

Effect of Monitor Size on User-Level QoS of Audio-Video Transmission over IP Networks in Ubiquitous Environments

Yoshihiro Ito and Shuji Tasaka

Department of Computer Science and Engineering,
Graduate School of Engineering, Nagoya Institute of Technology,
Nagoya, 466-8555, Japan
Email: {yoshi, tasaka}@nitech.ac.jp

Abstract— This paper investigates user-level (or perceptual) QoS of audio-video transmission over IP networks with terminals of different size monitors. As the first step of our research, we treat large-sized monitors. By experiment, we assess the user-level QoS for three monitor sizes (80, 60 and 40 inches) and perform QoS mapping. In our experiment, we changed the monitor size by using a projector. The user-level QoS is assessed by the method of successive categories, which is one of the psychometric methods. Moreover, we performed QoS mapping from application-level to user-level with multiple regression analysis. The experimental result indicates that when the degradation of the quality of the audio-video transmission is not noticeable, the user-level QoS of the large monitor size (i.e., 80 inch) becomes the highest. We also observe that when we can notice a low level of degradation of the transmission, the user-level QoS parameter of the large monitor size becomes the smallest. When the degradation is too much, we hardly notice any difference in the user-level QoS between the monitor sizes. We thus confirmed that the monitor size is one of important application-level QoS parameters we must consider.

I. INTRODUCTION

Recently, as many types of access networks such as IEEE 802.11 wireless LANs become available in the Internet, we can receive audio-visual application services in ubiquitous environments. For example, let us consider audio-visual broadcasting over the Internet. Some people may see the broadcasting with a huge display on the street, while the others may watch it via their PDAs with a wireless LAN. Moreover, in future mobile networks, cross-device handover will be realized [1]; it supports seamless handover between various types of terminals which treat audio-video transmission. Thus, we may watch the same audio-video stream with different terminals according to our situation.

In audio-video transmission over a ubiquitous IP network (i.e., an IP network in the ubiquitous environment), we use various terminals; the monitor size differs from terminal to terminal. The difference in monitor size between terminals can affect our subjectivity of audio-video transmission. Therefore, in order to achieve desirable user-level QoS in a ubiquitous IP network, we must understand the effect of the difference in monitor size between terminals on user-level QoS of audio-video transmission. However, it has not been clarified

how the difference affects the user-level QoS of audio-video transmission over a ubiquitous IP network.

The user-level QoS is perceptual one and the most important in audio-video transmission because the last recipient of the service is a user. Since the Internet has a layered structure, QoS also has a layered structure. For example, reference [2] identifies six levels of QoS: *physical-level*, *node-level*, *network-level*, *end-to-end-level*, *application-level* and *user-level*.

Many researches have indicated the effect of monitor size on video subjectivity; for instance, see [3], [4] and [5]. However, they were studied in local environments; that is, video is not transmitted over networks. On the other hand, there are some publications which study the effect of various aspects of video on the user-level QoS of audio-video transmission. For example, Steinmetz treated audio-video streams in three different views: head, shoulder, and body [6]. He investigates the subjective tolerance of skew, which is difference between audio delay and video one. Frowein *et al.* used two spatial resolutions, CIF and QCIF, and assessed the potential effectiveness of videotelephony as an adjunctive aid to speech reception by the hard of hearing [7]. Nakazono studied the effect of video frame rate on speech perception in multimedia communication [8]. However, references [7] and [8] do not treat network delay and delay jitter. In the literature, we can find no study that treats the effect of the monitor size on the user-level QoS of audio-video transmission over a ubiquitous IP network.

This paper quantitatively assesses the effect of the monitor size on the user-level QoS of audio-video transmission over a ubiquitous IP network. As the first step of our research, we treat large-sized monitors to confirm the existence of the effect of the monitor size clearly. First, we assess the user-level QoS with three monitor sizes (80, 60 and 40 inches). In order to assess the user-level QoS quantitatively, we adopt the *method of successive categories*, which is one of the *psychometric methods* [9]. Then, we perform QoS mapping from application-level to user-level. By QoS mapping, we quantitatively investigate the effect of the monitor size on the user-level QoS. In our research, we use *multiple regression*

analysis as a QoS mapping method.

The rest of the paper is organized as follows. Section II discusses the assessment of the user-level QoS with the method of successive categories. Section III describes a QoS mapping method with multiple regression analysis. Sections IV and V show our experiment and its results, respectively.

