PAPER

Metadata-Based Quality-Estimation Model for Tile-Based Omnidirectional Video Streaming

Yuichiro URATA[†a)], **Masanori KOIKE**[†], **Kazuhisa YAMAGISHI**[†], ***and*** **Noritsugu EGI**[†], ***Members***

SUMMARY In this paper, a metadata-based quality-estimation model is proposed for tile-based omnidirectional video streaming services, aiming to realize quality monitoring during service provision. In the tile-based omnidirectional video (ODV) streaming services, the ODV is divided into tiles, and the high-quality tiles and the low-quality tiles are distributed in accordance with the user's viewing direction. When the user changes the viewing direction, the user temporarily watches video with the low-quality tiles. In addition, the longer the time (delay time) until the high-quality tile for the new viewing direction is downloaded, the longer the viewing time of video with the low-quality tile, and thus the delay time affects quality. From the above, the video quality of the low-quality tiles and the delay time significantly impact quality, and these factors need to be considered in the quality-estimation model. We develop quality-estimation models by extending the conventional quality-estimation models for 2D adaptive streaming. We also show that the quality-estimation model using the bitrate, resolution, and frame rate of high- and low-quality tiles and that the delay time has sufficient estimation accuracy based on the results of subjective quality evaluation experiments.

key words: *tile-based VR videos, quality monitoring, quality-estimation models, QoE*

1. Introduction

Omnidirectional video (ODV) streaming services have received a lot of attention because they can provide highly immersive viewing experiences. ODV streaming is expected to provide a sense of presence in sports, music, etc. However, the ODV has a much larger data volume than conventional 2D video. To reduce the amount of the data, tile-based ODV streaming has been proposed [1], [2]. In tile-based ODV, the entire video frame is divided into tiles, and tiles with multiple quality levels are distributed in accordance with the viewing direction. When the user changes the viewing direction, the low-quality tile is displayed. As a result, users might stop watching video due to the degradation. Therefore, service providers need to use a quality-estimation model to perform in-service quality monitoring at clients.

Quality-estimation models are classified on the basis of input parameters: signal-based, bitstream-based, and metadata-based quality-estimation models. When the quality to be estimated is the quality of experience (QoE), even if the input is any type, some researchers call the quality-estimation model the "QoE model" or "QoE estimation model." To unify the terminology, the model is called the "quality-estimation model" in this paper. Signal-based quality-estimation models [3]–[8] take video signal as input and estimate quality. They are classified in full-reference (FR) [3]–[7], reduced-reference (RR), and no-reference (NR) [8], [9] quality-estimation models. FR signal-based quality-estimation models [3]–[7] take source and degraded video signals, RR signal-based models take degraded video signals and features derived from source video signals, and NR signal-based models [8] take degraded video signals. In general, these models can assess the impact of source and codec on video quality. However, it is not feasible to use a signal-based model at the clients because computational resources are needed to calculate quality on the basis of video signals. Bitstream-based quality-estimation models [10], [11] take bitstream as input and estimate quality using parameters such as the quantization parameter parsed from bitstream. These models can assess the impact of source on video quality, where to take into account the impact of codec on video quality, model coefficients are generally switched depending on codec. However, bitstream-based quality-estimation models are not used for monitoring quality at the clients because bitstream is encrypted just after encoding video. Metadata-based quality-estimation models [12]–[15] take metadata such as bitrate, resolution, and framerate as input and estimate quality using these parameters. These models cannot assess the impact of source on video quality, where to take into account the impact of codec on video quality, model coefficients are generally switched depending on codec. Metadata-based quality-estimation models can be used for monitoring quality at the clients because metadata is not encrypted and high performance computational resources are not needed. Therefore, metadata-based quality-estimation models are suitable for monitoring quality at the clients.

An example of tile-based ODV streaming [1], [2] is shown in Fig. 1. In [1], [2], two kinds of tile are used: the entire video (blue rectangle in Fig. 1) is encoded at a low bitrate as a low-quality tile (omnidirectional tile) and divided tiles (red rectangle in Fig. 1) are encoded at high bitrates as high-quality tiles. High-quality tiles for a viewing direction and low-quality tiles are distributed to clients. A wide variety of tiles can be represented by the Omnidirectional MediA Format (OMAF) [16], which is developed by the Moving Picture Experts Group (MPEG) to facilitate interoperability between devices and media system components by different vendors. In [17], Fuente et al. evaluated the case where the ODV is divided into 24 (6×4) tiles, encoded in two

Manuscript received June 30, 2022.
Manuscript revised October 12, 2022.
Manuscript publicized November 15, 2022.

[†] The authors are with NTT Network Service Systems Laboratories, NTT Corporation, Musashino-shi, 180-8585 Japan.
a) E-mail: yuichiro.urata@ntt.com

DOI: 10.1587/transcom.2022EBP3109

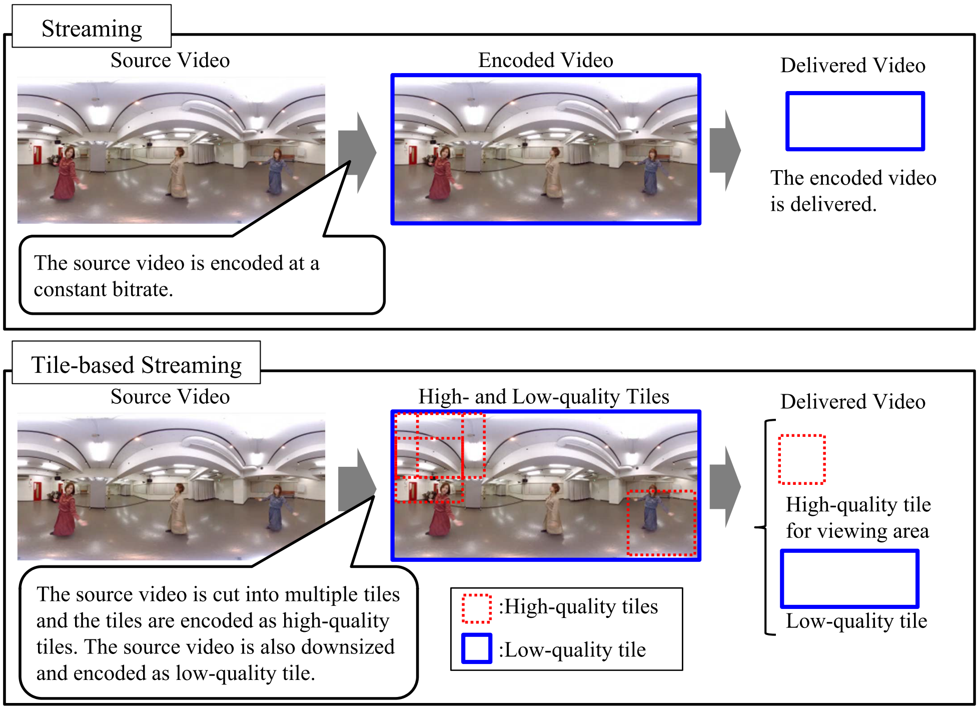

Streaming
Source Video
Encoded Video
Delivered Video
The encoded video is delivered.
The source video is encoded at a constant bitrate.
Tile-based Streaming
Source Video
High- and Low-quality Tiles
Delivered Video
High-quality tile for viewing area
Low-quality tile
The source video is cut into multiple tiles and the tiles are encoded as high-quality tiles. The source video is also downsized and encoded as low-quality tile.
:High-quality tiles
:Low-quality tile

