PAPER *Special Section on Multimedia QoS Evaluation and Management Technologies*

Subjective Quality Estimation Model for Video Streaming Services with Dynamic Bit-Rate Control

Takanori HAYASHI[†a)], **Ginga KAWAGUTI**[†*], **Jun OKAMOTO**[†], *and* **Akira TAKAHASHI**[†], *Members*

SUMMARY This paper proposes a subjective model for estimating the quality of video streaming services with dynamic bit-rate control. In a subjective quality assessment test, we clarify users' perceptions of distributed video signals whose quality is time-variant due to dynamic bit-rate control. Using this result, we constructed an estimation model considering the following three characteristics: 1) the influence of the video section where quality degradation is large will strongly affect the overall quality, 2) the impression of a past quality weakens with the passage of time, and 3) the range of evaluation scores becomes wider when the time duration of an evaluation is longer. We found that the proposed model enables the accuracy of estimating overall subjective quality to be dramatically improved compared with that of a model that averages segmental quality. The estimation error of the proposed model is less than the statistical reliability of the subjective score even for verification data. We also show that our findings are applicable to QoS design/management issues for video streaming services with dynamic bit-rate control.

key words: *video streaming, dynamic bit-rate control, QoS, quality estimation, subjective assessment*

1. Introduction

With the bandwidth of access lines to the Internet becoming broader, various video distribution services are being developed and provided. To maintain customer satisfaction at high levels, it is becoming more important to offer high-quality video distribution services by managing the quality of service (QoS).

The methods of distributing video on IP networks are classified into "downloading" and "streaming." The former transmits a video file from a server to a user terminal and plays the video after the data transmission has been completed. The latter plays the video while the video data is downloading. Since the quality of IP networks is not necessarily guaranteed, degradation in network quality, such as that caused by IP packet losses, IP packet transmission delays, and delay variations, is caused. This network quality degradation negatively affects the perceived QoS of video distribution services. For example, the downloading time, i.e., the waiting time, increases in the former, and video quality deteriorates due to block distortion and freezes in the latter.

This paper focuses on subjective quality for video streaming services because the video quality is independent of network performance in video downloading services. This is also because that there have been many studies about the relationship between network performance and the waiting time, i.e., TCP throughput. Some approximation formulas for computing TCP throughput from the packet loss ratio and delay (round trip time: RTT) have been proposed [1], [2].

Video streaming services generally implement a mechanism for dynamically changing their bit rates to offer real-time services using limited network and/or terminal resources. A terminal sends information about the amount of received/lost data, the time variation of data arrival, and the amount of terminal data buffering to a server, so the server can evaluate the degree of network congestion from this information. If a server judges the network to be congested, it decreases the video coding rate of the transmitted video. Serious quality degradation due to packet loss, such as block distortion and freezes, can be avoided by this mechanism although video quality slightly deteriorates when the coding rate becomes lower. However, it is not clear how the user perceives video quality that is time variant due to dynamic bit-rate control and what QoS management methods are suitable for grasping the quality that each user perceives.

Although subjective quality assessments usually examine short-term video sequences [3], we study on video quality of long-term sequences with dynamic quality change. There have been some related studies on temporal variations of video quality assessed by the single stimulus continuous quality evaluation (SSCQE) method [3]. However, these studies have mainly focused on how well the time-variant quality diagram of a video sequence could be evaluated [4], [5]. Little is known about overall quality of long-duration video sequences [6]. However, there have not been studies of what parameters are effective for managing video QoS.

In this paper, we propose an estimation model of subjective quality for video streaming services with dynamic bit-rate control. The video sequence to be assessed was three minutes long. The proposed model for estimating subjective overall video quality is discussed considering short-term qualities of 15-second inputs because the SSCQE method to the range of bit-rate and kind of applications is still under investigation [3].

