CUSTOMIZATION OF INTERACTIVE SERVICES FOR QOE ENHANCEMENT IN AUDIO–VIDEO TRANSMISSION OVER BANDWIDTH GUARANTEED IP NETWORKS

Shuji Tasaka and Norihito Misaki
Department of Computer Science and Engineering,
Graduate School of Engineering, Nagoya Institute of Technology, Nagoya 466–8555, Japan
Email: tasaka@nitech.ac.jp

ABSTRACT
This paper proposes a method of customizing interactive audiovisual communication services over bandwidth guaranteed IP networks typified by NGN in order to maximize QoE (Quality of Experience) on an individual users’ basis. The method is based upon an extension of a QoE guarantee architecture GPSQ (Guarantee of Psychologically Scaled Quality), which the authors proposed previously. GPSQ extended for interactive services is referred to as interactive GPSQ, which focuses on a salient feature associated with the interactive services, i.e., the existence of the optimum playout buffering time for absorbing delay jitter. The customization method is “semi–tailored” in the sense that it provides the users with a means of adjusting the buffering time so that it can maximize his/her own QoE. This paper examines the feasibility of the basic idea for interactive GPSQ and the customization method through a minimum number of conversational opinion tests, i.e., two tasks with different visual effects. For that purpose, we have built a simple experimental system of interactive GPSQ and made experiment on QoE assessment with and without the customization method. The experiment showed that the interactive GPSQ is feasible and that the customization improves QoE.

Keywords— QoE, interactive audiovisual communications, NGN, IP networks, service customization

1. INTRODUCTION
Customization is a natural demand in various areas of our society. As services and products become popular, the users’ demand for customization increases. We have seen this in the history of service industries such as dining, transportation and hotels as well as industrial products like clothes, cars, electric appliances and computer software.

We have witnessed a rapid and wide penetration of network services typified by the Internet in the past few decades. The network services are now at the stage where customization should be seriously considered and some mechanism for it is implemented into the networks.

The NGN (Next Generation Network) [1], which is an enhanced IP-based network with QoS (Quality of Service) guarantee, can be a platform for customization. Within the NGN, there is an increased emphasis on service customization by the service providers [1]. NGN is composed of the transport stratum and the service stratum. In the transport stratum, the NACFs (Network Attachment Control Functions) of the transport control functions have a database of transport user profile, which stores users’ information and other control data [2]. In the service stratum, the SCF (Service Control Functions) accommo-date databases of service user profiles; the application support functions and service support functions also work in conjunction with SCF to provide end–users and applications with the NGN services they request [2].

The customization of network services is closely related to QoE (Quality of Experience). This is because QoE represents the overall acceptability of an application or service, as perceived subjectively by the end–user [3]. QoE can be identified as user–level QoS from a layered network architectural point of view.

The network service should be customized ideally in the following three steps:
1. QoS guarantee at the network–level
2. QoE guarantee on the QoS guaranteed network
3. QoE–based customization

Step 1 can be reached through the QoS–guarantee mechanism of NGN, while Step 2 is not necessarily feasible only by means of QoS specifications on NGN since QoE is beyond the scope of the specifications.

The QoS guaranteed by NGN is described as network QoS by ITU–T Recommendation Y. 1540 [4] in terms of QoS parameters including IPTD (IP packet transfer delay), IPDV (IP packet delay variation), IPLR (IP packet loss ratio), and IPER (IP packet error ratio). Furthermore, the objective values of the QoS parameters are specified by Rec. Y. 1541 [5], which gives some guidance for it by classifying typical applications into eight classes. It should be noted here that the network QoS in Rec. Y. 1541 corresponds to the network–level QoS in the IP network; QoE is located above it in its hierarchy. Consequently, there is no guarantee that the objective values of network QoS thus specified realize QoE the users desire in efficient ways. Some additional mechanism is required for QoE guarantee on the basis of the network QoS guarantee.