Fig. 1 Image of tile-based streaming.

different resolutions, and delivered with 1/3 (8 tiles) in high resolution and the remaining 2/3 (16 tiles) in low resolution. Schatz et al. [18] divided ODVs into 24 (6×4) tiles, encoded them in three qualities (i.e., quantization parameter (QP) = 46, 32, 22), and investigated the case where the outside of the viewing area was delivered as low-quality tiles or not (represented as grey). Li et al. divided ODVs into 4×4, 6×6 and 8×8 tiles and encoded them into four levels of resolution [13]. They applied pyramid scheme to choose qualities with a gradually decreasing quality in accordance with the user's viewing direction.

In these tile-based ODV streaming, the client requests high-quality tiles in accordance with the viewing area, so users basically watch only high-quality tiles. However, when users change the viewing direction, the received high-quality tiles and the viewing area may differ. Thus, there is a case where high-quality tiles and low-quality tiles coexist as tiles presented to the user at the same time, and there is also a case where only low-quality tiles are presented. Furthermore, as in [18], there is a case where gray is presented instead of low-quality tiles. The quality of each high- and low-quality tile is affected by resolution, framerate and bitrate, like conventional 2D video streaming. In addition, the time required for switching from low- to high-quality tiles (hereafter, the delay time) also affects the perceived quality. Next, the client buffer depletes the video data depending on the throughput fluctuation. As a result, users perceive the stalling event. It is well-known that perceived quality is affected by stalling length and frequency [19]. Duan et al. [20] showed the effect of the combination of bitrate, framerate, and resolution on ODV quality. Fuente et al. [17] showed the effect of the delay time (display time of the low-resolution tiles) on tile-based ODV quality. Schatz et al. [19] showed the effect of stalling length and frequency on ODV quality. From these investigations, a quality-estimation model needs to evaluate the impact of these quality factors (i.e., resolution, framerate, and bitrate of both tiles, the delay time and stalling events).

Urata et al. [15] showed the effect of stalling events in tile-based OVDs can be estimated using the quality-estimation model for conventional 2D videos (2D model). In [15], the proposed model exhibited high quality-estimation accuracy even when there was stalling but low accuracy when there was only quality changes without stalling. Under the conditions with stalling, the quality of tile-based ODV could be estimated with relatively high accuracy because the effect of stalling events can be evaluated by using the 2D model. Therefore, this study focuses on the effect of short term quality factors (i.e., resolution, framerate, bitrate and the delay time) excluding the effect of stalling events.

In this paper, a metadata-based quality-estimation model is proposed for monitoring quality at clients for ODVs such as music and sports. We extended the 2D models in three different ways to evaluate the improvement of the quality-estimation accuracy by taking into account the quality factors (i.e., resolution, framerate, bitrate and the delay time). The difference between the three extended models is the influence of the quality factors taken into account. In the first model (Model A) and second model (Model B), the weighted sum of the quality of both tiles is used to estimate the overall video quality. The weight is calculated from the delay time and resolution in Model A, and the weight is a fixed value derived from experimental results in Model

B. The simplest third model (Model C) does not estimate the quality of both tiles, but estimates the overall quality by 2D model using total bitrate, which is the sum of the bitrates of both tiles, instead of each bitrate. To verify the estimation accuracy of these models, we conducted three subjective quality assessment experiments with [1], [2] and compared the outputs of these models with subjective scores. The P.1203 model [21] is a quality-estimation model for 2D video streaming, including sports and music, which is the target of ODV streaming, and its accuracy was targeted. The overall accuracy of quality-estimation was checked, and it was higher than the accuracy described in Recommendation P.1203. In addition, the accuracy of quality estimation for each quality factor was also checked.

First, related work is described in Sect. 2. The proposed quality-estimation models for tile-based ODV streaming are presented in Sect. 3. Subjective quality assessment experiments are described in Sect. 4. Section 5 shows that the proposed model (Model A) has sufficient estimation accuracy based on the results of these experiments. Finally, a summary and potential future work are presented in Sect. 6.

2. Related Work

As described in Sect. 1, metadata-based quality-estimation models are suitable because high performance computational power is not needed and they can be applied even when the bitstream is encrypted. Therefore, this section describes metadata-based quality-estimation models.

2.1 Conventional Metadata-Based Quality-Estimation Models for ODVs

Conventional metadata-based quality-estimation models for ODVs are proposed [12]–[15], [22], [23]. In [12], quality is estimated using the bitrate and stalling information. However, the model cannot be applied to the tile-based streaming because multiple tiles are not used in this experiment. The model in [13] estimates quality by using the tiling, quality levels, and stalling. For each tiling, a Gaussian function with latitude as a parameter is fitted to the subjective quality assessment results. However, since the relationship between the coefficients of each Gaussian function and bitrate, resolution, etc. has not been clarified, the model cannot be adapted to the case of different bitrate ladders. In addition, the delay of quality switching is not taken into account. Costa et al. proposed a model on the basis of the bitrate for tile-based streaming in [14], but the quality-estimation accuracy is unclear because it is not compared with the results of subjective quality assessment experiments. In addition, the size of high-quality tiles and the delay time of quality switching are not taken into account. Zhang et al. [22] proposed a deep reinforcement learning based ODV streaming system, with a focus on efficiently utilizing the limited bandwidth resources to improve the QoE of users when watching the viewport-and tile-based ODVs. The proposed system optimizes various QoE objectives that are based on the bitrate, stalling time, and viewport temporal variations. However, the delay time of quality switching is not taken into account. Kan et al. [23] designed a new QoE metric by introducing a penalty term for the large buffer occupancy to reduce the possible delay time of quality switching. The QoE metric is based on the bitrate, the buffer occupancy, and predicted viewport. The metric takes into account the delay time of quality switching but not the size of high-quality tiles.

Urata et al. [15], proposed extending quality-estimation models for 2D adaptive streaming to tile-based ODV streaming using the P.1203 model [21]. The models can estimate the impact of stalling events on tile-based ODV quality. However, the number of coding conditions and source contents were limited in the test of the quality-estimation accuracy. Since a single size of high-quality tiles was used, the impact of the size on quality is not taken into account.

From these investigations, the above models did not take some quality factors into account, and none can account for all of the impact of quality factors (high- and low-quality tiles, the delay time of quality switching, and the size of high-quality tiles) on quality comprehensively. Therefore, a model needs to be created that can take them all their impacts into account.

2.2 Extendable Conventional Metadata-Based Quality-Estimation Models to ODV

This subsection explains metadata-based quality-estimation models for conventional 2D videos.