First, we clarify the subjective quality estimation characteristics of distributed video signals whose quality is time-variant due to dynamic bit-rate control. Second, using this result, we construct an estimation model considering the fol-

Manuscript received April 27, 2005.
Manuscript revised August 12, 2005.
[†]The authors are with NTT Service Integration Laboratories, NTT Corporation, Musashino-shi, 180-8585 Japan.
[*]Presently, with Innovative IP Architecture Center, NTT Communications Corporation, Japan.
a) E-mail: t.hayashi@lab.ntt.co.jp
DOI: 10.1093/ietcom/e89–b.2.297

lowing three characteristics: 1) the influence of the video section where quality degradation is large will strongly affect the overall quality, 2) the impression of a past quality weakens with the passage of time, and 3) the range of evaluation scores becomes wider when the time duration of an evaluation is longer. Third, we verify the accuracy of the proposed model using the verification test results. The estimation error of the proposed model is less than the statistical reliability of the subjective score. Finally, we apply our findings to QoS design/management issues for video streaming services with dynamic bit-rate control.

The remainder of this paper is structured as follows. The factors affecting users' perceptions are considered based on a subjective quality experiment in Sect. 2. An estimation model of subjective video quality considering the above subjective quality characteristics is proposed and the accuracy of the model is discussed in Sect. 3. The proposed model is applied to the verification data and its validity is clarified in Sect. 4. Quality design/management issues for video streaming services with dynamic bit-rate control are discussed in Sect. 5. Finally, in Sect. 6, we summarize our findings.

2. Factors Affecting Subjective Video Quality

In video streaming services with dynamic bit-rate control, the video coding rate, i.e., video quality, is dynamically changed according to network congestion. It is well known that the influence of the video section where quality degradation is large will strongly affect the overall quality [7]. Therefore, the overall video quality cannot be expressed by averaging the short-term quality. It is also well known that the impression of a past quality weakens with the passage of time [6], [8]. In this section, we quantify the above factors affecting the subjective video quality when it is changed according to various bit-rate alteration patterns.

2.1 Subjective Video Quality Assessment

In a subjective test, a music video was prepared as a source sequence. The sequence was three minutes long although subjective tests usually examine sequences that are about 10 seconds long [3]. This is because: 1) a test sequence has a changing bit-rate with a cycle of several seconds or more, 2) a certain amount of time (on the order of a minute) is required for subjects to understand the meaning of video content, and 3) a test sequence should include various scenes and movement in the scenes.

The test sequences were produced as follows. First, a source sequence was coded at 2 Mbit/s, 768 kbit/s, and 256 kbit/s using an MPEG4-based codec. Then, each test sequence was played on a PC monitor, converted to an HDTV signal, and recorded on a VTR. Finally, video signals in the VTR were edited as the test sequences whose bit-rate was dynamically changed. In a test, edited signals were reconverted and played on a PC monitor. Note that the sound quality was not degraded in any of the test sequences.

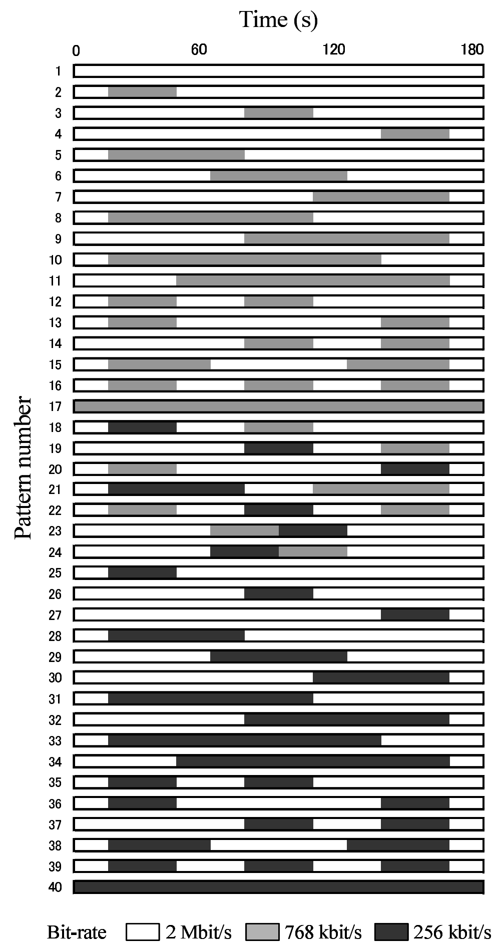


Fig. 1 Patterns of bit-rate alteration.

