QoE Enhancement of Multimedia Communications by User–Assistance

Shuji Tasaka, Nagoya Industrial Science Research Institute, Nagoya, Japan, Email: tasaka@nisri.jp
Toshiro Nunome, Nagoya Institute of Technology, Nagoya, Japan, Email: nunome@nitech.ac.jp

Abstract

This paper proposes a method of enhancing *QoE* (*Quality of Experience*) in audiovisual IP communications; it is based upon user's assistance through *GUI* (*Graphical User Interface*) for *QoS* (*Quality of Service*) control. The rationale behind the proposed method is the *accommodation of the user's inclination*, which is one of the most difficult factors affecting QoE in QoE/QoS control; the accommodation is by adjusting QoS control parameter values through GUI by the user. In order to explain the meaning of the proposal, we first discuss the relation between QoE and QoS; we then point out that traditional QoS control is not capable enough to achieve QoE reflecting individual users' inclinations. We present a simple example of the implementation of the proposed method into an interactive audiovisual communication system and demonstrate that the proposed method improves QoE in terms of a QoE measure (the *psychological scale*) by subjective experiment.

1 Introduction

The ultimate goal of network services is to provide satisfactory *QoE* (*Quality of Experience*), which is the quality experienced by the end–user. In the context of the IP networks, which are becoming increasingly important in practice as the current Internet and as the *Next Generation Network* (*NGN*), QoE can be identified as being on the top of the *QoS* (*Quality of Service*) hierarchy of the IP networks: *application–level*, *transport–level*, and *network–level*.
QoE is directly related to human subjectivity; consequently, conventional QoS–guarantee mechanisms do not necessarily realize QoE the end–users desire in efficient ways. This is typified by the QoS specifications in NGN. ITU–T Recommendation Y. 1540 describes the QoS guaranteed by NGN as *network QoS* in terms of QoS parameters including IP packet transfer delay, delay variation, loss ratio, and error ratio. Furthermore, Rec. Y. 1541 specifies the objective values of the QoS parameters. Since the relationship between the objective values and QoE has not been clarified sufficiently, it is unknown how high QoE the objective values realize.
As seen above, the traditional quality–guarantee of network services has been "*network–centric*". However, NGN has various functions for service customization. This will lead us to an advanced stage of network services where the paradigm is shifted from the "network–centric" to "*human–centric*".
Paying attention to application services in the IP networks, we notice that audio–video transmission supports many popular services over the current Internet and will remain important in NGN. A Cisco White Paper [1] tells us that Internet video traffic was 57 percent of all consumer Internet traffic in 2012 and will be 69 percent in 2017, not including the amount of video exchanged through P2P file sharing.
This paper discusses how QoE of audio–video IP transmission can be enhanced and/or assured from a human–centric point of view. Researches of this type have appeared recently in the literature; e.g., see [2], [3], [4], [5], and [6]. These researches adopt network architectural approaches to QoE enhancement and/or assurance in the context of NGN and/or the current Internet. This paper, on the other hand, is featured by a proposal of human–network interaction through *GUI* (*Graphical User Interface*) for QoE enhancement according to *individual users' inclinations*; this is a quite different approach to the QoE management problem from the previous studies.
In order to elaborate upon the meaning of the proposal, we first briefly review definitions of QoE in Sec. 2 and discuss the relation between QoE and QoS in Sec. 3; we then point out that it is difficult to achieve QoE desired by the user only by means of traditional QoS control since QoE itself depends on individual users' inclinations, which is a non-technical human factor. Note that QoS is based on technical performance while QoE depends on end–user behavior. As a method of coping with this inclination problem, Sec. 4 proposes an idea of *user–assisted* QoE enhancement. Section 5 presents a simple example of the implementation of the method into an interactive audiovisual communication system and show how much the proposed method improves QoE in terms of a QoE measure (the *psychological scale*) by subjective experiment.

