PAPER Special Section on Multimedia QoS Evaluation and Management Technologies

Opinion Model Using Psychological Factors for Interactive Multimodal Services

Kazuhisa YAMAGISHI[†a)] ***and*** **Takanori HAYASHI**[†], ***Members***

SUMMARY We propose the concept of an opinion model for interactive multimodal services and apply it to an audiovisual communication service. First, psychological factors of an audiovisual communication service were extracted by using the semantic differential (SD) technique and factor analysis. Forty subjects participated in subjective tests and performed point-to-point conversational tasks on a PC-based video phone that exhibited various network qualities. The subjects assessed those qualities on the basis of 25 pairs of adjectives. Two psychological factors, i.e., an aesthetic feeling and a feeling of activity, were extracted from the results. Then, quality impairment factors affecting these two psychological factors were analyzed. We found that the aesthetic feeling was affected by IP packet loss and video coding bit rate, and the feeling of activity depended on delay time, video packet loss, video coding bit rate, and video frame rate. Using this result, we formulated an opinion model derived from the relationships among quality impairment factors, psychological factors, and overall quality. The validation test results indicated that the estimation error of our model was almost equivalent to the statistical reliability of the subjective score.

key words: *quality of service, subjective assessment, semantic differential technique, factor analysis*

1. Introduction

With the advances in broadband IP technology, interactive multimodal services using audio and video communication such as instant messaging and teleconferencing have become popular as promising rich multimedia applications. Because quality is not generally guaranteed in an IP network, methodologies for evaluating the quality of such rich services are indispensable.

The prime criterion for the quality of communication services is the user's perceptual evaluation of service quality. This can be measured through subjective quality assessment. The most widely used metric is the mean opinion score (MOS). In order to assess the quality of rich services, we need to compare services not only on a one-dimensional scale, like the MOS, but also on a multidimensional scale that characterizes the QoS in a way that considers psychological factors for the richness of the services. Moreover, for network quality planning and management, it is important to establish an opinion model [1] that estimates the overall audiovisual communication quality taking into account interactivity and multimodality.

In this paper, we propose the concept of an opinion model expressed on a multidimensional quality scale for interactive multimodal services and apply it to an audiovisual communication service. Using our concept, we evaluate psychological factors and express the overall quality using psychological factors even if the overall quality is the same. That is, we can properly design applications and networks taking into account these psychological factors.

First, we analyze and determine the relationship between psychological factors that represent the richness and quality impairment factors for an assessment of interactive multimodal services. We conducted subjective quality experiments based on the semantic differential (SD) technique [2] to investigate the relationship. The SD technique can measure sensory impressions for an individual. As a result, we found that two psychological factors, i.e., an aesthetic feeling and a feeling of activity, were extracted. We showed that the aesthetic feeling was affected by IP packet loss and video coding bit rate, and the feeling of activity depended on delay time, video packet loss, video coding bit rate, and video frame rate. In addition, we formulate an opinion model using the psychological factors. Then, we demonstrate that overall audiovisual quality as it affects communication services can be expressed by two psychological factors and estimated with sufficient accuracy. Finally, the validation test results indicated that the estimation error of our model was equivalent to the statistical reliability of the subjective score.

There have been some studies of psychological factors that affect high-quality audio [3] and video [4], [5]. These studies have focused on single-modal, i.e., audio or video, services. Little is known about psychological factors for multimodal services. The psychological impressions of one-way audiovisual perception have been studied [6], but those for an audiovisual communication service have not. An opinion model for an audiovisual communication service has never been studied.

The remainder of this paper is structured as follows. We propose the concept of an opinion model for interactive multimodal services in Sect. 2. A method of performing subjective quality experiments is described in Sect. 3. Psychological factors in an audiovisual communication service are extracted using the SD technique and we determine the relationship between these psychological factors and quality impairment factors in Sect. 4. We apply the concept to an audiovisual communication service by formulating factor scores with quality impairment factors in Sect. 5. Finally, in Sect. 6, we summarize our findings and suggest possible

Manuscript received April 28, 2005.
Manuscript revised August 9, 2005.
[†]The authors are with NTT Service Integration Laboratories, NTT Corporation, Musashino-shi, 180-8585 Japan.
a) E-mail: yamagishi.kazuhisa@lab.ntt.co.jp
DOI: 10.1093/ietcom/e89–b.2.281

directions for future studies.

2. Proposal of Opinion Model for Interactive Multimodal Services

2.1 Conventional Opinion Model

A computational model that estimates the overall communication quality using a combination of quality factors is called the "opinion model." The "E-model" [7] standardized as ITU-T Recommendation G.107 is a typical example of the opinion model for telephone services. The E-model is based on the concept that the amount of psychological degradation can be accumulated on a psychological scale.