In [6], the authors have proposed a method of QoE guarantee over bandwidth guaranteed IP networks, which is referred to as GPSQ (Guarantee of Psychologically Scaled Quality). This method is a simple trial of Step 2 mentioned earlier; it can be a launch pad toward Step 3. Although the concept of GPSQ is applicable to both interactive and noninteractive services, the study on GPSQ in [6] is restricted to noninteractive streaming services, i.e., one–way audio–video streaming services, for simplicity of discussion.

This paper proposes a QoE–based customization method for interactive audiovisual communications over bandwidth guaranteed IP networks, which are supposed to be NGN, utilizing GPSQ as a foundation of QoE guarantee.

This paper has two original contributions to audiovisual IP...
communications in NGN. One is to propose and implement a QoE guarantee mechanism for interactive services. This is an extension of GPSQ in [6]; we refer to the extended version of GPSQ as interactive GPSQ. The other is the introduction of a customization method for maximizing QoE on an individual users’ basis along with showing its effectiveness by experiment.

In the customization, we focus on a salient feature of QoE in interactive audiovisual IP communications, i.e., a QoE tradeoff relationship between fidelity and latency\(^2\).

The relationship arises from playout buffering of audio/video MUs \((\text{media units})^3\) at the receiver for absorbing delay jitter of MUs. As the buffering time increases, the fidelity of output audio/video improves because packets arriving later can be output, while the latency increases, which leads to the degradation of the responsiveness. On the other hand, when the buffering time becomes shorter, the responsiveness improves at the expense of degraded quality of output media caused by packet loss and jitter. Since fidelity and latency affect QoE of interactive services in the opposite way, there exists the optimum playout buffering time, which maximizes QoE, as we will see in Figs. 3 and 4 later. Note that this feature does not appear so clearly in noninteractive services.

The proposed method maximizes QoE by adjusting the playout buffering time according to individual users’ inclination for output media quality and responsiveness.

The remainder of the paper is organized as follows. Section 2 describes related work to this study. Section 3 extends GPSQ to accommodate interactive services and proposes a customization method. Section 4 deals with the experimental methodology. Section 5 presents a preliminary experiment for the extension of GPSQ. Section 6 proposes an implementation method of the customization and then demonstrates the effectiveness by experiment. Section 7 concludes the paper.

2. RELATED WORK

Since there are a variety of issues related to audiovisual IP communications, we focus here on (1) QoE/QoS assessment of interactive services, (2) QoE metrics, and (3) QoE guarantee in NGN. We can find no study on QoE–based customization of audiovisual IP communications in the literature.

2.1. QoE/QoS assessment of interactive services

To the best of the authors’ knowledge, only Reference [7] investigates the QoE of interactive audiovisual communication services over bandwidth guaranteed IP networks by explicitly showing the QoE tradeoff relationships.

Reference [7] explores the effects of the guaranteed bandwidth between the terminals, video encoding bit rate, the playout buffering time and task on QoE. It has shown that there exists the optimum pair of the video encoding bit rate and playout buffering time, which maximizes QoE, for a given amount of the video guaranteed bandwidth and the audio guaranteed bandwidth; the optimum pair depends on the kind of task. In the current paper, we utilize this result for QoE guarantee and QoE–based customization.

2.2. QoE metrics

As the QoE metric, this paper adopts the psychological scale \(^8\) instead of MOS (Mean Opinion Score), which is used in ITU–T/R Recommendations and many technical papers. In the context of psychometric theory \(^9\), the psychological scale is an interval scale, whereas MOS is an ordinal scale; this means that the psychological scale can represent human subjectivity more accurately than MOS. The interval scale can be calculated by means of the method of successive categories, which is composed of two steps: the rating–scale method and the law of categorical judgment. For further details, see [8].

Note that the term “Psychologically Scaled Quality” in GPSQ originates from this metric.

2.3. QoE guarantee in NGN

Regarding QoE guarantee, as far as the authors know, only GPSQ in [6] treats the subject directly in a quantitative way from a viewpoint of network architecture, especially NGN.

GPSQ provides a framework for QoE guarantee by session control with SIP (Session Initiation Protocol) \([10]\). The basic principle of GPSQ is setting of system parameters including reservation of necessary bandwidth by session control to achieve specified QoE.