Table 1 Five-grade quality scale.

Score	Rating scale
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

Forty bit-rate alteration patterns consisting of three quality-level sections are shown in Fig. 1. White, gray, and black sections denote video sequences at 2 Mbit/s, 768 kbit/s, and 256 kbit/s, respectively. The video quality at the rate of 2 Mbit/s is almost equivalent to the original quality, so sections with bit-rates other than 2 Mbit/s are called "quality-degraded sections" in this paper.

The absolute category rating (ACR) method [9] was used in the subjective assessment test. Subjects watched a video sequence for three minutes and evaluated the video quality using the five-grade quality rating scale [9] shown in Table 1. The subjective quality for every 15 seconds (12 sections in the 3-minute test sequence) was also evaluated using the same method. Moreover, acceptability of video quality for three minutes was evaluated using a binary acceptance scale: yes or no. The subjective quality was represented as a mean opinion score (MOS) calculated by averaging the scores of 40 subjects, 20 males and 20 females, ranging in age from 20 to 35. The subjects were non-experts not directly concerned with video quality as part of their work, and were therefore not experienced assessors. All subjects had experience with Internet services via a broadband network. The PC monitor size was 17 inches and the video to be assessed was displayed in the full-screen mode. The ratio of viewing distance to picture height was four (about 90 cm).

2.2 Experimental Results

The effects of the percentage (P) of quality-degraded sections at the same bit-rate (768 kbit/s or 256 kbit/s) in three minutes are shown in Fig. 2 in terms of the MOS. The results where the positions of a quality-degraded section in the video differed were averaged, and there was one quality-degraded section in all cases. The MOS decreased monotonically with increasing P and was a convex function. This result indicates that the quality-degraded section strongly affected overall quality.

The effects of the position of a quality-degraded section are shown in Fig. 3. The quality of patterns 4, 7, 27, and 30 was a little poorer than that of patterns 2, 5, 25, and 28, respectively. The quality-degraded sections of the latter are located earlier than that of the former. This phenomenon is the *recency effect* described in [8]. There was no statistically significant difference caused by the number of quality-degraded sections.

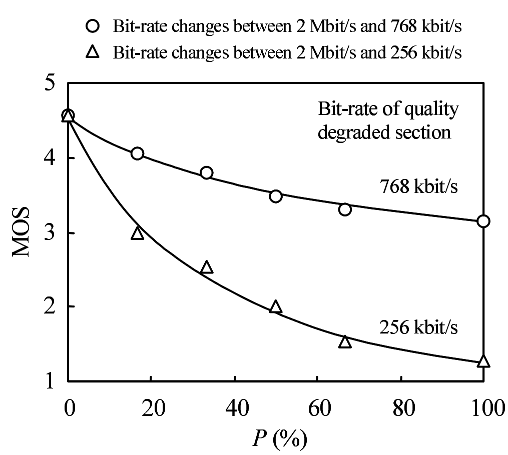


Fig. 2 Effects of percentage of quality-degraded sections.

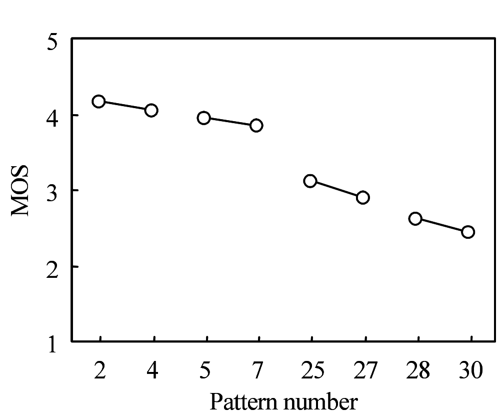


Fig. 3 Example of "recency effect."

These results suggest that P is a major factor affecting subjective quality for video streaming services with dynamic bit-rate control.

3. Estimation Model of Subjective Video Quality

3.1 Proposed Model

In this section, the model for estimating subjective overall video quality is discussed considering short-term qualities of 15-second inputs. Temporal variations of the short-term qualities for each bit-rate are shown in Fig. 4. Although each short-term quality at the same bit-rate is slightly different, we suppose that these qualities are the same. That is, the quality corresponds to the bit-rate. Therefore, we construct the estimation model of subjective video quality by using an averaged MOS.

