EBU

OPERATING EUROVISION AND EURORADIO

TECHNICAL REVIEW

A lightweight QoE evaluation model for OTT media services

FEBRUARY 2019

Dr Francesco De Angelis, EBU

FOREWORD

The main purpose of an EBU Technical Review is to critically examine new technologies or developments in media production or distribution. All Technical Reviews are reviewed by one (or more) technical experts at the EBU or externally and by the EBU Technical Editions Manager. Responsibility for the views expressed in this article rests solely with the author(s).

To access the full collection of our Technical Reviews, please see: tech.ebu.ch/publications

If you are interested in submitting a topic for an EBU Technical Review, please contact: tech@ebu.ch

ABSTRACT

This article introduces a model to evaluate the Quality of Experience (QoE) for Over-The-Top (OTT) audio-visual services. The model is based on several cutting-edge methods presented in the literature and includes further generalisations to better approximate the Quality of Experience perceived by users in Dynamic Adaptive Streaming over HTTP (DASH) sessions.

The model is flexible and modular and its input parameters can be customised to consider several Key Performance Indicators (KPIs) for Quality of Service. Its reduced complexity makes it a suitable candidate for lightweight assessment of QoE on heterogeneous devices such as TV sets and mobile devices, for example by implementing it in the form of a JavaScript library to integrate with off-the-shelf video players.

INTRODUCTION

Over the last couple of decades, average home and mobile Internet access speeds have been increasing sufficiently to allow access to high quality audio-visual streamed content on fixed and handheld devices.

The availability of affordable, powerful devices with high-speed Internet access, together with new trends in media consumption, have ensured optimal conditions for the emergence of new streaming platforms such as Netflix, YouTube and Amazon Prime (and many other online services delivered over the open Internet), which currently drive the growth of media consumption in the Internet.

The level of satisfaction achieved in consuming this audio-visual content plays a fundamental role in the success of streaming services. In this context, the term Quality of Experience (QoE) indicates the degree of satisfaction or annoyance perceived by users during the playout process.

Quantification of QoE is an important activity for offline and run-time optimisation of distribution chains and content, both in managed and unmanaged networks such as IPTV networks and the open Internet. Assessing the QoE perceived by the audience gives useful information on which parameters of the delivery chain and content to ‘tune’ in order to maximize the user's satisfaction. In particular, run-time optimisation can be carried out while the audio-visual service is active, which demands that the assessment process must enable the measurement of QoE during the consumption of content.

In the media services domain, image parameters have always been considered the reference point to evaluate the overall quality level perceived by users. However, image quality is only one fundamental aspect of QoE; for example, delays and stalls[1] above a certain tolerance threshold can significantly affect the user's perception even in case of high-definition artefact-free images. As a consequence, Quality of Images (QoI) turns out to be a fundamental aspect of Quality of Experience, but such terms cannot be considered synonymous as several additional factors can affect the whole experience.

Subjective QoE evaluations such as surveys are the most effective methodologies to assess Quality of Experience, as they try to consider all the fundamental elements affecting user's experience by gathering scores and opinions provided by human observers. These techniques usually generate more representative results compared to pure technical indicators; unfortunately, they also represent the most expensive and time consuming way to assess audio-visual services.

[1] A **stall** is when the playout process is blocked due to insufficient data in the internal buffer of the player.

Some limitations can be overcome by resorting to crowd-based testing but even this technique presents several disadvantages. Moreover, strict time constraints are introduced when the QoE assessment is performed for run-time adaptation.

Remote objective QoE evaluations can be carried out algorithmically on users' devices. Instantaneous QoE estimates are obtained for on-the-fly adaptation of distribution chains and adaptive media playout on users' devices.

During the last decade, several objective QoE methods have been proposed in the literature, having different degrees of complexity and interoperability. Nowadays the main efforts in this domain head toward the development of algorithms compatible with heterogeneous devices such as TV sets and mobile devices, possibly promoting integration with off-the-shelf video players in order to avoid expensive and invasive deployments.