![Fig. 1. Configuration of GPSQ.](image)

Figure 1 displays essential ingredients of GPSQ: a SIP server, a QoS manager and a bandwidth–controllable network, which consists of bandwidth–controllable routers, in addition to media terminals, which send/receive audio–video streams.

First, the user specifies QoE he/she desires, which we call the target QoE. The target QoE is usually specified by selecting one from among a finite set of quality–levels; e.g., excellent, good and fair. It is often the case that the price of the service is proportional to the quality–level selected.

Then, using an INVITE request message of SIP, the media terminal notifies the SIP server of the user’s request and the terminal’s capability concerning media encoding/decoding and display; the request contains information on the target QoE along with other kinds of necessary information.

Resource allocation and management are carried out by the QoS manager in cooperation with the SIP server. The SIP server extracts control information necessary for the QoS manager from the INVITE request and delivers the information to the QoS manager, which allocates necessary bandwidth to achieve the target QoE. The QoS manager keeps a database of information which expresses each target QoE as a function of system parameters including guaranteed bandwidth for audio and video; therefore, it can calculate the guaranteed bandwidth necessary for the target QoE. Then, the QoS manager sets up routers to guarantee the calculated bandwidth.

The SIP server plays a role of SCF in NGN, while the QoS manager has a function of RACF (Resource and Admission Control Functions) \([11]\) in the transport control functions along with databases of transport user profiles and service user profiles.

The RACF part of the QoS manager instructs bandwidth–controllable routers on the guarantee of required bandwidth, which is commonly realized with packet schedulers.

GPSQ in [6] supposes only one–way audio–video streaming services from a media sender to a media recipient for simplicity of discussion. The media sender stores a collection of audio–video streams along with their attributes and statistics; the attribute includes the name, content type and encoding schemes, while the statistics here imply the ones of audio and video bit rates such as the average, variance and maximum.
The construction of the database in the QoS manager is a key to the success of GPSQ; the database in [6] is a set of representative regression lines each of which expresses a target QoE for a content type, which means a group of contents with similar features, as a function of system parameters such as the guaranteed bandwidth for audio and video, and the statistics of audio and video. The regression line is an equation which estimates QoE from the system parameters. Since QoE depends on a variety of factors, there can be an infinite number of combinations of the factors, which leads to an infinite number of regression lines.

The approach of representative regression lines is a method of avoiding the explosion of the number of regression lines. See [6] for details.

3. INTERACTIVE GPSQ

In this section, we extend GPSQ in [6] to interactive GPSQ, since interactive services have different features of QoE from noninteractive ones; this implies the necessity for the construction of the database in the QoS manager so that it can reflect the features of interactive services.

For this purpose, this paper utilizes the consideration on this issue in [7]. There are many possible factors affecting QoE. For simplicity of discussion, however, reference [7] focuses on the following ones: (1) the guaranteed bandwidth between the terminals, (2) the video encoding bit rate, (3) the playback buffering time, and (4) the task. The first parameter is a factor dominating the network QoS, and the second one corresponds to the sender–part at the application–level of the terminal, while the third one is the receiver–part. The task represents the user’s behavior in interactive services [13].

Note that unlike noninteractive (streaming) GPSQ, interactive GPSQ must take account of the third factor (i.e., the QoE tradeoff relationship) explicitly in constructing the database of the QoS manager. See [6] and [7] for details.

3.1. Database in the QoS manager

The database for interactive services is a set of representative mapping functions, each of which expresses a target QoE for a task type, which means a group of tasks with similar features, as a function of system parameters listed up earlier, namely, the factors affecting QoE. The representative mapping functions and the task types here are the interactive counterparts of representative regression lines and content types for streaming services. The implications of the former pair are parallel to those of the latter in the previous section; the aim is to avoid the explosion of the database size.

As in the streaming case, we are faced with two important technical problems: (a) how should tasks be classified into types? and (b) how should the representative mapping function for a type be selected?

