Framework and Standardization of Quality of Experience (QoE) Design and Management for Audiovisual Communication Services

Akira Takahashi[†]

Abstract

The quality of rich broadband services that use audio and visual media such as Internet protocol television (IPTV) and video telephony, which are expected to become more widespread over the Next Generation Network (NGN), should be evaluated subjectively by users. This is referred to as quality of experience (QoE). This article introduces the technical framework of QoE evaluation methodologies for quality design and management of audiovisual communication systems. It also describes the latest status of international standardization activities for these technologies.

1. QoE assessment methodologies

Rich broadband services such as Internet protocol television (IPTV) and video telephony are expected to become more widespread over the Next Generation Network (NGN). The quality of these services, which use both audio and visual media, should be evaluated subjectively by users. This is referred to as the quality of experience (QoE). Designing the QoE and managing it while a service is being provided is indispensable to provide users with high-quality services. This requires assessment methodologies that can quantify the QoE. This Special Feature focuses on the QoE of audio and visual media, excluding other QoE factors such as availability and service usability.

The quality of audio and visual media should be evaluated in subjective terms. This is called subjective quality assessment. However, subjective quality assessment, in which human subjects evaluate the quality of various testing conditions, is time-consuming and expensive. In addition, special assessment facilities such as professional audio-visual devices and soundproof chambers are required. Therefore, it would be desirable to develop a way to estimate subjective quality solely from the physical characteristics of a system under test. This is called objective quality assessment.

Objective quality assessment is an efficient means of assessment, and it enables in-service realtime QoE management, which cannot be done by subjective quality assessment. In-service QoE management is very important for maintaining the service quality since the qualities of IP networks and media coding vary with time.

This Special Feature introduces NTT's recent R&D achievements in audio and visual objective quality assessment. This first article gives an overview of subjective and objective quality assessment methods and their standardization activities.

2. International standardization organizations for QoE assessment

ITU-T (International Telecommunication Union, Telecommunication Standardization Sector) and

[†] NTT Service Integration Laboratories Musashino-shi, 180-8585 Japan

	Absolute rating		Relative rating		
	Category rating	Continuous rating	Category rating	Continuous rating	
Speech	P.800 (ACR)	BT500 (SSCQE)	P.800 (DCR) P.800 (DCR)		
Audio	BS.1284 (ACR)		BS1284 (DCR) BS1284 (CCR)		
			BS1116 (DBTS-HR)		
Video	P.910 (ACR)		BT500 (DCR)	BT500 (DSCQS)	
	BT500 (SS)		BT500 (SC)	BT500 (SDSCE)	
			P.910 (DCR)		
Multimedia	P.911 (ACR)* P.920 (ACR)*	P.911 (SSCQE)	P.911 (DCR)		

Table 1.	Categories (of subjective	quality	assessment	methods
		,			

*P.911 is a one-way evaluation, whereas P.920 is a bidirectional one.

ITU-R (International Telecommunication Union, Radiocommunication Sector) are responsible for the standardization of audio and visual QoE assessment methods. ITU-T SG12 (Study Group 12) studies the performance, quality of service (QoS), and QoE of telecommunications services and is the lead SG for these study items. The work of SG12 is coordinated with that of other study groups in ITU-T. SG12 has standardized various quality assessment methodologies mainly for speech, and these are now applied to conventional PSTN/ISDN (public switched telephone network and integrated services digital network) services as well as newly developed IP telephony services.

SG9 studies television services over cable networks. Its responsibilities include the standardization of video QoE assessment methods. On the other hand, ITU-R SG6 is responsible for broadcasting services such as radio and television services. QoE aspects are studied in WP6C (Working Party 6C).

A Joint Rapporteurs' Group on Multimedia Quality Assessment (JRG-MMQA) was established to harmonize the work in ITU SG9 and SG12. Another group, called VQEG (Video Quality Experts Group), performs technical investigations of the validity of objective quality assessment methods proposed to ITU and proposes the best one(s) to SGs.

3. Subjective quality assessment

Subjective quality assessment, in which subjects judge the quality of media such as video, has been studied for a long time, and ITU has standardized various methods for different purposes. It is a psychoacoustic/visual experiment, which is the most fundamental and reliable way to quantify users' QoE.

Subjective quality assessment methods are categorized into absolute and relative ratings. On the other hand, they can also be differentiated as category and continuous ratings. The opinion rating, which has long been used as a quality assessment method for PSTN services, is the ACR (Absolute Category Rating) method defined in ITU-T Recommendation P.800. This article focuses mainly on objective quality assessment, so please refer to the associated ITU Recommendations if you want to know more about the methods introduced in **Table 1**.