FROM QUALITY OF IMAGES TO QUALITY OF EXPERIENCE

Historically, objective evaluations of image quality have been divided into the following categories:

(i) full reference (FR) metrics, when the original non-distorted content is available for comparison with the perceived one;

(ii) no reference (NR) metrics, when the analysis is effected without access to the reference content and,

(iii) reduced reference (RR) metrics, acting on a portion of data extracted or generated from the original content.

Clearly, full reference methods are better to assess image quality at the expense of cost and flexibility. They are applicable in a limited number of controlled scenarios (e.g. laboratories) due to the fact that a high fidelity copy of the original media resource must be available during the test. In broadcasting scenarios the Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), Video Quality Metric (VQM) and Structural Similarity (SSIM) [3, 4] are the most important video quality metrics; they are easy to calculate on modern devices but they can detect a restricted number of distortions compared to the Human Visual System (HVS). This is due to the fact the HVS has a high level of sensitivity to several image parameters, such as colours, local image contrast, light level and eccentricity. In this respect PSNR, MSE, VQM and SSIM paved the way for the development of advanced metrics such as Visual Information Fidelity (VIF) [5], able to take into account more aspects of the HVS during the assessment.

All the aforementioned video quality metrics are still used in software and devices for quality analysis; nonetheless, they do not represent good perceptual models for the assessment of QoE in IP-based distribution chains and even less for Over-The-Top (OTT) media services. In fact, nowadays a relevant amount of video traffic is delivered through reliable transport protocols (e.g. in Dynamic Adaptive Streaming over HTTP -DASH-, based on TCP), which ensure that frames are not disrupted during transfers. As a consequence, traditional artefacts such as blockiness, blurring, jerkiness and slicing [6] are usually solved by error recovery algorithms at the expense of longer delays and stalls. Moreover, context plays a key role in Quality of Experience: for example, the same undesired distortion effects could be perceived differently at different temporal instants (e.g. during goals in football games) even though images have equal PSNR values [7].

The Mean Opinion Score (MOS) [1] is a widely accepted way of quantifying the level of QoE: in MOS, a predefined scale ranging from 1 *bad)* to 5 (*excellent*) is used to assign a score value to the considered experience; the arithmetic mean over all values is interpreted as the overall quality of the experience as perceived by the audience.

Objective QoE evaluation methods try to replicate the results of subjective analysis by establishing relationships between Quality of Service (QoS) and Quality of Experience, i.e. by considering one or more Key Performance Indicator (KPI) having statistical correlation with QoE. For this reason, they are also called QoE prediction models, in order to highlight the capability of forecasting the perceived level of QoE (under statistical errors and approximations) starting from the quantification of a collection of variables. Such KPIs, as well as their internal analysis processes, represent the most important factors to characterise QoE evaluation models.

QOE EVALUATION MODELS FOR IP-BASED DELIVERY

Nowadays, several categories of QoE evaluation models for IP-based delivery exist in the scientific literature [8], but two macro-categories can be easily identified:

- Bitstream-based QoE evaluation models: they perform an analysis of the video IP-stream without decoding the sequence of frames, i.e. they are agnostic to the information of media resource conveyed in terms of images and audio. Some models, named parametric packet-layer models, infer QoE levels on the basis of packet-header analysis, i.e. avoiding payload inspection. Clearly such methods can analyse a limited number of features, which may be further reduced if the stream is encrypted. Even though employed in lightweight evaluation models, their precision can increase when equipped with functionalities to infer information about the global playout process (e.g. switches among quality representations of media resources, duration and position of stalls, etc.).
- Image-based QoE evaluation models: such methods perform an analysis of decoded video frames as they are reproduced during the consumption phase. Similarly to what introduced in the previous section, the following sub-categories can be identified:
    - *Full reference models:* they compare the decoded frames against a reference resource (typically the original source). The ITU J.341 VQuad [9] is a full reference image-based video quality method that is able to output two MOS scores, one for the quality of the compression and a second one for the quality related to the degradation due to the transmission.

- *No reference:* such methods analyse the decoded frames looking for traditional distortion problems such as stalling, slicing, blockiness, noise and blur. For practical reasons, no-reference models are widely used in real systems; this is also reflected in the high number of no-reference assessment methods proposed in the literature [13].
- Reduced reference models: these methods usually resort to the original signal to extract useful information about image properties, which may then be combined with further parameters (e.g. delays) typical of no reference models. This is the case, for example, of several models based on artificial intelligence techniques (e.g. [8, 14]) that performs offline learning process to avoid expensive computations on users' devices.