II. USER-LEVEL QoS ASSESSMENT

A. Assessment with psychometric methods

The authors proposed the utilization of psychometric methods for user-level QoS assessment in multimedia transmission [10], [11]. The psychometric methods were proposed to measure human subjectivity quantitatively in the psychological field [9]. In [10], the authors use the method of paired comparisons and Thurstone's law of comparative judgment [9]. Moreover, in [11], the method of successive categories [9] is utilized.

With the psychometric methods, the human subjectivity can be represented by a measurement scale. We can define four basic types of the measurement scales according to the mathematical operations that can be performed legitimately on the numbers obtained by the measurement; from lower to higher levels, we have *nominal*, *ordinal*, *interval* and *ratio* scales [9]. Since almost all the statistical procedures can be applied to the interval scale and the ratio scale, it is desirable to represent the user-level QoS by an interval scale or a ratio scale. With the psychometric methods used in [10] and [11], we can represent user-level QoS by an interval scale. The method of paired comparisons and Thurstone's law of comparative judgment can give more accurate values of the interval scale but takes longer experimental time than the method of successive categories.

In this paper, we adopt the method of successive categories to assess user-level QoS quantitatively. An outline of this method will be given in the following subsection.

B. Method of successive categories

In the method of successive categories, a subjective score is measured by the *rating-scale method* [9]. In the method, subjects classify each stimulus into one of a certain number of categories. Here, a stimulus means an object for evaluation, such as audio and video. Each category has a predefined number. For example, "excellent" is assigned 5, "good" 4, "fair" 3, "poor" 2 and "bad" 1. However, since the assigned number is an ordinal scale, we cannot use the assigned number for obtaining the user-level QoS parameter in the strict sense.

In order to obtain an interval scale as the user-level QoS parameter, we first measure the frequency of each category with which the stimulus was placed in the category by the rating-scale method. With the *law of categorical judgment* [12], we can translate the frequency obtained by the rating-scale method into an interval scale. We can apply almost all the operations to the scale.

C. The law of categorical judgment

The law of categorical judgment makes the following assumptions. Let the number of the categories be $m + 1$. When

a series of stimuli is presented to a subject, there occurs some kind of process (discriminal process) in him/her by which he/she reacts differently according to a quantitative or qualitative attribute of each stimulus. The quantitative attribute defines what we call the *psychological continuum* [9]. The psychological continuum is an interval scale. When stimulus j ($j = 1, \dots, n$) is presented to a subject, the discriminative process gives a psychological value designated by s_j on the interval scale (psychological continuum). For the $m + 1$ categories, their *boundaries* have values on the interval scale. We denote the upper boundary of category g ($g = 1, \dots, m + 1$) by c_g and define $c_0 \triangleq -\infty$ and $c_{m+1} \triangleq +\infty$. The subject sorts n stimuli into the $m + 1$ categories ($n > m + 1$) by comparing s_j with c_g . If $c_{g-1} < s_j \leq c_g$, then stimulus j is classified into category g . The categories can be arranged in a rank order, in the sense that each stimulus in category g is judged to have a psychological value which is "less than" the one for any stimulus in category $g + 1$. This statement holds for all values of g from 1 to m . The variable c_g is normally distributed with mean t_g and standard deviation d_g . Also, the variable s_j is normally distributed with mean R_j and standard deviation σ_j . Then, we can consider R_j as 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. Mosteller proposed a method of testing the goodness of fit for a scale calculated with Thurstone's law [13]. The method can be applied to a scale obtained by the law of categorical judgment. In this paper, we use Mosteller's method to test the goodness of fit.

III. QoS MAPPING

A. QoS mapping with multiple regression analysis

Since we cannot control user-level QoS directly, it is desirable to control QoS at lower levels so as to keep user-level QoS high. To do this, we need to clarify the relationship between user-level QoS and QoS at lower levels. We call the relationship *QoS mapping*. QoS mapping is also useful for investigating lower-level QoS parameters which affect user-level QoS. As a method of QoS mapping, the authors have proposed the utilization of multiple regression analysis [10], [11]. In this method, we consider the user-level QoS parameter as a criterion variable, and QoS parameters at lower levels as predictor variables.

In this paper, we adopt multiple regression analysis as the QoS mapping method and consider application-level QoS parameters as predictor variables since the application-level is beneath the user-level. By QoS mapping, we can quantitatively study the effect of the monitor size on user-level QoS of audio-video transmission.

B. Application-level QoS parameters

To perform multiple regression analysis, we must select some application-level QoS parameters as predictor variables. In this paper, we regard measures of media synchronization quality as candidates of the predictor variables.

