Generalized Westerink-Roufs Model for Predicting Quality of Scaled Video

Nabajeet Barman*, Rahul Vanam†§, and Yuriy Reznik‡
*Brightcove UK Ltd, London, United Kingdom, nbarman@brightcove.com
†Amazon Prime Video, Seattle, USA, rahulv@u.washington.edu
‡Brightcove Inc, Seattle, USA, yreznik@brightcove.com

***Abstract*—Resolution is a fundamental property of encoded video. Understanding the impact of resolution on quality as an independent parameter can help design better, more efficient systems, such as selecting optimum rendition in adaptive video streaming applications. One known quality model that considers resolution for predicting the perceived picture quality is the Westerink and Roufs (WR) model, which establishes the relationship between subjective quality and two parameters of viewing setup: angular resolution and viewing angle. This paper first validates the WR model on recent datasets and shows that it is reasonably accurate. We then propose a generalization of this model, allowing operation in a broader range of parameters and with more graceful saturation in extended regions. We then validate the performance of the proposed Generalized WR model on the new datasets and show that the proposed model achieves even a better fit to the recent datasets. We also demonstrate that the proposed Generalized model can account for the differences in scaling algorithms, including more advanced ML-based methods such as super-resolution. We conclude with a discussion of several possible applications of this model, including its use to guide the rendition selection decisions in streaming players and adapt that decision logic based on the upsampling algorithms used at the player.**

***Index Terms*—Adaptive Streaming, Video Streaming, QoE, Video Quality Estimation, Rendition Selection, Upsampling.**

I. Introduction

Adaptive streaming, where the playback is adapted based on the changing network conditions, is the fundamental technology that has enabled the recent growth of over-the-top (OTT) streaming services such as YouTube, Netflix, Hulu, and others. The two of the most widely used formats for adaptive streaming are HTTP Live Streaming (HLS) [1], and Dynamic Adaptive Streaming over HTTP (DASH) [2]. To create multiple representations video is encoded with several distinct resolution-bitrate parameters. The streaming client (player), depending on the available network throughput, buffer status, and player size, selects the appropriate rendition for playback [3]. The player at the end-user device typically upscales the lower resolution videos to fit the player/window resolution. Understanding the impacts of the encoded video resolution and scaling on the perceived quality is essential for such applications to allow them to select the optimum rendition [4].

Most of the existing full-reference video quality metrics (PSNR, SSIM, VMAF, etc.) either ignore the effects of video resolution altogether or treat it jointly with codec-introduced noise [5]. Such metrics are of no help to streaming clients. Some other metrics, such as Belmudez and Möller model [6], or ITU-T P.1203 metric [7], treat contribution to quality by video resolution and codec-introduced noise (bitrate) as independent parameters. The Westerink and Roufs (WR) model [8] is the best-known model that predicts the perceived quality directly based on video resolution [9]. According to this model, the perceived quality of still pictures projected to the screen depends on two parameters: (1) the angular span (in degrees) and (2) the angular resolution (in cycles per degree) of the projected picture.

The WR model, since its publication, has been validated and confirmed by other means and techniques, such as Barten's SQRI method [10], [11], and others [12] but such validations were done many years ago on limited datasets. However, despite its many advantages, the WR model suffers from certain limitations, primarily due to the limited nature of the subjective test parameters (limited range of viewing angles and angular resolutions) used in the model's design. Also, the model has no parameters to consider the effects of various upsampling methods or differences in HDR and SDR content, resulting in limitations on the model's applicability to more advanced display devices.

This paper proposes an extension of the WR model to address these limitations. Specifically, the proposed model adapts to a different dynamic range of the reproduction device and the upsampling method. We also discuss possible applications where such a generalized model can be used to achieve higher gain and increased end-user QoE in the design of streaming clients.

The rest of the paper is organized as follows. Section II introduces the original WR model and tests its accuracy by using six modern datasets. Section III presents the Generalized WR model and tests its accuracy. Section IV presents the results of tuning the Generalized WR model to HDR content and different upsampling methods. Section V discusses potential applications. Section VI presents conclusions and future work.

