Fundamental relationships between subjective quality, user acceptance, and the VMAF metric for a quality-based bit rate ladder design for over-the-top video streaming services

Andreas Kah[a], Christopher Friedrich[a], Thomas Rusert[b], Christoph Burgmair[b], Wolfgang Ruppel[a], and Matthias Narroschke[a]

[a]RheinMain University of Applied Sciences, Wiesbaden, Germany
[b]Joyn GmbH, Munich, Germany

ABSTRACT

A quality-based bit rate ladder design for over-the-top video streaming services is presented. Following the design criterion of maximizing subjective quality under the constraint of minimizing storage costs, the bit rate ladder is defined by three parameters. The first parameter is the lowest VMAF score at which a video signal is on average subjectively indistinguishable from the original video signal. Following the international recommendation ITU-R BT.500, extensive subjective tests were carried out to evaluate the fundamental relationships between the subjective quality and the VMAF score using a 4K OLED TV environment. Based on the test results, this VMAF score is set to 95. The second parameter is the lowest VMAF score being accepted on average by more than 50 % of the users for watching video signals of free streaming services. Additional tests yield in setting this VMAF score to 55. The third parameter is the maximum difference of two VMAF scores, for which the associated subjective qualities are approximately the same on average. In a third test, this difference is determined to be 2. This results in an ideal bit rate ladder providing each video signal in 21 qualities associated to the VMAF scores 95, 93, …, 57, 55. This bit rate ladder design can be applied to complete video signals occurring in per-title encoding strategies or to individual scenes of video signals occurring in per scene or shot-based encoding strategies. Applications using less than 21 renditions for this range, may suffer from impaired subjective quality.

Keywords: Video streaming, VMAF, subjective quality, bit rate ladder, ABR, OTT

1. INTRODUCTION

As the use of over-the-top (OTT) video streaming services continues to soar, the importance of adaptive bit rate (ABR) streaming is growing continuously. In ABR streaming, digital video signals are encoded at various bit rates $R_1, \dots, R_k, \dots, R_K$ and associated qualities $Q_1, \dots, Q_k, \dots, Q_K$, and provided to a plurality of end user devices via the internet using content delivery networks (CDNs). An encoded digital video signal of one particular bit rate and associated quality is denoted as a rendition (R_k, Q_k), the set of all K renditions $(R_1, Q_1), \dots, (R_K, Q_K)$ as a bit rate ladder.

Typical operational quality bit rate curves for digital video signals are shown in Figure 1. A first major property of digital video signals is that the quality Q increases with the bit rate R. A second major property is that the bit rate R of high complex content is higher than the bit rate of low complex content at the same quality Q.

In the following, the K renditions are ordered with respect to their associated bit rates, such that $R_1 < \dots < R_k < \dots < R_K$ applies. As a consequence of the first major property, the associated qualities are also ordered, i.e. $Q_1 < \dots < Q_k < \dots < Q_K$. Based on the renditions provided by the bit rate ladder, each end user device can request and stream the content from the CDN via the internet at a bit rate suiting the individual transmission rate T of the user's internet connection. There are various rendition selection strategies, whereas one simple strategy is to select the rendition (R_p, Q_p) for playout, which has the largest possible bit rate smaller than T, i.e.:

$$R_p(T) = \max_{k=1,\dots,K \mid R_k \leq T} R_k \ . \tag{1}$$

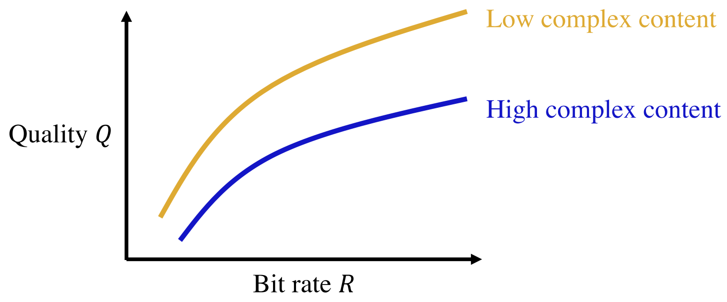

Figure 1: Example operational quality bit rate curves for coded video signals of low and high complex content.

With the selection strategy according to equation (1), the client avoids re-buffering during content playback and thus enables continuous playback of the video. More sophisticated strategies may alternate between the two renditions (R_p, Q_p) and (R_{p+1}, Q_{p+1}) trying to utilize the transmission rate more effectively. For these strategies the quality also alternates between Q_p and Q_{p+1}. In the following, the considerations in this paper are first focused on the simple strategy according to equation (1) and subsequently extended to these more sophisticated strategies.