2 Definitions of QoE

Since QoE is directly related to the end–users, it is affected by many factors between the information source and the recipient; e.g., media capturing process, media encoding schemes, networks, media decoding schemes and media output schemes. In addition to these technical factors, there are non–technical ones like service contents, pricing and customer support. Since the users become a part of the truly end–to–end total system, they should be taken into consideration; human factors such as age, gender, personal experiences, cultural background and socioeconomic status are also important factors as well as contents/tasks and tariff of the services. From a telecommunication point of view, the network is the most influential factor owing to its provision of nonzero delay and finite throughput.
Because of the multiple facets of QoE, we can find many definitions of QoE and closely related concepts in the literature; e.g., [7], [8], [9], [10].
From among them, we pick up the definitions of ITU-T Recommendation [7], *ETSI* (*European Telecommunications Standards Institute*) [8], and *QUALINET* (COST Action IC 1003) [9]. They are given below:

[ITU-T Recommendation's QoE]: The overall acceptability of an application or service, as perceived subjectively by the end-user.

Note (1) QoE includes the complete end-to-end system effects (client, terminal, network, service infrastructure, etc.).

Note (2) Overall acceptability may be influenced by user expectations and context.

[ETSI's QoE]: A measure of user performance based on both objective and subjective psychological measures of using an ICT service and product.

Note (1) It takes into account technical parameters (e.g., QoS) and usage context variables (e.g., communication task) and measures both the process and outcomes of communications (e.g., user effectiveness, efficiency, satisfaction and enjoyment).

Note (2) The appropriate psychological measures will be dependent on the communication context. Objective psychological measures do not rely on the opinions of the user. Subjective psychological measures are based on the opinion of the user (e.g., perceived quality of medium, and satisfaction with a service).

[QUALINET's QoE]: QoE is the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state.

3 QoS and its Relation to QoE

QoS affects QoE, and the relation between the two can be expressed in several ways. For example, QoS deals with performance aspects of physical systems, while QoE deals with the user's assessment of system performance [9].

We can find two typical definitions of QoS in the literature: the definitions in ITU–T Rec. (E.800 and E.802) and those in *layered network architecture* models, which are QoS in the *OSI* (*Open Systems Interconnection*) *reference model* and QoS in the *TCP/IP reference model*.

This paper adopts the layered network architecture model, where a network is composed of a certain number of *layers* or *levels*, each one built upon its predecessor. In particular, we suppose the *QoS hierarchy* of the TCP/IP networks, where we can identify the following five levels: *physical–level QoS*, *data link–level QoS*, *network–level QoS*, *transport–level QoS*, and *application–level QoS*. Although the physical layer and data link layer are not defined in the TCP/IP reference model, the two layers must exist in actual networks; therefore, we have included the corresponding two types of QoS. QoE is considered on the top of application–level QoS.

Because of the QoS hierarchy, QoS at a level is affected by that at the lower level in turn. This implies that QoS at a level is controllable by QoS at the lower level and that QoE can be affected and estimated from QoS at application and/or lower levels.

As we have seen so far, QoE has a variety of definitions given by many authors and many possible measures; it is not practical to treat all the measures in a single paper.

The aim of this paper is to propose technological methods for QoE enhancement and show their feasibility quantitatively. Therefore, for simplicity of discussion, this paper picks up *satisfaction, effectiveness*, and *efficiency* from among the measures; it mainly focuses on satisfaction.

This paper takes the QoS approach of the layered network architecture model; therefore, the standpoint of this paper for QoE and QoS can be illustrated as shown in Fig. 1.

The effectiveness and efficiency of some applications and services are technically measurable in quantitative ways. This implies that they may be regarded as application–level QoS parameters; therefore, in Fig. 1, QoE of this portion overlaps with application–level QoS.