II. Westerink and Roufs (WR) Model and its Validation

1) Viewing Angle and Angular Resolution: Consider a video of resolution $w \times h$ (pixels), streamed to a display of size $W \times H$ (pixels) of effective pixel density ρ (pixels per

§This work was conducted by Rahul Vanam while being affiliated with Brightcove Inc.

TABLE I: CHARACTERISTICS OF VIEWING SETUPS AND RESOLUTIONS TESTED FOR THE SIX DIFFERENT DATASETS CONSIDERED IN THIS WORK.

Characteristic* / Dataset	ITU TV	AVT-VQDB-UHD-1	NFLX	GamingVideoSET	ITU Tablet	ITU Mobile
Display size	75”	65”	24”	24”	9.7”	5”
Viewing Distance	1.5H	1.5H	3H	3H	18”	12.67”
Display pixel size	3840x2160	3840x2160	1920x1080	1920x1080	1920x1080	1920x1080
Viewing angle	61.3	61.3	33	33	29.3	19.48
Display Nyquist [cpd]	28.272	28.272	28.28	28.28	32.08	48.8018
Video resolutions tested (resolution -> cpd)	480x360 -> 3.53 960x540 -> 7.07 1280x720 -> 9.42 1920x1080 -> 14.14 3840x2160 -> 28.28	640x480 -> 4.71 1280x720 -> 9.42 1920x1080 -> 14.14 3840x2160 -> 28.28	384x288 -> 5.65 512x384 -> 7.54 720x480 -> 10.60 1280x720 -> 18.85 1920x1080 -> 28.28	640x480 -> 9.42 1280x720 -> 18.85 1920x1080 -> 28.28	1280x720 -> 21.39 1920x1080 -> 32.08	1280x720 -> 32.53 1920x1080 -> 48.80

*Notes: The characteristics of the datasets used in this work (AVT-VQDB-UHD-1 [13], Netflix [14], GamingVideoSET [15] ITU TV [16], ITU Tablet [16] ITU Mobile dataset [16]) are based on the information provided in the respective dataset. Parameter H in viewing distance refers to the height of the display in inches. In the absence of any required information, the values are assumed based on usage statistics as reported in [17] and [18].

inch). Let the playback take place over a player of size $w_p \times h_p$ (pixels) where the viewer is seated at a distance of d (inches) from the display. The viewing angle (in degrees) then can be defined as:

$$\phi = 2 \arctan\left(\frac{w_p}{2d\rho}\right) \approx \frac{180}{\pi}\left(\frac{w_p}{d\rho}\right). \tag{1}$$

Angular resolution is the inverse of an observation angle capturing the span of a 2-pixel interval in a video frame projected to the screen. Such a 2-pixel interval represents the length of a "cycle" of a waveform with the highest spatial frequency that may be present in the video. The inverse of it becomes "cycles per degree", and that is a unit in which angular resolution is typically measured and is calculated as:

$$\mu = \frac{1}{2 \arctan\left(\frac{w_p}{w\,\rho\,d}\right)} \approx \frac{\pi}{360}\,\frac{w\,\rho\,d}{w_p}. \tag{2}$$

We note that in both formulae 1 and 2 the viewing distance d appears in form of a product with pixel density ρ. This means that a single relative value of a viewing distance, such as e.g. viewing distance expressed in display heights:

$$\eta = \frac{d\rho}{H} \tag{3}$$

is generally sufficient for derivation of both viewing angle (ϕ) and angular resolution (μ) parameters.

A. WR Model

Westerink and Roufs [8] found that at a constant viewing distance the subjective quality of still pictures was influenced independently by both angular resolution and the size of the displayed picture. Therefore, even if correlated, angular resolution and image/display size represent two different dimensions. We define next the necessary terms and the model.

Let ϕ be the viewing angle (in degrees) and μ be the effective angular resolution of the projected video (in cycles per degree (cpd)). The perceived picture quality of the video, $Q(\phi,\,\mu)$ as estimated by the WR model [8] is:

$$Q(\phi,\,\mu) = 3.6\,\log\left(\phi\,\frac{\pi}{180}\right) + 2.9 + 4.6\log(\mu) + 2.7(\log(\mu))^2 - 1.7(\log(\mu))^3 \tag{4}$$

This model is henceforth referred to as the Original WR model. In the subjective test performed for the design of the mode, the range of viewing angles (ϕ) was from about 2.526 to 18.026 degrees and the range of angular resolutions (μ) used was from about 2.7 cpd to 38 cpd, which will, henceforth, be referred to as the "operating range" of the model.