Regarding Problem (a), ITU-T Rec. P.920 [13] can be a guideline. It states that the tasks should be structured so as to represent the applications as regards of the rate of information exchange and the degree of audio and video signal utilization.

On the basis of this policy, P.920 gives several examples of the task.

In Problem (b), a representative mapping function should be selected so that it can be as accurate as possible for all possible tasks in the type; achievable accuracy depends on what tasks are included in the type. Therefore, Problem (b) is closely related to Problem (a).

Since this paper is a first step trial of interactive GPSQ research, studies on Problems (a) and (b) are left as future work. Instead, this paper intends to examine the feasibility of the basic idea for interactive GPSQ through a minimum number of conversational opinion tests; we develop two simple tasks whose visual effect on QoE is different from each other. We will describe the two tasks in the next section.

The number of the target QoE can be either infinite, e.g., a case in which QoE metric can take continuous values, or finite. However, it is practical to restrict ourselves to the finite case, which we suppose in this paper.

Below we give an example of the target QoE with six levels: excellent, which is denoted as E, the boundary between excellent and good (EG), good (G), the boundary between good and fair (GF), fair (F), and the boundary between fair and poor (FP). “Poor” and “bad” are excluded since the user usually do not select them even if the price is low. In this example, the database has six representative mapping functions for each task type.

Thus, the information carried by the INVITE request for the QoS manager includes the user’s ID, the ID of the partner, the supposed type of task and the selected target QoE. These pieces of information enable the QoS manager to decide how much bandwidth should be allocated to the user’s session, referring to the database.

Note that GPSQ has adopted a new paradigm of “user–centricity” for quality control; it sets values of system parameters including the guaranteed bandwidth and encoding bit rates so that they can realize the target QoE. On the other hand, traditional (“network–centric”) quality control methods try to satisfy specified QoS parameter values as typified by ITU-T Rec. Y.1541, or they attempt to maximize the utilization of a given amount of bandwidth; in both cases, however, it is uncertain how high QoE is achieved.

We should recall here again that the price of the service is usually set to be proportional to the quality–level selected. When the user selects higher quality–level, he/she is ready to pay more money; otherwise, he/she chooses a moderate level of quality as in many other services. The user does not always select the top–level quality because of financial reasons.

3.2. QoE–based customization

Before starting the discussion on this issue, it should be noted that each representative mapping function in the database is obtained through statistically processed data of many subjects such as the psychological scale. Therefore, the mapping function does not necessarily relate the target QoE to the system parameters in the optimum way for individual users; the system parameter values selected by the mapping function may or may not maximize his/her own QoE. Note that individual users’ inclination varies from person to person.

On the basis of this observation, we propose a QoE–based customization method as follows. When the session is established, the QoS manager selects initial values (default values) of the system parameters, referring to the database by using the information conveyed by the INVITE request. Then, the QoS manager notifies the routers of the bandwidth to be guaranteed, while it informs the terminals of the video encoding bit rate and playback buffering time. During the session, the terminal provides some means (e.g., GUI (Graphical User Interface)) of changing the values with the user so that he/she can adjust them to his/her own taste; for example, the user selects different values from the initial ones from among a finite number of candidate values through GUI. This is the method we adopt in this paper as seen later. Note that the same idea is applicable to streaming services.
We utilize SIP to customize the system parameter values; the changed parameter values are conveyed in the Event header of the SUBSCRIBE request, and then they are confirmed by the NOTIFY request when the partner permits them.

Another possible method of the customization is to prepare the representative mapping functions for each user. This is quite simple in concept, but the implementation is very hard because of the huge number of the functions. Note that in addition to the difference in taste among the users, a user’s taste itself is not always the same; it varies according to the environment and his/her physical and mental conditions. This is the reason why we have adopted the proposed method.

A brief comment on the selection of the customization method follows. Comparing customization methods to clothes making, the GPSQ interactive service without customization is regarded as “ready made” clothes, while the second method of preparing the user’s own mapping functions is “fully tailored”. In the same context, the proposed method is “semi–tailored”, which saves the cost while keeping the quality approximately the same or slightly lower than “fully tailored”.