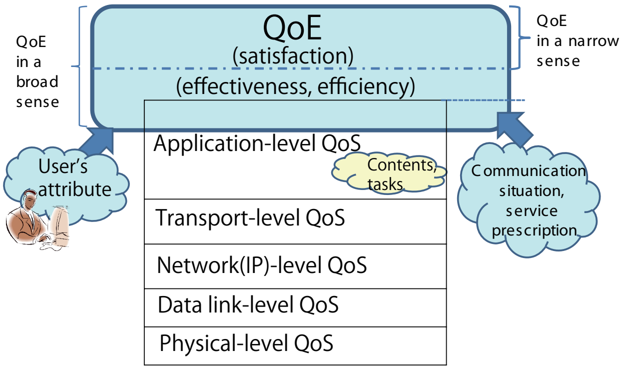

Figure 1: Standpoint of this paper for QoE and QoS.

It is apparent that the effectiveness and efficiency affect the satisfaction and vice versa; these three measures are not independent of each other.

Figure 1 also indicates that QoE is affected by the *user's attribute* such as age, gender, audio–visual capability, preference, experienced or not in the application/service, and mental conditions; also, it is affected by *communication situation* (e.g., user tasks, user groups such as business people and elderly people, and user environments like conference rooms) and *service prescription* (e.g., audio call, videophone, and IPTV, in addition to terminal types) [8], [9]. Furthermore, application–level QoS depends on the contents or tasks of the application/service, which therefore influence QoE.

In this paper, we refer to the portion of QoE for the satisfaction as *QoE in a narrow sense*, while the whole QoE, which also contains the effectiveness and efficiency, is called *QoE in a broad sense*. When we use the term QoE without any specific comments in this paper, it simply means "QoE in a narrow sense".

The most popular measure of "QoE in a narrow sense" is *MOS* (*Mean Opinion Score*). MOS is the average of the scores given by assessors in the subjective assessment of the degrees of user satisfaction like "excellent", "good", "fair", "poor" and "bad"; each degree of the satisfaction is expressed in terms of an integer score between 5 through 1, for example.

4 Methods for QoE Enhancement

Focusing on the technical aspects of QoE, we notice that three methods of QoE enhancement are available: *QoS control, utilization of multimedia*, and *user interface*.

QoS control is a traditional method of communication quality improvement; it includes bandwidth guarantee, media synchronization control and traffic control, for instance. We can expect that the QoS improvement by QoS control leads to QoE enhancement; however, quantitative relationship between the two has not been clarified sufficiently. Therefore, the QoS control employed may require too many resources to satisfy the QoE desired by the user or it may not realize the desired QoE.

Utilization of multimedia implies not only traditional audiovisual communications (senses of hearing and sight) but also communications dealing with other human senses like touch, smell and taste. It also means some extension of a sensory communication; e.g., from the conventional single-view video to multi-view video (MVV) or free-view

TV (FTV), which includes 3D TV as a special case, and from monophonic audio to stereophonic audio.

Furthermore, the user interface is important to QoE enhancement. This is because QoE heavily depends upon the users eventually; consequently, QoE enhancement requires smooth and appropriate interaction between the users and the system.

In order to assure QoE desired by users, we must design some mechanism to cope with all possible combinations of user's attributes, contents, tasks, communication situations, and service prescriptions; the mechanism should be implemented into some part of the networks, and necessary data should be saved in some database in the network like *transport user profile* of *NACF* (*Network Attachment Control Function*) and *service user profile* of *SCF* (*Service Control Function*) in NGN [11]. This is the main approach taken by the previous studies on QoE assurance and/or enhancement [2], [3], [4].

Note that the size of the databases and the complexity of the mechanism will be tremendous owing to a huge number of the combinations (say "*state space explosion*"), even if we restrict ourselves to a moderate number of supposed applications and situations. Thus, assuring QoE this way (namely, by some mechanism inside the network only) will be very difficult, though it is not impossible.