In general, media synchronization is classified into *intra-stream synchronization* and *inter-stream synchronization*. The former keeps the continuity of a single stream (audio or video), while the latter is synchronization between an audio stream and the corresponding video stream.

In order to represent media synchronization quality, reference [10] uses nine application-level QoS parameters. We also use them in this paper. First, we adopt the *coefficient of variation of output interval*, which is defined as the ratio of the standard deviation of the MU output interval of a stream to its average. MU stands for “media unit”, which indicates the information unit for media synchronization. This parameter is denoted by C_a for audio and by C_v for video. Second, we use the *average MU rate* for audio R_a and that for video R_v ; this is defined as the average number of (either audio or video) MUs output in a second at the destination. Third, we treat the *MU loss rate* for audio L_a and that for video L_v ; this is the ratio of the number of lost MUs to the total number of generated MUs. Finally, we adopt the *mean square error of intra-stream synchronization*, which is defined as the average square of the difference between the output interval of MU at the destination and the generation one at the source. We denote it by E_a for audio and by E_v for video. These eight parameters indicate the intra-stream synchronization quality.

The QoS parameter for the inter-stream synchronization is the *mean square error* E_{int} , which is defined as the average square of the difference between the output-time difference of the audio and corresponding video MUs and their timestamp difference.

IV. EXPERIMENT

A. Experimental method

In our experiment, subjects assess user-level QoS of audio-video transmission with three different monitor sizes. In order to change the monitor size, we use an LCD projector and a screen. Figure 1 shows our experimental environment, where the screen and two speakers are set up ahead of the projector, and three subjects sit at the back of it. The size of the screen is $1.8\text{m} \times 1.8\text{m}$. The three subjects simultaneously assess audio-video streams whose video is projected on the screen. The distance between the screen and the subjects is 4.0 m. The projector is EPSON ELP-5000. We adjust the distance between the projector and the screen so that the monitor size on the screen becomes a predefined size. In our experiment, we treat three monitor sizes which are equivalent to 40, 60 and 80 inch monitors. By using the projector, we can change the monitor size without varying the application-level QoS of audio-video transmission.

To produce audio-video streams for user-level QoS assessment, we simulated audio-video transmission. The network simulator we used is ns [14]. We set up a network configuration shown in Fig. 2. In this network, the media sender is connected to the media receiver via two routers and transmits a pair of audio and video streams to the media receiver. Table I shows the media specifications of the audio and video streams which the media sender transmits. The line speed between the

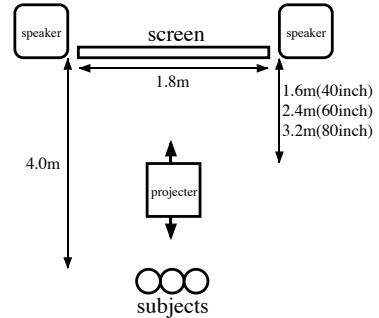


Fig. 1. Experimental environment.

two routers and the one between a router and a terminal are all 10.0 Mbps. The load generator generates UDP messages of 1472 bytes each at exponentially distributed intervals as load traffic. We changed the average amount of load traffic from 2.8 to 3.5 Mbps by 0.1 Mbps. While the media sender sends a pair of audio and video streams, the load generator transmits the load traffic. We recorded the audio-video streams that the media receiver output and regard them as stimuli for user-level QoS assessment. Thus, the number of stimuli for all monitor sizes becomes $3 \times 8 = 24$.

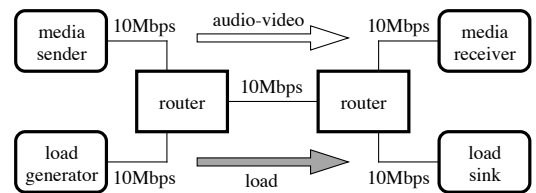


Fig. 2. Experimental network.

TABLE I
SPECIFICATIONS OF AUDIO AND VIDEO STREAMS.

	audio	video
coding scheme	linear PCM	MPEG1
image size[pixels]	-	640 × 480
picture pattern	-	IPPPPP
average MU rate[MU/s]	50	30
average bit rate[kbps]	1536	5000

B. Subjective assessment

In the rating-scale method, we used five categories (i.e., $m = 4$) of impairment: “imperceptible” assigned integer 5, “perceptible, but not annoying” 4, “slightly annoying” 3, “annoying” 2, and “very annoying” 1. The subjects were non-experts in the sense that they were not directly concerned with audio and video quality as a part of their normal work. The number of the subjects is 26. They are male, and their ages were between 20 and 25. It took about fifteen minutes for a subject to finish all the kinds of assessment.