2.2.1 P.1203 Model

The P.1203 model, which is used to estimate quality of 2D adaptive streaming, has been standardized in ITU-T. The model contains three modules (i.e., video-quality-estimation, audio-quality-estimation, and integration modules) and four modes (i.e., Modes 0 to 3). As mentioned above, this paper targets developing a video-quality-estimation module. Therefore, the details of the video-quality-estimation module are described.

The module is classified into four modes in accordance with the type of input information. However, as described in Sect. 1, since the metadata-based model is suitable for monitoring purposes, only Mode 0 is described. The Mode 0 model takes metadata such as codec, resolution, framerate, and bitrate as input and outputs video quality per second. The video quality $O.22$ in Mode 0 is estimated as follows.

$$O.22 = MOSfromR(100 - \min(D, 100)), \tag{1}$$

$$D = Dq + Du + Dt, \tag{2}$$

$$Dq = 100 - RfromMOS(MOSq), \tag{3}$$

$$MOSq = q_1 + q_2 \cdot \exp(q_3 \cdot quant), \tag{4}$$

$$\begin{aligned} quant &= a_1 + a_2 \cdot \ln(a_3 \\ &\quad + \ln(br) + \ln(br \cdot bpp + a_4)), \end{aligned} \tag{5}$$

$$\begin{aligned} Du &= u_1 \cdot \log_{10}(u_2 \\ &\quad \cdot (scaleFactor - 1) + 1), \end{aligned} \tag{6}$$

$$scaleFactor = \max\left(\frac{disRes}{codRes}, 1\right), \tag{7}$$

$$Dt = \begin{cases} 0 & (fr \geq 24) \\ \frac{(100 - Dq - Du) \cdot (t_{1v} - t_{2v} \cdot fr)}{t_{3v} + fr} & (fr < 24) \end{cases} \tag{8}$$

$MOSfromR$ and $RfromMOS$ convert the mean opinion score (MOS) from/to the psychological value R of $0 - 100$. The details of these two functions are described in Annex E of ITU-T Recommendation P.1203.1. The variable Dq is the amount of quality degradation related to the quantization calculated from $quant$, which is a variable related to the quantization as (3), (4). The parameter $quant$ is calculated from the bit amount per pixel bpp and bitrate br in Mode 0 as (5). The variable Du is the amount of quality degradation related to upscaling. The parameter $scaleFactor$ is calculated by dividing display resolution $disRes$ by coding resolution $codRes$. These resolutions mean the number of pixels. The variable Dt is the amount of degradation related to the frame rate fr. The coefficients q_{1-3}, a_{1-4}, u_{1-2}, and t_{1v-3v} are constant for each codec.

2.2.2 Y-Model

The Y-model [24], [25] also has modules to calculate audio quality, video quality, and media session quality for 2D video streaming, like the P.1203 model. The input of Y-model is the same as that of the P.1203 Mode 0 model.

The video-quality-estimation module calculates video quality $O.22$ by using video bitrate br, coding resolution $codRes$, and frame rate fr as follows.

$$O.22 = X + \frac{1 - X}{1 + (br/Y)^{v_1}}, \tag{9}$$

$$X = 1 + \frac{4 \cdot (1 - \exp(-v_3 \cdot fr) \cdot codRes)}{v_2 + codRes}, \tag{10}$$

$$Y = \frac{v_4 \cdot codRes + v_6 \cdot \log_{10}(v_7 \cdot fr + 1)}{1 - \exp(-v_5 \cdot codRes)} \tag{11}$$

The coefficients $v_1 - v_7$ are constant.

3. Proposed Models

The proposed models for tile-based ODV streaming are described. Especially, how to extend conventional 2D video quality-estimation models to tile-based ODV streaming is described. The two base models are used to check if the extension method can be applied independently of the base models.

3.1 Extended Model for Tile-Based ODVs

The P.1203 model and the Y-model are extended to tile-based ODV streaming [1], [2]. The overall video quality for tile-based ODV is affected by the quality of the omnidirectional tile (lower quality tile) and divided tile (higher quality tile) for the tile-based ODV. Therefore, the overall video quality is integrated by the quality of both tiles, which can be estimated by using either the P.1203 model or the Y-model. Since the impacts of omnidirectional tiles on the overall video quality depend on the delay time and the size of divided tiles as mentioned above, the delay time and the size of the area need to be taken into account to estimate the overall video quality.

To evaluate the improvement of the quality-estimation accuracy by taking into account these quality factors, the P.1203 model and the Y-model are extended in three different ways. In the first model (Model A), the quality of omnidirectional and divided tiles is estimated using either the P.1203 model or the Y-model, and the weighted sum is used to estimate the overall video quality. The weight ratio between the quality of the omnidirectional and divided tiles is calculated from the delay time and the size of the area in (12)–(14). The weight ω is limited to the range 0 to 1 in (13). In (14), the parameter $codRes$ is the resolution (number of pixels) of the divided tiles and $overallRes$ is the resolution of the entire sphere surface, so the ratio means the share of divided tiles on the sphere surface. The weight is calculated taking into account the share of the divided tile, since the divided tile will continue to be displayed when it is large enough and the change in viewing area is small. In addition, because a shorter delay time results in the prompt display of divided tile, the effect of the delay time is evaluated as a power of the delay time in (14). In the second model (Model B), the overall video quality is the weighted sum of the quality of omnidirectional and divided tiles as in Model A. In Model B, it can be considered that the impacts of the delay time and the size are ignored and the weight ratio (ω_c) is set to a fixed value derived from the results of subjective quality assessment experiments. The simplest third model (Model C) does not estimate the quality of both tiles, but estimates the overall quality by 2D model using the common resolution, common framerate and total bitrate, which is the sum of the bitrate of both tiles. Model C does not take into account the difference in quality between omnidirectional and divided tiles.

A) Based on the delay and the resolution

$$O.22 = \omega \cdot O.22_H + (1 - \omega) \cdot O.22_L, \tag{12}$$

$$\omega = \min(\max(\omega', 0), 1), \tag{13}$$

$$\omega' = \left(\omega_1 \cdot \log\left(\frac{codRes}{overallRes}\right) + \omega_2\right) \cdot delay^{-\omega_3} \tag{14}$$

B) Weighted sum

$$O.22 = \omega_c \cdot O.22_H + (1 - \omega_c) \cdot O.22_L \tag{15}$$

C) Total bitrate

$$O.22 = O.22_T \tag{16}$$

where $O.22_H$ and $O.22_L$ are the video quality of the divided tile and the omnidirectional tile calculated on the basis of either the P.1203 model or the Y-model, and $O.22_T$ is the

Fig. 2 SRCs in Experiment 1.

video quality calculated using the total bitrate instead of each bitrate. The coefficients ω_{1-3} are derived using the results of the subjective quality assessment experiments shown in the next section.

4. Subjective Experiment

To investigate the quality-estimation accuracy of the extended quality-estimation models for tile-based ODV, three subjective quality assessment experiments were conducted. Experiments 1 and 2 are used for deriving the models' coefficients (training), and Experiment 3 is used to investigate the quality-estimation accuracy (testing). The experiments were conducted using the tile-based ODV streaming system explained in Fig. 1.

