The Effect of Bandwidth Allocation Methods on QoE of Multi-View Video and Audio IP Transmission

Toshiro Nunome and Keita Furukawa
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 compares two allocation methods of bandwidth for each video traffic in multi-view video and audio transmission over bandwidth guaranteed IP networks. The methods are a main viewpoint prioritized allocation method and an equally allocation method. We assume simultaneous video transmission of multiple viewpoints. When bandwidth is insufficient for simultaneously transmitted multiple video streams, we investigate the effect of two allocation methods on QoE. We conduct an subjective experiment with the combination of bandwidth allocation methods and playout buffering time. As a result, we see that the appropriate allocation method and playout buffering time change by available bandwidth and contents.**

***Index Terms*—MVV, bandwidth guarantee, audio-video IP transmission, QoE, bandwidth assignment**

I. Introduction

Owing to high-speed and broadband IP networks, multimedia communication services have been popularized. Such network services require high *QoS (Quality of Service)*. However, best-effort IP networks cannot guarantee QoS and then cause packet losses and delays. On the other hand, bandwidth guaranteed IP networks can provide QoS for flows on the networks. The guaranteed networks are suitable for delay critical services.

As a multimedia communication service, we can consider *MVV (Multi-View Video)*, which treats multiple video streams obtained from plural cameras [1]. In [2], *QoE (Quality of Experience)* [3] of *MVV-A (MVV and Audio)* IP transmission is assessed multidimensionally. The paper considers that the server transmits only one video stream selected by the user. In this case, the viewpoint change response will be quick as the playout buffering time decreases. However, short buffering time cannot absorb network delay jitter sufficiently, and then the output quality of audio and video degrades. In addition, the viewpoint change response is affected by end-to-end delay between the server and the client. The degradation of output quality and viewpoint change response leads to the degradation of QoE.

From the results in [2], short buffering time is effective to achieve high user satisfaction for small load traffic, and long buffering time is well for large load traffic. The experiment in [2] is performed on the best-effort IP network. In bandwidth guaranteed networks, appropriate buffering time can differ from that in the best-effort network.

Reference [4] assesses the effect of simultaneous transmission methods on QoE. When the user wants to change his/her viewpoint, if the user's terminal already receives the requested video stream, the viewpoint change can be done quickly. Otherwise, it needs to request the video to the server. As MVV transmission schemes which exploit a tradeoff between the viewpoint change response and image quality determined by video encoding bitrate of each viewpoint, the paper considers three simultaneous transmission methods of video streams. It also assumes the best-effort IP network.

In the bandwidth guaranteed networks, we can prepare dedicated bandwidth for video and audio streams. It can avoid the effect of network congestion by interference traffic and then achieve high output quality of audio and video. In [5], Ito and Tasaka study the effect of the video GOP, video output scheme and content type on the optimum bandwidth allocation between audio and video. However, the paper does not consider MVV-A.

In this paper, we apply the bandwidth guaranteed IP network for simultaneous transmission of MVV-A in which the user watches only one viewpoint selected by him/her. When bandwidth is insufficient for simultaneously transmitted multiple video streams, how to allocate the bandwidth for each stream can affect the output quality of MVV-A. Hence, we assess the effect of bandwidth allocation methods, audio sampling rate, and playout buffering time on QoE.

When we threat multiple viewpoints, we can consider not only viewing the selected viewpoint but also viewing plural viewpoints at once. Reference [6] deals with an MVV-A system with four cameras and employs three transmission and presentation methods: the one-view one-stream method, the one-view four-streams method and the four-views four-streams method. Among them, the four-views four-streams method transmits and displays all the viewpoints with low bitrate. Through QoE assessment, the paper shows that when the user cannot change the viewpoint quickly, the user feels that the usability of the four-views four-streams method is good. However, in this paper, as a first step of the study for the bandwidth guaranteed IP network, we employ the one-view method.

The rest of the paper is structured as follows. Section II defines the bandwidth allocation methods in this paper. Section III outlines methods of the experiment. We present results of the experiment in Section IV, and Section V concludes this paper.