In consideration of the subjective quality characteristics described in the previous section, we propose the following subjective quality-estimation model.

$$Q_{est} = \frac{\int_{-180}^{0} Q(t)\, e^{-wQ(t)}\, e^{t/T}\, dt}{\int_{-180}^{0} e^{-wQ(t)}\, e^{t/T}\, dt}, \tag{1}$$

where

$$Q(t) = \ln\left\{\frac{MOS(t) - 1}{5 - MOS(t)}\right\}. \tag{2}$$

- Q_{est}: estimated overall quality in an interval scale
- $MOS(t)$: short-term MOS of 15-second at time t obtained by the subjective test
- w, T: constants

Here, to construct the quality-estimation model, we used an interval scale Q obtained by converting the MOS [10] because the MOS is on an ordinal scale and its values do not

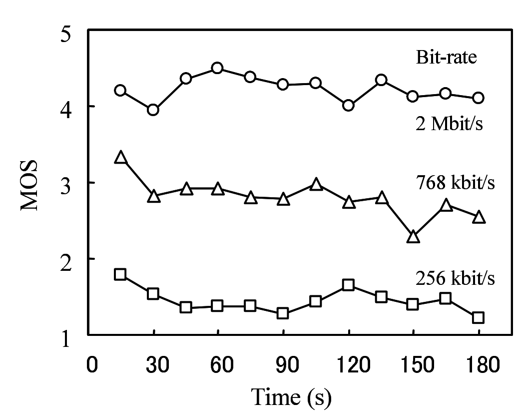


Fig. 4 MOSs of every 15-second for each bit-rate.

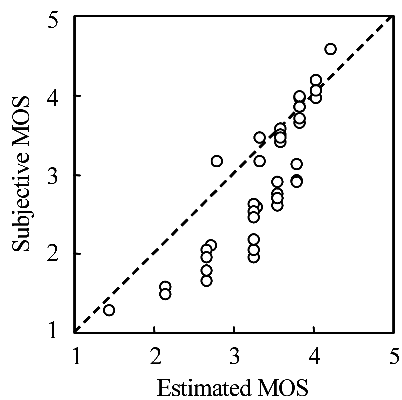


Fig. 5 Quality estimation by averaging segmental quality.

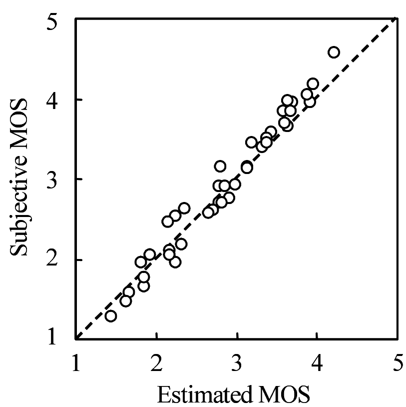


Fig. 6 Quality estimation using proposed model (without compensation for range of MOS).

necessarily linearly represent the subject's sensory impressions [11]. In Eq. (1), the term $e^{-wQ(t)}$ represents a weight for $Q(t)$ expressing strong influence of the video section where degradation is larger for overall quality. Moreover, the term $e^{t/T}$ shows a weight for $Q(t)$ reflecting the recency effect. Estimated overall quality in an interval scale is calculated by weighted averaging for $Q(t)$. Constants w and T can be determined from the subjective test results.