4.1 Source Reference Circuits

As shown in Figs. 2, 3, and 4, six source reference circuits (SRCs) for Experiments 1 and 2 and 27 SRCs for Experiment 3 were selected. Since we aim to monitor the quality of music, sports, etc. content, we have prepared SRCs that include such content. The duration of the SRCs is 20 seconds, the resolution is 7680×3840, and the frame rate is 30 fps. SRCs were characterized in terms of their Spatial Information (SI) and Temporal Information (TI) [26]. SI and TI are indices of spatial and temporal complexity of SRCs, respectively. Since the quality after encoding differs depending on SI/TI even at the same encoding bitrate, this needs to be confirmed with experimental data that has those variations. For ODVs, it is necessary to consider that planar representations (equirectangular, cube-map, etc.) change the characterization of the content because of warping, discontinuities, etc. To determine the average feature of SRCs, the average SI and TI of all frames in SRCs are calculated in the spherical domain [27]. Figure 5 shows the averaged spherical SI and TI in the experiments, and the SRCs are found to have different motions and edge features. In particular, SRCs in Experiment 3, which is used to test the accuracy of quality-estimation, have low to high SI and TI, indicating that there is variation in spatial and temporal characteristics. To compare the evaluation results between these experiments, six SRCs for Experiment 1 were also used as common videos in Experiments 2 and 3. For watching ODV naturally, the stereo channel audio was used in these experiments.

Fig. 3 SRCs in Experiment 2.

Fig. 4 SRCs in Experiment 3.

4.2 Experimental Conditions and Processed Video Sequence

In these experiments, divided (high-quality) and omnidirectional (low-quality) tiles were encoded by FFmpeg encoder v3.0 with H.265/high-efficiency video coding (HEVC) (Main Profile/Level 5, GOP: M = 3, N = 15, preset: medium, format: yuv420p). The segment size was set to 0.5 seconds.

The resolutions of the divided and omnidirectional tiles

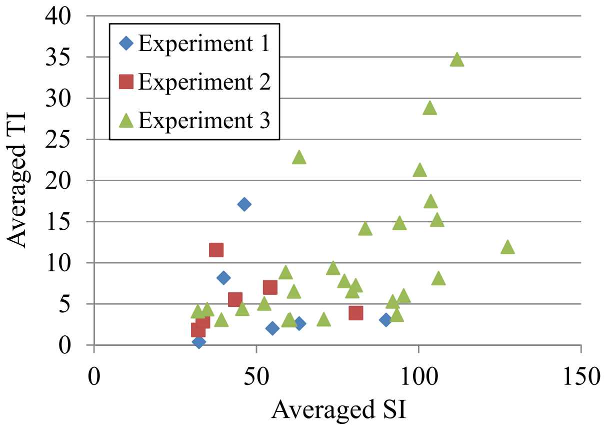


Fig. 5 Averaged SI and TI for the SRCs.

were set to five levels: 3840×3840, 2560×2560, 1920×1920, 1280×1280, and 960×960. The resolutions of both tiles were the same, and the divided tiles were encoded by being cropped from each SRC with 7680×3840, while the omnidirectional tiles were encoded by downscaling the entire sphere surface for each SRC. When playing back the omnidirectional tiles, the video was upscaled from the resolution given as the condition to the original resolution of 7680×3840. The viewable sphere surface is created by overlaying upscaled omnidirectional tile with divided unscaled tile, and the user can freely move around the viewing area, including the area where both tiles coexist. All divided tiles are overlapped to avoid quality degradation due to small changes in the user's viewing direction. All divided tiles are arranged at equal intervals over the entire horizontal or vertical space, and the degree of overlap depends on the resolution. Twelve divided tiles were placed at equal intervals in the horizontal direction. Since the resolution of the entire sphere is 7680×3840, the interval (the difference of x-coordinates of horizontally adjacent tiles) is 7680/12. The number of pixels of horizontal direction overlap is $side\_of\_tile - (7680/12)$, where $side\_of\_tile$ represents the length of the side of the square divided tiles, i.e., 3840, 2560, 1920, 1280, or 960 in the experiments. Five divided tiles were placed at equal intervals in the vertical direction except for 3840×3840. The 3840×3840 divided tiles are the same size as the entire sphere in the vertical direction, so they are not divided in the vertical direction. Since the vertical top and bottom are not connected, the interval is $(3840 - side\_of\_tile)/(5 - 1)$, so the number of pixels of vertical direction overlap is $side\_of\_tile - (3840 - side\_of\_tile)/(5 - 1)$. The number of divided tiles per frame was set to 12×1 when the resolution was 3840×3840, and 60 (=12×5) for the other resolutions.

The delay time was set in the range of 1 to 10 seconds. The client sends a request for the next divided and omnidirectional tiles on the basis of buffer remaining. When the divided tile to be downloaded changes due to a user's change of view direction, the next request will reflect the new position. In other words, tiles that have already been requested at that time are downloaded at the old position. Therefore, the delay time of quality switching is approximately equal to the buffer length when bandwidth is sufficient. In the experiments, the delay time is controlled by the buffer length setting. A minimum buffer of at least one second is required for stable playback, even when bandwidth is sufficient, and the minimum delay is set to one second.

Table 1 HRCs in Experiment 1.

Side of tile (Resolution)	Bitrate divided/omnidirectional	Delay
3840	40 Mbps / 40 Mbps	5 s
3840	40 Mbps / 20 Mbps	1 s
3840	40 Mbps / 10 Mbps	3 s
3840	10 Mbps / 5 Mbps	10 s
1920	40 Mbps / 40 Mbps	1 s
1920	40 Mbps / 20 Mbps	1 s, 3 s, 5 s, 10 s
1920	40 Mbps / 10 Mbps	1 s, 5 s
1920	10 Mbps / 10 Mbps	1 s
1920	10 Mbps / 5 Mbps	1 s, 5 s
1920	10 Mbps / 2.5 Mbps	1 s, 10 s
1920	2 Mbps / 2 Mbps	1 s, 10 s
1920	2 Mbps / 1 Mbps	1 s
1920	2 Mbps / 500kbps	1 s, 3 s
1280	40 Mbps / 20 Mbps	1 s, 5 s
1280	40 Mbps / 10 Mbps	10 s
1280	10 Mbps / 10 Mbps	3 s
1280	2 Mbps / 2 Mbps	5 s
1280	2 Mbps / 1 Mbps	3 s, 10 s
960	40 Mbps / 20 Mbps	1 s, 10 s
960	10 Mbps / 2.5 Mbps	5 s
960	2 Mbps / 1 Mbps	3 s

The encoding bitrate and delay time are shown in Tables 1 and 2 in Experiments 1 and 2. The experimental conditions, which are called hypothesis reference circuits (HRCs), were designed to investigate the impact of the resolution, quality of divided tiles, quality of omnidirectional tiles, and the delay time. In addition, to check the impact of SRCs, three SRCs are assigned to each HRC in Experiment 1, and two source videos were assigned to each HRC in Experiment 2. The ranges of encoding bitrate are shown in Tables 1 and 2 in Experiments 1 and 2. In Experiment 3, for testing, lots of HRCs (113 HRCs) were set up and only one SRC was assigned to each HRC. The encoding bitrate and delay time are shown in Fig. 6.