The E-model takes as inputs 20 parameters that represent the terminal, network, and environmental quality factors. Its output is called the *R-value*, which is a psychological scale that acts as an index of overall quality.

Because this model assumes only conventional telephone services, however, it can not evaluate the overall audiovisual communication quality taking into account images. In addition, it can not represent the characteristics of high-quality services determined by the contribution of individual psychological factors that form the overall audiovisual quality. Therefore, we introduce psychological factors into the model.

2.2 Concept of Proposed Model

Figure 1 shows the concept of the opinion model for an audiovisual communication service targeted in this study. The model first takes as input the physical features such as codec and coding bit rate in accordance with the system requirements and service tasks of audiovisual communication service. The features are quality factors that represent the terminal, network, and environmental characteristics, and some of them are the same as input parameters to the E-model. Using the physical features, it then computes the perceptual quality index (e.g., noisy is computed from the video codec, coding bit rate, frame rate, packet loss, and so on.) Finally, it computes psychological factor indices using perceptual quality indices (e.g., aesthetic feeling is computed from noisy, blurry, and so on). Each psychological factor index expresses how users are impressed by the systems and/or tasks that they can perform. Finally, it outputs an overall quality index using these psychological factor indices.

Since the quality evaluation score is affected by the interactivity of communication services, the opinion model should be adapted for different types of interactivity. We propose adaptively optimizing our computational models according to the communication purpose such as chatting or holding a business meeting, when calculating the psychological indices and the overall quality index. Since the model can describe the characteristics of the service from various psychological viewpoints, a service planner can, for example, design various service qualities while keeping a constant overall quality taking into account the intentions of the users; he/she can understand in detail the strong and/or weak points, which are based on the psychological human feelings, of the communication services.

In the following sections, we formulate an opinion

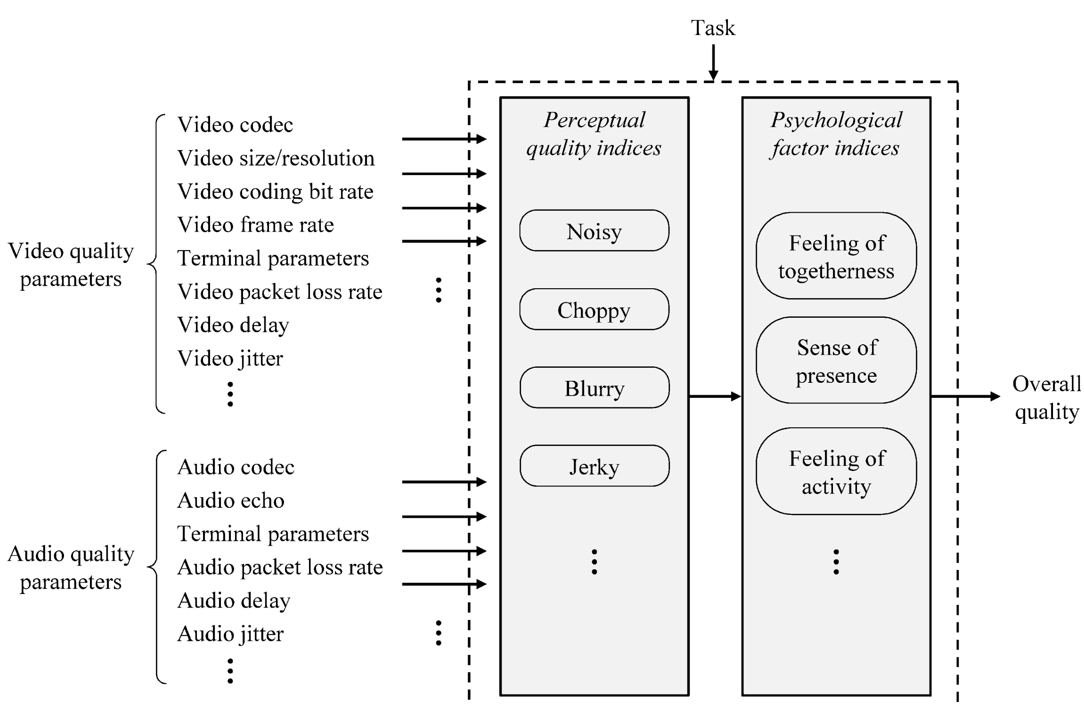


Fig. 1 Concept of opinion model for interactive multimodal services.

model for an audiovisual communication service applying this concept and validate the model using verification data.

