Multidimensional QoE Assessment of Multi-View Video and Audio (MVV-A) IP Multicast Transmission Methods

Kazunori Sako†, and Toshiro Nunome††
Graduate School of Engineering, Nagoya Institute of Technology, Nagoya 466–8555, Japan
†*Email:sako@inl.nitech.ac.jp* ††*Email:nunome@nitech.ac.jp*

***Abstract*—In this paper, we deal with Multi-View Video and Audio (MVV-A) IP transmission in which a server multicasts the audio and video data to multiple users. We consider two transmission methods; one is the all viewpoints transmission method, while the other is the requested viewpoint transmission method. The former transmits all the viewpoints data. The latter transmits only the requested viewpoint data by the user. This paper compares the two methods in terms of QoE. We conduct multidimensional assessment with eight pairs of polar terms by the SD (Semantic Differential) method, which can assess an object from many points. As a result, we investigate how the multicast transmission methods affect the user's satisfaction.**

***Keywords*-audio-video IP transmission, multicast, QoE, multidimensional assessment**

I. INTRODUCTION

Multimedia communications treating audio and video become popular with the acceleration of transmission speed of wired and wireless IP networks. However, IP networks are generally best-effort; audio and video packets can be lost during transmission and can be affected by network delay jitter. These impairments deteriorate the output quality of audio and video; then, *QoS (Quality of Service)* becomes lower. It leads to deterioration of *QoE (Quality of Experience)* [1] in many cases.

In [1], QoE is defined as *the overall acceptability of an application or service, as perceived subjectively by the end-user*. Enhancement of QoE in best-effort networks is important for many network services.

We can use multicast in order to transmit the same data to multiple users. Multicast reduces the amount of data to be transmitted against the case of using unicast because the multicast transmits the same data to multiple terminals at once.

As a new multimedia service on IP networks, *MVV (Multi-View Video)* is greatly expected [2]. MVV has some types (e.g., free viewpoint television [3], 3D tele-immersive video [4], and viewpoint selectable video [5]). We focus on the viewpoint selectable video; the user can watch video from an arbitrary viewpoint from various viewpoints taken by plural cameras. In this study, we deal with the transmission of *MVV-A (Multi-View Video and Audio)* [6], which is MVV accompanied by audio.

Reference [6] focuses on the behavior of users, and evaluates the effect of playout buffering time on QoE quantitatively when the users use the MVV-A system with two types of *GUI (Graphical User Interface)* for viewpoint change. However, the study is for a unicast environment with a single client. QoE studies in the case that multiple users view the video at the same time as live video streaming services like Ustream [7] are important.

As studies on MVV using multicast, Fujihashi *et al.* have proposed a transmission method which combines unicast and multicast [8]. In [9], two transmission methods which consider interactivity and reduction of the traffic have been employed. These studies focus on reduction of the amount of traffic. However, these studies have not considered users' overall satisfaction.

A content of MVV-A has multiple videos of various viewpoints. We can consider several video transmission methods. For example, one method transmits the videos of all the viewpoints, and another method transmits the videos of the requested viewpoint. The difference in the video transmission methods exist on the amount of data to be transmitted and the response when the user switches viewpoints.

In multicast, there are multiple users enjoying the same content. Their behavior can affect users' QoE, especially in the viewpoint change response.

In this study, we investigate the effect of multicast transmission methods for MVV-A on QoE. Hereby, we discuss a methodology for improvement of QoE. In the MVV-A system, not only the quality of video and audio but also the viewpoint change response affect QoE. Therefore, we assess QoE with the *SD (Semantic Differential)* method [10] that evaluates the object in multiple pairs of adjectives.

The rest of this paper is structured as follows. Section II describes MVV-A with multicast. Section III explains the method of the experiment. We show results of the experiment in Section IV, and Section V concludes this paper.

II. MULTI-VIEW VIDEO AND AUDIO WITH MULTICAST

In the MVV-A system, a user can watch video from an arbitrary viewpoint from various viewpoints taken by plural cameras.

As a method to multicast video in MVV-A, we can consider a method which assigns a multicast address to each viewpoint of a content. However, when a company provides plural contents with multicast, it is desirable that the multicast addresses for one content is a few. Therefore, we compare transmission methods that assign a multicast address for a content in this study. That is, multiple viewpoints are transmitted on the same address.

In this study, we consider two transmission methods.

One is the *all viewpoints transmission method*. The method transmits the videos of all the viewpoints. With this method, the receiver terminals can receive all the viewpoints without viewpoint change requests. Therefore, when switching the viewpoint, each receiver terminal only has to switch received streams inside the terminal; this method has an advantage on viewpoint change response. However, a disadvantage of this method is the large amount of traffic.