3.2 Estimation Accuracy of Proposed Model

The relationships between subjective and estimated MOS are shown in Figs. 5 and 6. The estimated MOSs were calculated by simply averaging the short-term MOSs of 15 seconds as shown in Fig. 5 and applying the proposed estimation model shown in Fig. 6. Here, from the subjective test results, constants w and T in Eq. (1) were set to 0.355 and 441, respectively. These figures indicate that the proposed model has good quality estimation accuracy. However, MOS is overestimated and underestimated in the areas where the subjective MOS is bad and good, respectively. This is the same phenomenon described in [12] in which the range of MOSs becomes wider when the time duration of an evaluation sequence is longer. This change in the range was therefore corrected by the following cubic function whose

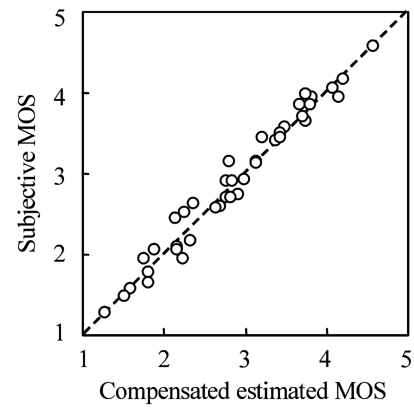


Fig. 7 Quality estimation using proposed model (with compensation for range of MOS).

origin corresponds to MOS=3.

$$Q_{est\_c} = 0.117\, Q_{est}^3 + 0.145\, Q_{est}^2 + 1.049\, Q_{est}. \qquad (3)$$

- Q_{est_c}: compensated estimated quality

Here, the analysis of variance (ANOVA) result showed that all terms of the cubic function were indispensable except for a constant term because partial regression coefficients in Eq. (3) were significant at the 5% level.

The relationship between subjective and compensated estimated-MOS is shown in Fig. 7. The root mean square errors (RMSEs) in Figs. 5, 6, and 7 were 0.64, 0.18, and 0.15, respectively. The half value of the averaged 95% confidence interval in a subjective test was 0.25, so the quality estimation accuracy of the proposed model with/without compensation is sufficiently good.

4. Verification Test

In the previous section, the coefficients of the model were determined from the subjective test results. To verify the generality of the model, it is necessary to examine the estimation accuracy when the model is applied to non-training data, i.e., different evaluation results. Therefore, we validate the proposed model using verification data.

4.1 Subjective Video Quality Assessment

In a verification test, a movie and a music video were prepared as source sequences, where the music video was the same as that used in the test described in Sect. 2. Each sequence was three minutes long. The movie video sequence was coded at 2 Mbit/s, 1 Mbit/s, 512 kbit/s, and 192 kbit/s. The music video sequence was coded at 2 Mbit/s, 768 kbit/s, and 256 kbit/s. Thirty five bit-rate alteration patterns are shown in Fig. 8. Patterns 1 to 25 are of the movie video and patterns 26 to 35 are of the music video. The test sequences were produced by the same method as that described in Sect. 2. The rating scale and the other assessment environment were also the same as those used in Sect. 2. Forty subjects who were different from those in Sect. 2 participated in this test.

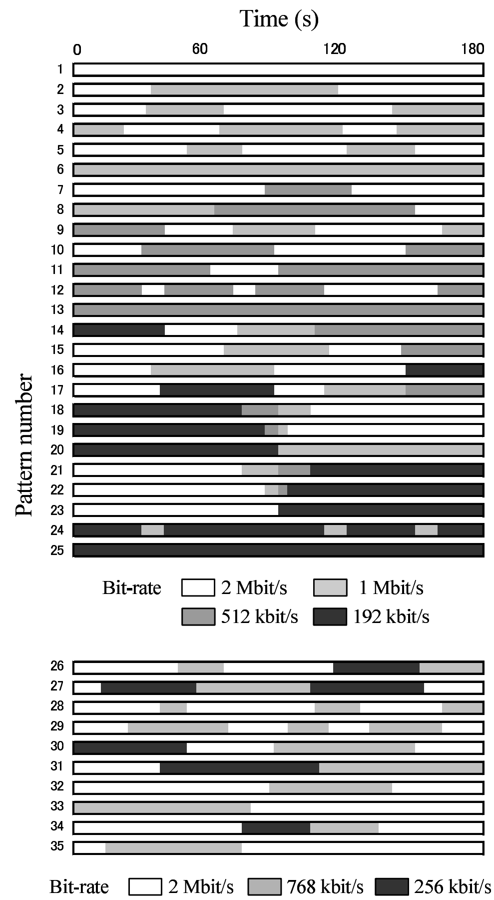


Fig. 8 Patterns of bit-rate alteration (verification data).