4. Objective quality assessment

Objective quality assessment is defined as a means for estimating subjective quality solely from objective quality measurement or indices. Objective quality assessment methods can be mapped into five categories, as listed in **Table 2**.

4.1 Planning model

ITU-T standardized a quality-planning tool for telephony services as Recommendation G.107, which is also called the "E-model". The output of the E-model is transmission rating scale "R". The Japanese government adopted the R-value as a quality index for IP telephony services. Therefore, service providers are required to evaluate the R-value of their services to obtain telephone numbers from the government for their customers.

SG12 extended the scope of G.107 so that it can cover so-called wideband (100–7000 Hz) speech communication services and provided G.107 Amendment 1. It also standardized G.113 Amendment 1,

Media-layer model	Parametric packet-layer model	Parametric planning model	Bitstream layer model	Hybrid model			
Media signal	Packet header information	Quality design parameters	Packet header and payload information	Combination of any			
Quality benchmarking	In-service non- intrusive monitoring (e.g., network probe)	Network planning, terminal/appli- cation designing	In-service non- intrusive monitoring (e.g., terminal- embedded operation)	In-service non- intrusive monitoring			
Existing standards and on-going projects in ITU							
ITU-T P.862	ITU-T P.564	ITU-T G.107	_	ITU-T P.CQO			
ITU-R BS1387	ITU-T P.NAMS (IPTV)	ITU-T G.1070 (videophone) ITU-T G. OMVS (IPTV)	ITU-T P.NBAMS (IPTV)	-			
ITU-T J.144 (SD)							
ITU-T J.vqhdtv (HD)				ITU-T J.bitvam			
ITU-T J.mm** (PC)				0.bitrqiii			
(ITU-T J.148)				_			
	Media-layer model Media signal Quality benchmarking ITU-T P.862 ITU-R BS1387 ITU-T J.144 (SD) ITU-T J.vqhdtv (HD) ITU-T J.mm** (PC) (ITU-T J.148)	Media-layer modelParametric packet-layer modelMedia signalPacket header informationQuality benchmarkingIn-service non- intrusive monitoring (e.g., network probe)Indards and on-going projects in ITUITU-T P.862ITU-T P.862ITU-T P.564ITU-T J.144 (SD) ITU-T J.vqhdtv (HD)ITU-T P.NAMS (IPTV)ITU-T J.148)ITU-T P.NAMS	Media-layer modelParametric packet-layer modelParametric planning modelMedia signalPacket header informationQuality design parametersQuality benchmarkingIn-service non- intrusive monitoring (e.g., network probe)Network planning, planning, erminal/appli- cation designingndards and on-going projects in ITU-ITU-T P.862ITU-T P.564ITU-T G.1070ITU-T J.144 (SD) ITU-T J.vqhdtv (HD) (ITU-T J.148)ITU-T P.NAMS (PTV)ITU-T G.1070 (Videophone) ITU-T Q.1070 (PTV)	Media-layer modelParametric packet-layer modelParametric planning modelBitstream layer modelMedia signalPacket header informationQuality design parametersPacket header and payload informationQuality benchmarkingIn-service non- intrusive monitoring (e.g., retwork probe)Network planning, erminal/appli- cation designingIn-service non- intrusive monitoring (e.g., terminal- embedded operation)ndards and on-going projects in ITUITU-T P.862ITU-T P.564ITU-T G.1070 videophone) ITU-T J.144 (SD)ITU-T J.vqhdtv (HD) ITU-T J.rum** (PC) (ITU-T J.148)ITU-T P.NAMS (IPTV)ITU-T G.1070 videophone) ITU-T G. OMVS 			



SD: standard definition

HD: high definition

PC: personal computer



Fig. 1. Block diagram of G.1070 model.

which provides quality assessment data for wideband speech codecs. The data was obtained through collaboration among the University of Tsukuba, NTT, Deutsch Telekom, and France Telecom.