The ITU J.343.1 VMon [10] is an example of a hybrid no-reference video quality method combining imaged-based and bitstream-based inspection techniques that can be used in live TV services. It should be noted that the Rohde & Schwarz SwissQual J.343.1 standard is based on ITU J.343.1 and it can be implemented on Android platforms.

Compared to bitstream-based evaluation models, image-based models can also better evaluate factors highly correlated with Quality of Experience, such as scene complexity (the number of elements in a scene) and video motion (an index referring to the degree of motion of elements in a scene) [12].

Typical functions used to map QoS KPIs to the MOS domain are cubic, logarithmic, exponential, logistic and power functions [21, 22, 8]. In QoE models based on machine learning methods, explicit functions are usually unknown as statistical relationships between KPIs and QoE are embedded into data structures (e.g. [14, 15]) produced at the end of the learning phase.

QoE models for DASH streaming sessions can leverage the existence of a reliable transport protocol involved in the delivery process, which prevent images from being affected by several artefacts at the expense of delays and stalls due to the presence of buffers [23]. In terms of KPIs, the Streaming Video Alliance [16] introduced the following metrics:

- Video start-up time: amount of time between the triggering of a play event and the rendering of the first frame.
- Re-buffer ratio: ratio between the total playing time and the sum of total playing time plus the re-buffering time.
- Average media bitrate: ratio between the total amount of received information (bits) and the total playing time.
- Video start failure: Boolean condition satisfied when the first chunk of a video is not received within a specific cut-off time.

Additional KPIs are proposed by the DASH Industry Forum [17] with the intent to introduce a rigorous terminology for the evaluation of Quality of Experience.

A LIGHTWEIGHT QOE EVALUATION MODEL

We now introduce a non-intrusive QoE objective model that can easily be integrated with most of DASH video players.

The model is a synthesis of several methodologies and results existing in the literature (e.g. [26, 27, 28]) together with additional elements used to improve the accuracy of results; moreover, it does not require any low-level inspection of data packet, complex analysis of frames nor data structures to be generated before its execution (as in several models based on machine learning techniques).

In consequence, it can easily be implemented in the form of a library for native and embedded video players. Most HTML5-compatible web-players provide JavaScript APIs to track the state of common QoS variables during the streaming playout, or at least to monitor the basic events used to infer such values [18].

ARCHITECTURE AND QOE EVALUATOR FUNCTIONS

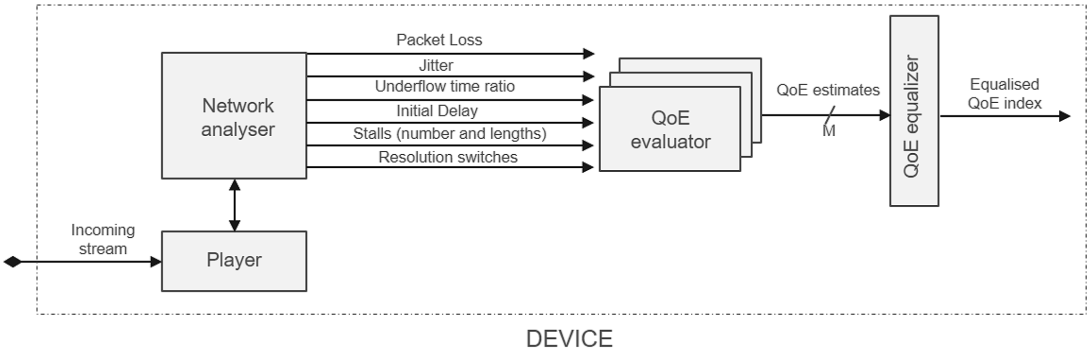

Figure 1: Abstract architecture of the QoE model

Figure 1 illustrates the abstract architecture of the model, which generates MOS estimates based on the following QoS Key Performance Indicators:

- *Packet loss*: percentage of IP packets lost during the playout session.
- *Jitter*: average difference between the mean latency and the latency values of the sample.
- *Initial delay*: delay between the request of a media resource and the instant when the first frame is reproduced by the player.
- *Underflow time ratio*: defined as $\frac{S}{S+R}$, where S is the cumulative duration of stalls (not including the initial delay) and R is the duration of the media resource being played.
- *Number of stalls*: number of times the playout process is blocked due to insufficient data in the internal buffer of the video player.
- *Duration of stalls*: delays experienced by the user when the playout is blocked due to insufficient data in the aforementioned buffer.
- *Resolution switches*: transitions between two different segment quality representations.