In the past, ABR streaming providers have generally used fixed bit rate ladders[1], i.e. a set of predefined bit rates across all their video contents, irrespective of the video quality resulting from the bit rates. Nowadays, this approach is increasingly substituted by a quality-based bit rate ladder design taking the individual content into account. Such content dependent bit rate ladder designs are often denoted as per-title encoding[2], per-scene encoding, or shot-based encoding[3]. Hereby, per-title encoding refers to a bit rate ladder optimization across full content assets while per-scene and shot-based encoding refer to a more granular bit rate ladder optimization based on individual scenes in each content asset. Thus, for per-title, per-scene and shot-based encoding, the renditions $(R_1, Q_1), \dots, (R_K, Q_K)$ and often also the number of renditions K are set individually for each title or scene.

Due to the use of a set of discrete bit rates $R_1, \dots, R_k, \dots, R_K$, the quality Q_p of the video played out by the client is typically lower than the quality $Q(T)$ that would be possible for the client's transmission rate T if a rendition with $R_k = T$ was available in the bit rate ladder. In this paper, the difference is denoted as the quality loss:

$$\Delta Q(T) = Q(T) - Q\left(R_p(T)\right). \qquad (2)$$

It may happen, that the video received at the quality $Q\left(R_p(T)\right)$ is subjectively worse than the video at the quality $Q(T)$, which the client could receive if it was available in the bit rate ladder. This happens in particular, if the bit rates R_p and R_{p+1} and thus the associated qualities Q_p and Q_{p+1} differ significantly. Based on the qualities Q_p and Q_{p+1}, the upper limit of the quality loss ΔQ_{max} is defined as:

$$\Delta Q_{max}(Q_p, Q_{p+1}) = Q_{p+1} - Q_p\,. \qquad (3)$$

In Figure 2 below, the quality loss ΔQ and the maximum quality loss ΔQ_{max} are illustrated.

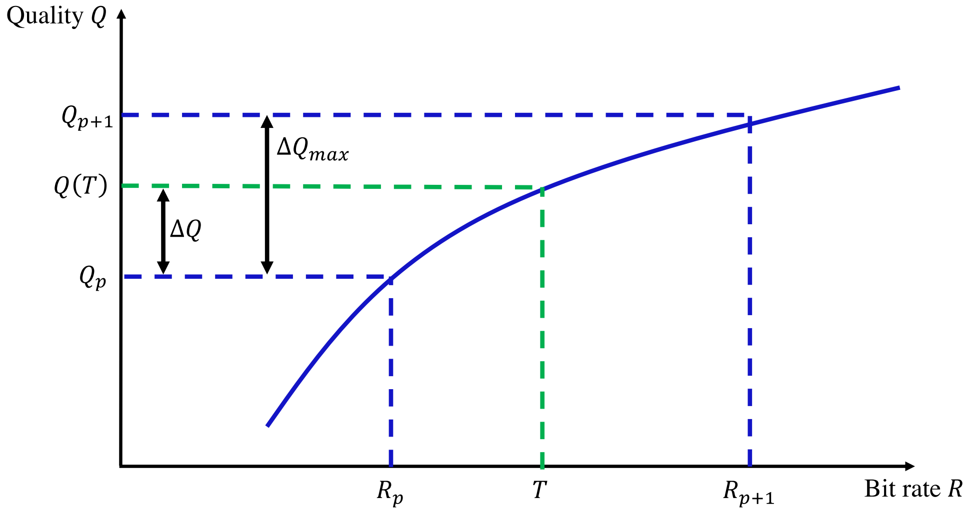

Figure 2: Illustration of the quality loss $\Delta Q(T)$, see equation (2), and the maximum quality loss ΔQ_{max}, see equation (3).

Each service provider aiming at delivering video of the highest possible quality should keep the maximum quality loss ΔQ_{max} as small as possible. As clients may be connected with transmission rates T ranging from very small ones, e.g. for mobile clients connected via cellular networks, to very large ones, e.g. clients connected via fiber channel, a small ΔQ_{max} for all these clients can only be realized by providing a large number K of renditions. However, a large number of renditions also results in high encoding and storage costs for the streaming provider. Consequently, the upper limit ΔQ_{max} needs to be chosen as a compromise of quality loss and these costs considering the subjective quality perceived by end users.

In the following Section 2, a quality-based bit rate ladder design for over-the-top video streaming services is presented, which follows the above principles. The parameters of the presented bit rate ladder design are set according to fundamental relationships between the subjective quality, the user acceptance, and the VMAF metric[4] determined in subjective tests as described in Section 3. The paper closes with conclusions in Section 4.

