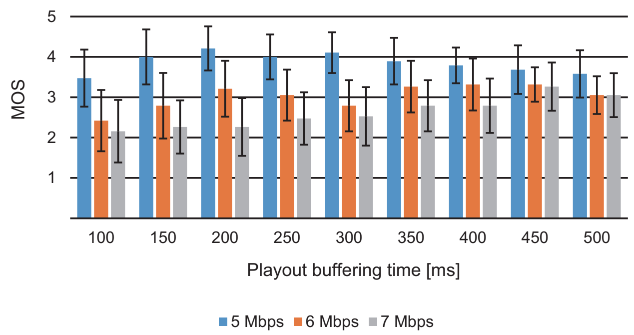

Fig. 15. MOS of viewpoint change response in Case 2

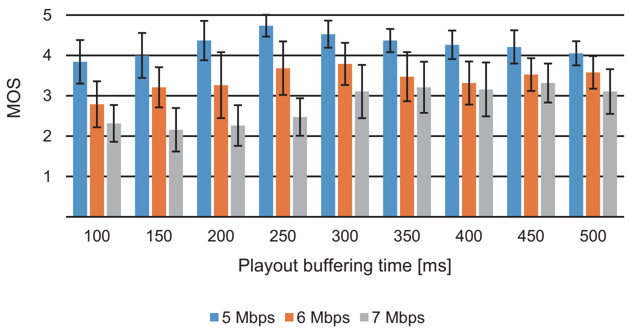

Fig. 16. MOS of overall quality in Case 1

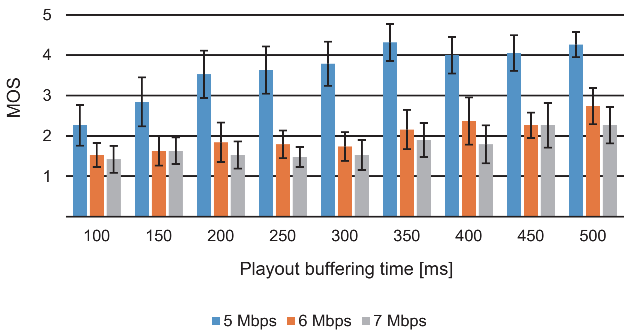

Fig. 17. MOS of overall quality in Case 2

This is because of the tradeoff of quality and response; as the playout buffering time increases, the viewpoint change response degrades while the output quality enhances.

On the other hand, when the average load is 7 Mbps, the MOS of overall quality increases as the playout buffering time increases. Under this condition, enhancing audiovisual quality by extending playout buffering time has a more significant impact than degrading viewpoint change response.

We find in Fig. 17 that the MOS of overall quality in Case 2 with 6 or 7 Mbps load is not affected by the playout buffering time. This is because the effect of playout buffering time on audiovisual quality and viewpoint change response is small owing to the viewpoint change strategy, as we see in Figs. 13 and 15.

VI. Conclusions

In this paper, we experimented with multi-view video and audio transmission over ICN/CCN, in which two recipients had different content request start timing. We then evaluated the effect of viewpoint change strategies, the playout buffering time, and the amount of load traffic on application-level QoS and QoE. As a result, we found that QoE can be enhanced through cache utilization when the viewpoint change behavior of two consumers is the same. As in the traditional IP networks, we saw a tradeoff between the output quality and the viewpoint change response by the playout buffering time.

In future studies, we need to devise appropriate parameter settings in more diverse network conditions. We also have to evaluate different viewpoint change strategies.

References

[1] D. Trossen, M. Sarela, and K. Sollins, "Arguments for an information-centric internetworking architecture," ACM SIGCOMM Computer Communication Review, pp. 26-33, Apr. 2010.
[2] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, "Networking named content," Proc. ACM CoNEXT 2009, pp. 1-12, Dec. 2009.
[3] K.-J. Oh, M. Kim, J. S. Yoon, J. Kim, I. Park, S. Lee, C. Lee, J. Heo, S.-B. Lee, P.-K. Park, S.-T. Na, M.-H. Hyun, J. W. Kim, H. Byun, H. K. Kim, and Y.-S. Ho, "Multi-view video and multi-channel audio broadcasting system," Proc. 3DTV Conference 2007, May 2007.
[4] ITU-T Rec. P.10/G.100, "Vocabulary for performance, quality of service and quality of experience," Nov. 2017.
[5] Y. Sun, S. K. Fayaz, Y. Guo, V. Sekar, Y. Jin, M. A. Kaafar, and S. Uhlig, "Trace-driven analysis of ICN caching algorithms on video-on-demandworkloads," Proc. ACM CoNEXT, pp. 363-376, Dec. 2014.
[6] M. N. Sadat, R. Dai, L. Kong, and J. Zhu, "QoE-aware multi-source video streaming in content centric networks," IEEE Trans. on Multimedia, vol. 22, issue 9, pp. 2321-2330, Sept. 2020.
[7] D. Stohr, T. Kalle, A. Mauthe, A. Rizk, R. Steinmetz, and W. Effelsberg, "Towards improved DASH adaptation in NDN: An emulative analysis," Proc. IEEE LCN 2018, pp. 385-392, Oct. 2018.
[8] K. Kobayashi and T. Nunome, "The impact of ICN/CCN cache decision policies on video and audio transmission QoE," Proc. ITNAC 2022, pp. 202-207, Nov./Dec. 2022.
[9] E. Jimenez Rodriguez, T. Nunome and S. Tasaka, "QoE assessment of multi-view video and audio IP transmission," IEICE Trans. on Commun., vol. E93-B, no. 6, pp. 1373-1383, June 2010.
[10] S. Tasaka and T. Nunome, "QoE enhancement of multimedia communications by user-assistance," Proc. World Telecommunications Congress 2014, SA.5.3, June 2014.
[11] Y. Maehara and T. Nunome, "Multidimensional QoE assessment of a simultaneous transmission method in multi-view video and audio transmission with MPEG-DASH," Proc. IEEE CIT 2017, pp. 101-106, Aug. 2017.
[12] E. Kurutepe and T. Sikora, "Feasibility of multi-view video streaming over P2P networks," Proc. 3DTV Conference 2008, May 2008.
[13] H. Asaeda, A. Ooka, K. Matsuzono, and R. Li, "Cefore: Software platform enabling content-centric networking and beyond," IEICE Trans. Commun., vol. E102-B, no. 9, pp. 1792-1803, Sept. 2019.
[14] CeforeWeb, https://cefore.net/.