V. EXPERIMENTAL RESULTS

A. User-level QoS assessment

Table II shows the number of subjects who classified the stimulus into each category by the rating-scale method. In Table II, the leftmost column represents stimulus j ($j = 1, \dots, 24$).

TABLE II
RESULT OF THE RATING-SCALE METHOD.

j	stimulus		category g				
	Load [Mbps]	Monitor size [inch]	1	2	3	4	5
1	2.8	80	0	0	0	2	24
2	2.9	80	0	0	0	2	24
3	3.0	80	0	0	0	5	21
4	3.1	80	0	0	4	15	7
5	3.2	80	0	6	14	5	1
6	3.3	80	0	13	12	1	0
7	3.4	80	14	9	2	1	0
8	3.5	80	19	7	0	0	0
9	2.8	60	0	0	0	3	23
10	2.9	60	0	0	0	3	23
11	3.0	60	0	0	0	4	22
12	3.1	60	0	0	1	18	7
13	3.2	60	0	7	14	5	0
14	3.3	60	1	13	9	3	0
15	3.4	60	10	11	5	0	0
16	3.5	60	15	11	0	0	0
17	2.8	40	0	0	0	3	23
18	2.9	40	0	0	0	3	23
19	3.0	40	0	0	0	3	23
20	3.1	40	0	0	1	19	6
21	3.2	40	0	8	15	3	0
22	3.3	40	1	17	8	0	0
23	3.4	40	9	15	2	0	0
24	3.5	40	18	7	1	0	0

We translate the data measured by the rating-scale method into an interval scale with the law of categorical judgment. In the law, we can consider four conditions, conditions A, B, C and D, which differ in assumptions, approximations, and degree of simplification [12]. In this paper, we try condition D.

Let the probability that s_j is less than c_g be p_{jg} . We regard the proportion of times that s_j is less than c_g to the number of subjects as the observed value of the probability. Under condition D, the law of categorical judgment can be represented by

$$t_g - R_j = Z_{jg} \quad (1)$$

where Z_{jg} is the *normal deviate*, which is defined as the distance from the mean of a normal distribution with zero mean and unit variance for p_{jg} .

In the law of categorical judgment, we assume that there exist the true values for the quantities introduced in the previous subsection and represent the corresponding observed and estimated values from observed data by attaching the prime ($'$) and the double prime ($''$), respectively, to the true values.

From Table II, we can get the probability p'_{jg} , which provides the estimated value Z''_{jg} . Using Z''_{jg} , we calculate the estimated values t''_g and R''_j for t_g and R_j , respectively. If an observed probability p'_{jg} is 0 or 1, Z''_{jg} becomes negative or positive infinity. In this case, Z''_{jg} is considered as a *missing entry*. If some missing entries exist, we have the alternative of finding t''_g first or R''_j first. We select the former here; that is, we calculate t''_g first and then use the obtained values to determine R''_j .

Table III shows the normal deviate Z''_{jg} corresponding to p'_{jg} . We consider Z''_{jg} whose p'_{jg} is 0 or 1 as a missing entry and represent it by “-” in Table III. Note that p'_{j5} always becomes 1 since all stimuli belong to any one of categories 1 through 5. Consequently, Z''_{j5} is always regarded as a missing entry. We then omit the column of Z''_{j5} from Table III.

TABLE III
NORMAL DEViate Z''_{jg} CORRESPONDING TO p'_{jg} .

j	stimulus		Z''_{j1}	Z''_{j2}	Z''_{j3}	Z''_{j4}
	Load [Mbps]	Monitor size [inch]				
1	2.8	80	-	-	-	-1.426
2	2.9	80	-	-	-	-1.426
3	3.0	80	-	-	-	-0.869
4	3.1	80	-	-	-1.020	0.615
5	3.2	80	-	-0.736	0.736	1.769
6	3.3	80	-	0.000	1.769	-
7	3.4	80	0.097	1.198	1.769	-
8	3.5	80	0.615	-	-	-
9	2.8	60	-	-	-	-1.198
10	2.9	60	-	-	-	-1.198
11	3.0	60	-	-	-	-1.020
12	3.1	60	-	-	-1.769	0.615
13	3.2	60	-	-0.615	0.869	-
14	3.3	60	-1.769	0.097	1.198	-
15	3.4	60	-0.293	0.869	-	-
16	3.5	60	0.194	-	-	-
17	2.8	40	-	-	-	-1.198
18	2.9	40	-	-	-	-1.198
19	3.0	40	-	-	-	-1.198
20	3.1	40	-	-	-1.769	0.736
21	3.2	40	-	-0.502	1.198	-
22	3.3	40	-1.769	0.502	-	-
23	3.4	40	-0.396	1.426	-	-
24	3.5	40	0.502	1.769	-	-