II. Bandwidth allocation methods

The bandwidth guaranteed network can ensure network resource. However, it takes costs for ensuring the resource. Thus, rich multimedia services such as MVV-A cannot always achieve enough bandwidth. For allotting limited bandwidth for each video stream, we can consider several strategies.

If we allot much bandwidth for the main viewpoint (i.e., the selected viewpoint), we can keep the output quality high, while the video frame loss of sub viewpoints (i.e., unselected viewpoints) becomes often. In this case, when the user changes viewpoint, the receiver cannot exploit simultaneously transmitted sub viewpoints effectively; it develops large viewpoint change delay or disturbance of video output. On the other hand, if we allot rich bandwidth for sub viewpoints, the viewpoint change delay can decrease, while the output quality of the main viewpoint can degrade.

In this paper, we compare the following two bandwidth allocation methods.

- Equally allocation method
  The method equally allocates bandwidth for each video stream.
- Main viewpoint prioritized allocation method
  The method allocates larger bandwidth for the main viewpoint. The rest bandwidth is equally divided for each sub viewpoint.

Here, we define

N: number of viewpoints in the system
BW_T: total bandwidth for video and audio
BW_A: bandwidth for audio
BW_{MV}: bandwidth for main viewpoint video
BW_{SV}: bandwidth for each sub viewpoint video

Then, the bandwidth for each sub viewpoint video in the equally allocation method is Eq. (1)[1].

$$BW_{SV} = BW_{MV} = \frac{BW_T - BW_A}{N} \qquad (1)$$

That in the main viewpoint prioritized allocation method is Eq. (2).

$$BW_{SV} = \frac{BW_T - BW_A - BW_{MV}}{N-1} \qquad (2)$$

III. Methodology of experiment

A. Experimental system

Figure 1 shows the experimental network in this paper. The system consists of Media Server, Media Client, Load Sender, Load Receiver, and two routers. The two routers are Cisco 7301. All the communication links in the network are 100 Mb/s duplex Ethernet.

[1]In the experiment, as we consider the standard deviation of video encoding bitrate, the bandwidth is slightly different for each video stream.

TABLE I
Specifications of video and audio

	video	audio
coding scheme	H.264 (704×480)	Linear PCM
average MU rate [MU/s]	30	25
average encoding bitrate [kbps]	3000	128, 512
picture pattern	I frame only	-

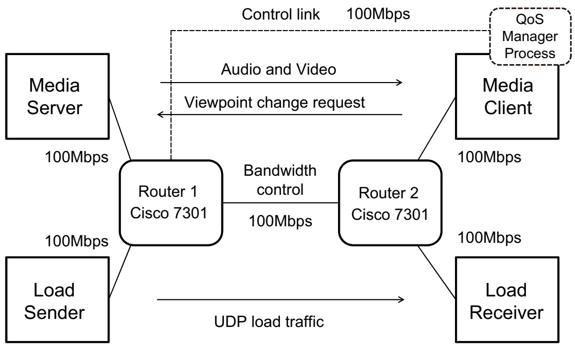

Fig. 1. Experimental system

For bandwidth management, we assume session control by means of SIP (Session Initiation Protocol). We put a QoS manager between Media Server and Media Client. For simplicity, we house the QoS manager in Media Client. By using the dedicated link shown in Fig. 1, the QoS manager controls Router 1.

Media Server transmits audio and video streams to Media Client after establishment of the session. Router 1 performs bandwidth control by *LLQ (Low Latency Queuing)* [7]; it employs *PQ (Priority Queuing)* for audio and applies *CBWFQ (Class-Based Weighted Fair Queuing)* for video and load traffic.

In the main viewpoint prioritized allocation method, when a viewpoint change request occurs, Router 1 needs to change bandwidth allocation. The QoS manager knows the requested viewpoint by a SUBSCRIBE message for viewpoint change and then requests bandwidth rearrangement to Router 1.

For load traffic, we assign the residual bandwidth (i.e., 100 Mbps − audio and video bandwidth). Load Sender transmits load traffic (1480 bytes UDP packets generated with the exponentially distributed intervals) as the average bitrate equals to the assigned bandwidth. In LLQ, when surplus bandwidth exists, the bandwidth is re-allocated to CBWFQ classes. In order to consume the surplus bandwidth, we employ the load traffic.