In addition to this difficulty, there is another factor which impedes QoE assurance; that is the user him/herself. This is because QoE depends upon the user's physical and mental conditions at the usage time as well as the user's attribute. This implies that if the user him/herself is regarded as a finite state machine, the number of the possible states is extraordinary large; this again leads to the "state space explosion". In essence, the user him/herself becomes an uncertain factor which makes the QoE level supported by the mechanism inside the network unsatisfactory.

In this paper, we propose a partial solution to the "state space explosion" problem. The basic idea is to leave QoS control which largely affects QoE to the user. More precisely, we provide an appropriate user interface through which the user can exert the QoS control; this can accommodate the user's inclination. In short, *render unto Caesar the things which are Caesar's*. We refer to this method as *user–assisted enhancement of QoE*. In the next section, we will give an example of the method.

We can find similar ideas to the proposed one in networking applications; they have been implemented in a variety of forms. For example, many plug–ins of Web browsers are available to *individual users for their customization* of Web services; the user can activate/deactivate plug–ins through the browser GUI when he/she feels them necessary/unnecessary. Thus, the basic idea itself is popular, but it has not been clarified how the idea can improve the services quantitatively from a QoE point of view.

The novelty of this paper is to show the effectiveness of the idea through a simple example quantitatively in addition to explicitly identifying it as a possible partial solution to the QoE management problem along with its reasoning.

5 An Example of User–Assisted Enhancement of QoE

To give an example of the proposed method, we suppose interactive audiovisual communications over a bandwidth guaranteed IP network [14] and focus on *playout buffering control* at the receiver as the QoS control.

5.1 Principle

The playout buffering control is widely used in audiovisual IP communications to absorb delay jitter produced in the network; it improves *intra–stream media synchronization quality* and therefore enhances QoE.

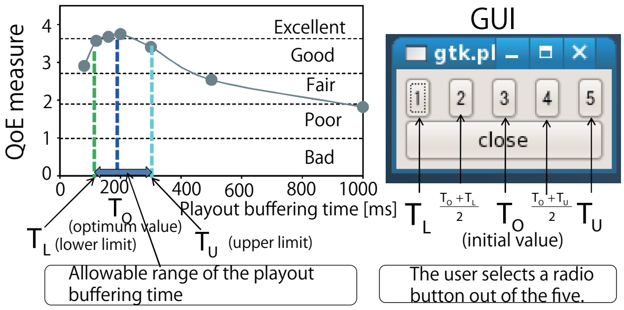

Figure 2: The optimum playout buffering time and GUI.

As shown on the left–hand side of Fig. 2, there exists the optimum playout buffering time T_O in the sense that T_O maximizes the QoE measure [14]. In other words, the playout buffering time produces a tradeoff relationship between *fidelity* and *latency*. As the buffering time increases, the fidelity of output audio and/or video improves because packets arriving later can be output (otherwise discarded) by absorbing larger delay jitter. At the same time, however, the longer buffering time increases the latency, which degrades the responsiveness. Shorter buffering time brings about the opposite effects on the output media quality and responsiveness. Thus, we have the tradeoff relationship. The value of T_O depends on the task performed. Consequently, we can anticipate that setting the playout buffering time to the T_O at the session setup time according to the task to be performed can achieve the maximum QoE.

However, it should be noted at this point that the optimum value T_O is usually the average of optimum values selected by many individual users; it may or may not be the optimum for each user. Some users prefer the fidelity of audio-video to the responsiveness, whereas other users may like the opposite. This implies that we need some "*customization*" mechanism that accommodates the *users' inclinations*; QoE can be improved further with this mechanism. This can be realized with a GUI (Graphical User Interface) which enables the users to select their favorite values of the playout buffering time T_b.

The right–hand side of Fig. 2 displays an example of GUI for the customization. The main part of the GUI is a set of five radio buttons 1 through 5. Button 3 corresponds to the initial value, which is set to T_O, while Buttons 1 and 5 are allocated to the lower limit and the upper one, respectively. Each limit is a value of T_b which achieves the QoE measure 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. 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.