2. QUALITY-BASED BIT RATE LADDER DESIGN FOR OVER-THE-TOP VIDEO STREAMING SERVICES

In order to automate the bit rate ladder generation, the subjective quality perceived by end users is estimated by the VMAF-metric[4]. Hereby, a score ranging from 0 to 100 is computed for each frame of a video signal, where 0 corresponds to low and 100 to high subjective quality. In the following, the mean value of the VMAF scores over all frames of a given video signal is defined as the VMAF score of that video signal, i.e. $Q = VMAF$. With this definition, the quality loss of equation (2) turns into

$$\Delta VMAF(T) = VMAF(T) - VMAF\left(R_p(T)\right). \tag{4}$$

A bit rate ladder, which limits the quality loss $\Delta VMAF(T)$ to $\Delta VMAF_{max}$ for all transmission rates T with $\mathrm{R}_1 \leq T \leq \mathrm{R}_\mathrm{K}$ is depicted in Figure 3. The limitation is achieved by constraining the maximum quality loss between each pair of renditions $(R_\mathrm{k}, VMAF_\mathrm{k})$ and $(R_{k+1}, VMAF_{k+1})$, with $k = 1, \ldots, K-1$, to $\Delta VMAF_\mathrm{max}$. $\Delta VMAF_\mathrm{max}$ is the first parameter of the presented design. It should ideally be set small enough so that the subjective quality of the video signal is the same for neighboring renditions k and $k+1$. This way, any potential quality loss due to not fully exploiting the available transmission rate T can be avoided. In addition, switching between neighboring renditions remains subjectively unnoticeable, so that the temporal consistency of the video playback is maximized.

The second parameter of the presented design is the maximum quality $VMAF_K$ provided by the bit rate ladder. Ideally, it should be set to the lowest possible VMAF score for which a video signal is subjectively indistinguishable from the original video signal. This minimizes storage costs while still ensuring optimal subjective quality.

The third parameter of the presented bit rate ladder is the minimum quality $VMAF_1$ provided by the bit rate ladder. Ideally, it should be set to the lowest VMAF score for which video is acceptable for watching by the users. Such strategy minimizes encoding and storage costs by avoiding renditions not being watched by users due to unacceptable subjective quality.

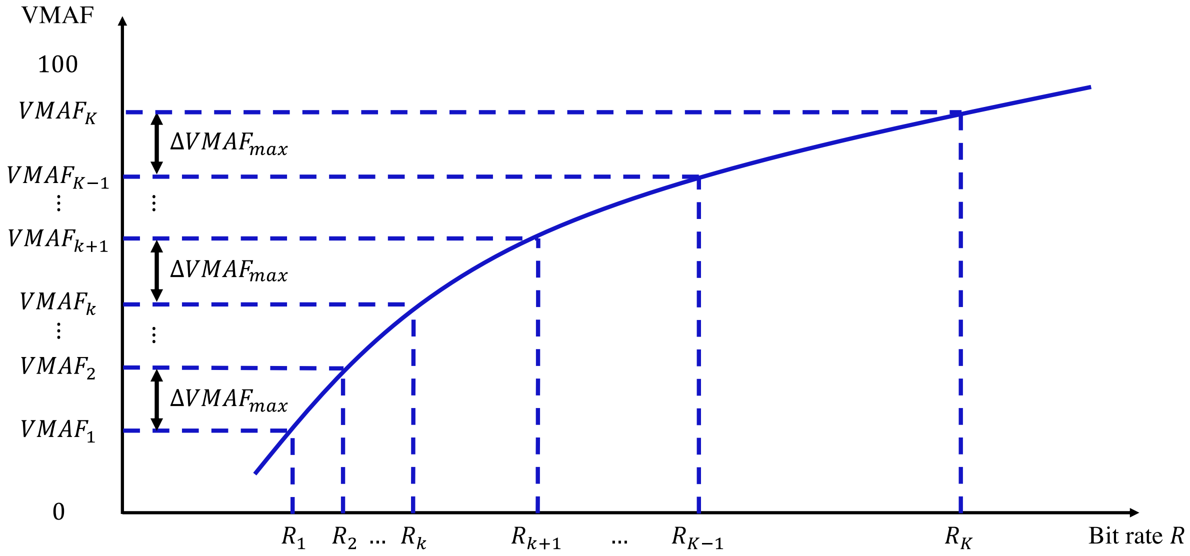

Figure 3: Bit rates $R_1, \ldots, R_K$ and associated VMAF scores $VMAF_1, \ldots, VMAF_K$ of a bit rate ladder limiting the quality loss according to equation (4) to $\Delta VMAF_{max}$.

3. DETERMINATION OF THE BIT RATE LADDER PARAMETERS BASED ON FUNDAMENTAL RELATIONSHIPS BETWEEN SUBJECTIVE QUALITY, USER ACCEPTANCE, AND THE VMAF METRIC

As outlined in this Section, subjective tests were designed and conducted in order to determine the fundamental relationships between subjective quality, user acceptance, and the VMAF metric, based on which the bit rate ladder parameters can be properly set.