3. Subjective Quality Experiments

Subjective quality experiments were conducted using a point-to-point audiovisual communication service to determine the relationship between psychological factors and quality impairment factors and that between psychological factors and overall audiovisual quality.

3.1 Experimental Setting

We built a system for a point-to-point audiovisual communication service using personal computers. G.722 and MPEG4 were used for audio and video coding, respectively. Packet loss concealment algorithms were not used for audio and video. A 10-inch image of the conversational partners was displayed on a 17-inch PC monitor. Microphones and headphones were used for audio input and output, respectively. The ambient noise was set to about 35 dB (A), the background room illumination to about 500 lux, and the viewing distance to 50–80 cm.

The experimental parameters were video coding bit rate, video frame rate, video packet loss rate, audio packet loss rates, and one-way absolute transmission delay (audio and video were synchronized), as shown in Table 1. The number of test conditions, which are the combinations of parameters in Table 1, was 46.

Table 1 Experimental settings.

Video bit rate (kbps)	640, 1024, 3072, 4096, NTSC (not coded)
Video frame rate (fps)	2, 5, 10, 30
Video packet loss rate (%)	0.0, 0.2, 1.0, 2.0
Audio packet loss rate (%)	0.0, 1.0, 6.0, 10.0
Delay (ms)	66, 165, 600, 1000, 1500

3.2 Task

It is known that the quality evaluation score in a conversational experiment is affected by the conversation task. In this investigation, we thought that a free conversation was important as the first target as well as E-model, and used the Name-Guessing Task defined in ITU-T Recommendation P.920 [8] in which subjects were required to guess the name of an object, such as furniture or food and freely talk about the question. The duration of a conversation was 60 s per condition. The interactivity of conversation in this task was similar to that of a free conversation.

3.3 Rating Scales

For subjective quality assessment, overall audiovisual quality was first evaluated using a five-grade quality scale (Table 2). Then, using the SD technique proposed by Osgood [2], the subject's impression of an audiovisual communication service was evaluated on the basis of 25 pairs of bipolar adjectives on a seven-grade comparison scale (Table 3). Rating scales and adjectives are given in Japanese.

The following procedures were conducted for selecting adjective pairs. First, 45 adjective pairs were selected by 7 experts who were directly concerned with multimedia quality assessment as part of their work. Next, by conducting preliminary experiments we selected 25 adjective pairs that subjects can evaluate easily.

Table 2 Five-grade quality scale.

Score	Quality scale (in Japanese)
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

Table 3 Example of seven-grade comparison scale.

Score	Comparison (in Japanese)	Adjective pairs (in Japanese)
3	Extremely	Fast
2	Quite	
1	Slightly	
0	Neutral	
−1	Slightly	Slow
−2	Quite	
−3	Extremely	

We added margins of +3 to −3 to the scores of each comparison scale obtained by the subjective quality experiment to express positive and negative impressions, respectively.

3.4 Subjects

Forty subjects, who had the ability to use the Internet, aged 20 to 39, participated in the experiment. They were nonexperts not directly concerned with multimedia quality as part of their work, and therefore not experienced assessors.

4. Analysis of Psychological Factors

Applying factor analysis to the subjective experimental data, we extracted two psychological factors. We then determined the relationship between psychological factors and quality impairment factors.

4.1 Psychological Factors

First, we verified that the scores of adjective pairs for each condition have a standard normal distribution at a 5% significance level using the D'Agostino-Pearson test. This result indicated the score of the adjective pair can be expressed by the mean value.

We applied a factor analysis using Varimax rotation to scores of adjective pairs and obtained two psychological factors. As a result, psychological factors were divided as shown in Table 4.

Table 4 Factor loading.

Aesthetic feeling (contribution: 54.3%)			
Positive	Negative	1st factor	2nd factor
quiet	clamorous	0.929	0.242
clear	cloudy	0.905	0.369
orderly	disturbed	0.880	0.451
delicate	rugged	0.872	0.383
beautiful	dirty	0.866	0.401
pleasant	unpleasant	0.839	0.524
distinct	dull	0.824	0.398
regular	irregular	0.803	0.552
relieved	uneasy	0.800	0.539
calm	restless	0.797	0.563
stable	unstable	0.785	0.584
rich	poor	0.774	0.576
unenclosed	enclosed	0.763	0.571
natural	unnatural	0.755	0.610
sharp	fuzzy	0.748	0.430
familiar	strange	0.724	0.660
tight	loose	0.722	0.579
warm	cold	0.702	0.499
Feeling of activity (contribution: 36.5%)			
Positive	Negative	1st factor	2nd factor
dynamic	static	0.215	0.899
fast	slow	0.375	0.886
light	heavy	0.434	0.832
strong	weak	0.550	0.802
fluid	erratic	0.590	0.735
soft	hard	0.621	0.731
smooth	rough	0.658	0.703