B. Specification of video and audio

We consider an MVV-A system with four cameras, i.e., the system has four viewpoints. In this study, we employ pre-recorded video and audio streams. The specifications of audio and video are shown in Table I. Here, an *MU (Media*

Unit) is a unit for media synchronization control. A video MU is a video frame and an audio MU is audio samples for 40 ms. The target encoding bitrate of video is 3000 kbps. We employ CBR (Constant Bit Rate) for video encoding; however, the actual bitrate of encoded video fluctuates. We employ two types of audio: high quality (32 kHz sampling, 512 kbps) and standard quality (8 kHz sampling, 128 kbps).

Media Client uses playout buffering control for absorbing network delay jitter. In order to investigate the effect of playout buffering time, we employ two values of playout buffering time: 100 ms and 500 ms. As we discussed earlier, in the MVV-A system, the playout buffering control brings tradeoff between the viewpoint change response and output quality [2].

Each MU is transmitted as a RTP/UDP packet. There is no retransmission mechanism. We employ frame skipping as the output method of video. That is, when some packets consisting of an MU is lost, output of the MU is skipped.

Table II shows the experimental parameters in this paper. The duration of each experimental run is 30 seconds. The total time of the experiment for an assessor is about 50 minutes.

C. Allocated bandwidth

When we send application data through the IP networks, header information is added in each layer of the Internet protocol stack. The bandwidth control is performed on the data-link layer. Thus, when we assign bandwidth for the average media encoding bitrate and the header size of each layer, the minimum bandwidth will be guaranteed; however, starvation can occur because of bitrate fluctuation.

In this study, we employ two cases for bandwidth allocation: *wide* and *narrow*. In the narrow case, we assign bandwidth according to the average encoding bitrate with header size for the four video streams and standard quality audio; it is minimum allocation.

When we transmit pre-encoded video like video-on-demand services, we can obtain specifications of video and audio in advance. In order to absorb the fluctuation of bitrate, we can assign bandwidth by peak bitrate [8]; it may waste bandwidth. For enough assignment, we consider the standard deviation of video encoding bitrate for bandwidth allocation.

As a preliminary experiment, we transmitted audio and video with bandwidth allocation by means of the following equation.

Allocated Bandwidth [kbps]
= Target encoding bitrate [kbps]
+ Size of headers [kbps]
+ Standard deviation of the encoding bitrate [kbps]

We then found that the MU loss scarcely occurs. Hence, we employ the equation for the bandwidth allocation for video. We assign the bandwidth for the main viewpoint in the main viewpoint prioritized allocation method and all the viewpoints in the wide case.

At first, we assign bandwidth which equals to the average bitrate for audio. We then allocate bandwidth for video.

TABLE II
EXPERIMENTAL PARAMETERS

parameter	value
playout buffering time	100 ms, 500 ms
bandwidth	wide, narrow
streams	4
allocation method	equally, main prioritized
audio sampling rate	8 kHz, 32 kHz
content	toy insects, lecture
duration	30 seconds
number of experimental runs	40 and four dummies

TABLE III
BANDWIDTH ASSIGNMENT

		VP 1	VP 2	VP 3	VP 4	audio
encoding bitrate		3000	3000	3000	3000	128
with header length		3080	3080	3080	3080	140
standard deviation		158	190	178	194	0
wide	equally	3238	3270	3258	3274	140
narrow	equally	3074	3104	3093	3108	140
	prioritize VP 1	3238	3049	3038	3053	140
	prioritize VP 2	3019	3270	3037	3052	140
	prioritize VP 3	3019	3049	3258	3053	140
	prioritize VP 4	3019	3049	3038	3274	140

VP: viewpoint, [kbps]

Table III shows an example of bandwidth allocation for audio and video when the audio sampling rate is 8 kHz. Note that there is no fluctuation in audio because we employ Linear PCM for audio codec.

D. QoE assessment methods

In the experiment, we employ two contents: *toy insects* and a *lecture*.