4. EXPERIMENTAL METHODOLOGY

4.1. Experimental system

As shown in Fig. 2, the experimental system consists of seven PC’s (Terminal 1, Terminal 2, SIP server & QoS manager6, Load Sender 1, Load Sender 2, Load Receiver 1, and Load Receiver 2), two bandwidth controllable routers (Routers 2 and 3: Cisco System’s 7301), and two ordinary routers (Routers 1 and 4: RiverStone’s RS3000). Each Terminal transmits an audio stream and the corresponding video stream as two separate UDP streams to the other Terminal. A real–time H.264 video encoding board (DSP Research Inc.) equipped with a video camera has been installed into each Terminal along with a microphone and headphones. The nominal error ratio of the average encoding bit rate of the board is less than 10%.

Load Sender 1 and Load Sender 2 transmit UDP load traffic; they generate UDP datagrams of 1472 bytes each at exponentially distributed intervals. Load Receiver 1 and Load Receiver 2 are the corresponding receivers.

Fig. 2. Configuration of the experimental system.

The links between the routers and ones between a router and a PC are all full duplex Ethernet channels. The transmission rate of the link between Router 2 and Router 3 is 10 Mb/s, while the others are 100 Mb/s. Therefore, the link of 10 Mb/s becomes a bottleneck.

The bandwidth control is exerted between Router 2 and Router 3: LLQ (Low Latency Queuing) is adopted as the packet scheduling algorithm. In LLQ, we can set a PQ (Priority Queuing) class and CBWFQ classes [12]: Each class has a dedicated buffer. Packets in the PQ class are served with high priority until its buffer becomes empty; then, the server goes down to the CBWFQ classes. The PQ class is assigned to the audio streams. The video streams and the UDP load traffic are treated as two separate CBWFQ classes.

For simplicity of experiment, the guaranteed bandwidth for audio is kept constant, while we try four values for video.

4.2. Tasks

Since the purpose of this paper is to examine the feasibility of interactive GPSQ, we give simple conversational opinion tests, referring to ITU–T Rec. P.920. We have designed two simple tasks in order to examine the effects of audio and video in a systematic way: 1) audio is dominant, and 2) audio is enforced by visual impacts [7]. Note that if we employed tasks in real use in subjective assessment, it would be very difficult to identify dominant factors affecting QoE because of a number of the factors and complicated relations among them.

task 1: One subject selects a number randomly from 1 through 5 and reads the numbers from 1 to the selected number aloud. Immediately after the reading, the other subject reads the same numbers aloud. This interaction is repeated by alternating the initiator during a predetermined interval.

task 2: Each subject reads numbers aloud in the same way as in task 1, but clapping on PGQ before reading. Note that task 1 is audio–dominant since the subjects exhibit only low motion, whereas task 2 has a visual impact on QoE because of clapping.

5. PRELIMINARY EXPERIMENT

This section presents a summary of a preliminary experiment for the database construction in the QoS manager, which is reported in [7], and actually constructs the database.

5.1. Method of experiment

The preliminary experiment was conducted on the two tasks under the condition that the guaranteed bandwidth for video and that for audio are allocated in advance; i.e., we did not carry out the session setup/termination and therefore consider the data transfer phase only.

The guaranteed bandwidth for audio is kept constant at 90 kb/s, while that for video is set to either 1, 2, 3, or 4 Mb/s. The remaining bandwidth is allocated to the UDP load traffic, whose average transmission rate is set to the allocated bandwidth so that the link between Router 2 and Router 3 can be congested.

Table 1. Specification of video.