In order to determine t''_g , we first estimate the average width of each category. Then, we regard one boundary as the origin and calculate boundaries of the others. An estimated value $t''_{g+1} - t''_g$ can be calculated by

$$t''_{g+1} - t''_g = \frac{1}{q_g} \sum_j^{q_g} (Z''_{j,g+1} - Z''_{jg}) \quad (2)$$

where $\sum_j^{q_g}$ means the summation for j for which both Z''_{jg} and $Z''_{j,g+1}$ are available, and q_g is the number of the available data for a given g [11]. Table IV shows $Z''_{j,g+1} - Z''_{jg}$ for $g = 1, 2$ and 3.

TABLE IV
ESTIMATE OF CATEGORY WIDTH.

j	stimulus		$Z''_{j2} - Z''_{j1}$	$Z''_{j3} - Z''_{j2}$	$Z''_{j4} - Z''_{j3}$
	Load [Mbps]	Monitor size [inch]			
1	2.8	80	-	-	-
2	2.9	80	-	-	-
3	3.0	80	-	-	-
4	3.1	80	-	-	1.635
5	3.2	80	-	1.473	1.033
6	3.3	80	-	1.769	-
7	3.4	80	1.102	0.570	-
8	3.5	80	-	-	-
9	2.8	60	-	-	-
10	2.9	60	-	-	-
11	3.0	60	-	-	-
12	3.1	60	-	-	2.384
13	3.2	60	-	1.485	-
14	3.3	60	1.865	1.102	-
15	3.4	60	1.163	-	-
16	3.5	60	-	-	-
17	2.8	40	-	-	-
18	2.9	40	-	-	-
19	3.0	40	-	-	-
20	3.1	40	-	-	2.505
21	3.2	40	-	1.701	-
22	3.3	40	2.271	-	-
23	3.4	40	1.822	-	-
24	3.5	40	1.266	-	-
Average			1.582	1.350	1.889

From Table IV, we can calculate $t''_{g+1} - t''_g$. By regarding the mean t''_1 of the upper boundary of category 1 as the origin of the obtained interval scale, we can obtain the mean of the upper boundary of each category. Thus, we have $t''_1 = 0.000$, $t''_2 = 1.582$, $t''_3 = 2.932$ and $t''_4 = 4.821$.

Next, R''_j can be obtained by

$$R''_j = \frac{1}{q_j} \sum_g^{q_j} (t''_g - Z''_{jg}) \quad (3)$$

where $\sum_g^{q_j}$ means the summation for g for which Z''_{jg} is available, and q_j is the number of the available data for a given j [11]. Table V shows R''_j in the rightmost column. The obtained R''_j is an interval scale, that is, the user-level QoS parameter.

TABLE V
SCALE VALUE R''_j OF THE STIMULUS j .

j	stimulus		t''_1 Z''_{j1}	t''_2 Z''_{j2}	t''_3 Z''_{j3}	t''_4 Z''_{j4}	average R''_j
	Load [Mbps]	Monitor size [inch]					
1	2.8	80	-	-	-	6.247	6.247
2	2.9	80	-	-	-	6.247	6.247
3	3.0	80	-	-	-	5.690	5.690
4	3.1	80	-	-	3.951	4.205	4.078
5	3.2	80	-	2.318	2.195	3.052	2.522
6	3.3	80	-	1.582	1.163	-	1.372
7	3.4	80	-0.097	0.383	1.163	-	0.483
8	3.5	80	-0.615	-	-	-	-0.615
9	2.8	60	-	-	-	6.019	6.019
10	2.9	60	-	-	-	6.019	6.019
11	3.0	60	-	-	-	5.841	5.841
12	3.1	60	-	-	4.700	4.205	4.453
13	3.2	60	-	2.197	2.062	-	2.219
14	3.3	60	1.769	1.485	1.733	-	1.662
15	3.4	60	0.293	0.712	-	-	0.053
16	3.5	60	-0.194	-	-	-	-0.194
17	2.8	40	-	-	-	6.019	6.019
18	2.9	40	-	-	-	6.019	6.019
19	3.0	40	-	-	-	6.019	6.019
20	3.1	40	-	-	4.700	4.084	4.392
21	3.2	40	-	2.084	1.733	-	1.909
22	3.3	40	1.769	1.079	-	-	1.424
23	3.4	40	0.396	0.155	-	-	0.276
24	3.5	40	-0.502	-0.187	-	-	-0.345

In order to test the obtained interval scale, we performed Mosteller's test. As a result of Mosteller's test, the null hypothesis that the obtained interval scale fits the observed data cannot be rejected at significance level 0.05. That is, if the hypothesis is right, the probability that the hypothesis is rejected by mistake is less than 0.05. Therefore, we consider that the obtained scale is appropriate for the user-level QoS parameter.