B. Validation of Original WR model on additional datasets

Since the WR model was initially proposed in 1989 and has since been validated and used in many works [9], [10], [19], one may still wonder about its suitability more than 30 years later given newer display technologies. Hence, we perform a more exhaustive validation on six new, open-source datasets which have been designed considering different display sizes, viewing distances and resolutions. The selected open-source datasets, their respective settings and characteristics are summarized in the Table I.

It can be observed that these datasets cover a diverse range of use cases – from QCIF (video conferencing) to HD and UHD types of experiences and viewing setups. Also, the datasets consist of different content types from gaming to animation and sports, which is representative of typical content streamed by any typical modern-era streaming system. The datasets are also representative of almost all major device types such as UHD TV to PC monitors to smaller screen devices such as Mobile and Tablets as used by consumers today. We also note that these data sets are also exhibiting a variety of distortions – such as codec noise and/or artifacts introduced by different up-sampling algorithms. No efforts were also made to post-process results accounting for differences in scores based on content. However, to minimize the codec noise and/or scaling artefacts, for each dataset, if multiple renditions for particular resolutions are available, we only consider the subjective quality (MOS) score corresponding to the highest encoded bitrate representation. However, since there are a lot of different contents, there will still exist a broad variation of MOS scores and hence, one should not expect a perfect fit. Since the perceived quality (Q) as estimated by the Original WR model can be unbounded, we have used a linear fitting function $\alpha + \beta x$ to fit the Q values to the MOS