5.2 QoE measure

As the QoE measure, this paper adopts the *psychological scale* [12] instead of MOS. In the context of *psychometric theory* [13], 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. An example of the relation between the psychological scale and MOS is displayed in Fig. 3. See [12] for more details.

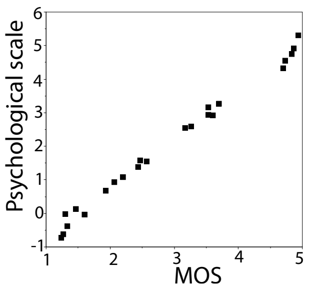

Figure 3: An example of psychological scale versus MOS.

5.3 Experimental system

In order to examine how effective the proposed method is, we constructed an experimental system for interactive audiovisual communications, which is shown in Figure 4.

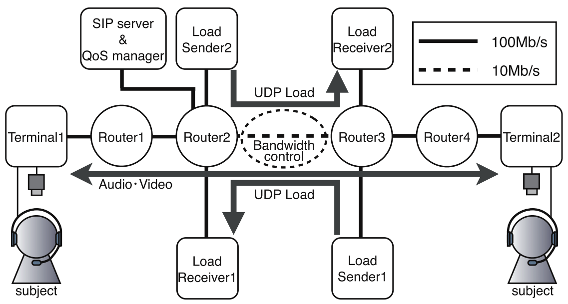

Figure 4: Configuration of the experimental system.

Resource allocation and management are carried out by the QoS manager in cooperation with the SIP [15] server. The SIP server extracts control information necessary for the QoS manager from the *INVITE* request message sent by the terminal and delivers it to the QoS manager, which sets up routers to guarantee the specified bandwidth.
We utilize SIP to customize the system parameter values (in particular, the playout buffering time); 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.
Each Terminal (PC) 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 1480 bytes each at exponentially distributed intervals. Load Receiver 1 and Load Receiver 2 are the corresponding receivers[1].
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 Routers 2 and 3; *LLQ* (*Low Latency Queueing*) [16] is adopted as the packet scheduling algorithm. In LLQ, we can set a *PQ* (*Priority Queueing*) class and *CBWFQ* (*Class–Based Weighted Fair Queueing*) classes: 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.

Table 1: Specification of audio and video.

audio encoding scheme	ITU–T G.711 μ–law
audio MU rate	25 [MU/s]
audio bit rate (guaranteed bandwidth)	64 [kb/s] (90 [kb/s])
video encoding scheme	H.264
image size	704 $\times$ 480 [pixel]
picture pattern	I
video encoding bit rate R_v [Mb/s] (B_G [Mb/s]: video guaranteed bandwidth)	when $B_G = 1$: 0.7, 0.8, 0.9, 1.0 when $B_G = 2$: 1.7, 1.8, 1.9, 2.0 when $B_G = 3$: 2.7, 2.8, 2.9, 3.0 when $B_G = 4$: 3.7, 3.8, 3.9, 4.0
video average MU rate	25 [MU/s]
monitor	17 inch LCD
monitor resolution	1280 $\times$ 1024 [pixel]

The specifications of audio and video used in the experiment are shown in Table 1, where an *MU* (*media unit*) means the transmission unit at the application layer, which corresponds to a constant number of audio samples or a video frame. The video frame skipping is employed for decoding of frames with packet loss and/or error.
The guaranteed bandwidth for audio is kept constant at 90 kb/s, while that for video[2] 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 fully utilized.
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.

5.4 Tasks

Since the purpose of this paper is to examine the feasibility of the user–assisted QoE enhancement method, we give simple conversational opinion tests, referring to ITU–T Rec. P.920 [17]. We have designed the following 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.

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. Note: task 1 is audio–dominant since the subjects exhibit only low motion.