The common PVSs (C1-C9), which are used for comparing the evaluation results of the experiments, are shown in Table 3. The total number of processed video sequences (PVSs) was 96 in Experiment 1, 99 in Experiment 2, and 113 in Experiment 3 (including the common PVSs).

Audio was encoded by AAC-LC. Bitrate and sampling rate were 128 kbps and 48 kHz, respectively. The audio conditions and the audio quality are the same for all PVS so that they do not affect the video quality.

4.3 Experimental Environment and Assessment Method

The experiments were conducted using a panoramic super-engine developed on the basis of the technologies [1], [2]. The participants wore a head-mounted display (HMD), which is HTC Vive Pro, and watched the tile-based ODVs in

Table 2 HRCs in Experiment 2.

Side of tile (Resolution)	Bitrate divided/omnidirectional	Delay
3840	40 Mbps / 5 Mbps	1 s
3840	30 Mbps / 3 Mbps	1 s
3840	20 Mbps / 10 Mbps	6 s
3840	10 Mbps / 10 Mbps	6 s
3840	8 Mbps / 8 Mbps	1 s
3840	3 Mbps / 6 Mbps	2 s
3840	2 Mbps / 2 Mbps	6 s
3840	2 Mbps / 1 Mbps	1 s
2560	8 Mbps / 6 Mbps	10 s
2560	8 Mbps / 4 Mbps	4 s
2560	8 Mbps / 1 Mbps	1 s
2560	5 Mbps / 5 Mbps	4 s
2560	4 Mbps / 4 Mbps	2 s
2560	4 Mbps / 1 Mbps	6 s
2560	2 Mbps / 1 Mbps	1 s
1920	30 Mbps / 30 Mbps	10 s
1920	30 Mbps / 5 Mbps	10 s
1920	16 Mbps / 16 Mbps	1 s
1920	16 Mbps / 8 Mbps	2 s
1920	10 Mbps / 20 Mbps	4 s
1920	10 Mbps / 8 Mbps	8 s
1920	10 Mbps / 6 Mbps	1 s
1920	10 Mbps / 2 Mbps	1 s
1920	8 Mbps / 12 Mbps	8 s
1920	8 Mbps / 8 Mbps	6 s
1920	6 Mbps / 1 Mbps	8 s
1920	5 Mbps / 10 Mbps	2 s
1920	5 Mbps / 2 Mbps	6 s
1920	4 Mbps / 2 Mbps	4 s
1920	4 Mbps / 1 Mbps	1 s
1920	4 Mbps / 500 kbps	6 s
1280	40 Mbps / 10 Mbps	8 s
1280	20 Mbps / 20 Mbps	6 s
1280	16 Mbps / 16 Mbps	6 s
1280	8 Mbps / 4 Mbps	1 s
1280	8 Mbps / 2 Mbps	2 s
1280	6 Mbps / 6 Mbps	4 s
1280	6 Mbps / 3 Mbps	10 s
1280	4 Mbps / 3 Mbps	4 s
1280	2 Mbps / 500 kbps	10 s
960	20 Mbps / 2 Mbps	8 s
960	10 Mbps / 4 Mbps	10 s
960	6 Mbps / 4 Mbps	2 s
960	6 Mbps / 2 Mbps	1 s
960	2 Mbps / 2 Mbps	8 s

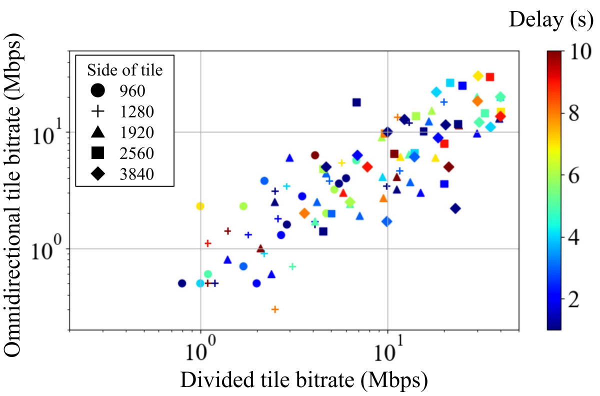


Fig. 6 Bitrates and delay in Experiment 3.

Table 3 Common PVSs.

PVS	SRC	Side of tile	Bitrate divided / omnidirectional	Delay
C1	1-01	1280	2 Mbps / 1 Mbps	10 s
C2	1-01	1920	40 Mbps / 40 Mbps	1 s
C3	1-02	3840	10 Mbps / 5 Mbps	10 s
C4	1-02	1920	2 Mbps / 2 Mbps	1 s
C5	1-03	1920	10 Mbps / 2.5 Mbps	10 s
C6	1-04	1920	40 Mbps / 20 Mbps	10 s
C7	1-05	1920	10 Mbps / 10 Mbps	1 s
C8	1-05	3840	40 Mbps / 10 Mbps	3 s
C9	1-06	1920	2 Mbps / 0.5 Mbps	3 s

Table 4 Sessions and playlists.

Experiment	1	2	3
#Sessions	5	5	4
#Playlists	18	18	10
#Playlists per session	3 or 4	3 or 4	2 or 3
#PVSs per playlist	6	6	12
Duration per playlist	3.5 min	3.5 min	6.5 min
Short break time	2 min	2 min	2 min
Break time	5 or 10 min	5 or 10 min	10 or 15 min

a room in accordance with the voice instructions of the operator, and evaluated the quality. The display device showed both eyes with pseudo-parallax. Since the participants could freely change the viewing area of the ODVs, the area to be viewed differed for each participant.

The participants were explained the instructions to understand the procedure and the object of the subjective experiment and took visual acuity and color vision tests. After these tests, participants took a practice session in which they learned how to wear the HMD and the method to evaluate the quality of the watched video. Participants watched 20-second practice videos 6 times in Experiments 1 and 2 and 12 times in Experiment 3. For each video, participants gave a score with a 5-point absolute category rating (ACR) method [26] using a controller. Three seconds after scoring, the next video in the practice session started playing.

After the practice session, the main sessions for evaluating all PVSs started. All PVSs consisted of multiple playlists. In Experiments 1 and 2, the playlists had 6 PVSs, and 3 or 4 playlists were evaluated in one session. In Experiment 3, the playlists had 12 PVSs and 2 or 3 playlists were evaluated in one session. The evaluated playlists took about 3.5 minutes in Experiments 1 and 2 and 6.5 minutes in Experiment 3. Participants had breaks for about 2 minutes after every playlist and about 5–15 minutes after every session. For each experiment, the number of sessions, the number of playlists, and the number of PVSs in a playlist are shown in Table 4. The total time from the explanation to the end of the evaluation was about 150 minutes, including visual acuity and color vision tests, instruction, practice, main sessions, and breaks. The presentation order of the PVSs was randomized.