Every KPI is used to generate a QoE estimate through the mathematical relationships reported in Tables 1 and 2 (called QoE evaluators); finally, all estimates are aggregated to output a single MOS value, which depicts the approximated level of Quality of Experience perceived by the user. The model is flexible and extensible: a proper subset of KPIs can be selected to generate the QoE estimates to aggregate; moreover, further mathematical relationships can be added to capture the effects of additional QoS KPIs.

Table 1: Evaluators for playout-oriented KPIs
(media resources having resolution: 1920 x 1080 pixels; frame rate 30 fps)

KPI	QoE evaluator		
	Encoder	**Factor**	
Underflow time ratio U	5	$MOS_U(U) = 5e^{-5.571\,U}$	
Initial delay I	5	$MOS_I(I) = 5e^{-0.0416\,I}$	
Number of stalling events N of length L	5	$MOS_S(N,L) = a(L)e^{-Nb(L)} + c(L)$ $a(L) = \begin{cases} 2.99, & L < 3 \\ 0.024483L + 2.916652, & 3 \le L \le 32 \\ 3.2, & L > 32 \end{cases}$ $b(L) = \begin{cases} 0.247625L + 0.247625, & L < 3 \\ 0.034861L + 0.885917, & 3 \le L \le 32 \\ 1.856, & L > 32 \end{cases}$ $c(L) = \begin{cases} 2.01, & L < 3 \\ -0.02448L + 2.083448, & 3 \le L \le 32 \\ 1.3, & L > 32 \end{cases}$	
Resolution switches[2] R	4.4875	$R = 240$	$MOS_R(R) = 2.0625$
		$R = 240 - 480 - 240$	$MOS_R(R) = 2.25$
		$R = 240 - 1080 - 240$	$MOS_R(R) = 2.4$
		$R = 480 - 1080 - 240$	$MOS_R(R) = 2.7375$
		$R = 480 - 240$	$MOS_R(R) = 2.8125$
		$R = 1080 - 240$	$MOS_R(R) = 2.8375$
		$R = 1080 - 480 - 240$	$MOS_R(R) = 2.8875$
		$R = 1080 - 240 - 480$	$MOS_R(R) = 3.1125$
		$R = 240 - 1080 - 480$	$MOS_R(R) = 3.03$
		$R = 480 - 240 - 1080$	$MOS_R(R) = 3.2625$
		$R = 480 - 240 - 480$	$MOS_R(R) = 3.315$
		$R = 240 - 480$	$MOS_R(R) = 3.425$
		$R = 1080 - 240 - 1080$	$MOS_R(R) = 3.475$
		$R = 240 - 480 - 1080$	$MOS_R(R) = 3.5125$
		$R = 480 - 1080 - 480$	$MOS_R(R) = 3.7875$
		$R = 480$	$MOS_R(R) = 4$
		$R = 1080 - 480$	$\mathrm{MOS}_R(R) = 4.0025$
		$R = 1080 - 480 - 1080$	$MOS_R(R) = 4.3375$
		$R = 480 - 1080$	$MOS_R(R) = 4.35$
		$R = 1080$	$MOS_R(R) = 4.4875$

[2] Values "X", "X − Y" and "X − Y − Z" must be intended as resolution values and resolution switches. For example, "480 − 240 − 1080" identifies a sequence of two transitions, one from resolution "720x480" to "320x240" and a second one from "320x240" to "1920x1080".

Table 2: Evaluators for network-oriented KPIs
(media resources having resolution: 1920 x 1080 pixels; frame rate: 30 fps).