In the toy insects, the number of insects are four, and each insect has different color (red, blue, green, and black). Figure 2 shows the camera arrangement for the toy insects. The four cameras surround the course in which the insects move. The microphone is placed on the center of the course. The user follows one of the four insects by changing viewpoint with GUI shown in Fig. 3; it also indicates color of the target insect. This content assumes objects moving high speed; it needs high viewpoint change response.

In the lecture, we take a lecture with slides by four cameras. Figure 4 shows the camera arrangement for the lecture. The four cameras watch the slides (Camera 1: slide), the lecturer (Camera 2: presenter), the view from the audience area (Camera 3: wide), and the audience (Camera 4: audience). The user freely switches the viewpoint every about three seconds by using GUI shown in Fig. 5. The content is audio dominant and has small movement.

In the experiment, we perform multidimensional QoE assessment with six adjective pairs of polar terms. The pairs of polar terms in the subjective experiment are shown in Table IV. Note that the experiment was performed with the Japanese language. This paper has translated the used Japanese terms into English. Therefore, the meanings of

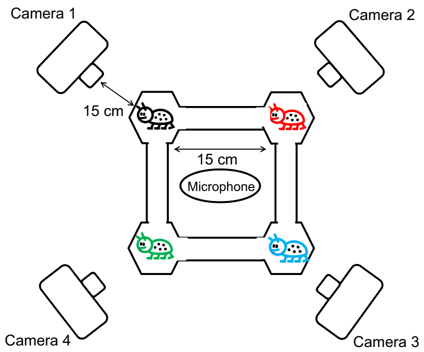

Fig. 2. Contents (toy insects)

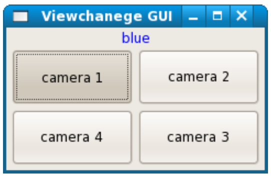

Fig. 3. Viewpoint change GUI (toy insects)

adjectives or verbs written in English here may slightly differ from those of Japanese ones.

For each criterion, a subjective score is measured by the *rating scale method*. In the method, an assessor classifies the stimuli into a certain number of categories; here, each criterion is evaluated to be one of five grades. The best grade (score 5) represents the positive adjective (the right-side one of each pair), while the worst grade (score 1) means the negative adjective (the left-side one). We then calculate *mean opinion score (MOS)* for each criterion as the quantitative measure of perceptual quality. The assessors are 15 male students in their twenties.

IV. Experimental Results

A. Application-level QoS

1) Video MU loss ratio: Figures 6 and 7 show the video MU loss ratio. The MU loss ratio is the ratio of lost or

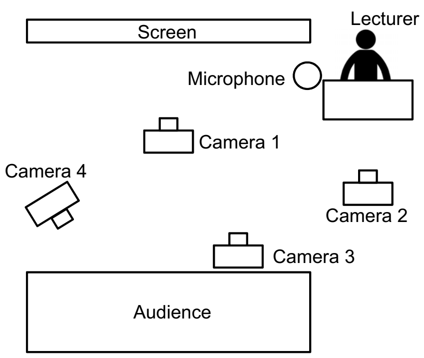

Fig. 4. Contents (lecture)

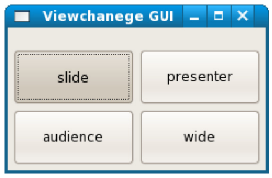

Fig. 5. Viewpoint change GUI (lecture)

TABLE IV
ADJECTIVE PAIRS

category	adjective pair
video	video is rough - smooth
audio	audio is dull - clear
response	viewpoint change is slow - fast
stability	viewpoint change is unstable - stable
synchronization	audio and video are out of synchronization - in synchronization
overall satisfaction	bad - excellent

discarded video MUs at Media Client to the transmitted MUs from Media Server. The main viewpoint loss ratio means the MU loss ratio for the main viewpoint. The sub viewpoint loss ratio means the average of the MU loss ratio of three sub viewpoints.

We find in Fig. 6 that there is little deference between the contents on the MU loss ratio. When the bandwidth is wide, the MU scarcely drops because sufficient bandwidth is assigned for video streams.