4.4 Participants

In each experiment, 36 participants took part: 18 males and 18 females with visual acuity of 1.0 or more with contact lenses or the naked eye. All the participants passed the visual acuity and color tests. They were naive participants who had not participated in subjective quality assessment experiments of ODV streaming in the previous six months. The ages of participants were 18 to 31 years old (average 21.0) in Experiment 1, 18 to 25 years old (average 20.9) in Experiment 2, and 19 to 27 (average 21.4) in Experiment 3.

5. Results

In this section, results of subjective quality assessment are first described to show the stability and the bias in MOS among experiments. Then, quality-estimation accuracies of models are described.

5.1 Statistics of the Scores and Common PVSs Analysis

The stability of subjective quality assessment experiments are shown in terms of the 95% confidence intervals (CI) for MOS. Table 5 shows the mean, standard deviation, minimum, and maximum of CI. The stability in these experiments can be said to be high enough because these mean CIs were almost the same as the mean CI (0.23) for 2D videos in Tominaga et al. [28].

To check for bias in MOS among experiments, MOSs of nine common PVSs are compared, as shown in Fig. 7, where the error bars show the CIs. The MOSs of the common PVSs in each experiment are almost the same, so the bias in MOS among experiments is small.

5.2 Quality-Estimation Accuracy

To investigate the quality-estimation accuracy of the proposed models (extending Model A on the basis of two conventional models: the P.1203 model and the Y-model), the coefficients were optimized by using Experiments 1 and 2 results and Microsoft Excel Solver. In addition, to investigate how much the accuracy of quality-estimation improved by addition of low-quality terms, the resolution, and the delay time, the remaining four models (two base models × two extensions: Models B and C), the coefficients were optimized as in the proposed models. Table 6 shows the Root Mean Squared Errors (RMSEs) and Pearson Correlation Coefficients (PCCs) for training data (Experiments 1 and 2) and test data (Experiment 3). In both proposed models, the quality-estimation accuracy for the test data is better than the quality-estimation accuracy described in Recommendations P.1203 (RMSE: 0.465, PCC: 0.814 in mode 0) and P.1204 (RMSEs: 0.421 in bitstream-based model and 0.444 in FR signal-based[†] model) [29]. Therefore, the proposed models achieve the target quality-accuracy and have sufficiently useful performance. By comparing extended models A, B, and C in Table 6, quality-estimation accuracy of model A is highest. Figures 8(a)–9(b) show scatter plots of the MOSs obtained from the experiments for training and test, and the

Table 5 Summary statistics of the confidence intervals.

Experiment	Mean	Std.	Min.	Max.
1	0.244	0.039	0.143	0.345
2	0.261	0.035	0.178	0.363
3	0.255	0.049	0.056	0.372

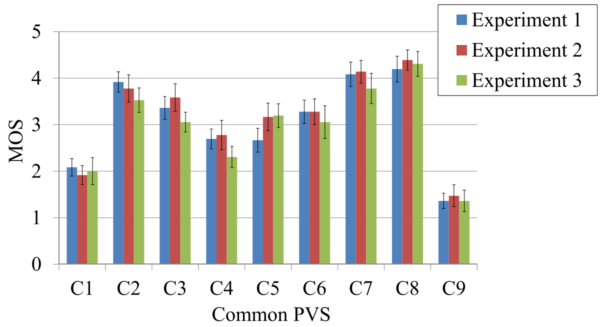
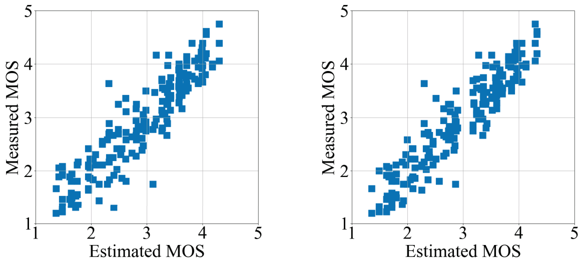


Fig. 7 MOSs of common PVSs.

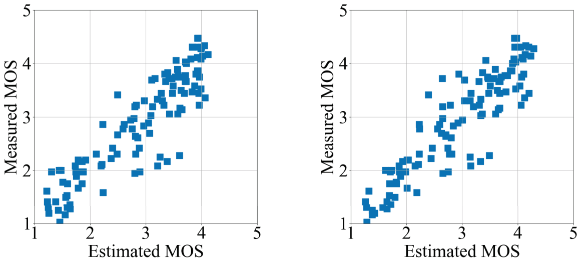
Table 6 RMSEs and PCCs.

		Training		Test	
		(Experiment 1 & 2)		(Experiment 3)	
Base model	Extending	RMSE	PCC	RMSE	PCC
P.1203	A	0.352	0.915	0.412	0.902
	B	0.399	0.889	0.419	0.900
	C	0.443	0.864	0.430	0.897
Y-model	A	0.321	0.930	0.408	0.904
	B	0.398	0.891	0.423	0.898
	C	0.407	0.884	0.428	0.895



(a) P.1203 model (b) Y-model

Fig. 8 Estimated MOS by using Model A for training data.



(a) P.1203 model (b) Y-model

Fig. 9 Estimated MOS by using Model A for test data.

[†] The "signal-based" term in this paper is "pixel-based" in the Recommendation P.1204.

Table 7 RMSEs for conditions in training data (Experiments 1 and 2).

	P.1203			Y-model		
Resolution	A	B	C	A	B	C
960	0.294	0.317	0.383	0.259	0.361	0.323
1280	0.288	0.305	0.313	0.274	0.304	0.312
1920	0.315	0.379	0.442	0.302	0.380	0.437
2560	0.562	0.604	0.629	0.557	0.629	0.652
3840	0.437	0.491	0.422	0.322	0.442	0.329
	P.1203			Y-model		
BR ratio	A	B	C	A	B	C
< 0.4	**0.306**	**0.398**	**0.458**	**0.299**	**0.420**	**0.455**
≥ 0.4	0.369	0.400	0.412	0.329	0.388	0.385
	P.1203			Y-model		
Delay time	A	B	C	A	B	C
< 3	0.376	0.447	0.450	0.327	0.451	0.432
≥ 3	0.327	0.352	0.392	0.309	0.348	0.378

Table 8 RMSEs for conditions in test data (Experiment 3).

	P.1203			Y-model		
Resolution	A	B	C	A	B	C
960	0.339	0.320	0.359	0.236	0.256	0.257
1280	0.353	0.378	0.378	0.331	0.354	0.378
1920	0.403	0.417	0.434	0.410	0.417	0.434
2560	0.498	0.466	0.493	0.447	0.480	0.505
3840	0.450	0.487	0.468	0.531	0.537	0.502
	P.1203			Y-model		
BR ratio	A	B	C	A	B	C
< 0.4	**0.428**	**0.451**	**0.462**	**0.446**	**0.500**	**0.517**
≥ 0.4	0.405	0.407	0.418	0.394	0.392	0.391
	P.1203			Y-model		
Delay time	A	B	C	A	B	C
< 3	0.392	0.398	0.413	0.391	0.406	0.419
≥ 3	0.416	0.418	0.432	0.401	0.410	0.415

MOSs estimated by the extended Model A based on P.1203 and Y-model. As shown in Figs. 8(a)–9(b), the shapes of scattered plots for both models are almost the same. From these results, it is shown that quality-estimation accuracy is improved by adding low-quality terms, the resolution, and the delay time. It is also shown that the proposed extension method can be effectively extended from either base model, so the proposed extension method is versatile.