The other is the *requested viewpoint transmission method*. This method transmits only the video of requested viewpoint by the user. Thereby, instead of the time lag between transmitting viewpoint change request and receiving the new viewpoint, the amount of transmission data is smaller than that of the all viewpoints transmission method. In this experiment, the viewpoint is transmitted with the same multicast address. Hereby, each user also receives the viewpoint data which the other users requested. Therefore, the user can change the viewpoint to already received one without transmitting the viewpoint change request. Since the media transmission terminal needs to grasp the viewpoint which each user requests, transmission of the viewpoint change request is performed even if the receiver already receives the viewpoint.

The number of transmitted viewpoints in the requested viewpoint transmission method is directly related to the number of users. If the number of users is fewer than the number of viewpoints, the amount of transmission data is smaller than that of the all viewpoints transmission method. On the other hand, if the number of users is larger than the number of viewpoints, and all the viewpoints are watched by the users, this method is equivalent to the all viewpoints transmission method.

III. Experimental Method

A. System

Figure 1 shows the network configuration in this experiment. *MS (Media Server)* transmits media streams for the MVV-A application. *MR (Media Receiver)* receives the media streams. *LS (Load Server)* is the server of the load traffic, and *LR (Load Receiver)* is the client. Netem, which is a PC, is laid out between the routers. This PC delays packets going through routers 1 and 2 by using netem [12]. Both router 1 and router 2 are Cisco's 7301. All terminals

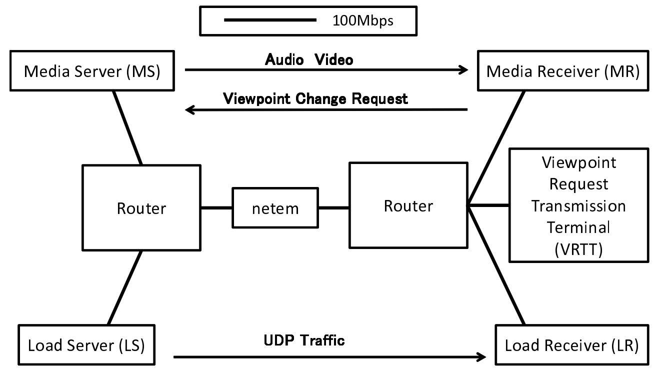

Figure 1. System configuration

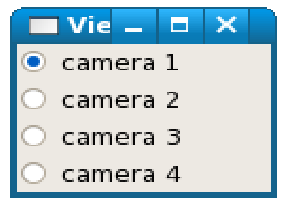

Figure 2. GUI

and routers are connected by a full-duplex Ethernet line of 100 Mbps[1].

Four SONY HDR-CX170 video cameras with the standard definition mode are connected to MS, which is equipped with two real-time H.264 encoding boards by DSP Research; each board accommodates two cameras. MS multicasts video and audio to MR. In the experiment, we use the video which is recorded in advance for simplicity of experiment and easiness to regulate the experimental condition.

MVC (Multi-view Video Coding) [13] is an efficient coding method for MVV. However, in the experiment, we do not adopt MVC because of low correlation among the viewpoints.

When the user changes the viewpoint with a GUI in Figure 2, MR transmits the viewpoint change request to MS with the SUBSCRIBE method in SIP (Session Initiation Protocol).

VRTT (Viewpoint Request Transmission Terminal) emulates a couple of media receiver terminals; it transmits viewpoint change requests automatically for emulating other users. In the experiment, we assume the network service which has plural users. Therefore, the viewpoint change requests are generated by the plural users. In this paper, we define a user which perform subjective assessment in the experiment as a subject.

[1]In this paper, we employ the wired network for simplicity. We can include wireless networks in the experimental framework. This is future work.

Table I
AUDIO AND VIDEO SPECIFICATIONS

	video	audio
Coding method	H.264	ITU-T G.711 μ-law
Average MU rate [MU/s]	30	25
Average bit rate [kbps]	4000	64
Picture pattern	I	-
Image size [pixels]	704 $\times$ 480	-
Playing time [s]	20	

Table II
EXPERIMENTAL PARAMETERS

Fixed additional delay [ms]	0, 75, 150
UDP load traffic [Mbps]	54, 80
Playout buffering time [ms]	100

LS generates UDP packets of 1480 bytes each with exponentially distributed interval and send them to LR for arising network congestion.

B. Parameters

Table I shows the specifications of the audio and video. We refer to the transmission unit at the application-level as an *MU (Media Unit)*. A video MU is a video frame and an audio MU consists of 320 audio samples. Each MU is transmitted as a UDP datagram. We employ frame skipping as the output method of video. In addition, we use the picture pattern, which consists of I-frames only, for simplicity to discuss fundamental characteristics of QoE in MVV-A. If all the packets of an MU are not correctly received in time for output, the MU is not output.