The first factor, with a 54.3% contribution, can be interpreted as an aesthetic feeling because adjective pairs such as quiet-clamorous, clear-cloudy, delicate-rugged and distinct-dull had high loadings. The second factor, with a 36.5% contribution, can be interpreted as a feeling of activity because adjective pairs such as dynamic-static and fast-slow had high loadings. The name of each psychological factor was determined based on these adjective pairs. Because these two factors account for 90.8% of the psychological impression of an audiovisual communication service, they are extremely important factors for assessing service quality.

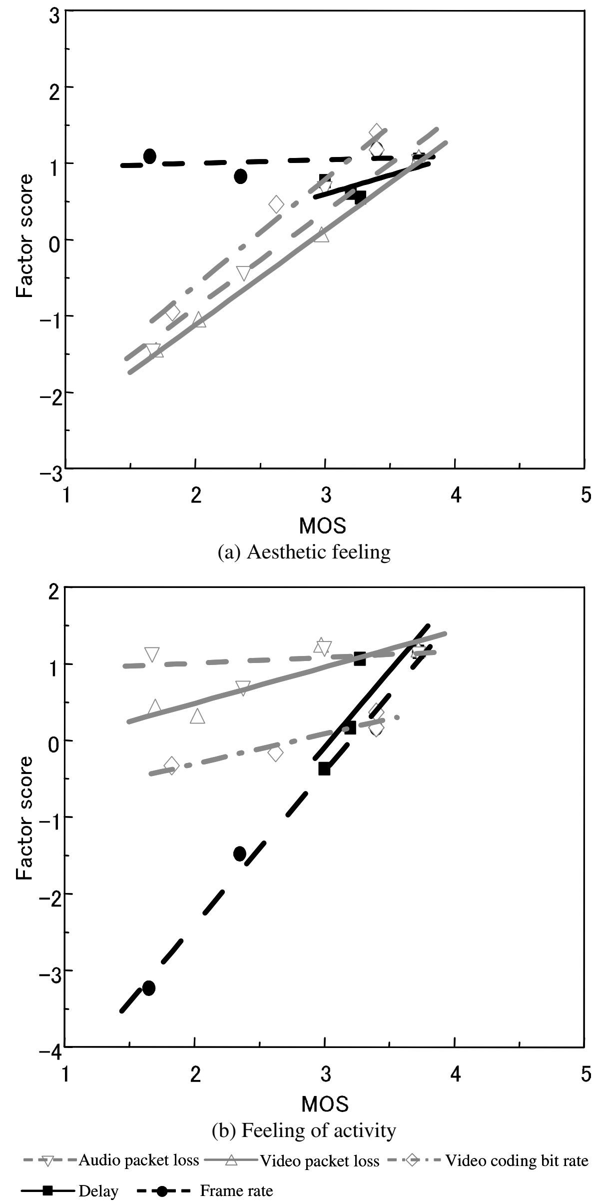


Fig. 2 Relationship between psychological factors and factor scores.

4.2 Relationship between Psychological Factors and Quality Impairment Factors

We assumed that the two psychological factors are independent and examined the relationship between these factors and quality impairment factors. The two most popular methods for estimating factor scores are those of Thomson's regression method [9] and Bartlett's weighted least squares method. Here we use Thomson's regression method:

$$f_{ik} = \sum_{j=1}^{n} a_{ij} x_{jk}, \tag{1}$$

where f_{ik} is a factor score, a_{ij} is the transposed matrix of the factor loads of an adjective pair, x_{jk} represents subjective experimental data standardized so that the mean value is set to 0, the standard deviation is set to 1, i specifies a factor, j specifies an adjective, k is the number of conditions, and n is the number of adjectives.

The overall quality and factor scores plotted against each quality impairment factor are shown in Fig. 2. In Fig. 2, calculating the slope and correlation of all lines, aesthetic feeling was not affected by frame rate and delay because the slope of frame rate was small, delay was not linearly correlated, and feeling of activity was not affected by audio packet loss because the slope of audio packet loss was small. In an aesthetic feeling, factor scores of video packet loss, audio packet loss, and video coding bit rate decrease monotonically as the MOS decreases, and in a feeling of activity, factor scores of frame rate, video packet loss, video coding bit rate, and transmission delay decrease monotonically as the MOS decreases. We found that aesthetic feeling was affected by IP packet loss and video coding bit rate, and the

Table 5 Relationship between psychological factors and quality impairment factors.