KPI	Encoder	QoE evaluator	
		Cofactor	Factor
Packet loss ratio L (%)	VP9	4.094	$MOS_L(L) = 2.96e^{-1.379L} + 1.134e^{-0.0496L}$
	H.265	4.2368	$MOS_L(L) = 3.661e^{-1.561L} + 0.5758e^{-0.0579L}$
Jitter J (ms)	VP9	16.028	$MOS_J(J) = 11.62e^{-3.386J} + 4.408e^{-0.3477J}$
	H.265	4.50899	$MOS_J(J) = 4.509e^{-0.3713J} - 2.09e^{-16}e^{6.729J}$

Every QoE evaluator is associated with an element of type $\frac{factor}{cofactor}$, where $factor$ is an n –variable scalar valued function $MOS\ (X_1, \dots, X_n)$ and $cofactor$ is a number such that the ratio is contained in the range $(0; 1]$. $X_1, \dots, X_n$ are the QoS input variables affecting the QoE estimate generated by the evaluator. The mathematical relationships reported in Table 1 can be combined to evaluate DASH-based streaming sessions. In particular:

- All the predictors have one maximum at the QoS value that represents the "best scenario" (e.g. $MOS_I(I)$ has a maximum at $I = 0$, i.e. when there is no initial delay).
- The formulae for the underflow time ratio, the initial delay and the formula used to combine all QoE evaluators are introduced in [26]. The original model provides good results also in the context of adaptive media playout [30].
- The evaluator for resolution switches is inferred from the results presented in [27].
- The evaluator for stalling events is inferred from the data reported in [27]. More specifically, the formulas identify a family of functions of type $\{a(L)e^{-Nb(L)} + c(L)\}$, where N is the number of stalls and L is their duration. Such functions are based on the general solution of the differential equation $\frac{\partial QoE}{\partial N} = -(QoE - c(t))$ introduced in [32]. Coefficients $a(L)$, $b(L)$ and $c(L)$ are approximated from the results published in [27] by assuming that they are in a linear relationship with L. By making such hypothesis, it can be proved that the following important properties hold for every function in $\{a(L)e^{-Nb(L)} + c(L)\}$:

(i) $\forall N \forall L\ (N \in \mathbb{N} \wedge L \in [1,3) \Rightarrow f(N,1) \le f(N,L) \le f(N,3))$

(ii) $\forall N \forall L\ (N \in \mathbb{N} \wedge L \in [3,32) \Rightarrow f(N,1) \le f(N,L) \le f(N,32))$

(iii) $\forall N \forall L\ (N \in \mathbb{N} \wedge L \in [32,\infty) \Rightarrow f(N,L) \le f(N,32))$

(iv) $\forall L\ (L \in [1,\infty) \Rightarrow f(0,L) \le 5)$

Properties I - III ensure that every $f(N,L)$ generates values compatible with the statistical results published in [11, 27]; property IV ensures that the evaluator generates the maximum MOS value (i.e. 5) if and only if there are no stalls in the playout process. It is important to notice that, during the assessment, it is unlikely to obtain two or more stalls having exactly the same duration. To improve the precision of the model, alternatives functions such as $f\left(N, \left\lfloor\frac{L}{\Delta}\right\rfloor \Delta\right)$ and $f\left(N, \left\lfloor\frac{L}{2\Delta}\right\rfloor \Delta + \Delta\right)$ with $\Delta \in \mathbb{R}^{+}$ could be employed to round L off.

Many mappings of Table 1 involve one or several negative exponential functions: this is in line with the models reported in [22, 32], which refer to statistical relationships between QoS KPIs (such as initial delays, stalls, etc.) and QoE that are well-accepted in the literature (e.g. [24, 25]). Table 2 reports two evaluators for network-oriented performance indicators, based on the results published in [29]. In this case, the information about encoders is exploited to give better approximation of Quality of Experience.

The final QoE evaluator related to a set of functions

$$\left\{\frac{MOS_1(X_{11}, \dots, X_{1n_1})}{cofactor_1}, \dots, \frac{MOS_m(X_{m1}, \dots, X_{mn_m})}{cofactor_m}\right\}$$

is obtained as follows:

$$5 \cdot \frac{MOS_1(X_{11}, \dots, X_{1n_1})}{cofactor_1} \cdots \frac{MOS_m(X_{m1}, \dots, X_{mn_m})}{cofactor_m} \tag{1}$$

EXAMPLES

We can define a MOS evaluator by taking into consideration the effect of initial delay and underflow time ratio. By combining the corresponding functions reported in Table 1, according to equation (1) we obtain:

$$MOS(U, I) = 5 \cdot \frac{MOS_U(U)}{5} \frac{MOS_I(I)}{5} \tag{2}$$

Figure 2 shows the graph of $MOS(U, I)$.