<table>
<thead>
<tr>
<th>encoding scheme</th>
<th>H.264</th>
</tr>
</thead>
<tbody>
<tr>
<td>image size [pixel]</td>
<td>704 × 480</td>
</tr>
<tr>
<td>picture pattern</td>
<td>1</td>
</tr>
<tr>
<td>encoding bit rate [Mb/s]</td>
<td></td>
</tr>
<tr>
<td>when $B_G = 1$: 0.7, 0.8, 0.9, 1.0</td>
<td></td>
</tr>
<tr>
<td>when $B_G = 2$: 1.7, 1.8, 1.9, 2.0</td>
<td></td>
</tr>
<tr>
<td>when $B_G = 3$: 2.7, 2.8, 2.9, 3.0</td>
<td></td>
</tr>
<tr>
<td>when $B_G = 4$: 3.7, 3.8, 3.9, 4.0</td>
<td></td>
</tr>
<tr>
<td>average MU rate [MU/s]</td>
<td>25</td>
</tr>
<tr>
<td>monitor</td>
<td>17 inch LCD</td>
</tr>
<tr>
<td>monitor resolution [pixel]</td>
<td>1280 × 1024</td>
</tr>
</tbody>
</table>

The audio is encoded at 64 kb/s by ITU–T G.711 μ–law; its MU rate is 25 MU/s.

The video is H.264 whose specification is given in Table 1: The video encoding bit rate is set to various values according to the guaranteed bandwidth ($B_G$). The video frame skipping is employed for decoding [14].
Each Terminal carries out playout buffering control on received MU’s in audio and video streams independently. We take nine values of the playout buffering time [ms]: 40, 80, 120, 160, 200, 300, 400, 500 and 1000. The same value is chosen at both Terminals in an experimental run.

In order to obtain the mapping function for each task, we carried out subjective experiment for all combinations of the video encoding bit rate and playout buffering time for each video guaranteed bandwidth.

After an experimental run, each subject assessed the stimuli, which are audio–video streams output at his/her own Terminal during the run, by the rating–scale method, where the Absolute Category Rating (ACR) with the following five–level quality scale is used: “excellent” assigned score 5, “good” 4, “fair” 3, “poor” 2 and “bad” 1. The subjects are 24 Japanese male and female students in their twenties.

5.2. Experimental results

Figure 3 plots the psychological scales for task 1 as a function of the playout buffering time for the video guaranteed bandwidth 1 Mb/s, 2 Mb/s, 3 Mb/s and 4 Mb/s. For each video guaranteed bandwidth, we have chosen the video encoding bit rate which maximizes QoE from among the four rates and display the corresponding curve only; the optimum rate is also shown in the box in the figure. Note that the optimum rate in this case is equal to the video guaranteed bandwidth. Figure 4 is a counterpart of task 2.

From Figs. 3 and 4, we find that there exists the optimum playout buffering time, which maximizes QoE, for each video guaranteed bandwidth. We also notice that the kind of task influences the sensitivity of the maximum psychological scale value to the playout buffering time; for the wider bandwidth (say, 3 Mb/s and 4 Mb/s), task 2 is more sensitive than task 1. This is due to the difference in visual effects between the two tasks: task 1 is audio–dominant in the sense that the video shows only low motion, whereas the video motion of task 2 is high because of the clapping, which gives a visual impact.

More detailed discussion on the experimental results can be found in [7].

5.3. Database construction

Utilizing the results in the previous subsection, we construct a database for the QoS manager of the experimental system. We denote the video guaranteed bandwidth by $B_G$, as before and the playout buffering time by $T_b$. Then, for each task, we first find the maximal achievable value of the psychological scale value for each value of $B_G$; we use it as the target QoE, which is denoted by $Q_T$. Then, for each $Q_T$, we identify a pair of the video encoding bit rate and the playout buffering time that achieves $Q_T$, which are denoted by $R_v$ and $T_O$ (i.e., the optimum of $T_b$), respectively. Thus, the mapping function for the task has a form such as $Q_T = f(B_G, R_v, T_O)$.

It is difficult to derive an explicit expression of the mapping function, we implement it as a lookup table. Now, let us find $Q_T$, $B_G$, $R_v$, and $T_O$ for task 1 in Fig. 3. We then notice that $Q_T = E$ is realized by $B_G = 4$ Mb/s, $E G$ by 3 Mb/s, $G$ by 2 Mb/s and $F P$ by 1 Mb/s, all when $R_v = B_G$. Also, from Fig. 3, we can find $T_O$ for each $Q_T$; $T_O$ is used as the default (initial) value of $T_b$ for the Terminals. Similarly, we can identify the values for task 2 in Fig. 4. The results thus obtained are summarized in Table 2.