First, a set of eight video signals was composed on which the subjective tests were carried out. This composition considers that the content is of different complexity comprising various genres including sports, animation, and comedy. Six of the video signals originate from the set of video signals being used in the international ISO and ITU standardization[5] and two of them are proprietary video signals. Spatial resolutions of 1920 x 1080 (HDTV) and 3840 x 2160 (4K) luminance samples with a 4:2:0 subsampling of the chrominances were used. Each of the video signals has a duration of 10 seconds. Table 1 gives an overview of the set together with corresponding properties.

In order to determine the parameter $VMAF_K$ of the bit rate ladder design presented in Section 2, the set of video signals was encoded resulting in the VMAF scores 92, 93, 94, 95, 96, and 97. Based on a subjective test with 20 observers following the Double Stimulus Impairment Scale test method of JVET-H1002[6] and ITU-R BT.500[7], a mean opinion score (MOS) along with a corresponding 95 % confidence interval was determined for the video signals of each of these VMAF scores as well as for the original video signals. In Figure 4, the basic test cell of this test is shown. An 11-grade numerical scale ranging from 0 (very disturbing impairments) to 10 (imperceptible impairments) was used to evaluate the video signals. None of the observers was rejected as an outlier resulting from applying the screening method described in ITU-R BT.500[7], and for none of these observers, color blindness was detected when applying an Ishihara color test[8].

A	Original video signal	B	Video signal to be voted	Vote #
0.5 sec.	10 sec.	0.5 sec.	10 sec.	4.5 sec.

Time

Figure 4: Structure of the basic test cell based on the Double Stimulus Impairment Scale test method defined in JVET-H1002[6].

For coding, the HEVC reference coder in the version 16.22[9] was used. The Random Access configuration defined by the international standardization groups of the ISO and the ITU[5] was applied as it represents a typical configuration for OTT video streaming services. In order to precisely achieve the desired VMAF score for a video signal, the quantization parameter QP and a one-time increase of it per encoding was set as described in JVET-H1002[6]. For calculation of the VMAF score, the implementation[4] was used in the version 1.5.3. For HDTV signals, the HDTV-VMAF-model for 4K signals, the 4K-VMAF-model was applied.

The viewing environment was set up according to ITU-R BT.500[7] and JVET-H1002[6]. A Panasonic 4K OLED TV TX-55HZW2004 with 55" diagonal was used as display. All decoded HDTV signals were scaled to 4K resolution prior to displaying using bicubic interpolation. The room was carefully protected from external visual and acoustic disturbances, and the walls were colored in dark gray, using chromaticity of D65. During the test, the general light was turned off and a uniform light was placed behind the display. The ratio of the background luminance and the peak luminance of the display was adjusted to approximately 0.15. Following Corona virus related regulations, only one observer participated at a time. In accordance with JVET-H1002[6], the distance from the observer to the display was adjusted two times the height of the active part of the display.

Table 1: Selected video signals and their properties for the subjective tests.

Video signal	Source	Spatial Resolution [Luminance samples]	Frame rate [fps]
BasketballDrive	ISO / ITU	1920x1080	50
Animation	Proprietary	1920x1080	25
Comedy	Proprietary	1920x1080	25
MarketPlace	ISO / ITU	1920x1080	60
Campfire	ISO / ITU	3840x2160	30
CatRobot1	ISO / ITU	3840x2160	60
FoodMarket4	ISO / ITU	3840x2160	60
ParkRunning3	ISO / ITU	3840x2160	50

Figure 5 shows the mean opinion scores with the corresponding 95 % confidence intervals determined for each of the considered VMAF scores and for the original video signals. As the number of observers is relatively small, the calculated mean opinion scores are associated with the uncertainty that the true mean opinion score may be different. The uncertainty is expressed by the confidence interval defining a range, in which the true mean opinion score is located with a probability of 95 %. It can be seen that for all VMAF scores lower than 95, the confidence intervals do not overlap with the confidence interval for the original video signals. Thus, the difference of the subjective qualities of the corresponding video signals is noticeable in average. For all VMAF scores larger or equal to 95, the confidence intervals overlap with the confidence interval for the original video signals. This indicates that it is uncertain that the true and the derived mean opinion scores and thus the subjective qualities are different or equal. However, this uncertainty of the MOS is 0.8 in maximum as can be derived from the difference between the upper limit of the confidence interval of the original signals, which 9.2, and the lower limit of the confidence interval for VMAF = 95, which is 8.4.