Psychological factor	Main quality impairment factors
Aesthetic feeling	Audio packet loss, video packet loss, and video bit rate
Feeling of activity	Delay, frame rate, video bit rate, and video packet loss

feeling of activity depends on delay time, video packet loss, video coding bit rate, and video frame rate. The relationship between psychological factors and the quality impairment factors is shown in Table 5.

5. Factor Score Estimation for Each Psychological Factor

In Sect. 4, we showed that an audiovisual communication service could be described by two psychological factors, which are users' impressions of such services. This section describes the determination of the relationship between overall quality and psychological factors, which is necessary to establish the opinion model for an audiovisual communication service, taking into account psychological aspects. First, we show that overall quality can be expressed based on two psychological factors. Then, we propose an opinion model derived from the relationship among quality impairment factors, psychological factors, and overall quality.

5.1 Overall Quality Estimation Using the Psychological Factors

We assume that the MOS, which expresses overall quality, can be formulated by a linear combination of the two psychological factor scores of an aesthetic feeling (S_1) and a feeling of activity (S_2). Applying multiple regression analysis [10], we obtained the following regression equation with $S_k(k = 1, 2)$ as explanatory variables and the MOS as a criterion variable.

$$MOS = 2.45 + 0.61S_1 + 0.35S_2 \tag{2}$$

The coefficient of determination (R^2), which indicates the explanation rate of the variation in the explanatory variables, was 0.949, indicating that a linear combination of these two factor scores can explain the overall quality. The analysis of variance (ANOVA) result also supports this conclusion. Partial regression coefficients were significant at the 1% level, so $S_k(k = 1, 2)$ and the constant term were all indispensable terms in Eq. (2). The relationship between the subjective MOS and the estimated MOS obtained by using the model shown in Eq. (2) is shown in Fig. 3. The figure indicates that the proposed model estimated the overall quality with sufficient accuracy.

5.2 Factor Score Estimation Using the Quality Impairment Factors

Factor score estimation is performed as follows. We estimate factor scores of an aesthetic feeling (S_1) and a feeling of activity (S_2) from quality impairment factors. First, we estimate the factor scores from each quality impairment factor shown in Step 1. The equations in Step 1 mean that each quality impairment factor affects aesthetic feeling or feeling of activity. Second, we estimate the factor scores of audio, video, and delay using multiple regression analysis, where the criterion variable is the factor score and the explanatory variables are the equations obtained from Step 1 and their linear interaction shown in Step 2. The equations in Step 2 mean that audio, video, and delay affect aesthetic feeling or feeling of activity. Third, we estimate the factor score using multiple linear regression analysis, where the criterion variable is factor score and the explanatory variables are the equations of each medium obtained from Step 2 and their linear interaction shown in Step 3. The equation in Step 3 means that aesthetic feeling and feeling of activity are affected by quality impairment factors that are relevant to each feeling. In Steps 2 and 3, we chose the terms that satisfy the 1% significance level.

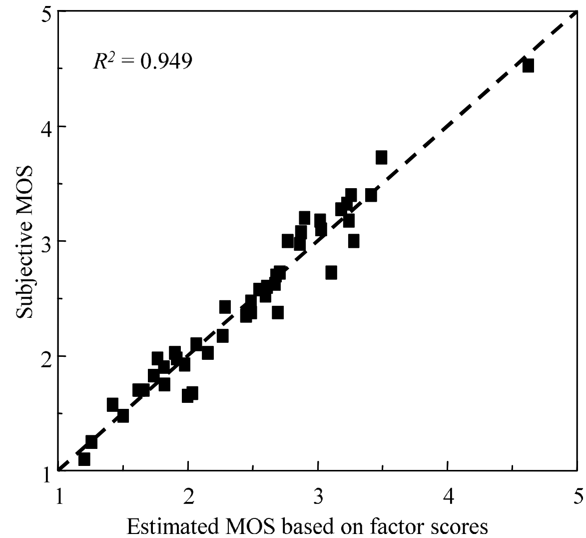


Fig. 3 Estimated overall quality based on factor scores.

In our experimental analysis, first, we estimated the factor score of aesthetic feeling (S_1) applying the above steps to quality impairment factors (audio packet loss Pa, video packet loss Pv, and video coding bit rate Br) which are relevant to an aesthetic feeling.