Figure 3 plots the user-level QoS parameter versus the average load. In Fig. 3, we make the following observations. First, when the average load is less than 3.0 Mbps, the user-level QoS parameter value of the 80 inch monitor size takes the highest. In this case, it was difficult to notice the degradation of the quality of the media for all the monitor sizes. Therefore, it is considered that a large monitor size makes us judge higher quality than a small one when there is hardly any degradation of the media. Second, when the amount of the average load are 3.0 and 3.1 Mbps, the user-level QoS parameter values of the 80 inch monitor size become the smallest. This result suggests that we can notice a low level of degradation of media with the 80 inch monitor size, while we cannot notice it with the other two monitor sizes. Third, when the average load is 3.2 Mbps,

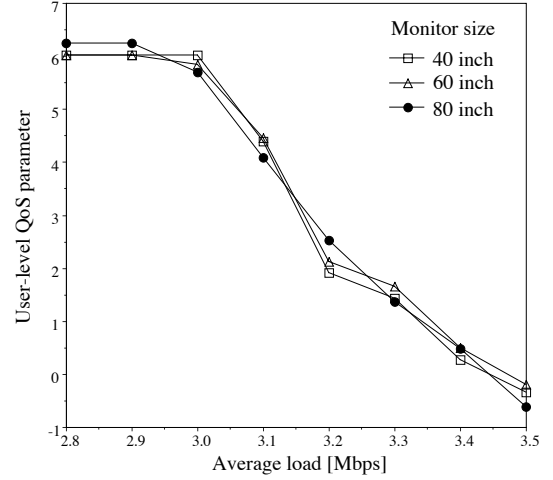


Fig. 3. User-level QoS parameter versus average load.

the user-level QoS value of the 80 inch monitor size becomes the highest again. This means that a large monitor makes us feel better when the degradation of media is obvious. Finally, when the average load is over 3.2 Mbps, we hardly notice any difference in the user-level QoS between the monitor sizes. This is because the degradation of media is too much.

B. Classification of application-level QoS parameters

In order to perform QoS mapping from the application-level QoS to the user-level one by multiple regression analysis, we must select some application-level QoS parameters as predictor variables. Then, we classify the application-level QoS parameters by the principal component analysis. As a result, we see that the cumulative contribution rate for the first two principal components is 93.5 %. This means that the first two principal components can present 93.5 % of information involved by the nine application-level QoS parameters. Therefore, we adopt the first and second principal components. Table VI displays the principal component loading of each principal component.

TABLE VI
PRINCIPAL COMPONENT LOADING.

	First principal component	Second principal component
R_a	-0.995	0.018
L_a	0.995	-0.017
E_a	0.647	0.642
C_a	0.877	-0.377
R_v	-0.991	-0.005
L_v	0.991	0.004
E_v	-0.564	0.677
E_{int}	0.953	0.260
C_v	0.986	0.076

From Table VI, we find that the nine parameters can be classified into two groups:

- group a)** R_a , L_a , C_a , R_v , L_v , E_{int} and C_v
- group b)** E_a and E_v

The parameters in group a) highly correlate with the first principal component. On the other hand, the parameters in

group b) slightly correlate with the first and second principal components.

C. QoS mapping

By multiple regression analysis, we perform QoS mapping from the application-level QoS parameters to the user-level QoS parameter for each monitor size. That is, we consider the application-level QoS parameters and the user-level QoS parameter as predictor variables and the criterion variable, respectively. Then, we calculate a multiple regression line.

Before multiple regression analysis, we select some out of the nine application-level QoS parameters as predictor variables. In order to avoid the effect of multi-collinearity, we select one application-level QoS parameter from each group described in the previous subsection. Consequently, the number of the combination of the application-level QoS parameters becomes $2 \times 7 = 14$. In this paper, we first perform multiple regression analysis with all combinations of the application-level QoS parameters as predictor variables for each monitor size. Then, we select a combination which achieves the highest contribution rate adjusted for degrees of freedom. Tables VII through IX present contribution rates adjusted for degrees of freedom for each monitor size.