Assuming equal subjective quality, the parameter $VMAF_K$ needs to be set to

$$VMAF_K \geq 95\,. \tag{5}$$

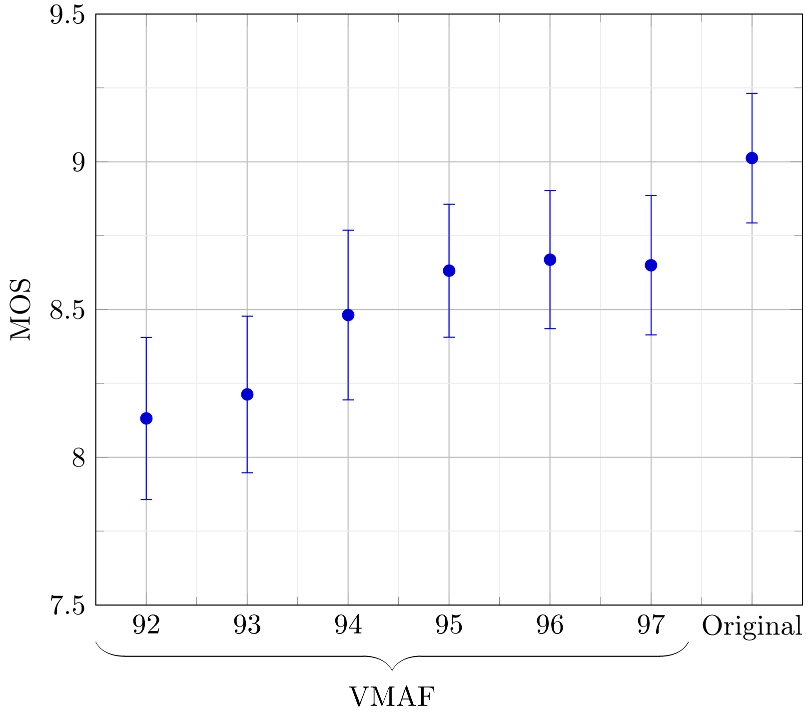

Figure 5: Measured mean opinion score MOS along with the 95 % confidence intervals for the VMAF scores 92, 93, 94, 95, 96, and 97 as well as for the original video signals.

In order to determine the parameter $VMAF_1$ of the bit rate ladder design presented in Section 2, the set of video signals was encoded resulting in the VMAF scores covering almost the full range, i.e. 20, 25, 30,…, 95. For each video signal and each VMAF score as well as for the original video signal, another 20 observers, which passed the Ishihara color test, were asked to select one out of the following three choices:

1. Subjective quality is acceptable for permanent viewing.
2. Subjective quality is acceptable only for a temporary impairment of approx. 30 seconds.
3. Subjective quality is not acceptable at all.

Permanent viewing implies a continuous playout at the same video quality, temporary impairment as playout with a temporary impaired video quality, which may occur due to a reduction of the transmission rate for a short period. Each observer was asked to perform the selection twice, first under the assumption of a free streaming service and second under the assumption of a paid streaming service that charges typical subscription fees. For each observer, all video signals were displayed in an individualized random order. In Figure 6, the basic test cell for the acquisition of one selection is illustrated.

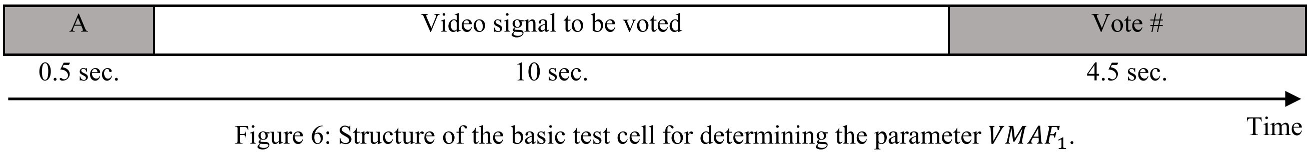

Figure 6: Structure of the basic test cell for determining the parameter $VMAF_1$.

The same viewing environment and the same HEVC reference coder was used for the determination of the parameter $VMAF_K$. However, in order to be able to generate video signals resulting in low VMAF scores, additional spatial subsampling at the encoder side in combination with bicubic interpolation at the decoder side was applied for some of the video signals, using FFmpeg[10].

From the selections of the observers, acceptance rates AR were derived for every considered VMAF score. Here, the four scenarios resulting from any combination of permanent viewing and temporary impairment as well as of free and paid video streaming services were distinguished. For this purpose, the value 1 was assigned to the selection “acceptable subjective quality” and the value 0 for the choice “not acceptable subjective quality”. Subsequently, the acceptance rate was determined as the mean value of the assignments.