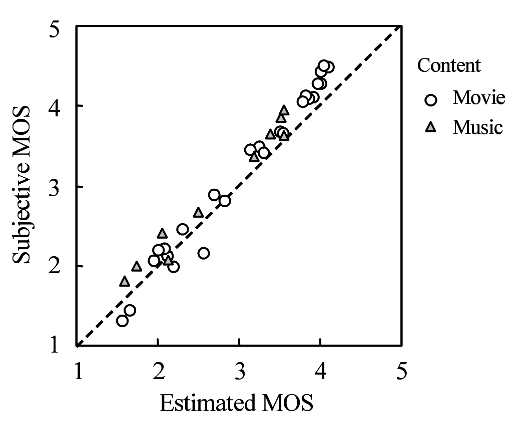


Fig. 9 Quality estimation using proposed model (without compensation, verification data).

4.2 Estimation Accuracy of Proposed Model Using Verification Data

The estimation accuracy of the proposed model is demonstrated without compensation in Fig. 9 and with compensation in Fig. 10. The coefficients of the model used in the test were the same as those determined in Sect. 3.

The RMSE in Fig. 9 was 0.25 and that in Fig. 10 was

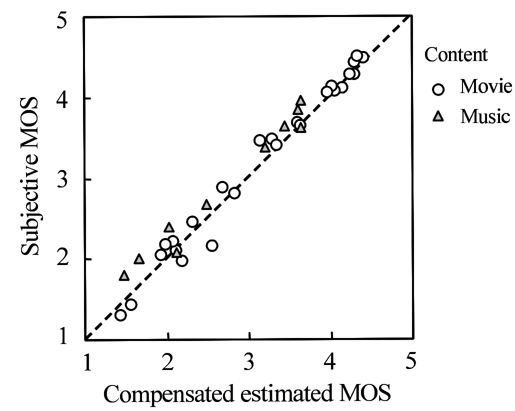


Fig. 10 Quality estimation using proposed model (with compensation, verification data).

0.19. The half value of the averaged 95% confidence interval in the subjective test was 0.23, so the quality estimation accuracy of the proposed model with compensation is sufficiently good for the verification data. Therefore, overall quality can be accurately estimated by this model independent of MOS input values and the kind of video content. We found that all coefficients in Eqs. (1) and (3) were unaffected by the difference in video contents used in the tests. Applying the proposed model to other kinds of video contents is for further study.

5. Application to Quality Design and Management

In this section, we discuss quality design/management issues for video streaming services with dynamic bit-rate control. This is based on the proposed model and our findings in the previous section.

5.1 QoS Design

Our subjective video quality-estimation model suggests that the percentage P of the duration of the quality-degraded sections is a major factor affecting subjective video quality. When the effect of "irregularity" of deterioration was strong, the quality for $P < 100\%$ became lower than that for $P = 100\%$ as reported in [13] which assessed video quality degradation due to irregular frame-rate reduction. In our experiments, however, for deterioration due to dynamic bit-rate change, video quality for $P < 100\%$ was not lower than that for $P = 100\%$. Here, we describe a guideline of QoS design that makes good use of these characteristics.

First, the video quality can be maintained more than that at the guaranteed bit-rate by allocating a minimum guaranteed bandwidth to each video flow. (In this paper, a "flow" is defined as streaming video traffic distributed from a server to a user's terminal.) This means that the lowest quality of the distributed video can be guaranteed. DiffServ AF [14]–[16], which is standardized in IETF, is one of the methods for guaranteeing a minimum usable bandwidth. If this technique and a video distribution technique are combined, we

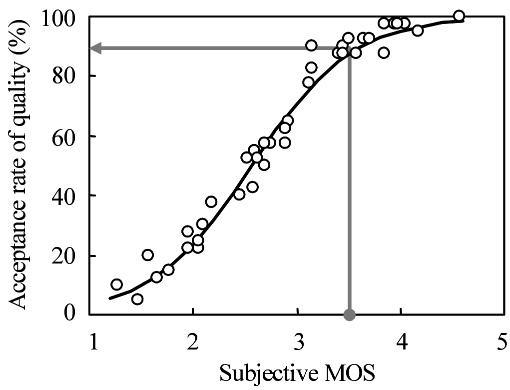


Fig. 11 Subjective MOS vs. acceptance rate of quality.

