

Recent Activities of QoE-related Standardization in ITU-T SG12

Kazuhisa Yamagishi and Yoichi Matsuo

Abstract

This article introduces recent standardization activities related to the evaluation of the quality of experience (QoE) of speech and video services, focusing on the activities of ITU-T SG12 (International Telecommunication Union - Telecommunication Standardization Sector, Study Group 12), which is responsible for standardization work on performance, quality of service, and QoE.

Keywords: adaptive bitrate streaming, crowdsourcing, gaming, quality of experience, virtual reality

1. Introduction

The International Telecommunication Union - Telecommunication Standardization Sector, Study Group 12 (ITU-T SG12) is a lead study group on network performance and quality of service (QoS) and quality of experience (QoE). In January 2017, SG12 was restructured by incorporating two questions on quality assessment, which had been studied in SG9. ITU-T SG12 is the leader in the worldwide standardization of speech and video quality evaluation, taking into account achievements in regional standardization bodies such as ETSI (European Telecommunications Standards Institute) and ATIS (Alliance for Telecommunications Industry Solutions). Standardization work on network performance parameters is carried out in various standardization organizations and all these organizations have confirmed that their work matches that of SG12.

2. Full-band and super-wideband E-model (G.107.2)

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 the R-value as a transmission rating scale. Q.15/12 (Parametric and E-model-based planning, prediction and monitoring of conversational speech quality) extended the scope of G.107 so that it can cover

super-wideband (50–14,000 Hz) and full-band (20–20,000 Hz) speech communication services and standardized Recommendation G.107.2. This enables us to calculate the quality of super-wideband and full-band speech encoded by EVS (Enhanced Voice Services).

3. Quality-estimation model for adaptive bitrate streaming (P.1203 and P.1204)

An important application of QoE estimation methods is in-service non-intrusive quality monitoring. For such a scenario, parametric quality models, which calculate QoE on the basis of packet-header information or metadata such as bitrate, should be developed due to the limited computational resources of end-clients.

Q.14/12 (Development of models and tools for multimedia quality assessment of packet-based video services) has been working on models for adaptive bitrate streaming, namely P.1203 and P.1204. P.1203 can be used to estimate the quality of adaptive bitrate streaming with high-definition resolution video encoded by H.264/AVC (Advanced Video Coding). The P.1203 model consists of video- and audio-quality-estimation modules (P.1203.1 and P.1203.2) and an integration module (P.1203.3). The video- and audio-quality-estimation modules calculate video and audio quality per second, and the integration module takes video and audio quality and stalling information

to calculate overall audiovisual quality. In addition, the video-quality-estimation module (P.1203.1) has four modes. The module takes metadata such as bitrate, framerate, and resolution in mode 0; frame-level information in addition to the input of mode 0 in mode 1; 2% of bitstream information in mode 2; and full bitstream information in mode 3.

Recently, 4K resolution videos encoded by H.265/HEVC (High Efficiency Video Coding) and VP9 have become popular. Therefore, in P.1204, Q.14/12 extended the scope of P.1203 so that it can cover these applications. In other words, the extension of the video-quality-estimation module has been studied. In P.1204, there are five models: a mode-0 model (P.1204.1), mode-1 model (P.1204.2), mode-3 model (P.1204.3), full-reference and reduced-reference pixel-based model (P.1204.4), and hybrid model (P.1204.5). P.1204.3, P.1204.4, and P.1204.5 have been standardized, but P.1204.1 and P.1204.2 are still being studied.

4. QoE-influencing factors and subjective evaluation for 360-degree video (G.1035 and P.360-VR)

As the fifth-generation mobile communication system (5G) is being launched, higher-speed and lower-latency video streaming services are expected. Since virtual reality (VR) is expected as one of the most promising services, Q.13/12 (Quality of experience (QoE), quality of service (QoS) and performance requirements and assessment methods for multimedia) has been studying subjective evaluation methodology to assess the quality of VR services.

In VR video streaming services, many users become nauseous due to motion sickness while watching VR video. Therefore, QoE-influencing factors for VR services are defined in detail in G.1035.