Figure 7 shows the derived acceptance rates for all four scenarios. From Figure 7, the following findings can be derived for the permanent viewing:

- The VMAF score for paid streaming services must be 10 – 15 higher than for free streaming services in order to achieve the same acceptance rate.
- To achieve an acceptance rate of at least 50 % for permanent viewing, it is necessary to play out video signals of a VMAF score larger or equal than 70 for free services and larger or equal than 85 for paid services. For these VMAF scores, the lower limits of the confidence intervals are all indicating acceptance rates larger than 50 %.
- For paid video services, the acceptance rate for video signals of VMAF scores below 55 is approximately zero. For free services, this threshold is around 30.
- An acceptance rate of at least 80 % is reached by a VMAF score of 80 for free and of 95 for paid services.

For temporary impairment, the same acceptance rates as for permanent viewing are achieved with VMAF scores being approximately 10 – 15 lower.

In order to achieve an acceptance rate of more than 50 % for both situations, permanent viewing and temporary impairment, the parameter $VMAF_1$ needs to be set to

$$VMAF_{1,AR50,Free} = 55 \qquad (6)$$

for free video services, and to 70 for paid ones. These values equal the lowest VMAF scores in the right diagram of Figure 7, for which the lower limit of the confidence interval indicates acceptance rates larger than 50 %.

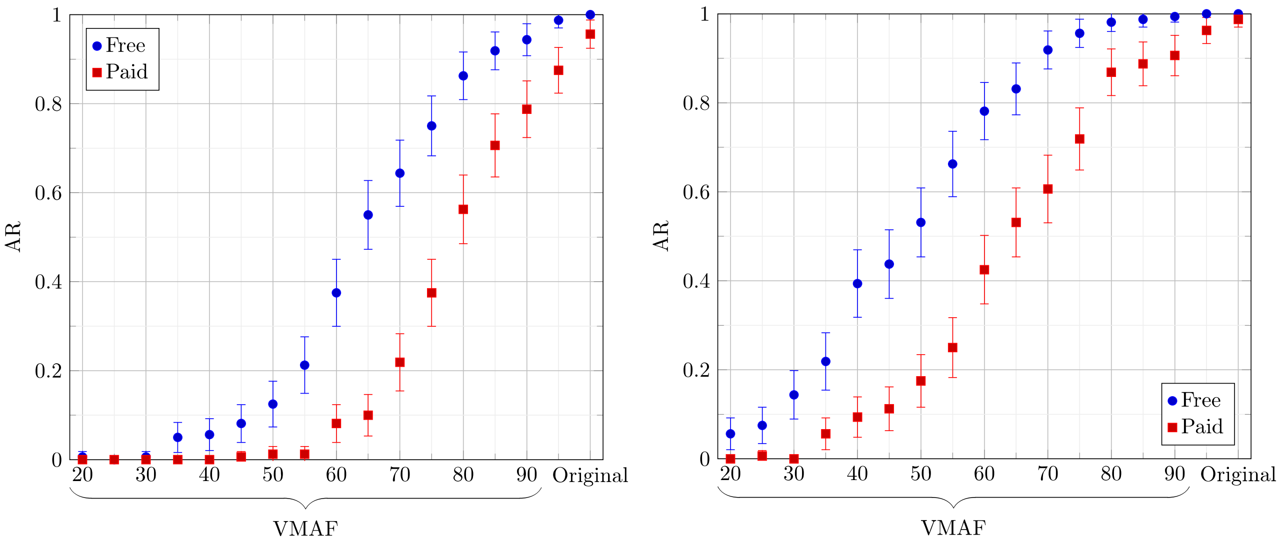

Figure 7: Acceptance rates AR for free and paid video streaming services.
Left: Permanent viewing. Right: Temporary impairment of approximately 30 seconds.

In order to determine the parameter $\Delta VMAF_{\mathrm{max}}$ of the bit rate ladder design presented in Section 2, three operating points were considered, a first one at low quality of $VMAF_{\mathrm{OP1}} = 60$, a second one of mid quality of $VMAF_{\mathrm{OP2}} = 75$, and a third one of high quality of $VMAF_{\mathrm{OP3}} = 90$. These three operating points were set in accordance with the range given by $VMAF_{1,AR50,Free} = 55$ of equation (6) and $VMAF_K = 95$ of equation (5). For each operating point $VMAF_{\mathrm{OPi}}$, with i = $1, \ldots, 3$, the set was encoded resulting in seven VMAF scores of $VMAF_{\mathrm{OPi}} - 5, VMAF_{\mathrm{OPi}} - 2, VMAF_{\mathrm{OPi}} - 1, VMAF_{\mathrm{OPi}}, VMAF_{\mathrm{OPi}} + 1, VMAF_{\mathrm{OPi}} + 2, VMAF_{\mathrm{OPi}} + 5$. Following the same test procedure as for the determination of $VMAF_K$, see above, a mean opinion score along with a corresponding 95 % confidence interval was determined for the video signals of each of these VMAF scores. The basic test cell was also the same for this test, see Figure 4. Additional 20 observers were asked to conduct the test. Due to the large amount of video signals to be voted, the total length of the test exceeded the maximum duration for such a test of half an hour as defined in ITU-R BT.500[7]. Following the advice of ITU-R BT.500 for such situations, appropriate breaks were inserted for regeneration.