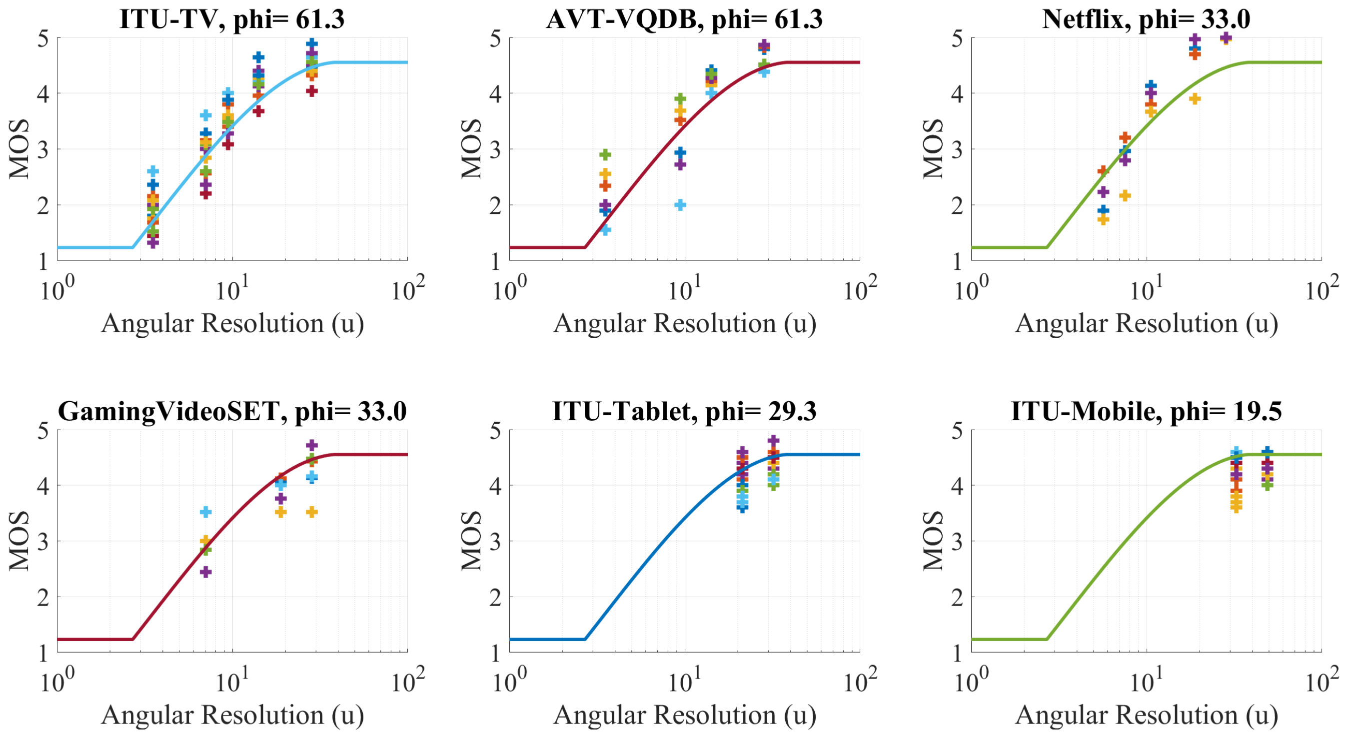


Figure 1: MOS vs Angular resolution (cpd) plot for the six datasets. The fitted line is the perceived picture quality (Q) scores as predicted by the Original WR model. The colours of the markers in each plot represents a particular video sequence.

scores (1-5). The global fitting parameters, obtained for all six datasets combined are $\alpha = -1.0739$, $\beta = 0.67015$. The fit of the Original WR model to six different datasets is shown in Figure 1. As can be observed, despite the fact that these six datasets capture reproductions with different viewing angles (ϕ), the shape of the original WR model does not change much. This is due to the fact that the original model was only verified in the range of ϕ = [2.526 to 18.026] and we hence have to clip ϕ in model computations. Still, we observe that the fit to newer datasets is quite good, with the "goodness of fit" scores in terms of RMSE for each dataset summarized in Table II.

III. Generalized WR Model and its Validation

As briefly mentioned earlier, one of the major limitations of the WR model is that it works well only for a narrow range of resolutions and viewing angles, and does not scale well outside the test range. Extrapolating the WR model beyond the operating range of angular resolution ($\mu = [2.7, 38]\, cpd$), one can observe that the MOS scores predicted by the WR model may turn negative at low resolutions or decay with resolutions beyond 40 cpd. Additionally, with an increase in viewing angle, the values of perceived picture quality, Q increases unbounded, which (according to our knowledge of human visual system) should rather saturate at higher resolutions and decay to 0 when resolutions are low. To overcome the above-mentioned shortcomings, we next propose a modified version of the WR model, offering due saturations at broader ranges of parameter values. We first note that the WR model works with logarithms of ϕ and μ as main variables. This means that their weighted sum translates to a geometric (or power) weighted average under the logarithm:

$$Q(\phi, \mu) = \lambda \log(f(\phi)) + \eta \log(g(\mu)) = \log(f(\phi)^{\lambda} g(\mu)^{\eta}) \quad (5)$$

where $f(.)$ and $g(.)$ are certain functions.

When either of the variables is approaching 0, the power average turns to 0 as well. This makes a lot of sense in the context of the problem: if either picture angle ϕ or angular resolution μ is approaching 0, the same should happen to overall quality expression. However, the subsequent application of the logarithm turns this 0 to $-\infty$. Also, in the Original WR model with an increase in viewing angle, the values of perceived picture quality, Q increases unbounded. In order to address these issues, we introduce α and β parameters,

$$Q(\phi, \mu) \approx \log\left(\alpha + \beta\left(f(\phi)^{\lambda} g(\mu)^{\eta}\right)\right) \quad (6)$$

where $\alpha \geq 1$.

We next introduce saturations. For this purpose, we will use L_p norm with negative power parameter:

$$f\left(x, x_s, k\right) = \left(x^{-k} + x_s^{-k}\right)^{-\frac{1}{k}} = x_s \left(1 + \left(\frac{x}{x_s}\right)^{-k}\right)^{-\frac{1}{k}} \quad (7)$$

For example, with $k = 1$, this turns into a harmonic mean between x and x_s, and when $k \rightarrow \infty$, it turns into $min(x, x_s)$.