[1]Routers 1 and 4 are laid out for possible connections of the Load Senders/Receivers in order to generate various patterns of the load traffic.
[2]Strictly speaking, this is the *minimum guaranteed bandwidth* because of the CBWFQ scheduling; the actually allocated bandwidth may be wider.

task 2 : Each subject reads numbers aloud in the same way as in task 1, but *clapping once for each number*.
Note: task 2 has a visual impact on QoE because of clapping.

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.

5.5 Method of experiment

In order to examine the effectiveness of the customization method, we conducted two kinds of subjective experiment: experiments without and with customization. The former keeps the playout buffering time constant to T_O throughout the session, while the subject in the latter tries the five values of T_b during a certain period of learning after the session establishment and finally chooses his/her favorite playout buffering time before the actual conversation; the value is kept the same until the session termination.
We easily notice that the user's interaction with GUI can impose a burden on him/her; this may degrade his/her QoE, contrary to the aim of QoE enhancement. Therefore, we specify that this method does not oblige the user to interact with GUI during not only the actual conversation period but also even the learning period. If the user takes no action, the value is kept to the initial default one T_O during the whole session. Whether he/she interacts or not with GUI during the learning period is solely up to him/herself.
Before we conducted the two kinds of experiment for the customization, we had made a preliminary experiment.

5.6 Preliminary experiment

In each task, we carried out subjective experiment for all combinations of R_v and T_b for each value of B_G to find out T_O (i.e., the initial default value of T_b for the customization) and the corresponding value of R_v, which maximizes the psychological scale value.
In each 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 for this experiment are 24 Japanese male and female students in their twenties.
We calculate the psychological scale by applying *Thurstone's law of categorical judgment* [13] to the 288 stimuli and give *Mosteller's test* for the goodness of fit on the hypothesis that the observed value equals the calculated one.
For each task, we calculated the psychological scale as a function of T_b for all combinations of B_G and R_v (16 pairs according to Table 1). After removing 38 stimuli with large errors of Mosteller's test, we found that the hypothesis cannot be rejected at a significance level of 0.05.

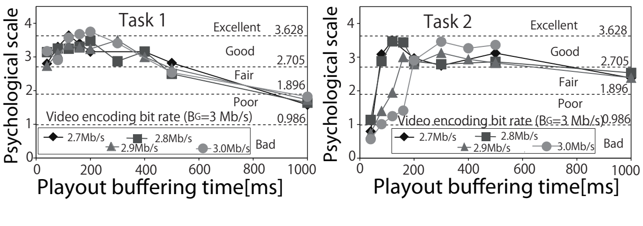

Figure 5: Psychological scale versus playout buffering time for task 1 and task 2 when B_G=3 Mb/s.

As examples, we show the results when B_G=3 Mb/s for task 1 and task 2 in Fig. 5, where four curves corresponding to the four values of R_v are plotted for each task; the minimum value of the psychological scales for the 250 stimuli is set to the origin. Note that we can select an arbitrary origin in the interval scale [13]. In Fig. 5, the dashed lines parallel to the abscissa represent the lower boundaries of the ACR categories.
Then, for each value of B_G, we chose the value of R_v which maximizes the psychological scale value (e.g., R_v=3.0 Mb/s for task 1 and R_v=2.7 Mb/s for task 2 in Fig. 5); for that pair, we also identified the optimum playout buffering time T_O, the lower limit T_L and the upper limit T_U for GUI as illustrated in Fig. 2. Table 2 summarizes the results thus obtained. More detailed description of the preliminary experiment can be found in [14].