To clarify the impacts to take into account the size of divided tiles (resolution), the quality of both tiles, and the delay time of quality switching, we investigated the impact of the resolution, the bitrate ratio (omnidirectional tile bitrate/divided tile bitrate), and the delay time on estimation accuracy. Tables 7 and 8 show the RMSEs of each extended model for the conditions in the training data and the test data. In all aspects of the resolution, the bitrate ratio, and the delay time, Model A achieved better quality-estimation accuracy than Models B and C for training data. The improvement of quality-estimation accuracy by using Model A is especially large when the bitrate ratio is less than 0.4 for training data and test data. When the bitrate ratio is low, the difference in quality between divided and omnidirectional tiles is large, and the influence of resolution and delay time is strong, so Model A, which calculates the weights based on resolution and delay time, is more accurate than Model B, which is extended by a simple weighted sum.

Table 9 Mean errors for SRCs in Experiments 1 and 2.

	P.1203			Y-model		
SRC	A	B	C	A	B	C
1-01	0.184	0.166	0.207	0.141	0.124	0.152
1-02	0.039	0.016	0.064	−0.01	−0.02	0.031
1-03	0.237	0.263	0.247	0.22	0.23	0.209
1-04	−0.098	−0.068	−0.075	−0.125	−0.094	−0.116
1-05	0.006	0.052	0.052	−0.026	0.021	−0.0
1-06	−0.08	−0.104	−0.071	−0.123	−0.144	−0.12
2-01	−0.24	−0.168	−0.157	−0.216	−0.124	−0.162
2-02	**0.55**	**0.521**	**0.505**	**0.537**	**0.549**	**0.52**
2-03	−0.113	−0.142	−0.158	−0.126	−0.114	−0.143
2-04	−0.071	−0.166	−0.156	−0.005	−0.133	−0.087
2-05	0.099	0.171	0.182	0.122	0.215	0.177
2-06	**−0.503**	**−0.598**	**−0.588**	**−0.436**	**−0.564**	**−0.518**

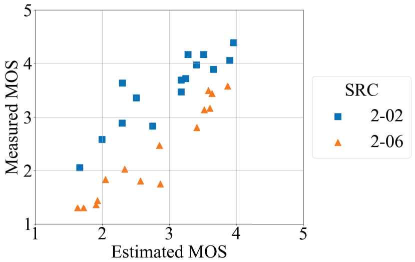


Fig. 10 Estimated MOS by using Model A of Y-model for the two SRCs.

Next, to examine the possibility of improving the quality-estimation accuracy of the proposed model, we investigated the quality factors that increase the quality-estimation error of the proposed model. Even if the HRCs are the same, the impact of coding degradation differs among SRC. Table 9 shows the average of errors ($measuredMOS - estimatedMOS$) for each SRC in Experiments 1 and 2 whose SRCs have lots of HRCs. All models have different error values, but the trends are consistent. For example, there is a tendency to estimate lower for SRC2-02 and higher for SRC2-06. Figure 10 shows the estimation results for ‘SRC:2-02’ and ‘SRC:2-06’ by using Model A of Y-model. The values of coordinates on the horizontal axis and the vertical axis are the estimated MOS and the measured MOS, respectively. The proposed model is a metadata layer, so it cannot take into account the feature of SRCs the same way as 2D models [24]. As a result, there is a discrepancy between the two contents. To address this issue, signal-based or bitstream-based models need to be studied. In general, these models can assess the impact of source on video quality. However, as described in Section 1, these models cannot feasibly be used at the clients because computational resource is needed to calculate and bitstream is encrypted just after encoding video. Therefore, to solve the problem, we have to study methods such as calculating the content features at the headend and sending them to the client for further computation.

6. Conclusion

To monitor the normality of tile-based omnidirectional video (ODV) services that provide music, sports, etc., we proposed extension methods of the two conventional 2D models (P.1203.1 mode 0 model and a model proposed by Yamagishi et al.) to tile-based ODV streaming services. To evaluate its quality-estimation accuracy, we conducted three subjective quality assessment experiments for training and testing. The verification results based on these experiments show the quality-estimation accuracy can be improved by taking into account the quality of divided (higher quality) and omnidirectional (lower quality) tiles, the resolution, (the size of the area of the divided tiles), and the delay time. In particular, where the bitrate ratio (i.e., the bitrate of divided tile/the bitrate of omnidirectional tile) is low, the difference in quality between divided and omnidirectional tiles is large, and the effects of resolution and the delay are significant, so accounting for these factors improves the estimation accuracy. Due to the limitations of the metadata layer models, the accuracy of quality estimation is degraded depending on the source contents.

In this paper, quality-estimation models of ODV streaming services were studied. There are also interactive services that use virtual reality (VR) technology, such as video that enables movement within the VR space (i.e., 6DoF video) and VR chat, and the proposed models cannot be applied to these services because their quality factors are not limited to video quality. In the future, we will study methods such as calculating the content features at the headend and sending them to the client for further computation to take into account the impact of SRCs. We will also study quality-estimation technologies for VR services with such interactivity.

References

[1] H. Kimata, D. Ochi, A. Kameda, H. Noto, K. Fukazawa, and A. Kojima, “Mobile and multi-device interactive panorama video distribution system,” Proc. 1st IEEE Global Conference on Consumer Electronics 2012, pp.574–578, Oct. 2012.

[2] D. Ochi, Y. Kunita, A. Kameda, A. Kojima, and S. Iwaki, “Live streaming system for omnidirectional video,” Proc. of 2015 IEEE Virtual Reality (VR), pp.349–350, March 2015.

[3] Y. Sun, A. Lu, and L. Yu, “Weighted-to-spherically-uniform quality evaluation for omnidirectional video,” IEEE Signal Process. Lett., vol.24, no.9, pp.1408–1412, June 2017.

[4] Y. Zhou, M. Yu, H. Ma, H. Shao, and G. Jiang, “Weighted-to-spherically-uniform SSIM objective quality evaluation for panoramic video,” Proc. 2018 14th IEEE International Conference on Signal Processing, pp.54–57, Aug. 2018.

[5] X. Liu, P. An, C. Meng, C. Yang, and X. Huang, “Multiscale WS-SSIM for panoramic video quality assessment,” Proc. SPIE/COS Photonics Asia, pp.96–101, Oct. 2020.

[6] Y. Guan, C. Zheng, X. Zhang, Z. Guo, and J. Jiang, “Pano: Optimizing 360° video streaming with a better understanding of quality perception,” Proc. ACM Special Interest Group on Data Communication, pp.394–407, Aug. 2019.

[7] P. Gao, P. Zhang, and A. Smolic, “Quality assessment for omnidirectional video: A spatio-temporal distortion modeling approach,” IEEE Trans. Multimedia, vol.24, pp.1–16, 2022.