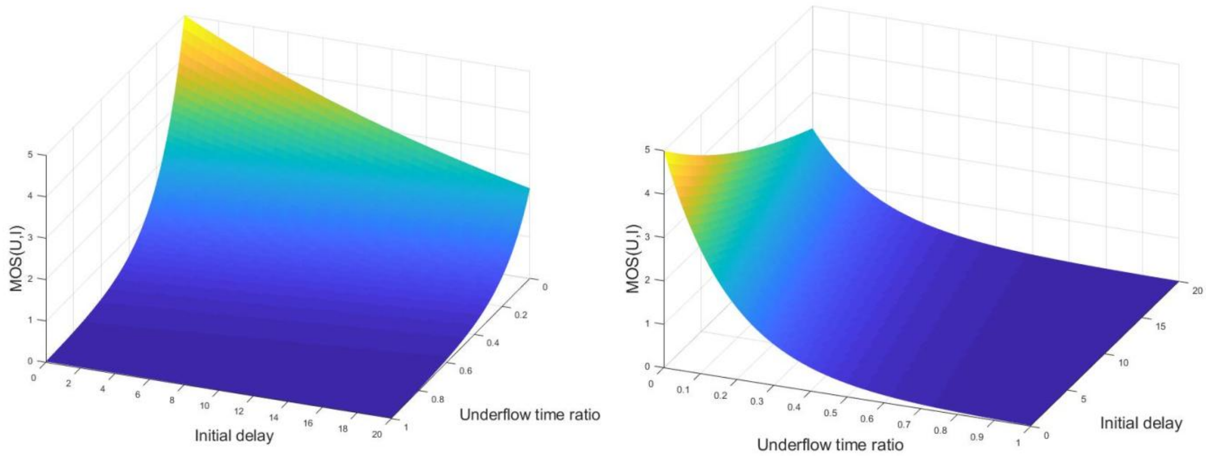

Figure 2: Qualitative graph of the QoE evaluator of Equation (2) based on underflow time ratio and initial delay.

Within a time window, several stalls and resolution switches may take place. In this case, computing the minimum value of the corresponding evaluators can give a lower approximation of Quality of Experience, which is a better "conservative" estimate for certain scenarios. For example, defining S_T and R_T respectively as the datasets for MOS_S and MOS_R over a period T, the lower approximation for MOS can be refined as follows:

$$MOS(S_T, R_T) = 5 \cdot \min_{(N,L)\in S_T} \left\{\frac{MOS_S(N_i, L_i)}{5}\right\} \cdot \min_{R\in R_T} \left\{\frac{MOS_R(R)}{4.4875}\right\}$$

Both examples highlight how the QoE evaluation model can be obtained by easily combining the formulae in Table 1 and Table 2. This also implies that the effort for the implementation of the global MOS evaluator is reduced and essentially consists in the encoding of the aforementioned tables, provided that the input QoS variables can be exported from a video player (e.g. through JavaScript APIs).

REFINEMENT OF THE EVALUATOR FOR INITIAL DELAY

In their tests, the authors of [26] consider videos lasting about 64 seconds but their original definition of the delay evaluator does not depend on the total duration of media resources. Nonetheless, it is intuitive to consider that the global effect of two identical initial delays should be perceived differently when the total cumulative time spent consuming the audio-visual service is substantially different. If such an evaluator is invariant with respect to the ratio between initial delay and media resource duration then we can easily obtain a simple variant overtaking this limitation.

Formally, we look for a formula of type $MOS_I(I, T) = 5e^{-\alpha_T \cdot I}$.

Given that $MOS_I(I) = 5e^{-0.0416\, I} = MOS_I(I, 64)$, we obtain $\alpha_{64} = 0.0416$.