Table 2: The optimum parameters and the playout buffering time for GUI.

task	B_G [Mb/s]	R_v [Mb/s]	T_b [ms]		
			T_O	T_L	T_U
task 1	1	1.0	200	160	300
	2	2.0	120	80	400
	3	3.0	200	120	300
	4	4.0	160	120	300
task 2	1	1.0	400	300	1000
	2	1.9	300	200	400
	3	2.7	120	80	200
	4	3.7	160	120	200

5.7 Setup of experiment

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 video guaranteed bandwidth B_G was set to one selected from among 1.0, 2.0, 3.0 and 4.0 Mb/s for each task. For each value of B_G, we set the initial value of T_b and the value of R_v to T_O and the corresponding value of R_v, respectively, referring to Table 2. In both kinds of the experiment, the value of R_v remained the same throughout the whole session. Regarding T_b, on the other hand, the value in the experiment without customization was kept at the initial value T_O (not changed) throughout the session, while the experiment with customization allowed the users to change the value of T_b only during the learning period.
Thus, we have 16 stimuli because of the four values of the video guaranteed bandwidth B_G, 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.

5.8 Experimental results

We obtained the psychological scale for the 16 stimuli; Mosteller's test on the scale indicated that the hypothesis cannot be rejected at a significance level of 0.05.
Figure 6 shows the psychological scales with and without customization as a function of the video guaranteed bandwidth for tasks 1 and 2; the minimum value of the psychological scales for the 16 stimuli is set to unity (i.e., 1).
In the figure, we first see that the customization method improves QoE for every guaranteed bandwidth as expected; a Wilcoxon signed rank test with a 0.05 significance level

has rejected the null hypothesis representing no difference in the psychological scale between with customization and without one. We also confirmed this with MOS.

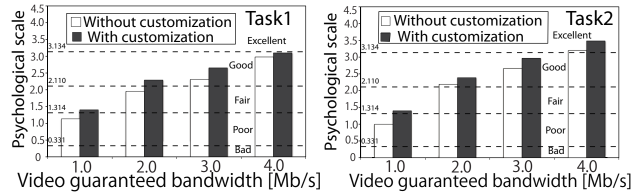

Figure 6: QoE enhancement by customization.

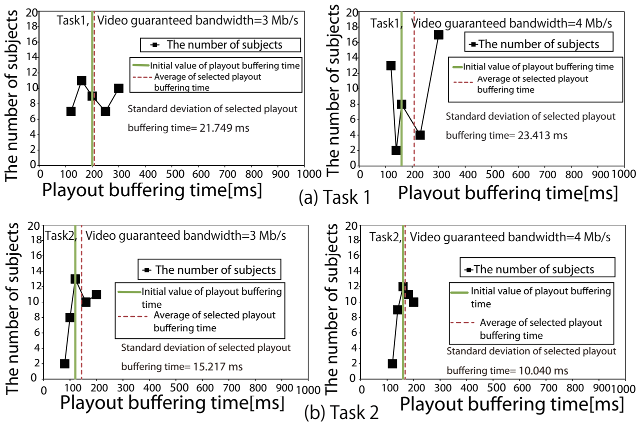

Figure 7: Histogram of selected playout buffering time for (a) task 1 and (b) task 2 (B_G=3 Mb/s, 4 Mb/s).

Comparing the result for task 1 and that for task 2 in Fig. 6, we find that for the video guaranteed bandwidth of 3.0 Mb/s and 4.0 Mb/s, actually achieved QoE of task 2 is higher than that of task 1, while for 1.0 Mb/s and 2.0 Mb/s, they are close to each other. This implies that the visual effect enhances QoE when the video quality is high.

Now, let us give further consideration to the implication above by examining the distribution of selected playout buffering time through the customization. Figure 7 displays histograms of the selected playout buffering time for task 1 and task 2 when the video guaranteed bandwidth is 3 Mb/s and 4 Mb/s.

Comparing task 1 and task 2 in Fig. 7, we easily notice that Fig. 7(b) presents convex curves centered round the initial value (i.e., convergent patterns), whereas Fig. 7(a) shows rather divergent patterns (larger standard deviations than task 2). This is because task 1 is audio–dominant, while task 2 has a stronger visual effect due to clapping, with which the subjects can recognize the difference in video quality more clearly in task 2 than in task 1. This may make the subjects select the button of GUI with more confidence. Regarding this issue, we need more detailed analysis of the subjects' behavior, which is left as future work.