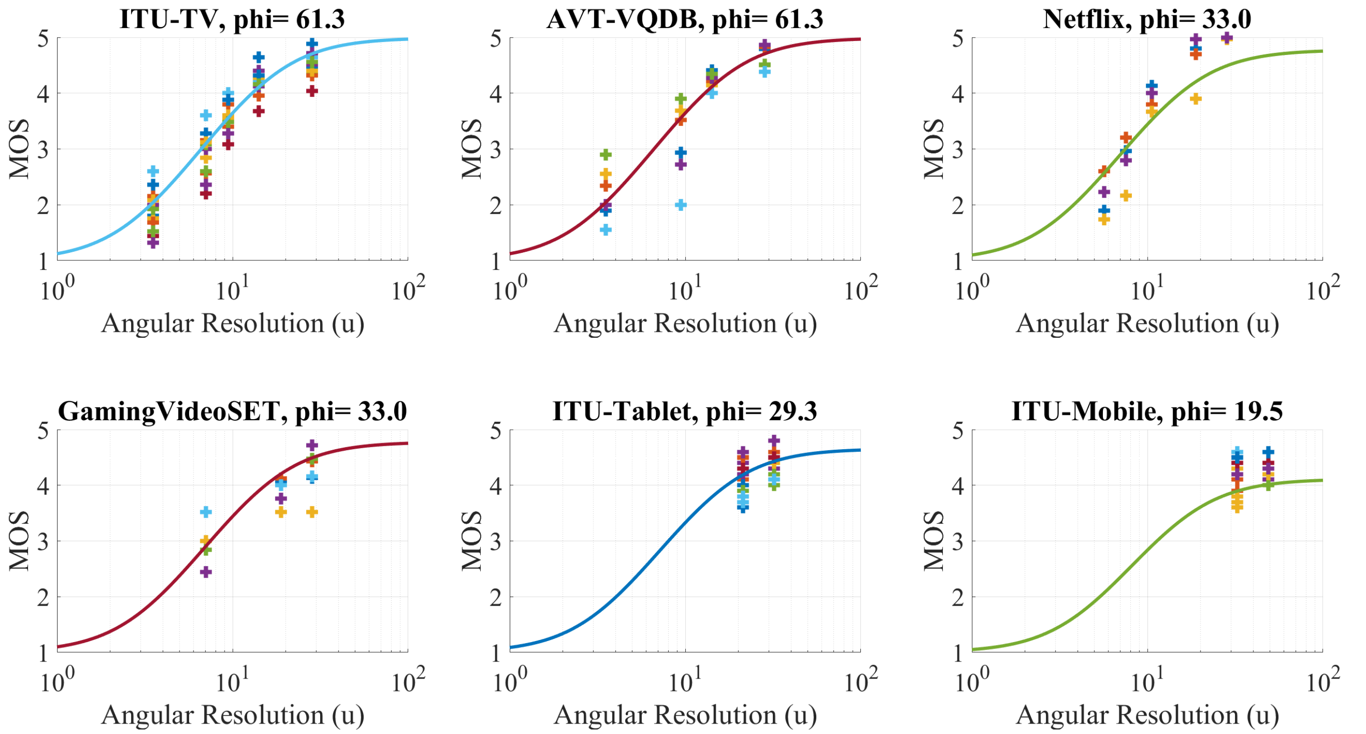


Figure 2: MOS vs Angular resolution (cpd) plot for the six datasets. The fitted line is the perceived picture quality (Q) scores as predicted by the GWR model. The colours of the markers in each plot represents a particular video sequence.

By using these techniques and moving everything under the logarithm, we next arrive at the following, generalized version of the Westerink & Roufs model:

$$Q(\phi,\, \mu) = \log\left(\alpha + \beta\left(1+\left(\frac{\phi}{\phi_s}\right)^{-k}\right)^{-\frac{\gamma}{k}}\left(1+\left(\frac{\mu}{\mu_s}\right)^{-l}\right)^{-\frac{\delta}{l}}\right) \quad (8)$$

where γ, δ, k, l, ϕ_s and μ_s are model parameters controlling the behaviour with respect to viewing angle and angular resolution. This model will henceforth be referred to as *GWR* model.