In addition, SG12 has also been studying similar models for audiovisual communication services. In May 2007, Recommendation G.1070, which provides a quality-planning model for video-telephony applications, was standardized. This standard adopts a model that was studied and proposed by NTT. A block diagram of the G.1070 model is shown in **Fig. 1**. The input parameters to the model are as fol-

lows:

- Speech and video coding schemes

- Video resolution (QCIF (176×144 pixels), CIF

 $(352 \times 288 \text{ pixels})$, and VGA $(640 \times 480 \text{ pixels}))$

- Key frame interval of video coding
- Speech and video delays
- Speech and video bitrates
- Speech and video packet-loss rates
- Video frame rate

- Speech echo attenuation (TELR: talker-echo loudness rating)

Moreover, SG12 started investigating such a model

for IPTV services. This project is provisionally called G.OMVAS (opinion model for video and audio streaming applications). Standardization is expected in 2010. The model that NTT has been developing is introduced in "Planning Model for Audiovisual Communication Services" in this Special Feature.

4.2 Media-layer model

Media-layer models estimate the audio and/or visual QoE by using media signals such as the speech waveform and video pixel data. Depending on the application, there are three different approaches. Full-reference (FR) models require the original and the degraded signals to quantify the degree of degradation through a comparison. No-reference (NR) models take only a degraded signal as input, enabling quality monitoring at the site where the original signal is not available. Reduced-reference (RR) models are intermediate ones; that is, they use limited information about the original signal sent over networks for measurement, in addition to the degraded signal.

For speech, ITU-T SG12 standardized Recommendation P.862 (PESQ: perceptual evaluation of speech quality) as an FR model and Recommendation P.563 as an NR model. It also standardized an extension of PESQ for wideband speech as Recommendation P.862.2.

For audio, ITU-R standardized Recommendation BS.1387 (PEAQ: perceptual evaluation of audio quality) as an FR model. The effects of audio coding are evaluated by BS1387. However, BS1387 cannot be used to evaluate the effects of packet loss, which is one of the most important quality factors in recent applications.

For video, ITU-T SG9 standardized Recommendations J.144 and J.247 as FR models. Recommendation J.144 is a method used for evaluating MPEG-2 SD (Motion Picture Experts Group, standard definition) coding distortion, but it is not applicable to packet-loss degradation. On the other hand, Recommendation J.247 is used for various video codecs for QCIF, CIF, and VGA, and the effects of packet-loss degradation can be evaluated. SG9 also standardized Recommendation J.246 as an RR model. One of the J.247 models, which is provided in Annex A of that Recommendation, was developed by NTT. This technology is introduced in "Media-layer Objective Video Quality Assessment Technology for Video Communication Services (ITU-T J.247)" in this Special Feature.

4.3 Packet-layer model

One of the important applications of QoE evaluation methods is in-service non-intrusive quality monitoring. In such a scenario, it is almost impossible to use media signals such as audio and video to predict perceived quality because of the processing load limitation. This led us to develop packet-layer objective quality models, which estimate QoE solely based on packet-header information.

For IP telephony services, in 2006, ITU-T SG12 standardized Recommendation P.564, which determines the performance criteria for objective quality models. That is, models that estimate the QoE more accurately than the criteria are all regarded as "P.564-compliant" models.

SG12 is now working on models for IPTV, which are provisionally called P.NAMS (non-intrusive parametric model for the assessment of performance of multimedia streaming). P.NAMS utilizes only packet-header information (e.g., IP-MPEG-2-TS (TS: transport stream), which is useful when the processing load is very limited, e.g., monitoring inside a settop box. The model proposed by NTT is introduced in "Packet-layer Model for End-user QoE Management" in this Special Feature.

This project in SG12 is being conducted in close collaboration with VQEG. SG12 is responsible for determining the Terms of Reference, while VQEG will perform the technical evaluation and select candidate models.

5. Conclusion

This article gave an overview of the QoE assessment of audiovisual communication services and its standardization activities. The following articles in this Special Feature introduce NTT's recent research in this field.



Akira Takahashi

Senior Research Engineer, Supervisor, Service Assessment Group, NTT Service Integration Laboratories.

Laboratories. He received the B.S. degree in mathematics from Hokkaido University, Hokkaido, the M.S. degree in electrical engineering from California Institute of Technology, USA, and the Ph.D. degree in engineering from the University of Tsukuba, Ibaraki, in 1988, 1993, and 2007, respectively. He joined NTT Laboratories in 1988 and has been engaged in the quality assessment of audio and visual communications. He is a Vice-chairman of ITU-T SG12. He has been a co-Rapporteur of ITU-T Question 13/12 on Multimedia QoE and its assessment since 2005. He received the Telecommunication Technology Committee Award in Japan in 2004, the ITU-AJ Award in Japan in 2005, the Best Tutorial Paper Award from IEICE Com. Soc. (IEICE: Institute of Electronics, Information and Communication Engineers of Japan), and the Telecommunication Advancement Foundation Award in Japan in 2008.