6 Conclusions

We proposed an idea of *user–assisted enhancement of QoE* by the accommodation of *individual users' inclinations* and gave a simple example of the implementation with GUI for interactive audiovisual communications; the usage of the method is not mandatory but optional in order not to impose a burden on the users. We then showed that the proposed method can improve QoE. It should be noted that the proposed method is intended not to replace the QoE/QoS provisioning mechanism inside the network but to fill a gap. This is a sort of *truly end–to–end argument*.

The method proposed in this paper is just a first–step trial of study on user–assisted enhancement of QoE; in order to implement practical systems based on this idea, we have to solve many problems which are left as future work. An extension of the proposed method to a QoE guarantee architecture called *interactive GPSQ* can be found in [18].

Acknowledgment: The authors thank N. Misaki for his assistance in the experiment. This work was supported by the Grant-In-Aid for Scientific Research of Japan Society for the Promotion of Science under Grant 25420365.

References

[1] Cisco White Paper: Cisco Visual Networking Index: Forecast and Methodology, 2012–2017, http://www.cisco.com. Cisco Inc., May 2013.

[2] Volk, M.: Sterle, J.: Sedlar, U.: Kos, A.: An approach to modeling and control of QoE in next generation networks, *IEEE Commun. Mag.*, vol.48 No.8 pp.126–135, Aug. 2010.

[3] Sterle, J.: Volk, M.: Sedlar, U.: Bester, J.: Kos, A.: Application–based NGN QoE controller, *IEEE Commun. Mag.*, vol.49 No.1 pp.92–101, Jan. 2011.

[4] Zhang, J.: Ansari, N.: On assuring end–to–end QoE in next generation networks: Challenges and a possible solution, *IEEE Commun. Mag.*, vol.49 No.7 pp.185–191, July 2011.

[5] Special issue on improving QoE for network services, *IEEE Networks*, vol.24 no.2 Mar./Apr. 2010.

[6] Special issue on QoE management in emerging multimedia services, *IEEE Commun. Mag.*, vol.50 No.4, Apr. 2012.

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

[8] Brooks, P.: Hestness, B.: User measures of quality of experience: Why being objective and quantitative is important, *IEEE Network*, vol.24, no.2, pp.8–13, Mar./Apr. 2010.

[9] Le Callet, P.: Möller, S.: Perkis, A., eds.,: Qualinet Whitepaper on Definitions of Quality of Experience (2012), European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), Lausanne, Switzerland,Ver. 1.2, Mar. 2013.

[10] Wu, W.: Arefin, A.: Rivas, R.: Nahrsted, K.: Sheppard, R. M.: Yang, Z.: Quality of experience in distributed interactive multimedia environments: Toward a theoretical framework, in *Proc. ACM Multimedia2009*, pp.481–490, Oct. 2009.

[11] ITU-T Rec. Y.2012: Functional requirements and architecture of the NGN release 1, Sept. 2006.

[12] Tasaka, S.: Ito, Y.: Psychometric analysis of the mutually compensatory property of multimedia QoS, in *Conf. Rec. IEEE ICC*, pp. 1880–1886, May 2003.

[13] Torgerson, W. S.: Theory and methods of scaling, John Wiley & Sons, Inc., 1958.

[14] Tasaka, S.: Misaki, N.: Maximizing QoE of interactive services with audio–video transmission over bandwidth guaranteed IP networks, in *Conf. Rec. IEEE GLOBECOM*, Dec. 2009.

[15] IETF RFC 3261: SIP: Session Initiation Protocol, June 2002.

[16] Vegesna, S.:IP Quality of Service, CiscoPress, 2001.

[17] ITU-T Rec. P.920: Interactive test methods for audiovisual communications, May 2000.

[18] Tasaka, S.: Misaki, N.: Customization of interactive services for QoE enhancement in audio–video transmission over bandwidth guaranteed IP networks, in *Proc. IEEE ICME*, pp.546–551, July 2010.