When the bandwidth is narrow, the MU loss occurs for the playout buffering time 100 ms. In the equally allocation method, both the main viewpoint and sub viewpoints have approximately the same MU loss ratio. On the other hand, in the main viewpoint prioritized allocation method, the main viewpoint merely has lost MUs, while the sub viewpoints in the method have larger MU loss ratio than those in the equally allocation method.

In Fig. 7, we see that as the audio sampling rate increases, the MU loss ratio of video increases. This is because the allocated bandwidth for video decreases as the audio bitrate

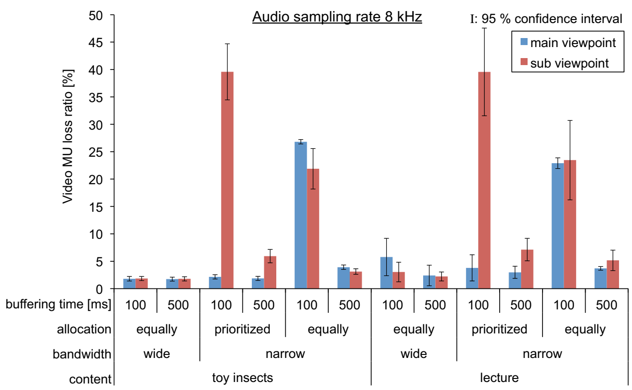

Fig. 6. Video MU loss ratio (audio: 8 kHz)

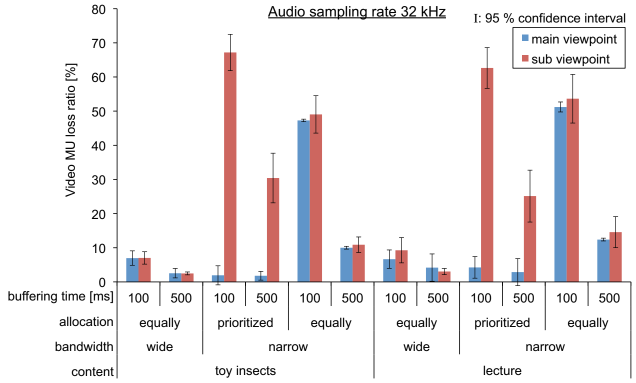

Fig. 7. Video MU loss ratio (audio: 32 kHz)