Table II shows experimental parameters. We employ a simple scheme of playout buffering control at the client to absorb network delay jitter. In the MVV-A system, the playout buffering control brings trade-off between output quality and responsiveness of viewpoint change. In the experiment, we set the buffering time to 100 ms on the basis of the results in the previous MVV-A study [14], which discusses unicast transmission of MVV and selectable audio.

We assume two kinds of the average amount of UDP load traffic: 54 Mbps and 80 Mbps. They are based on [15]. In [15], Cho *et al.* investigate residential per-customer traffic in one ISP by comparing traffic in 2005 and 2008, before and after the advent of YouTube and other similar services. Reference [15] reveals that the amount of daytime traffic is about 70 % of that of nighttime traffic. We have realized a situation in which congestion sometimes occurs between the two routers in Figure 1 on the nighttime traffic condition; considering this situation, we set the average amount of the UDP traffic to 80 Mbps. The amount of daytime traffic is selected to be 54 Mbps, which is about 70 % of 80 Mbps.

The netem software adds a constant delay, which can emulate a large scale network. We set three values of the delay: 0 ms, 75 ms, and 150 ms. We assume the value of 0 ms as communications delay inside a city, the values of 75 ms and 150 ms as the latency of international communications from Japan to U.S.A. and U.K., respectively. These values have been selected from [16], where the one-way delay from Japan to U.S.A. has a wide distribution from 60 ms to 150 ms, and the first peak of the distribution can be found at around 75 ms; as for the delay from Japan to U.K., the peak of the distribution is around 150 ms.

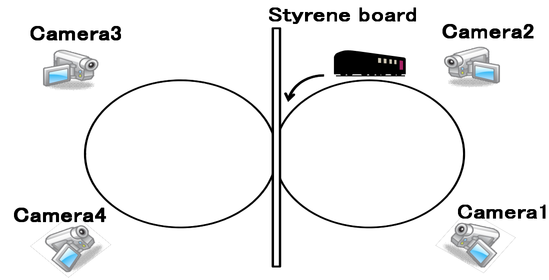

Figure 3. Camera arrangement of the train

Figure 4. Camera arrangement of the dog

C. Scenarios

In this experiment, we performed a subjective assessment for two scenarios. One is the scenario that the plural users watch almost the same viewpoint. In the other scenario, they watch different viewpoints. The difference of the user's behavior affects the number of transmitted viewpoints. Therefore, we deal with two scenarios which assume actual services.

As the scenario that plural users watch almost the same viewpoint, we employ the video of a running toy train. Figure 3 shows the cameras' arrangement in this scenario. On the content of train, the viewpoint in which the user can watch the train changes dynamically. It takes approximately four seconds from when the train enters a viewpoint to when it leaves the viewpoint. Moreover, as shown in Figure 3, a partition of styrene foam is put on the middle of the field so the user cannot see the other side of the field. When the user watch the video on Cameras 1 or 2, the user cannot see the left side in Figure 3 and vice versa. We assume the

Table III
PAIRS OF POLAR TERMS

ID	polar terms
v1	The video is smooth - The video is rough
v2	The video is steady - The video is unstable
v3	The video is easy to grasp - The video is hard to grasp
a1	The audio is natural - The audio is artificial
s1	The audio and video are in synchronization - The audio and video are out of synchronization
r1	The viewpoint change response is rapid - The viewpoint change response is slow
r2	The viewpoint change response is steady - The viewpoint change response is unstable
p1	Excellent - Bad

content as a large field like a sport (e.g., soccer, baseball) broadcasting.

In this scenario, VRTT transmits the viewpoint change request from 3.5 to 4.0 seconds interval based on uniform distribution in order to change the viewpoint which focuses on the train. Hereby, we actualize the scenario that the users excluding the subject change their viewpoints to watch the train. Viewpoints which are not watched can exist even if there are many users.

As the scenario that plural users watch different viewpoints, we employ the video of a toy dog walking forward and backward. Figure 4 shows the cameras' arrangement in this scenario. The toy dog can be seen from all the viewpoints. We assume the content as a situation that users can see a target from various angle like a figure skating broadcasting. In this scenario, VRTT transmits the viewpoint change request for arbitrary viewpoint from 3.0 to 5.0 seconds interval based on uniform distribution. That is, each user tends to watch a different viewpoint. Hereby, the more users exist, the more viewpoints are watched.

In this experiment, the initial viewpoint is Camera 1 in the figures.

D. Application-level QoS parameters

The *MU loss ratio* and the *average viewpoint change delay* are employed as the application-level QoS parameters for output quality of video. The MU loss ratio is the ratio of the number of MUs not output to the total number of MUs transmitted. The average viewpoint change delay is the average time in seconds from the moment the client sends a request for viewpoint change until the instant a new viewpoint is output at the client.