Under the aforementioned hypothesis we obtain $MOS_I(64{,}64) = MOS_I(T, T)$,

i.e. $\alpha_T = \frac{0.0416}{T} 64$.

We conclude that $MOS_I(I, T) = 5e^{-\frac{0.0416}{T} \cdot 64 \cdot I}$.

CONCLUSIONS AND FUTURE WORK

This article has proposed a lightweight evaluation model for Quality of Experience of audio-visual services that can be used for assessment. The model stems from the consolidation of several evaluators and predictors existing in the literature, together with additional functions introduced for a finer-grained approximation of QoE. This enhanced formulation remains coherent with statistical relationships between QoS KPIs and QoE already highlighted in several studies.

Moreover, the model can be further extended with supplementary metrics and its reduced complexity also makes it a suitable candidate for lightweight assessment of QoE in HTML5-enabled web browsers. This property allows the model to be used as support both for offline and run-time optimisation of distribution chains, including QoE-driven adaptation of DASH-like protocols on users' devices.

In spite of the significant number of QoS KPIs considered as input, the functions reported in Table 1 and Table 2 have been inferred for a few instances of codecs and resolutions. Indeed, the current literature does not provide an exhaustive overview of how the most typical configurations of such parameters affect QoE, which will be the subject of future researches.

To further increase the precision of the model, additional evaluators taking into consideration screen features (e.g. resolution, colour depth, etc.), spatiotemporal complexity and stall patterns [31] could be introduced, at the expense of a higher complexity for implementation and integration with video players.

REFERENCES

[1] Recommendation ITU-T P.800.1 (2016): *Mean Opinion Score (MOS) terminology*.

[2] M.H. Pinson, S. Wolf, "*A new standardized method for objectively measuring video quality*", IEEE Trans. on broadcasting, 2004.

[3] Z.Wang, H.R.Sheikh, A.C.Bovik, Objective Video Quality Assessment, "*The Handbook of Video Databases: Design and Applications*", 2003.

[4] E. A. Silva, K. Panetta, S. S. Agaian, "*Quantifying image similarity using measure of enhancement by entropy*", Mobile Multimedia/Image Processing for Military and Security Applications 2007.

[5] H.R.Sheikh, A.C. Bovik, "*Image information and visual quality*", IEEE Transactions on Image Processing, 2006.

[6] J. Berger, "*Quality Evaluation of Streaming Video on Mobile Networks*", Rohde & Schwarz, White Paper, August 2017.

[7] B. Lembke, C. Kunert, "*MOS in eMBMS*", IRT Project Report, May 2017.

[8] M. Alreshoodi, J. Woods, "*Survey on QoE\QoS Correlation Models for Multimedia Services*", International Journal of Distributed and Parallel Systems, Vol. 4, No. 3, May 2013.

[9] Recommendation ITU-T J.341 (2016): *Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference.*

[10] Recommendation ITU-T J.343 (2014): *Hybrid perceptual bitstream models for objective video quality measurements.*

[11] T. Hossfeld, M. Seufert, M. Hirth, T. Zinner, "*Quantification of YouTube QoE via Crowdsourcing*", IEEE International Symposium on Multimedia, ISM 2011, Dana Point, CA, USA, December 5-7, 2011.

[12] J. Hu, H. Wildfeuer, "*Use of context complexity factors in video over IP quality monitoring*", Quality of Multimedia Experience, QoMEx 2009, September 2009.

[13] M. Torres Vega, V. Sguazzo, D.C. Mocanu, A. Liotta, "*An experimental survey of no-reference video quality assessment methods*", International Journal of Pervasive Computing and Communications, April 2016.

[14] M. Torres Vega, E. Giordano, D.C. Mocanu, D. Tjondronegoro, A. Liotta, "*Cognitive*

no-reference video quality assessment for mobile streaming services", 7th International Workshop on Quality of Multimedia Experience (QoMex), May 2015.

[15] V. Menkovski, A. Oredope, A. Liotta, A. Cuadra Sanchez, "*Predicting Quality of Experience in Multimedia Streaming*", 7th International Conference on Advances in Mobile Computing and Multimedia, January 2009.

[16] *Key Network Delivery Metrics*, Quality Experience working group Streaming Video Alliance, May 2016.