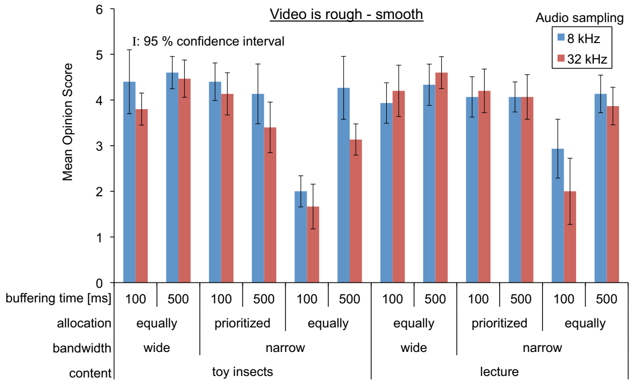

Fig. 9. Video is rough - smooth

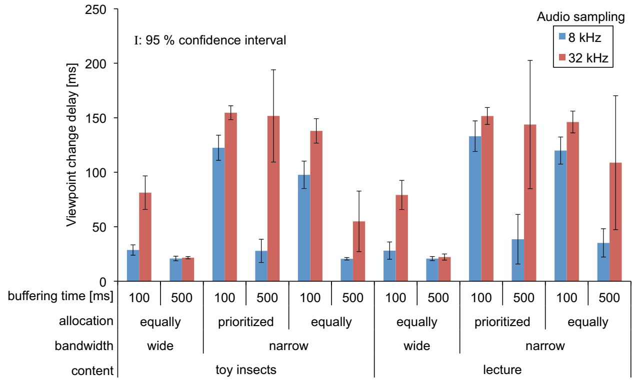

Fig. 8. Viewpoint change delay

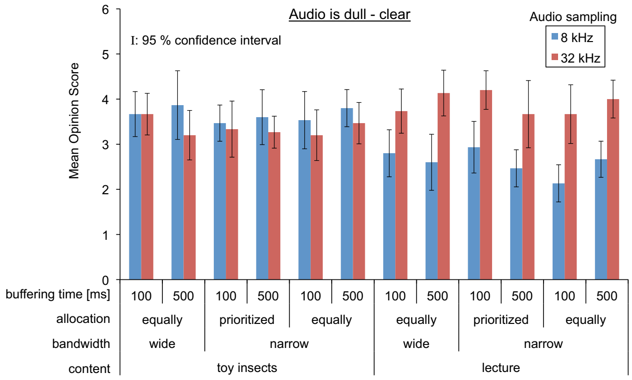

Fig. 10. Audio is dull - clear

increases.

2) Viewpoint change delay: Figure 8 shows the average viewpoint change delay. It is defined as the time in milliseconds from the moment the user inputs a request for viewpoint change by the GUI until the instant a new viewpoint is output at the client.

We notice in Fig. 8 that the viewpoint change delay is lower than 100 ms when the bandwidth is wide. This is because the MU loss ratio of sub viewpoints is small, and then the client can change viewpoint by using the received video for sub viewpoints.

When the bandwidth is narrow, we find that the viewpoint change delay for the main viewpoint prioritized allocation method is larger than that for the equally allocation method. This is because the allocated bandwidth for sub viewpoints in the main viewpoint prioritized allocation method becomes small. The MU loss ratio for sub viewpoints increases, and then the client cannot use the sub viewpoints for viewpoint change efficiently.

In addition, we see in Fig. 8 that the playout buffering time increases, the viewpoint change delay decreases in the main viewpoint prioritized allocation method. This is because the MU loss ratio for sub viewpoints decreases as the playout buffering time increases.

B. QoE

Figure 9 shows the MOS for “Video is rough - smooth”. Figure 10 depicts the MOS for “Audio is dull - clear”. Figures 11 and 12 present the MOS for “Viewpoint change is slow - fast” and that for “Bad - Excellent”, respectively.

1) Video is rough - smooth: We find in Fig. 9 that when the bandwidth is narrow, the main viewpoint prioritized allocation method has higher MOS values than the equally

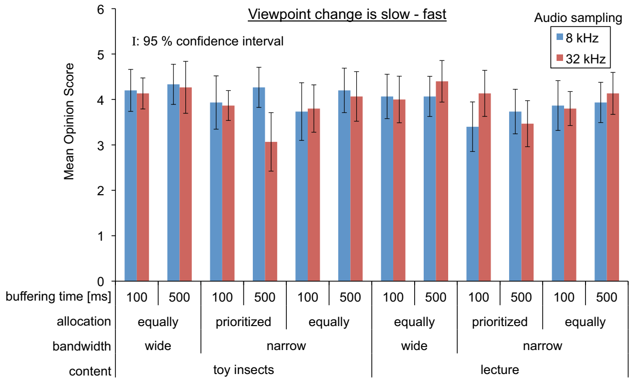

Fig. 11. Viewpoint change is slow - fast

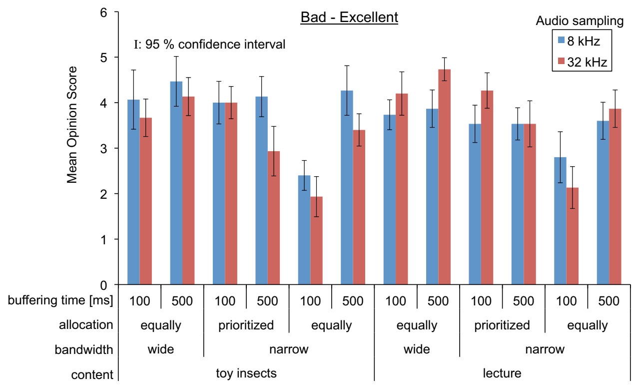

Fig. 12. Bad - Excellent

allocation method. This is because the equally allocation method has large MU loss ratio for the main viewpoint because of small allocated bandwidth for the main viewpoint.