can achieve the video distribution with a maintained minimum quality. For example, it is effective to set the minimum quality to satisfy MOS=3.5. Figure 11 shows the relationship between MOS and acceptance rate of video quality obtained from the subjective test in Sect. 2. The acceptance rate of quality is more than about 90% when the MOS is greater than 3.5. The video quality can be maintained at more than MOS=3.5 by allocating a minimum guaranteed bandwidth to satisfy MOS=3.5.

Second, we can design the number of simultaneous flows at which the minimum quality for video streaming services with dynamic bit-rate control can be maintained. Dividing the bandwidth of a bottlenecked link by the minimum bit-rate for achieving the required minimum quality gives the number of simultaneous flows. For another example, when the minimum quality is guaranteed for a certain percentage of flows and the duration of video streaming follows an exponential distribution, the possible number of simultaneous flows can be simply approximated by an M/M/S/S queuing model. The number of users accommodated in a bottlenecked link can be determined by taking into consideration how many users receive video-streaming service simultaneously.

5.2 QoS Management

We can manage QoS for video distribution services with dynamic bit-rate control by using the proposed model. By grasping the relationship between the bit-rate and subjective video quality (MOS) in advance, we can estimate the subjective video quality that a user perceives and manage it by monitoring the change in bit-rate. Since the video quality at the same coding rate varies with resolution and movement in the video, video quality at a certain rate should be determined by evaluating many video sequences. For instance, video sequences prepared in [17] can be used in subjective tests. From a subjective test result, the relationship between the distributed bit-rate and MOS can be represented by using the average quality characteristic. It is possible to manage quality with an appropriate safety margin using strictly assessed video sequence quality results.

As a result, the change in bit rate can be treated as a time series representing the video quality. In video distribution services, the video quality of any past three minutes can be estimated in measuring the distributed bit rate for each video flow.

In video on demand (VoD) services, we can preliminarily evaluate changes in video quality for each bit rate of a video sequence. Therefore, by using this database as the model input, distributed video quality can be managed more accurately because the time series for video quality can be obtained directly. Many objective video quality measurement techniques have now been developed [7], [18], so QoS management combined with those techniques is a topic for further study.

6. Conclusion

This paper proposed a model for subjectively estimating the quality of video streaming services with dynamic bit-rate control. This model can be used to estimate long-term video quality of three minutes from short-term qualities. In this model, the following factors affected the users' perceptions: 1) the influence of the video section where quality degradation is large will strongly affect the overall quality, 2) the impression of a past quality weakens with the passage of time, and 3) the range of evaluation scores becomes wider when the time duration of an evaluation is longer. Considering the above, the proposed model enables overall subjective quality to be accurately estimated and therefore dynamically improved compared with that of a model that averages quality segmentally. Moreover, we validated that the estimated error with the proposed model was less than the statistical reliability of a subjective score for verification data.

Based on the proposed model and our findings, we discussed QoS design/management issues for video streaming services with dynamic bit-rate control. The key points of the characteristics of the estimation model are that the percentage P of the duration of quality-degraded sections is a major factor and video quality is not lower than the quality for $P = 100\%$. This means that the video quality can be maintained at more than the guaranteed bit rate by allocating a minimum guaranteed bandwidth. This result lets us design the number of simultaneous flows at which the minimum quality for video streaming services with dynamic bit-rate control can be maintained. We found that QoS for video distribution services can be managed by monitoring the change in bit rate related to MOS.

Topics for further study include studying the effects of time durations of video sequences, a QoS management combined with an objective video quality evaluation technique, and a QoS control method based on the derived users' perceptions.

Acknowledgements

The authors are grateful to the subjects for quality assessment tests.

References

[1] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP throughput: A simple model and its empirical validation," Proc. ACM SIGCOMM'98, vol.28, pp.303–314, Sept. 1998.