Step 1

$$V_{11}(Br) = a_{11} \exp(-Br/b_{11}) + c_{11}, \tag{3}$$
$$V_{12}(Pv) = a_{12} \exp(-Pv/b_{12}) + c_{12}, \tag{4}$$
$$A_{11}(Pa) = a_{13} \exp(-Pa/b_{13}) + c_{13}, \tag{5}$$

Step 2

$$V_1(Br, Pv) = a_{10}V_{11}(Br) + b_{10}V_{12}(Pv) + c_{10}, \tag{6}$$
$$A_1(Pa) = A_{11}(Pa), \tag{7}$$

Step 3

$$S_1(Br, Pv, Pa) = S_1(V_1, A_1)$$
$$= \alpha_1 V_1(Br, Pv) + \beta_1 A_1(Pa)$$

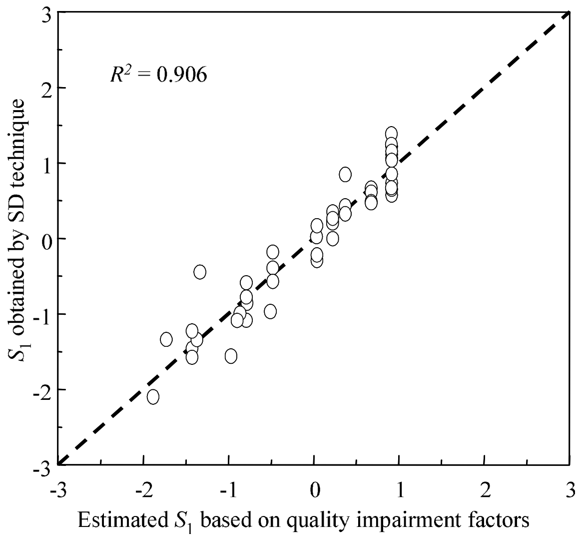


Fig. 4 Estimated factor score of an aesthetic feeling.

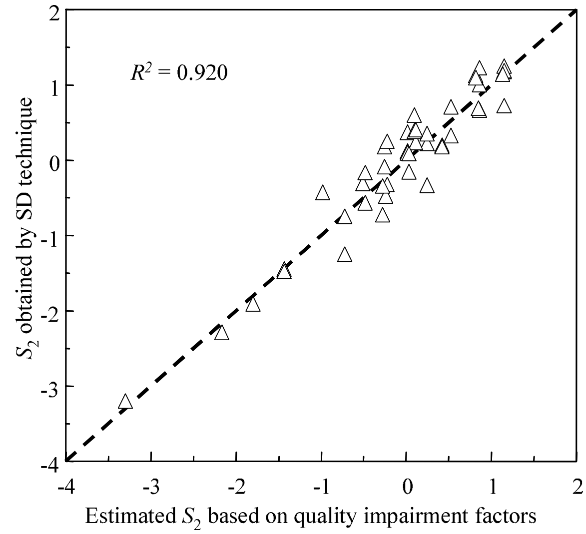


Fig. 5 Estimated factor score of a feeling of activity.

$$+ \gamma_1 V_1(Br, Pv)A_1(Pa) + \delta_1, \quad (8)$$

where $a_{1m}, b_{1m}, c_{1m} (m = 0, \ldots, 3), \alpha_1, \beta_1, \gamma_1$, and δ_1 are constants and positive real numbers, and $V_1(Br, Pv)$ means a video quality parameter that affects the aesthetic feeling.

Second, we estimated the factor score of a feeling of activity (S_2) from quality impairment factors (video packet loss Pv, video coding bit rate Br, frame rate Fr, and delay D), which are relevant to a feeling of activity, using the above steps.

Step 1

$$V_{21}(Br) = a_{21} \exp(-Br/b_{21}) + c_{21}, \quad (9)$$
$$V_{22}(Pv) = a_{22} \exp(-Pv/b_{22}) + c_{22}, \quad (10)$$
$$V_{23}(Fr) = a_{23} \exp(-Fr/b_{23}) + c_{23}, \quad (11)$$
$$D_{21}(D) = a_{24} \exp(-D/b_{24}) + c_{24}, \quad (12)$$

Step 2

$$V_2(Br, Pv, Fr) = a_{20}V_{21}(Br)V_{22}(Pv) + b_{20}V_{22}(Pv)V_{23}(Fr) + c_{20}, \quad (13)$$
$$D_2(D) = D_{21}(D), \quad (14)$$

Step 3

$$\begin{aligned} S_2(D, Fr, Br, Pv) &= S_2(V_2, D_2) \\ &= \alpha_2 V_2(Fr, Br, Pv) \\ &\quad + \beta_2 D_2(D) + \gamma_2, \end{aligned} \quad (15)$$

where $a_{2n}, b_{2n}, c_{2n} (n = 0, \ldots, 4), \alpha_2, \beta_2$, and γ_2 are constants and positive real numbers, and $V_2(Br, Pv, Fr)$ means a video quality parameter that affects the feeling of activity.