[8] H.G. Kim, H.T. Lim, and Y.M. Ro, “Deep virtual reality image quality assessment with human perception guider for omnidirectional image,” IEEE Trans. Circuits Syst. Video Techno., vol.30, no.4, pp.917–928, 2020.

[9] Y. Liu, H. Yu, B. Huang, G. Yue, and B. Song, “Blind omnidirectional image quality assessment based on structure and natural features,” IEEE Trans. Instrum. Meas., vol.70, pp.1–11, 2021.

[10] M. Huang, Q. Shen, Z. Ma, A.C. Bovik, P. Gupta, R. Zhou, and X. Cao, “Modeling the perceptual quality of immersive images rendered on head mounted displays: Resolution and compression,” IEEE Trans. Image Process., vol.27, no.12, pp.6039–6050, 2018.

[11] M.S. Anwar, J. Wang, W. Khan, A. Ullah, S. Ahmad, and Z. Fei, “Subjective QoE of 360-degree virtual reality videos and machine learning predictions,” IEEE Access, vol.8, pp.148084–148099, Aug. 2020.

[12] G. Dimopoulos, I. Leontiadis, P. Barlet-Ros, and K. Papagiannaki, “Measuring video QoE from encrypted traffic,” Proc. 2016 Internet Measurement Conference, pp.513–526, Nov. 2016.

[13] J. Li, R. Feng, Z. Liu, W. Sun, and Q. Li, “Modeling QoE of virtual reality video transmission over wireless networks,” Proc. 2018 IEEE Global Communications Conference, pp.1–7, Dec. 2018.

[14] R.I.T. da Costa Filho, M.C. Luizelli, M.T. Vega, J. van der Hooft, S. Petrangeli, T. Wauters, F. De Turck, and L.P. Gaspary, “Predicting the performance of virtual reality video streaming in mobile networks,” Proc. 9th ACM Multimedia Systems Conference, pp.270–283, June 2018.

[15] Y. Urata, M. Koike, K. Yamagishi, N. Egi, and J. Okamoto, “Extension of ITU-T P.1203 model to tile-based omnidirectional video streaming,” Proc. EI HVEI 2021, pp.161-1–161-6, Jan. 2021.

[16] MPEG, “Omnidirectional Media Application Format,” WD on ISO/IEC 23000-20 Omnidirectional Media Application Format.

[17] Y.S. de la Fuente, G.S. Bhullar, R. Skupin, C. Hellge, and T. Schierl, “Delay impact on MPEG OMAF’s tile-based viewport-dependent 360 video streaming,” IEEE Trans. Emerg. Sel. Topics Circuits Syst., vol.9, no.1, pp.18–28, 2019.

[18] R. Schatz, A. Zabrovskiy, and C. Timmerer, “Tile-based streaming of 8K omnidirectional video: Subjective and objective QoE evaluation,” Proc. 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), pp.1–6, June 2019.

[19] R. Schatz, A. Sackl, C. Timmerer, and B. Gardlo, “Towards subjective quality of experience assessment for omnidirectional video streaming,” Proc. 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), pp.1–6, IEEE, June 2017.

[20] H. Duan, G. Zhai, X. Yang, D. Li, and W. Zhu, “IVQAD 2017: An immersive video quality assessment database,” Proc. 2017 International Conference on Systems, Signals and Image Processing (IWSSIP), pp.1–5, IEEE, 2017.

[21] ITU-T, “ITU-T Recommendation P.1203 - Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport,” Oct. 2017.

[22] Y. Zhang, P. Zhao, K. Bian, Y. Liu, L. Song, and X. Li, “DRL360: 360-degree video streaming with deep reinforcement learning,” Proc. IEEE Conference on Computer Communications (INFOCOM), pp.1252–1260, May 2019.

[23] N. Kan, J. Zou, K. Tang, C. Li, N. Liu, and H. Xiong, “Deep reinforcement learning-based rate adaptation for adaptive 360-degree video streaming,” Proc. 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.4030–4034, May 2019.

[24] K. Yamagishi and T. Hayashi, “Parametric quality-estimation model for adaptive-bitrate-streaming services,” IEEE Trans. Multimedia, vol.19, no.7, pp.1545–1557, July 2017.

[25] P. Lebreton and K. Yamagishi, “Transferring adaptive bit rate streaming quality models from H.264/HD to H.265/4K-UHD,” IEICE Trans. Commun., vol.E102-B, no.12, pp.2226–2242, Dec. 2019.

[26] ITU-T, “ITU-T Recommendation P.919 - Subjective test methodologies for 360° video on head-mounted displays,” Oct. 2020.

[27] Youncy-Hu, “Spherical SI/TI,” https://github.com/Youncy-Hu/Spherical-SI-TI

[28] T. Tominaga, T. Hayashi, J. Okamoto, and A. Takahashi, “Performance comparisons of subjective quality assessment methods for mobile video,” Proc. 2010 Second International Workshop on Quality of Multimedia Experience (QoMEX), pp.82–87, June 2010.

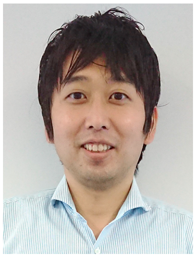
[29] ITU-T, “ITU-T Recommendation P.1204 - Video quality assessment

of streaming services over reliable transport for resolutions up to 4K," Jan. 2020.

Yuichiro Urata received his B.E. and M.E. degrees in engineering from University of Electro-Communications, Tokyo, Japan in 2009 and 2011. Since 2011, he has worked for NTT Laboratories in Tokyo, where has been engaged in the research of quality-estimation models for videos. He received the Research Encouragement Award (IEICE Technical Committee on Communication Quality) in Japan, in 2020.

Masanori Koike received his B.E. degree in mathematical engineering and information physics in 2015 and M.S. degree in information science and technology in 2017 from the University of Tokyo, Japan. Since joining NTT Laboratories in 2017, he has been engaged in research on quality of experience for VR video distribution.

Kazuhisa Yamagishi received his B.E. degree in electrical engineering from the Tokyo University of Science, Chiba, Japan, in 2001, and his M.E. and Ph.D. degrees in electronics, information, and communication engineering from Waseda University, Tokyo, Japan, in 2003 and 2013. He joined NTT Laboratories, Tokyo, Japan, in 2003. From 2010 to 2011, he was a visiting researcher with Arizona State University, Tempe, AZ, USA. His research interests include the development of objective quality-estimation models for multimedia telecommunications. He was a recipient of the Young Investigators' Award (IEICE) in Japan in 2007 and the Telecommunication Advancement Foundation Award in Japan, in 2008, the ITU-AJ Encouragement Awards in Japan, in 2017, and the TTC Award for Distinguished Service in Japan, in 2018.



Noritsugu Egi received B.E. and M.E. degrees from Tohoku University in 2003 and 2005. In 2005, he joined NTT Laboratories, where he has been engaged in the quality assessment of speech and audio. Currently, he is working on the quality assessment of video and web-browsing services over IP networks. He received the Young Researchers' Award (IEICE) in 2009.