A. *Validation of the GWR Model on New Datasets*

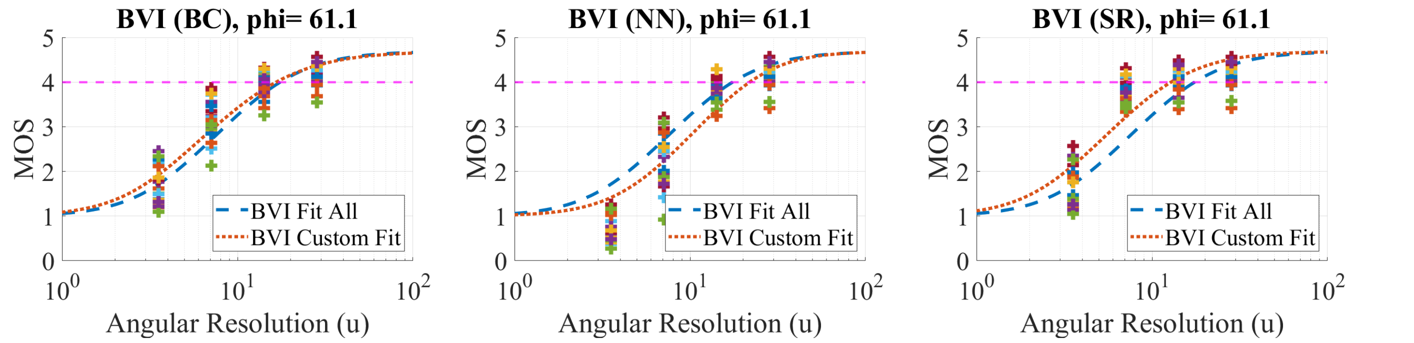
Similar to the Original WR model, we validate the proposed generalized model to the newer datasets. By fitting the Generalized model to the six modern datasets combined we obtain the model parameters as: α= 2.72, β=145.69, γ = 1.55, δ=2.12, k=6.01, l=2.11, ϕ_s=35.0, and μ_s=16.93. Figure 2 shows the individual fitting of the perceived picture quality (Q) as estimated using the GWR model on all six datasets. One can observe that the model fits quite nicely to all datasets allowing for a more graceful saturation in extreme regions (very high/low angular resolution values). Table II shows the "goodness of fit" in terms of RMSE scores for the Original and GWR model for the six datasets. It can be seen that the GWR model results in a better fit than the Original WR model, especially when comparing the RMSE scores for smaller screen devices such as Mobile and Tablet.

TABLE II: "Goodness of fit" of Original and Generalized WR model to the six different datasets in terms of RMSE score.

Dataset	**Original**	**Generalized**
Netflix	0.40	0.40
AVT	0.23	0.24
Gaming	0.28	0.22
ITU-TV	0.21	0.15
ITU-Mobile	0.36	0.23
ITU-Tablet	0.21	0.13
Average	**0.28**	**0.23**

IV. Tuning GWR model to cases of HDR content and different upsampling algorithms

So far we limited our analysis and validation of the models to SDR, 8-bit content using traditional upsampling algorithms such as *bicubic* or *lanzcos3*. However, 10-bit HDR content, as well as different upsampling algorithms, have different effects on the subjective scores, with MOS scores often reaching saturation at lower angular resolutions for AI/ML based upsampling algorithms as compared to the traditional algorithms such as Nearest Neighbour or Bicubic. To further study the effect of such content and upsampling algorithms, we use the MOS scores from the open-source *BVI* dataset [20] consisting of 24 10-bit five-second source reference videos sequences of 60 fps which were encoded in three different resolutions (1920×1080, 960×540 and 480×270). The encoded



(a) Bicubic Upsampling.

(b) Nearest Neighbour Upsampling.

(c) Super Resolution Upsampling.

Figure 3: MOS vs Angular resolution (cpd) plot considering three different scaling from the BVI dataset. Both the custom fit (adapting the u_s and l to each individual rescaling filter dataset) and the generic fit (to full BVI dataset) are shown. The colours of the markers in each plot represents a particular video sequence.