TABLE VII

CONTRIBUTION RATES ADJUSTED FOR DEGREES OF FREEDOM (80 INCH OF MONITOR SIZE)

	R_a	L_a	C_a	R_v	L_v	E_{int}	C_v
E_a	0.918	0.920	0.893	0.904	0.903	0.834	0.957
E_v	0.911	0.912	0.736	0.896	0.894	0.840	0.958

TABLE VIII

CONTRIBUTION RATES ADJUSTED FOR DEGREES OF FREEDOM (60 INCH OF MONITOR SIZE)

	R_a	L_a	C_a	R_v	L_v	E_{int}	C_v
E_a	0.905	0.907	0.870	0.889	0.888	0.818	0.950
E_v	0.917	0.918	0.760	0.898	0.896	0.816	0.959

TABLE IX

CONTRIBUTION RATES ADJUSTED FOR DEGREES OF FREEDOM (40 INCH OF MONITOR SIZE)

	R_a	L_a	C_a	R_v	L_v	E_{int}	C_v
E_a	0.894	0.896	0.860	0.877	0.876	0.807	0.944
E_v	0.908	0.909	0.753	0.887	0.885	0.805	0.955

From Table VII through IX, we see that two combinations, (E_a, C_v) and (E_v, C_v) , show high contribution rates adjusted for degrees of freedom for all the monitor sizes. Here, we select a parameter regarding audio and one concerning video. Since C_v is common to the two combinations and concerns video, we choose E_a as the other predictor variable.

Multiple regression analysis provides the following equations:

$$S_{80} = 1.197 \times 10 - 2.520 \times 10^{-2} E_a - 6.345 C_v \quad (4)$$

$$S_{60} = 5.460 + 1.150 \times 10^{-2} E_a - 6.220 C_v \quad (5)$$

$$S_{40} = 5.077 + 1.423 \times 10^{-2} E_a - 6.499 C_v \quad (6)$$

where $S_L (L = 40, 60, 80)$ is an estimate of the user-level QoS parameter when the monitor size is L inch. As shown in Tables VII through IX, the contribution rates adjusted for degrees of freedom of the obtained regression line are 0.957, 0.950 and 0.944 for 80, 60 and 40 inches of the monitor size, respectively.

We statistically test whether E_a and C_v make a significant contribution to the multiple regression line. The result of the statistical test [15] shows that the partial regression coefficient of C_v is statistically significant. However, E_a does not make any significant contribution to the multiple regression line. That is, only C_v is appropriate to the predictor variable of the regression line. Regression analysis with one predictor variable C_v gives the following equations:

$$S_{80} = 7.617 - 6.483 C_v \quad (7)$$

$$S_{60} = 7.449 - 6.157 C_v \quad (8)$$

$$S_{40} = 7.537 - 6.422 C_v \quad (9)$$

The contribution rates adjusted for degree of freedom of Eqs. (7), (8) and (9) become 0.964, 0.958 and 0.954, respectively.

Moreover, we used $1/C_v$ as the predictor variable instead of C_v . The result of multiple regression analysis is as follows:

$$S_{80} = -1.838 + 2.523 \cdot 1/C_v \quad (10)$$

$$S_{60} = -1.582 + 2.421 \cdot 1/C_v \quad (11)$$

$$S_{40} = -1.923 + 2.541 \cdot 1/C_v \quad (12)$$

The contribution rates adjusted for degree of freedom of Eqs. (10), (11) and (12) become 0.984, 0.991 and 0.992, respectively. These values are higher than those of Eqs. (7), (8) and (9), respectively. Therefore, we use Eqs. (10), (11) and (12) as a QoS mapping scheme.

In Eqs. (10) through (12), the values of intercept differ from each other, and this is also the case with those of regression coefficient of $1/C_v$. This makes us confirm that the monitor size affects user-level QoS. The regression coefficient of $\frac{1}{C_v}$ in Eq. (11) is smaller than those of Eqs. (10) and (12). That is, when the monitor size is 60 inch, the value of the user-level QoS parameter decreases more slowly than those for the other two sizes as C_v increases. We can consider two reasons for this. One is that a larger monitor size emphasizes the subjective degradation of video. The other is that a small monitor size decreases subjective quality of video. Consequently, there exists an appropriate monitor size, which may depend on the contents.