A new subjective evaluation methodology (P.360-VR) has been studied because a head-mounted display is worn to watch VR video streaming services, in contrast to regular two-dimensional video streaming services. A subjective evaluation methodology needs to be developed for VR video streaming services on the basis of the results of stability and reliability in many experiments. In addition, detailed procedures need to be described in the recommendation. In these tests, SG12 relies on the test results provided by VQEG (Video Quality Experts Group), and many tests have already been conducted. However, statistical analysis has not been completed. After the statistical analysis is conducted, the final draft of P.360-VR

will be submitted in September 2020.

5. QoE-influencing factors, subjective evaluation, and opinion model for gaming applications (G.1032, P.809, and G.1072)

Since gaming applications have spread rapidly, their QoE-influencing factors (G.1032) need to be identified, and a subjective assessment methodology (P.809) and quality-estimation model (G.1072) for them need to be developed. Like VR services, gaming applications have many QoE-influencing factors, so video, audio, and latency-related factors are described in detail in G.1032.

A subjective evaluation methodology for gaming applications is standardized in Recommendation P.809. In general, the five-point ACR (absolute category rating) is widely used in telecommunication. However, the seven-point continuous scale defined in P.851 is recommended because gaming applications have many QoE-influencing factors.

Like the E-model for telephony services, a quality-planning tool for gaming applications has been studied and standardized as Recommendation G.1072. In this recommendation, mathematical equations and parameters are defined. In other words, special software is not necessary. Therefore, this enables network operators and application developers to easily use the model, which takes parameters such as bitrate as input and calculates the quality of gaming applications.

6. Subjective evaluation with a crowdsourcing approach (P.808, P.CROWDV, P.CROWDG)

Subjective evaluation testing is generally conducted using special equipment and on the basis of expertise. However, to obtain quality from the actual services and obtain much high-quality data, subjective evaluation with a crowdsourcing approach has been studied and standardized in Recommendation P.808, which is used for evaluating speech quality. In addition, like speech quality, subjective evaluation with a crowdsourcing approach has been studied for video streaming and gaming applications because the demand to evaluate the quality of their applications with a crowdsourcing approach is increasing.

7. 2021–2024 study period

The structure of SG12 in the 2021–2024 study period has been discussed. Although the current

structure will basically be maintained, several questions will be closed and several others will be opened. Maintenance of recommendations under the responsibility of Q.3/12 (Speech transmission and audio characteristics of communication terminals for fixed circuit-switched, mobile and packet-switched Internet protocol (IP) networks) will be transferred to Q.5/12 (Telephonometric methodologies for handset and headset terminals) and Q.6/12 (Analysis methods using complex measurement signals including their application for speech and audio enhancement techniques). Q.18/12 (Measurement and control of the end-to-end quality of service (QoS) for advanced television technologies, from image acquisition to rendering, in contribution, primary distribution and secondary distribution networks) was closed because of a lack of contributions. However, maintenance of recommendations under the responsibility of Q.18/12 was transferred to Q.19/12 (Objective and subjective methods for evaluating perceptual audiovisual quality in multimedia and television services) during the

2017–2020 study period. In addition, a new question will be launched to study digital financial services, which were studied under Q.13/12

8. Outlook

This article described subjective assessment and quality-estimation models for speech, video streaming, and gaming applications. VR video streaming services and subjective evaluation with a crowdsourcing approach have been studied. Recently, SG12 has studied, for example, the analysis of quality-impairment factors and quality-estimation models using artificial intelligence technologies. Since many new services, such as telemedicine, are expected to be launched in the 5G era, more complicated issues related to QoS and QoE need to be addressed, for example, the demand of QoS and QoE planning and management. Therefore, it is important to investigate the recent activities of SG12.



Kazuhisa Yamagishi

Senior Research Engineer, NTT Network Technology Laboratories.

He received a B.E. in electrical engineering from Tokyo University of Science in 2001 and an M.E. and Ph.D. in electronics, information, and communication engineering from Waseda University, Tokyo, in 2003 and 2013. In 2003, he joined NTT, where he has been engaged in the development of objective quality-estimation models for multimedia telecommunications. He has been contributing to ITU-T SG12 since 2006. He was a co-rapporteur of Question 13/12 for the 2017–2020 study period.



Yoichi Matsuo

Research Engineer, NTT Network Technology Laboratories.

He received an M.E. and Ph.D. in applied mathematics from Keio University, Kanagawa, in 2012 and 2015. Since joining NTT in 2015, he has been engaged in research on network management. He has been contributing to ITU-T SG12 since 2019.