E. QoE assessment

QoE is assessed multidimensionally with the SD method. The SD method can assess an object for evaluation, which is called a *stimulus*, from many points of view with many pairs of polar terms. A pair of polar terms consists of one adjective and its opposite one, e.g., warm and cool. With this method, we can assess QoE in detail.

Table III shows the polar terms used in this paper. The polar terms are classified into five categories. In Table III, "v" means video, "a" audio, "s" synchronization, "r" response, and "p" psychology. Note that this experiment was performed in Japanese. This paper has translated the used Japanese terms into English. Therefore, the meaning of adjectives written in English here may slightly differ from those of Japanese one. For each selected pair of polar terms, a subjective score of an object for evaluation is measured by the rating scale method with five grades. The best grade (score 5) represents the positive adjective (left or upper side one in each pair in Table III). The worst grade (score 1) means the negative adjective (right or lower side one). The middle grade (score 3) is neutral. The scores 4 and 2 show slightly positive and slightly negative, respectively.

In the experiment of the requested viewpoint transmission method, we set the number of users including the subject to 2, 4, and 8 on the assumption that the number of users is smaller than, equal to, and larger than the number of viewpoints, respectively. In this paper, we employ two kinds of scenarios, the all viewpoints transmission method and the requested viewpoint transmission method with the three kinds of the number of users. In total, we consider 50 stimuli obtained by these combinations and additional dummies.

The subjects assess quality expressed by the pairs of polar terms with an assessment GUI, which is displayed on the screen whenever an experiment is finished. For our experiment, we employed 20 male students in their twenties as the subjects. The total assessment time of a subject is about 60 minutes.

The rating scale method is also used to measure *MOS (Mean Opinion Score)*, which is widely utilized for assessment of a single medium. In the rating scale method, assessors classify each stimulus into one of a certain number of categories. Each category has a predefined number, i.e., a score. However, the numbers assigned to the categories only have a greater-than-less-than relation between them; that is, the assigned number is nothing but an ordinal scale. When we assess the subjectivity quantitatively, it is desirable to use at least an *interval scale*. We can perform most of statistical procedures with the interval scale.

In order to obtain an interval scale from the result of the rating scale method, we first measure the frequency of each category with which the stimulus is placed in the category. With the *law of categorical judgment* [17], we can translate the frequency obtained by the rating scale method into an interval scale. Since the law of categorical judgment is a suite of assumptions, we must test goodness of fit between the obtained interval scale and the measurement result. This paper uses Mosteller's method [18] to test the goodness of fit. Once the goodness of fit has been confirmed, we refer to the interval scale as the *psychological scale*; it is a QoE metric.

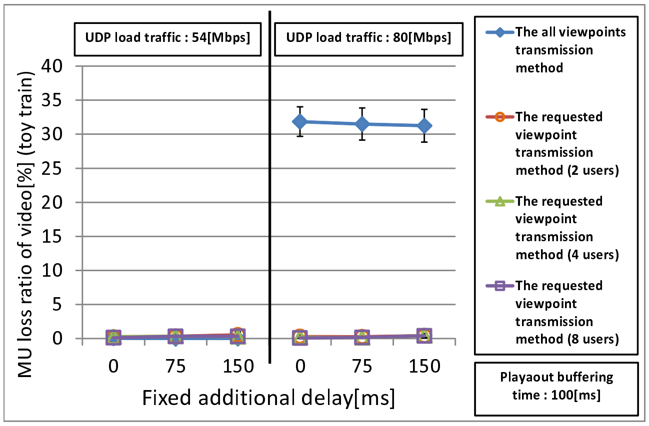

Figure 5. MU loss ratio (toy train)

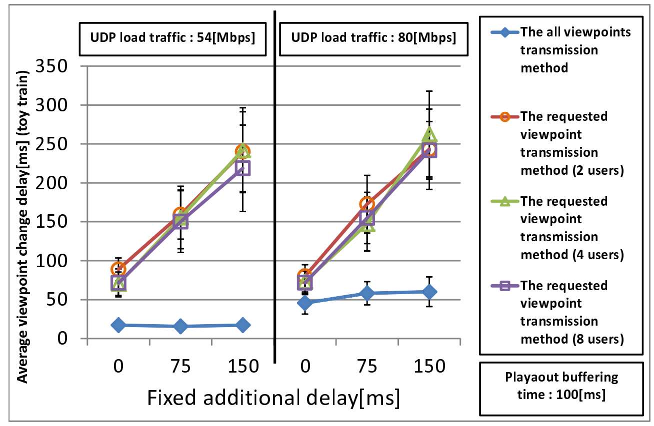

Figure 7. Viewpoint change delay (toy train)