In this paper, for simplicity of discussion, we suppose that only $T_b$ is changeable for the customization. For this reason, Table 2 displays the lower limit of the allowable $T_b$ and the upper one, which are denoted by $T_L$ and $T_U$, respectively. Each limit is a value of $T_b$ which achieves the psychological scale value approximately 10% lower than the maximum; note that unless such a limit is set, the user can change $T_b$ largely, and as a result, achieved QoE may diverge from his/her own optimum instead of converging to it.

The QoS manager notifies each Terminal of the values of $R_v$, $T_O$, $T_L$ and $T_U$ at the session setup.

### 6. CUSTOMIZATION

In this paper, we adopt a customization method where the value of the playout buffering time $T_b$ can be changed from the initial value $T_O$ to two higher or lower ones. Also, we simplify the implementation by restricting the terminal entitled to request the changes only to the session initiator.
We have implemented the method through a GUI with five radio buttons. Button 3 corresponds to the initial value, while Buttons 1 and 5 are allocated to the lower limit and the upper limit, respectively. Button 2 corresponds to the average of the initial value and the lower limit, and Button 4 to the average of the initial value and the upper limit. The user can select a value of $T_b$ by clicking the corresponding button.

### 6.1. Method of experiment

In order to examine the effectiveness of the customization method, we conducted two kinds of subjective experiment, i.e., experiment without and with customization. The former uses the playout buffering time of $T_{ij}$ throughout the session, while the subject in the latter tries the five values of $T_{ij}$ during a certain period of learning after the session establishment and finally chooses his/her favorite playout buffering time; the chosen value is used until the session termination.

Before starting the experiment, for each task, the subject went through a training in rating the quality by ACR under four combinations of the system parameter values $B_G$, $R_v$ and $T_b$, which realize various quality of output audio–video ranging approximately from “bad” to “excellent”.

In an experimental run, the target QoE was set to one selected from among $FP$, $G$, $EG$, and $E$ for each task. Thus, the number of stimuli is 16 because of the four target QoE’s, the two tasks, and the playout buffering with and without customization.

In addition to 22 people out of the 24 subjects in the preliminary experiment, other 22 people participated in this experiment; so, 44 subjects in total. They are Japanese male and female students in and around their twenties.

### 6.2. Experimental results

As in the preliminary experiment, we applied the law of categorical judgment to the 16 stimuli and then gave Mosteller’s test, which indicates that the hypothesis that the observed value equals the calculated one cannot be rejected at a significance level of 0.05.

Figures 5 and 6 show the psychological scales with and without customization as a function of the target QoE for tasks 1 and 2, respectively; the minimum value of the psychological scales for the 16 stimuli is set to unity (i.e., 1).

In these two figures, we first see that the customization method improves QoE for every target QoE as expected.

We next notice that actually achieved QoE is around the target QoE when using the customization, though the former tends to be slightly smaller than the latter, especially for task 1. This is due to the insufficient accuracy of the database in the QoS manager.

Comparing Figs. 5 and 6, we find that for the target QoE of $E$ and $EG$, actually achieved QoE of task 2 is higher than that of task 1, while for $G$ and $FP$, they are close to each other.

This implies that the visual effect enhances QoE when the video quality is high.

### 7. CONCLUSIONS

We introduced interactive GPSQ for audiovisual IP communication services and proposed a “semi–tailored” method of customizing the services so that it can maximize QoE on an individual users’ basis. We then learned that interactive GPSQ is feasible and that the customization enhances QoE.

Furthermore, we applied the method to the first–step trial of study on interactive GPSQ and the customization method; therefore, many problems to be solved are left as future work. It includes the classification of tasks into types, derivation of the representative mapping function for each task type, and study on the effects of the packet scheduling algorithms. We also need to make QoE assessment by subjects at various ages and in many language environments. Furthermore, it is interesting to apply the idea of customization to streaming services.

### 8. REFERENCES