On the other hand, for the playout buffering time 500 ms, we see that the difference between the two methods becomes small. This is because the buffering can absorb delay jitter due to bitrate fluctuation.

We notice in Fig. 9 that for the toy insects, the audio sampling rate 32 kHz has lower MOS for video quality than the audio sampling rate 8 kHz. This is because the allocated bandwidth for video decreases as the audio bitrate increases. However, the difference for the lecture is small. This is because the audio quality affects the perceived video quality in the content; this is a cross-modal effect [9].

2) Audio is dull - clear: We see in Fig. 10 that for the toy insects, the audio sampling rate 8 kHz has larger MOS values than the rate 32 kHz. The audio in toy insects is mechanical noise of toy insects. Thus, the assessors scarcely feel the difference of audio quality.

On the other hand, for the lecture, the audio sampling rate 32 kHz has larger MOS values. The lecture is an audio dominant content. Thus, the audio sampling rate directly affects the perceived audio quality.

3) Viewpoint change is slow - fast: In Fig. 11, we find that for the audio sampling rate 8 kHz, the both methods have high MOS values for the playout buffering time 500 ms when the bandwidth is narrow. This is because the buffering time can decrease lost MUs for sub viewpoints, and then the client can change the viewpoint by using received MUs for sub viewpoints.

4) Bad - Excellent: We notice in Fig. 12 that the equally allocation method has the highest MOS values for the toy insects in the audio sampling rate 8 kHz with the buffering time 500 ms when the bandwidth is narrow. This is because the buffering can absorb fluctuation of bitrate and then decreases lost MUs.

On the other hand, for the lecture, the main viewpoint prioritized allocation method with the buffering time 100 ms for the audio sampling rate 32 kHz has the highest MOS values when the bandwidth is narrow. This is because the content is audio dominant and is not sensitive for viewpoint change response.

In addition, the low audio sampling rate is better for the toy insects, while the high audio sampling rate is better for the lecture. We can consider that the appropriate bandwidth allocation is affected by the contents.

V. Conclusions

In this paper, we investigated the effect of the bandwidth allocation methods on QoE in simultaneous transmission of MVV-A over the bandwidth guaranteed IP network. As a result, we found that when the allocated bandwidth is narrow, for the content in which quick viewpoint change response is required, the equally allocation method with large buffering time is a good strategy. In addition, for the audio dominant content, the main viewpoint prioritized allocation method with high audio sampling rate is better.

For future study, we need to assess the methods in multicast communications.

References

[1] I. Ahmad, "Multiview video: Get ready for next-generation television," *IEEE Distributed Systems Online*, vol. 8, no.3, art. no. 0703-o3006, Mar. 2007.

[2] E. Jimenez Rodriguez, T. Nunome and S. Tasaka, "QoE assessment of multi-view video and audio IP transmission," *IEICE Trans. Commun.*, vol.E92-B, no.6, pp. 1373-1383, June 2010.

[3] ITU-T Rec.P.10/G.100, "Amendment 5: New definitions for inclusion in Recommendation", July 2016.

[4] F. J. Wilson Francis and T. Nunome, "QoE assessment of multi-view video and audio simultaneous IP transmission: The effect of user interfaces," *Proc. ICTC 2014*, pp. 466-471, Oct. 2014.

[5] H. Ito and S. Tasaka, "The optimum bandwidth allocation between audio and video by using temporal and spatial quality of video in bandwidth guaranteed audio-video IP transmission," Tech. Rep. IEICE, CQ2009-51, Nov. 2009. (in Japanese)

[6] T. Nunome and T. Ishida, "Video transmission and presentation methods for multi-view video and audio IP transmission," *Proc. ICSPCS 2014*, pp. 173-178, Dec. 2014.

[7] S. Vegesna, "IP Quality of Service, Cisco Press," 2001.

[8] ITU-T recommendation I.371, "Traffic control and congestion control in B-ISDN," 2004.

[9] S. Tasaka and Y. Ishibashi, "Mutually compensatory property of multimedia QoS," *Conf. Rec. IEEE ICC2002*, pp. 1105-1111, Apr./May 2002.