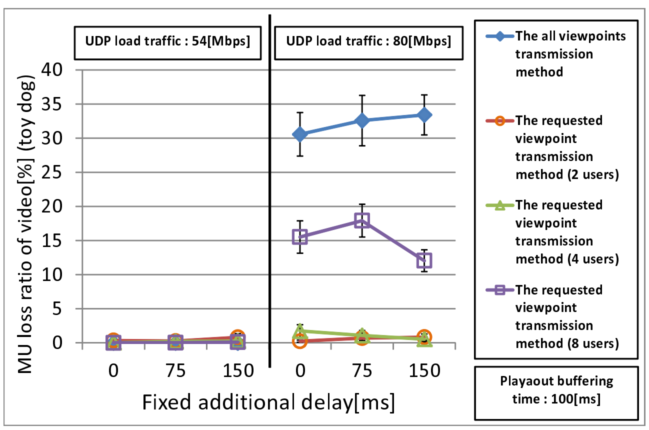

Figure 6. MU loss ratio (toy dog)

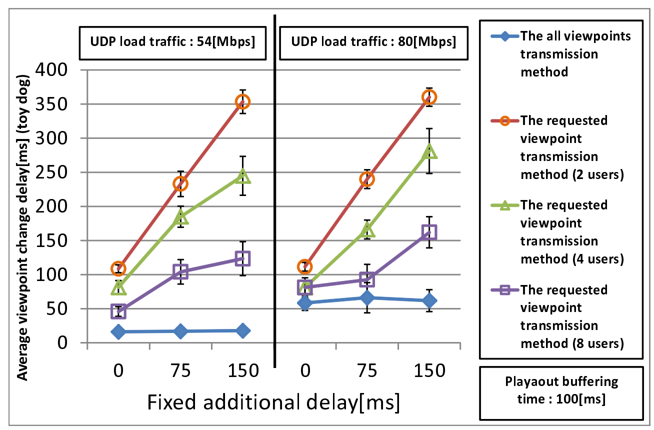

Figure 8. Viewpoint change delay (toy dog)

IV. Experimental Results

A. Application-level QoS parameters

Figures 5 and 6 show the MU loss ratio of the toy train and the toy dog, respectively, as a function of the fixed additional delay for each UDP load traffic condition. Each bar shows 95 % confidence interval.

In Figures 5 and 6, we do not find the difference arising from the transmission methods on the UDP load traffic 54 Mbps. On the other hand, the MU loss ratio of the all viewpoints transmission method is the largest on the UDP load traffic 80 Mbps. The reason is that the all viewpoints transmission method transmits all the viewpoints; that is, the amount of transmitted data becomes large.

The MU loss ratio of the requested viewpoint transmission method is small on the toy train. This is because almost all the users request the same viewpoint, and then the number of transmitted viewpoints get fewer.

On the requested viewpoint transmission method of the toy dog, the MU loss ratio becomes high as the number of users increases. This is because the number of requested viewpoints increases. Hereby, many viewpoints are transmitted, and the amount of traffic increases.

Figures 7 and 8 show the average viewpoint change delay of the toy train and the toy dog, respectively. In these figures, the average viewpoint change delay on the all viewpoints transmission method is low irrespective of the amount of fixed additional delay. This is because the client do not need to transmit the viewpoint change request in the all viewpoints transmission method. Therefore, the client only has to switch the selected viewpoint stream which is already received behind the played viewpoint stream.

In the requested viewpoint transmission method, as the fixed additional delay increases, the average viewpoint change delay grows. This is because the client needs to transmit the viewpoint change request to the server.

In Figure 7, we hardly find the difference of the number of users in the requested viewpoint transmission method. On the other hand, in Figure 8, we notice that as the number of users increases, the average viewpoint change delay reduces. The reason is that the number of transmitted viewpoints depends on the user's behavior. The number of transmitted

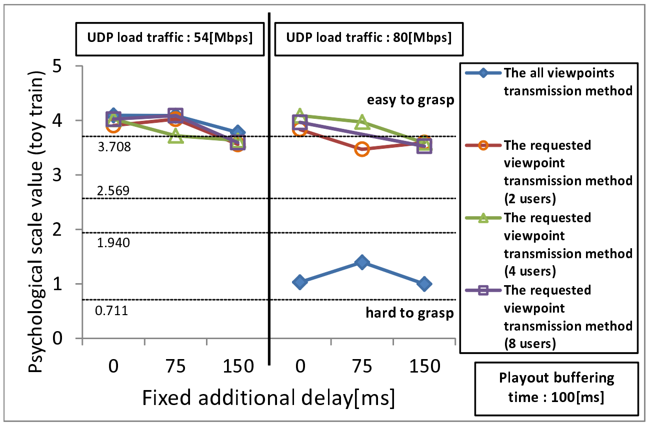

Figure 9. Difficulty of grasping video (toy train)