The same viewing environment and the HEVC reference coder were used as above. For some video signals, spatial subsampling was applied to reach the desired VMAF score.

Figure 8 summarizes the measurements. From this figure, it can be concluded that the dependency of the mean opinion score and the VMAF is approximately linear. This justifies applying the same $\Delta VMAF_{\mathrm{max}}$ over the whole quality range from $VMAF_{1,AR50,Free} = 55$ to $VMAF_K = 95$ as presented in Section 2, see Figure 3.

In order to derive the parameter $\Delta VMAF_{\mathrm{max}}$, all pairs of two different VMAF scores are evaluated. Denoting for each pair the lower score as $VMAF_{low}$ and the higher one as $VMAF_{high}$, the difference of both scores can be interpreted as the parameter $\Delta VMAF_{\mathrm{max}}$ under evaluation. From Figure 8 it can be drawn that the confidence intervals associated to $VMAF_{low}$ and $VMAF_{high}$ do not overlap for most pairs with differences $VMAF_{\mathrm{high}} - VMAF_{\mathrm{low}} > 2$, compare e.g. for $VMAF_{low} = 85$ and $VMAF_{high} = 88$. Consequently, for these pairs a difference of the subjective quality is definitely noticeable on average. For all pairs resulting in differences $VMAF_{\mathrm{high}} - VMAF_{\mathrm{low}} \leq 2$, the confidence intervals overlap. This indicates that the corresponding subjective qualities are the same on average up to an uncertainty. On the MOS scale, this uncertainty is 1.4 in maximum, measured between the lower limit of the confidence interval associated to VMAF = 58 and the upper limit of the confidence interval associated to VMAF = 60.

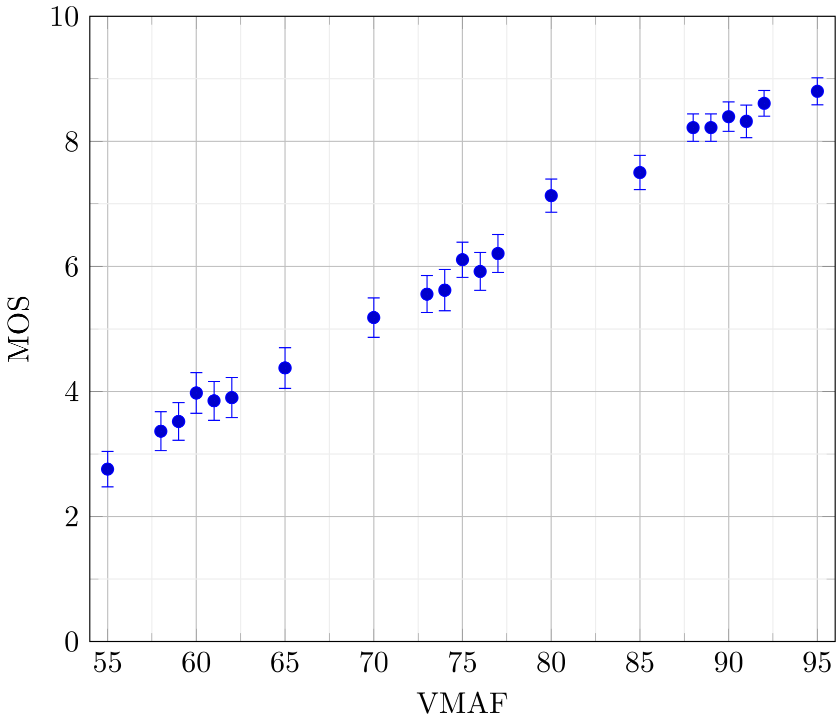

Figure 8: Measured mean opinion score MOS along with the 95 % confidence intervals versus VMAF score.

Assuming the same subjective quality, the parameter $\Delta VMAF_{max}$ needs to be set to

$$\Delta VMAF_{max} \leq 2\,. \tag{7}$$

Ideally setting $VMAF_{K,ideal} = 95$, i.e. as small as possible, and $\Delta VMAF_{max,ideal} = 2$, i.e. as large as possible, leads to minimum storage costs and a minimum number of required renditions:

$$K_{ideal} = \frac{VMAF_{K,ideal} - VMAF_{1,AR50,Free}}{\Delta VMAF_{max,ideal}} + 1 = \frac{95-55}{2} + 1 = 21\,. \tag{8}$$

Consequently, this bit rate ladder provides each video signal in 21 qualities associated to the VMAF scores 95, 93, …, 57, 55. To exactly reach the VMAF scores of this bit rate ladder, large encoding effort is typically necessary. In order to reduce the encoding effort, parameters $VMAF_K > 95$, or $\Delta VMAF_{max} < 2$ may be accepted while generating the bit rate ladder. However, this increases the number of renditions and thus the storage costs.