The estimated factor scores of on aesthetic feeing and a feeling of activity are shown in Figs. 4 and 5 where the coefficients of determination were 0.906 and 0.920, respectively. The factor scores calculated using Eqs. (8) and (15) were at the 1% significance level using ANOVA. Partial regression coefficients were significant at the 1% level.

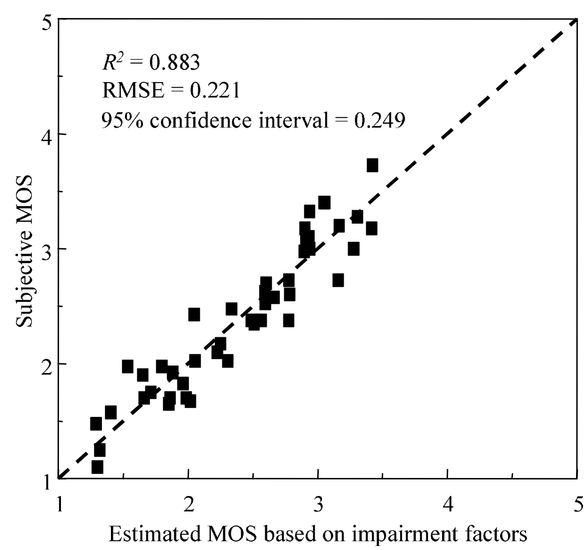


Fig. 6 Relationship between subjective MOS and estimated MOS (training data).

When free conversation is the communication mode in an audiovisual communication service, factor scores of an aesthetic feeling and a feeling of activity can be expressed based on quality impairment factors. The feeling of activity would be affected increasingly by delay, in such a case, because the degradation received from the delay increases when a highly interactive conversation is conducted. Verification of other tasks, such as highly interactive conversation, is for further study.

5.3 Overall Quality Estimation Using the Quality Impairment Factors

Substituting Eqs. (8) and (15) into Eq. (2), we can estimate overall quality (MOS) from quality impairment factors. The results are shown in Fig. 6. The coefficient of determination

Table 6 Experimental settings for verification test.

Video bit rate (kbps)	640, 768, 1024, 1536, 2048, 3072, 4096, NTSC (not coded)
Video frame rate (fps)	2, 5, 10, 15, 30
Video packet loss rate (%)	0.0, 0.1, 0.2, 0.5, 1.0, 1.5, 2.0
Audio packet loss rate (%)	0.0, 0.5, 1.0, 2.5, 6.0, 7.5, 10.0
Delay (ms)	66, 165, 363, 600, 825, 1000, 1188, 1500

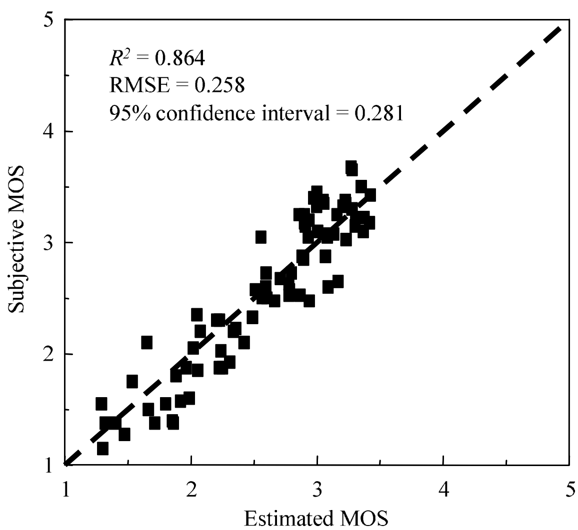


Fig. 7 Relationship between subjective MOS and estimated MOS (verification data).

was 0.883, the 95% confidence interval for the subjective MOS was 0.249, and the root mean square error (RMSE) was 0.221. Since the estimation error of the proposed model is almost equivalent to the statistical reliability of the subjective score, we can say that overall audiovisual quality can be estimated from quality impairment factors.

5.4 Validity of the Proposed Model

We validate the proposed model using verification data. Table 6 shows the verification test conditions. The test evaluated 82 conditions, which were combinations of the parameters in Table 6. The coefficients of the model used in the test were the same as described in Sect. 3. Task and rating scales used in the test were the same as determined in Sect. 3. We used forty subjects who were different from those in Sect. 3.4.