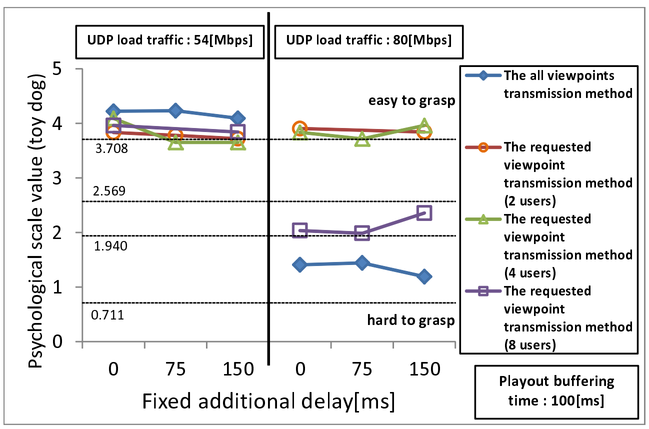

Figure 10. Difficulty of grasping video (toy dog)

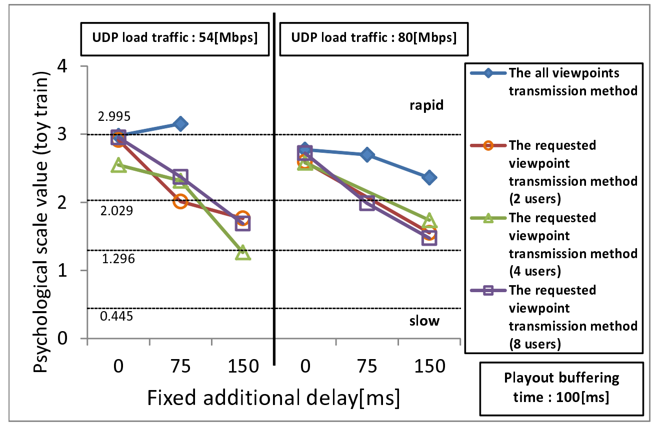

Figure 11. Viewpoint change response (toy train)

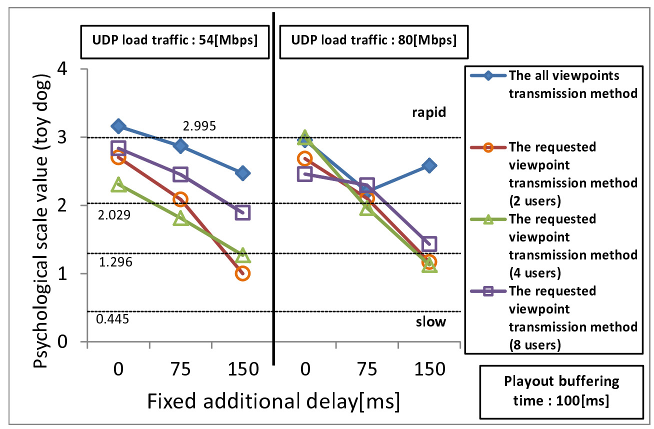

Figure 12. Viewpoint change response (toy dog)

viewpoints increases, then the user tends to switch the viewpoint to already received one. Accordingly, the average viewpoint change delay becomes low. Therefore, in the case that the watched viewpoints are diverse, i.e., the toy dog, as the number of users increases, the average viewpoint change delay becomes low.

On the UDP load traffic 80 Mbps, the average viewpoint change delay becomes higher than that of 54 Mbps, and the confidence interval spreads. The reason is as follows. When the first MU of a new viewpoint can not output owing to the MU loss, the viewpoint change delay increases. For detail, see [6].

B. Psychological scale

From among the pairs of polar terms shown in Table III, we focus on v3, r1, and p1 because they are related to the difference of the two transmission methods. We calculated the interval scale for each criterion. We then carried out the Mosteller's test. As a result, we have found that the hypothesis that the observed value equals the calculated one can be rejected with a significance level of 0.05. Therefore, we removed the stimuli which have large errors until the hypothesis cannot be rejected. In this paper, we use obtained values by these processes as the psychological scale.

Since we can select an arbitrary origin in an interval scale, for each criterion, we set the minimum value of the psychological scale to unity.

Figures 9 and 10 show the psychological scale for difficulty of grasping video (v3). Figures 11 and 12 present the psychological scale for viewpoint change response (r1), and Figures 13 and 14 depict that for overall satisfaction (p1).

Each of Figures 9 through 14 shows the results for the two UDP load traffic values: 54 Mbps and 80 Mbps. The ordinate is the psychological scale. The abscissa is the three fixed additional delay in netem: 0 ms, 75 ms, and 150 ms. The dotted lines in the figures indicate the lower limits of Category 2 through Category 5. Note that the results removed by the Mosteller's test are not plotted in the figures.