[17] "*DASH-IF position Paper: Proposed QoE Media metrics standardisation for segment media playback*", Dash Industry Forum, version 1.0, October 2016.

[18] B. Tullemans, "*Technical Practices for a Multi-CDN Distribution Strategy*", European Broadcasting Union, IBC October 2017.

[19] P. Goudarzi, "*A no-reference low-complexity QoE measurement algorithm for H.264 video transmission systems*", Scientia Iranica, Transactions D: Computer Science & Engineering and Electrical Engineering, January 2013.

[20] D. Ghadiyaram, C. Chen, S. Inguva, A. Kokaram, "*A no-reference video quality predictor for compression and scaling artifacts*", IEEE International Conference on Image Processing, IEEE, 2017.

[21] M. Fiedler, T. Hossfeld, P. Tran-Gia, IEEE Network: The Magazine of Global Internetworking - Special issue on improving quality of experience for network services archive, March/April 2010.

[22] M. Fiedler, T. Hossfeld, "*Quality of Experience-Related Differential Equations and Provisioning-delivery Hysteresis*", 21st ITC Specialist Seminar on Multimedia Applications - Traffic, Performance and QoE, Japan, 2010.

[23] O. Hohlfeld, "*Impact on Buffering on Quality of Experience*", PhD dissertation, Berlin 2014.

[24] V. D. Bhamidipati, S. Kilari, "*Effect of Delay / Delay Variation on QoE in Video Streaming*", Master Thesis in Electrical Engineering, School of Computing, Blekinge Institute of Technology, Sweden, May 2010.

[25] J. S. Mwela, O. E. Adebomi, "*Impact of Packet Losses on the Quality of Video Streaming*", Master Thesis in Electrical Engineering, School of Computing, Blekinge Institute of Technology, Sweden, June 2010.

[26] M. Li, C. Lee, "*A cost-effective and real-time QoE evaluation method for multimedia streaming services*", Telecommunication Systems, Vol. 59, Issue 3, pp 317-327, July 2015.

[27] A. Asan, W. Robitza, I. Mkwawa, L. Sun, E. Ifeachor, A. Raake, "*Impact of Video Resolution Changes on QoE for Adaptive Video Streaming*", IEEE International Conference on Multimedia and Expo (ICME), July 2017.

[28] T. Hossfeld, D. Strohmeier, R. Schatz, "*Pippi Longstocking Calculus for Temporal Stimuli Pattern on YouTube QoE: 1+1=3 and 1·4 ≠ 4·1*", ACM MoVid '13 Proceedings of the 5th Workshop on Mobile Video, Pages 37-42, Oslo, Norway, February 2013.

[29] D. Pal, V. Vanijja, "*Effect of network QoS on user QoE for a mobile video streaming service using H.265/VP9 codec*", 8^{th} International Conference on Advances in Information Technologies, IAIT2016, 19-22, December 2016, Macau, China.

[30] M. Li, C. Yeh, S. Lu, "*Real-Time QoE Monitoring System for Video Streaming Services with Adaptive Media Playout*", Hindawi International Journal of Digital Multimedia Broadcasting, Vol. 2018, Article ID 2619438.

[31] D. Ghadiyaram, A. C. Bovik, H. Yeganeh, R. Kordasiewicz, M. Gallant, "*Study of the Effects of Stalling Events on the Quality of Experience of Mobile Streaming Videos*", 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP), December 2014.

[32] M. Fidler, T. Hossfeld, P. Tran-Gia, "*A generic quantitative relationship between quality of experience and quality of service*", IEEE Network: The Magazine of Global Internetworking - Special issue on improving quality of experience for network services archive, Volume 24 Issue 2, March/April 2010.

AUTHOR BIOGRAPHY

Francesco De Angelis received his Master's Degree in Computer Engineering from the University of Rome "Tor Vergata" and a PhD in Information Systems from the University of Geneva.

Dr De Angelis is currently a Project Manager in Telecommunications in the Technology & Innovation department at the EBU. His main focus is on multicast media delivery protocols and Quality of Experience.

Published by the European Broadcasting Union, Geneva, Switzerland
ISSN: 1609-1469

Editor-in-Chief: Patrick Wauthier
E-mail: wauthier@ebu.ch