The estimation accuracy of the model is demonstrated in Fig. 7. Since the cross-correlation coefficient was 0.864, the 95% confidence interval for the subjective MOS was 0.281 and the RMSE was 0.258, overall audiovisual quality can be estimated by the proposed model with sufficient accuracy.

6. Conclusion

We proposed the concept of an opinion model that takes into account psychological factors for interactive multimodal services. We then applied the concept to an audiovisual communication service and derived an opinion model that estimates the overall audiovisual quality.

First, we found that the major psychological factors in an audiovisual communication service are an aesthetic feeling and a feeling of activity. The relationship between these psychological factors and quality impairment factors was also determined. The aesthetic feeling is affected by audio packet loss, video packet loss, and video coding bit rate. The feeling of activity depends on frame rate, video packet loss, video coding bit rate, and transmission delay.

Then, we showed that overall audiovisual quality can be estimated based on these two psychological factors. After that, we formulated an opinion model for an audiovisual communication service using factor scores calculated from quality impairment. Finally, we validated the model based on subjective testing. The relationship between psychological factor and MOS when packet loss concealment algorithms are used is not verified by conducting an experiment. Verification of these conditions is for further study.

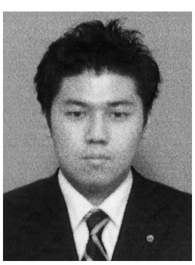
We believe that our study provides a basis for an opinion model such as the E-model for interactive multimodal services. The proposed model can be applied to design applications and networks taking into account psychological factors for various service scenarios. For instance, we can maintain a high level of aesthetic feeling for Internet shopping by controlling the video coding bit rate and audio/video packet loss. For business conferences, transmission delay, video packet loss, video coding bit rate, and video frame rate should be set appropriately to improve the feeling of activity.

The following issues call for further study. If we used packet loss concealment, the other video codec or conversational task, the relationship between feeling and them is different because quality degradation is different from above results. Because our model is applicable only to the same environment, we have to prepare the coefficient tables for each environment or change our model a little. The conventional targets of quality assessment have been point-to-point communications. With the advent of high-speed wired and wireless networks, multiparty communication services, such as instant messaging, teleconferencing, and distributed collaboration services, are being deployed. Such services are multipoint (users are geographically dispersed), real-time, and interactive. A lot of problems remain for the study of psychological factors in analyzing such services. The important issues in evaluating the quality of multiparty services are the heterogeneous communication environments of the users and synchronization.

References

[1] A. Takahashi, H. Yoshino, and N. Kitawaki, "Perceptual QoS assessment technologies for VoIP," IEEE Commun. Mag., pp.28–34, July 2004.
[2] C.E. Osgood, G.J. Suci, and P.H. Tannenbaum, The Measurement of Meaning, University of Illinois Press, 1957.
[3] H. Aoki and A. Takahashi, "Analysis of relationship between over-

all quality and psychological factors affecting high-quality speech communication services," ICASSP, pp.I-57–I-60, March 2005.
[4] T. Tachi, S. Iai, and N. Kitawaki, "Proposal of selection method of test pictures in HDTV subjective quality assessment," IEEE Multimedia'92, pp.67–68, 1992.
[5] O. Kitamura, S. Namba, and P. Matsumoto, "Factor analysis research of tone color," Proc. 6th International Congress on Acoustics, pp.A-5–A-11, 1968.
[6] Y. Ito and S. Tasaka, "Multidimensional assessment of user-level QoS for audio-video transmission over IP networks," IEICE Trans. Commun. (Japanese Edition), vol.J88-B no.3, pp.689–702, March 2005.
[7] ITU-T Recommendation G.107, "The E-model, a computational model for use in transmission planning," 2000.
[8] ITU-T Recommendation P.920, "Interactive test methods for audiovisual communications," 2000.
[9] G.H. Thomson, The Factorial Analysis of Human Ability, London University Press, 1951.
[10] R.M. Mickey, O.J. Dunn, and V.A. Clark, Analysis of Variance and Regression, 3rd ed., Wiley-Interscience, 2004.



Kazuhisa Yamagishi received the M.E. degree in electronics, information and communication engineering from Waseda University in Japan and joined NTT Laboratories in 2003. He has been engaged in subjective quality assessment of multimedia telecommunications and image coding. Currently, he is working on the quality assessment of interactive multimodal service over IP networks.

Takanori Hayashi received the M.E. degree in physical engineering from Tsukuba University in Japan and joined NTT Laboratories in 1990. He has been engaged in subjective quality assessment of multimedia telecommunications and network performance measurement methods. Currently, he is working on a multimodal quality assessment method over IP networks.