In today's applications, typically less than 21 renditions are used in order to reduce the encoding and storage costs. Often, only 9 - 12 renditions[1] are applied. For 9 renditions covering a range from 55 to 95, this bit rate ladder design requires

$$\Delta VMAF_{max,9\,renditions} = \frac{95-55}{9-1} = 5\,. \tag{9}$$

However, the discussions above reveal that using a $\Delta VMAF_{max,9\,renditions} = 5$ may lead to quality loss, which is subjectively visible on average. According to the measurements summarized in Figure 8, the difference in the MOS is 1.6 in maximum, compare the difference between the limits of the confidence intervals for the VMAF scores 75 and 80.

4. CONCLUSION

In this paper, a quality-based bit rate ladder design for over-the-top video streaming services is presented. Following the design criterion of maximizing subjective quality under the constraint of minimizing storage costs, the bit rate ladder is defined by three parameters. The first parameter is the lowest possible VMAF score at which a video signal is on average subjectively indistinguishable from the original video signal. Following the international recommendation ITU-R BT.500 and using a 4K OLED TV environment, extensive subjective tests were carried out to evaluate the fundamental relationships between the subjective quality and the VMAF score. Based on the test results, this VMAF score is set to 95. The second parameter is the lowest VMAF score being accepted on average by more than 50 % of the users for watching video signals. Additional test show that this VMAF score is 55 for free streaming services and 70 for paid ones. The third parameter is the maximum difference of two VMAF scores, for which the associated subjective qualities are approximately the same on average. In a third test, this difference is determined to be 2. For free video streaming services, this results in an ideal bit rate ladder providing each video signal in 21 qualities associated to the VMAF scores 95, 93, …, 57, 55. The presented bit rate ladder design can be applied to both, to complete video signals occurring in per-title encoding strategies or to individual scenes of video signals occurring in per scene or shot-based encoding strategies. If less than these 21 renditions are used to represent the range from VMAF = 55 to VMAF = 95, as may be done in certain applications, an impairment of the subjective quality may become noticeable on average.

ACKNOWLEDGMENTS

The authors thank the German Federal Ministry of Education and Research for funding this work via the program FHprofUnt 2018, contract number 13FH152PX8.

REFERENCES

[1] Apple Developer Documentation., "HLS Authoring Specification for Apple Devices," 22 June 2020, <https://developer.apple.com/documentation/http_live_streaming/hls_authoring_specification_for_apple_devices> (17 June 2021).
[2] Aaron, A., Li, Z., Manohara, M., De Cock, J. and Ronca, D., "Per-Title Encode Optimization," Netflix Technol. Blog, 14 December 2015, <https://netflixtechblog.com/per-title-encode-optimization-7e99442b62a2> (17 June 2021).
[3] Katsavounidis, I., "Dynamic optimizer — a perceptual video encoding optimization framework," Netflix Technol. Blog, 5 March 2018, <https://netflixtechblog.com/dynamic-optimizer-a-perceptual-video-encoding-optimization-framework-e19f1e3a277f> (26 May 2020).
[4] Netflix, Inc., "VMAF," GitHub, 2 June 2021, <https://github.com/Netflix/vmaf> (2 June 2021).
[5] Bossen, F., Boyce, J., Suehring, K., Li, X. and Seregin, V., "JVET-T2010: VTM common test conditions and software reference configurations for SDR video," 28 October 2020, <https://jvet-experts.org/doc_end_user/current_document.php?id=10545> (23 June 2021).
[6] Segall, A., Baroncini, V., Boyce, J., Chen, J. and Suzuki, T., "JVET-H1002: Joint Call for Proposals on Video Compression with Capability beyond HEVC," 8 November 2017, <https://jvet-experts.org/doc_end_user/current_document.php?id=3361> (2 June 2021).
[7] ITU-R., "Rec. BT.500-14: Methodologies for the subjective assessment of the quality of television images" (2019).
[8] Ishihara, S., [Tests for Colour Blindness], Handaya Hongo Harukich, Tokyo (1917).
[9] JVET., "HEVC HM reference software / HM-16.22," GitLab, <https://vcgit.hhi.fraunhofer.de/jvet/HM/-/tree/HM-16.22> (2 June 2021).
[10] FFmpeg., "FFmpeg," GitHub, 2 June 2021, <https://github.com/FFmpeg/FFmpeg> (2 June 2021).