1) "Video is easy to grasp - hard to grasp": In Figures 9 and 10, we can not find the difference of the transmission methods on the UDP load traffic 54 Mbps. On the UDP load traffic 80 Mbps with the toy train, the difference of the

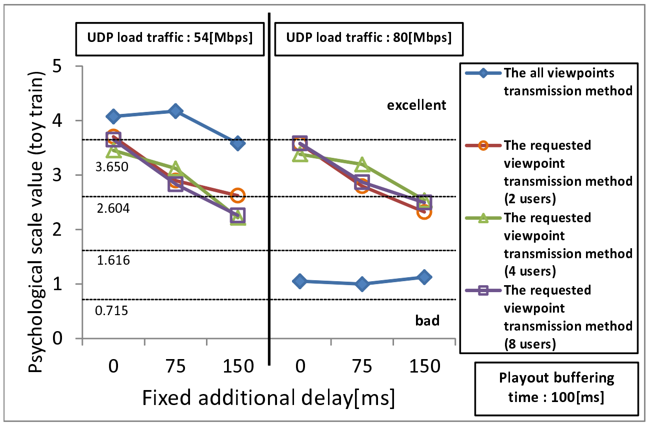

Figure 13. Overall satisfaction (toy train)

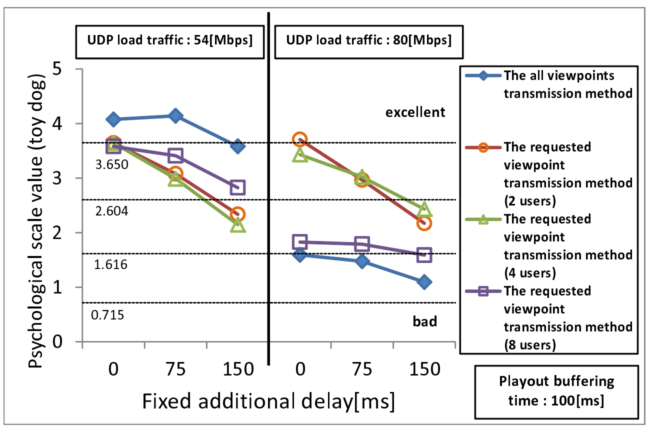

Figure 14. Overall satisfaction (toy dog)

number of users is trivial, because the users excluding the subject change the viewpoint to almost the same one.

Meanwhile, the psychological scale value on the all viewpoints transmission method is small on the UDP load traffic 80 Mbps, because all the viewpoints are transmitted.

On the toy dog in the requested viewpoint transmission method with 8 users on the UDP load traffic 80 Mbps, the psychological scale value is also as small as that on the all viewpoints transmission method. As the number of users increases in the requested viewpoint transmission method, the number of viewpoints which are watched by the users tends to be large. Accordingly, the amount of transmitted data becomes large, and the network tends to be congested.

2) "Viewpoint change response is rapid - slow": In Figures 11 and 12, we notice that the psychological scale value is high with the all viewpoints transmission method on the UDP load traffic 54 Mbps. As shown in Figures 7 and 8, the viewpoint change delay is small on the all viewpoints transmission method, because all the viewpoints are transmitted without viewpoint change requests.

On the UDP load traffic 80 Mbps, the psychological scale value of the all viewpoints transmission method and the requested viewpoint transmission method of 8 users are smaller than the UDP load traffic 54 Mbps because of the MU loss.

In Figure 12, the psychological scale value in the requested viewpoint transmission method of 8 users is higher than that of 2 and 4 users on the UDP load traffic 54 Mbps. All the viewpoints tend to be transmitted in the requested viewpoint transmission method of 8 users on the toy dog, and then MR tends to receive all the viewpoints data.

3) "Excellent - Bad": In Figure 13 and Figure 14, we notice that as the fixed additional delay increases, the psychological scale value of overall satisfaction decreases in the requested viewpoint transmission method. Hereby, we notice that deterioration of response affects the user's satisfaction. Therefore, the all viewpoints transmission method, which is not affected the fixed additional delay, shows the highest psychological scale values on the UDP load traffic 54 Mbps.

On the other hand, the all viewpoints transmission method shows the smallest psychological scale values on the UDP load traffic 80 Mbps. This is because the all viewpoints transmission method is inferior to the requested viewpoint transmission method on the video output quality.

C. Correlation coefficient

Table IV shows the correlation coefficient between each pair of polar terms and the overall satisfaction in descending order. We find that the satisfaction has high correlation with the pairs of polar terms for the video and audio. On the other hand, the satisfaction has small correlation with the pairs of polar terms for the viewpoint change response. From the above discussions, for the subjects in this scenario, the video and audio output quality is more important than the viewpoint change response. We also notice that the viewpoint change response affects the subjects' satisfaction a little.