VI. CONCLUSIONS

We quantitatively assessed the user-level QoS of audio-video transmission over a ubiquitous IP network with the method of successive categories by experiment. As the first step of our research, we adopted large-sized monitors for evaluation. We assessed the user-level QoS of the audio-video transmission with three kinds of monitor sizes (80, 60 and 40 inches). We changed the monitor size by using

a projector and a screen. The experimental result showed that the user-level QoS parameter of the 80 inch monitor size takes the highest value when we can hardly notice the quality degradation of the audio-video transmission. On the other hand, we also found that the user-level QoS parameter value of the 80 inch monitor size becomes the lowest when the load traffic causes a certain level of degradation of user-level QoS. However, the result also indicated that we hardly recognize any difference in the user-level QoS between the monitor sizes when the degradation is too much. Moreover, we performed QoS mapping between user-level and application-level by multiple regression analysis. From the QoS mapping, we confirmed that the monitor size affects user-level QoS of the audio-video transmission. That is, the monitor size is one of important application-level QoS parameters which affect user-level QoS.

Some important issues are left as future work. First, in this paper, we treated rather large monitor sizes. In order to consider the effect of the monitor size in ubiquitous environments, we must treat smaller monitor sizes, which are used in PDAs or cellular phones. However, the method used in this paper is applicable to investigation of the case of small-sized monitors. Second, in our experiment, we used a single type of contents. However, the effect of the monitor size on user-level QoS may depend on the contents. Therefore, we will examine the subjective effect by various contents. Third, we will investigate a QoS control scheme which takes into consideration the effect of the monitor size on user-level QoS. By utilizing the effect, we will be able to control the QoS of audio-video transmission more effectively.

ACKNOWLEDGMENTS

The authors thank Yuji Tahara for his assistance in the experiment. This work was supported by the Telecommunications Advancement Foundation and by the Grant-In-Aid for Scientific Research of Japan Society for the Promotion of Science under Grant 17360179.

REFERENCES

- [1] M. Hasegawa, H. Morikawa, M. Inoue, U. Bandara, H. Murakami and K. Mahmud, "Cross-device handover using the service mobility proxy," *The 6th International Symposium on Wireless Personal Multimedia Communications (WPMC'03)*, vol.2, pp.1033-1037, Oct. 2003.
- [2] S. Tasaka and Y. Ishibashi, "Mutually compensatory property of multimedia QoS," *Conf. Rec. IEEE ICC2002*, pp.105-1111, Apr./May 2002.
- [3] T. Simmons and M. Manahan, "The effects of monitor size on user performance and preference," *Proc. Human Factors and Ergonomics Society - 1999*, pp.1593-1597, 1999.
- [4] W. IJsselstein, H. De Ridder, F. J. Freeman, S. E. Avons and D. Bouwhuis, "Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence," *Teleoperators and Virtual Environments*, vol.10, pp.298-311, 2001.
- [5] A. Walker and S. Brewster, "Sitting too close to the screen can be bad for your ears" : a study of audio-visual location discrepancy detection under different visual projections," *Proceedings of the 2001 International Conference on Auditory Display*, July/Aug., 2001.
- [6] R. Steinmetz, "Human perception of Jitter and media synchronization," *IEEE J.Select. Areas Commun.*, vol. 14, no. 1, pp. 61-72, Jan. 1996.
- [7] H. W. Frowein, G. F. Smoorenburg, L. Pyters and D. Schinkel. "Improved speech recognition through videotelephony: experiments with the hard of hearing," *IEEE J.Select. Areas Commun.*, vol. 9, no. 4, pp.611-616, May 1991.
- [8] K. Nakazono, "Frame rate as QoS parameter and its influence on speech perception," *Multimedia Systems*, vol. 6, pp. 359-366, 1998.
- [9] J. P. Guilford, *Psychometric methods*, McGraw-Hill, N. Y., 1954.
- [10] Y. Ito and S. Tasaka. Quantitative assessment of user-level QoS and its mapping. *IEEE Trans. Multimedia*, vol. 7, no. 3, pp.572-584, June 2005.
- [11] S. Tasaka and Y. Ito, "Psychometric analysis of the mutually compensatory property of multimedia QoS," *Conf. Rec. IEEE ICC2003*, pp. 1880-1886, May 2003.
- [12] W. S. Torgerson, *Theory and methods of scaling*, J. Wiley, N. Y., 1958.
- [13] 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, pp. 207-218, June 1951.
- [14] URL <http://www.isi.edu/nsnam/ns/>
- [15] M. G. Bulmer, *Principles of statistics*, Dover publications, Inc., N. Y., 1979.