[2] M. Mathis, J. Semke, J. Mahdavi, and T. Ott, "The macroscopic behavior of the TCP congestion avoidance algorithm," ACM SIGCOMM Comput. Commun. Rev., vol.27, no.3, pp.67–82, July 1997.

[3] ITU-T Rec. P.911, "Subjective audiovisual quality assessment methods for multimedia applications," Dec. 1998.

[4] K. Tan, M. Ghanbari, and D. Pearson, "An objective measurement tool for MPEG video quality," Signal Proc., vol.70, no.3, pp.279–294, Nov. 1998.

[5] Y. Horita, M. Honda, T. Murai, and Z. Kawasaki, "Full reference quality evaluation model for MPEG-2 coded video," Proc. PCS2004, Dec. 2004.

[6] R. Hamberg and H. de Ridder, "Time-varying image quality: Modeling the relation between instantaneous and overall quality," SMPTE J., vol.108, no.11, pp.802–811, 1999.

[7] ITU-T Rec. J.144, "Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference," March 2001.

[8] R. Aldridge, J. Davidoff, M. Ghanbari, D. Hands, and D. Pearson, "Recency effect in the subjective assessment of digitally-coded television pictures," Proc. 5th IEE Intl. Conf. on Image Processing and its Applications, pp.336–339, Edinburgh, U.K., July 1995.

[9] ITU-T Rec. P.910, "Subjective video quality assessment methods for multimedia applications," Sept. 1999.

[10] R.D. Prosser, J.W. Allnatt, and N.W. Lewis, "Quality grading of impaired television pictures," Proc. IEE, vol.111, no.3, pp.491–502, March 1964.

[11] S. Tasaka and Y. Ito, "Psychometric analysis of the mutually compensatory property of multimedia QoS," Proc. ICC'03, pp.1880–1886, May 2003.

[12] M. Inoue, K. Kato, and K. Watanabe, "Influence of picture presentation period on subjective assessment," ITE Technical Report ID88-2, pp.9–16, 1988.

[13] T. Hayashi, S. Iai, and N. Kitawaki, "Video quality degradation due to field-repetition in video coding," IEEE Workshop on VSPC'93, pp.339–342, Sept. 1993.

[14] J. Heinanen, F. Baker, W. Weiss, and J. Wroclawski, "Assured forwarding PHB group," RFC2597, June 1999.

[15] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An architecture for differentiated service," RFC2475, Dec. 1998.

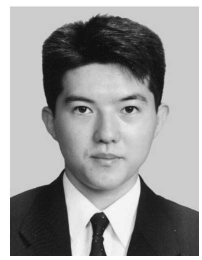
[16] D. Grossman, "New terminology and clarifications for DiffServ," RFC3260, April 2002.

[17] ITU-R Rec. BT.1210-3, "Test materials to be used in subjective assessment," April 2004.

[18] J. Okamoto, T. Hayashi, A. Takahashi, and T. Kurita, "An objective quality assessment method for arbitrary video streams," Proc. PCS2004, Dec. 2004.

Takanori Hayashi See this issue, p.281.

Ginga Kawaguti received B.E. and M.E. degrees in physical engineering from University of Tokyo in Japan in 1996 and 1998. He joined NTT Laboratories in 1998 and has been engaged in subjective quality assessment of multimedia telecommunications. In 2003, he transferred to Innovative IP Architecture Center of NTT Communications Corp. He is currently developing a quality measurement system for VoIP services.



Jun Okamoto received B.E. and M.E. degrees in electrical engineering from Science University of Tokyo in Japan in 1994 and 1996. He joined NTT Laboratories in 1996 and has been engaged in research into the quality assessment of video. He is currently working on the quality assessment of video streaming and video-communication services over IP networks.

Akira Takahashi received a B.S. degree in mathematics from Hokkaido University in Japan in 1988 and an M.S. degree in electrical engineering from the California Institute of Technology in 1993. He joined NTT Laboratories in 1988 and has been engaged in research into the quality assessment of speech and audio telecommunications. Currently, he is primarily working on the quality assessment of speech over IP networks. He has been contributing to ITU-T SG12 since 1994. He received the Telecommunication Technology Committee Award in Japan in 2004.