In this experiment, there are some limitations. For example, we deal with two contents; they simulate real contents but are not real ones. Moreover, the network configuration in this experiment is different from real network environment.

Table IV
CORRELATION COEFFICIENT WITH OVERALL SATISFACTION

ID	polar terms	co-efficient
s1	The audio and video are in synchronization - The audio and video are out of synchronization	0.868
v3	The video is easy to grasp - The video is hard to grasp	0.857
v1	The video is smooth - The video is rough	0.855
v2	The video is steady - The video is unstable	0.847
a1	The audio is natural - The audio is artificial	0.844
r2	The viewpoint change response is steady - The viewpoint change response is unstable	0.582
r1	The viewpoint change response is fast - The viewpoint change response is slow	0.407

Therefore, we need to perform experiments with more diverse situations.

V. Conclusions

In this paper, we compared QoE of two transmission methods with multicast. One is the all viewpoints transmission method. The other is the requested viewpoint transmission method.

In this experiment, we found that the quality of the video is more important than the viewpoint change response for the users. From this, in multicast transmission of MVV-A, the quality of video and audio is important for users' satisfaction. Moreover, deterioration of the viewpoint change response affects users' satisfaction. Therefore, we first needs to keep the quality of the video-audio and next improve the viewpoint change response.

As future work, we need to investigate the trade-off between the quality of the video and response on different networks with various contents, video encoders, and subjects. We also need to consider the new transmission method for QoE enhancement.

VI. Acknowledgment

We thank Emeritus Prof. Shuji Tasaka in Nagoya Institute of Technology for his support on the experimental system. We also thank Takuya Ishida for his support on the experiment.

References

[1] ITU-T Rec.P.10/G.100, Amendment 2, "New definitions for inclusion in Recommendation ITU-T P.10/G.100," July 2008.

[2] I. Ahmad, "Multi-view video: Get ready for next-generation television," *IEEE Distributed Systems Online 2007*, vol. 8, no. 3, pp. 6, Mar. 2007.

[3] M. Tanimoto and T. Fujii, "FTV — Free Viewpoint Television," ISO/IEC JTC1/SC29/WG11 M8595, July 2002.

[4] W. Wu, A. Arefin, G. Kurillo, P. Agarwal, K. Nahrstedt and R. Bajcsy, "Color-plus-depth level-of-detail in 3D tele-immersive video: a psychophysical approach," ACM Multimedia 2011, pp. 13-22, Nov 2011.

[5] 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," *3DTV Conference 2007*, May. 2007.

[6] 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.

[7] Ustream, *http://www.ustream.tv/*

[8] T. Fujihashi, Z. Pan and T. Watanabe, "Traffic reduction on multi-view video live streaming for multiple users," *IEICE Trans. Commun.*, vol.E96-B, no.7, pp. 2034-2045, July 2013.

[9] L. Zuo, J. Lou, H. Cai and J. Li, "Multicast of real-time multi-view video," *Proc. IEEE ICME 2006*, pp. 1225-1228, July 2006.

[10] C. E. Osgood, "The nature and measurement of meaning," *Psychological Bullentin*, vol. 49, no. 3, pp. 197-237, May 1952.

[11] IPv4_Multicast_Address_Space_Registry, *http://www.iana.org/assignments/multicast-addresses/multicast-addresses.xhtml*

[12] The Linux Foundation, "netem," *http://www.linuxfoundation.org/collaborate/workgroups/networking/netem*

[13] P. Merkle, A. Smolic, K. Muller and T. Wiegand, "Efficient prediction structures for multiview video coding," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 17, no. 11, Nov. 2007.

[14] T. Ishida and T. Nunome, "Multidimensional QoE assessment of multi-view video and selectable audio (MVV-SA) IP transmission," *Proc, IEEE ISM 2013*, pp. 534-535 Dec. 2013.

[15] K. Cho, K. Fukuda, H. Esaki and A. Kato, "Observing slow crustal movement in residential user traffic," *ACM CoNEXT2008*, Dec. 2008.

[16] T. Tejima, A. Tan, S. Watanabe and K. Yoshida, "Large-scale measurements of network quality using online game," *IEICE Trans. on. Commun.*, J92-B(10), pp. 1566-1578, Oct. 2009, in Japanese.

[17] J. P. Guilford, *Psychometric methods*, McGraw-Hill, N. Y., 1954.

[18] F. Mosteller, "Remarks on the method of paired comparisons: III. a test of significance for paired comparisons when equal standard deviations and equal correlations are assumed," *Psychometrika*, vol. 16, no. 2, June 1951.