TABLE III: MODEL PARAMETERS, μ_s AND l OBTAINED FOR THE GWR MODEL WHEN FITTED SEPARATELY TO EACH BVI UPSAMPLING ALGORITHM DATASET.

Upsampling Algorithm	μ_s	l
Nearest Neighbour (NN)	23.4	2.5
Bi-cubic (BC)	13.93	1.76
Super Resolution (SR)	12.24	2.06

video sequences were upsampled to the native resolution (2160p) using three different upsampling algorithms: BiCubic (BC), Nearest Neighbour (NN) and Super Resolution (SR) and then viewed on a display measuring 65.4×36.8 cm of BT.2020 colour space (full range) at a viewing distance of 1.5H. Using the (default) model parameters obtained by fitting the proposed GWR model to the six datasets earlier, we noticed that the fit is not that good. Hence, using all MOS scores from the BVI dataset, we allowed different values of α and β, as well as added a common scale factor, ϵ for γ and δ. By refitting, we obtain the new parameter values as: $\alpha = 2.72, \beta = 106.91, \epsilon = 1.08, \gamma = 1.55\epsilon, \delta = 2.12\epsilon$. We observed that these changes only affect the overall scale of the model, but not its dependency on parameters ϕ and μ. Hence more customized fit to each rescaling method is required.

We next fit the GWR model to BC, SR, and NN subsets of the BVI dataset corresponding to the upsampling method used (henceforth, BVI (BC), BVI (SR) and BVI (NN) respectively) by allowing only parameters μ_s and l to vary (since they control the model behaviour wrt μ), while using default values of all other parameters. The fit values of μ_s and l obtained for each upsampling algorithm are summarized in Table III.

Figure 3 shows the plot of MOS vs angular resolution (μ) for the BVI dataset for both the generic fit ("BVI Fit All") to all of the BVI dataset as well as the customized fit to each upsampling algorithm with new fitting parameters (μ_s) and l) ("BVI Custom Fit"). It can be observed that with the new μ_s and l values, a much better fit is achieved. Figure 3 also shows a line at MOS = 4, establishing a baseline that we could use for comparison. For example, one might be interested in understanding the differences in the required angular resolutions of the content to achieve "good" MOS score (for example, MOS=4) by the viewers? We observe that fit to BVI (BC), BVI (SR) and BVI (NN) datasets hits MOS=4 at μ = 16.6, 12.8 and 21.9 respectively. This can be used to compute relative savings in "encoding resolutions" that can be achieved by SR upsampling techniques as compared to the traditional upsampling at different quality levels or in the design of optimal encoding ladders.

V. APPLICATIONS

A. Adaptive Video Streaming

In adaptive streaming, the players switch renditions based on the available network bandwidth. Such adaptation allows continuous playback and prevents buffering. However, as discussed in [4], in the case of streaming videos embedded in web pages, the sizes of such embedded videos, or equivalently, the stretch factors of browser windows, become additional factors influencing the stream selection logic.

In such situations, web players typically select rendition with the nearest available resolution to match page viewport size. Such logic is suboptimal. A more advanced rendition selection logic should also consider the form factor of the device, dynamic range, and the upsampling method used, as they all influence QoE [21]. In this context, the proposed generalization of the WR metric could be used as a tool for driving optimal rendition choices by considering all such parameters. Our upcoming publication [21] offers additional details about the design of such rendition selection algorithms.

B. (Adaptive) Video Conferencing

Like web-streaming, video conferencing applications such as Skype, Google Meet, or Zoom also commonly present videos on devices with different form factors and different positions of video windows on the screen. They also dynamically choose video resolution for encoding and delivery to each screen. The proposed model can be used to guide such decisions optimally by accounting for the receiving device form factor, video window resolution, and quality of upsampling algorithms.

C. Adaptive Resolution Codecs

Some new codecs, such as VVC [22] allow dynamic resolution changes of the encoded video to achieve better quality-bitrate tradeoffs for a specific reproduction environment (e.g. ITU-R BT.500 [23]). The proposed model can help in guiding the resolution-based decisions in encoders utilizing such features of compression standards.

D. Clients using advanced upscaling algorithms

In Section IV we presented how the proposed model can be re-tuned to take into account more advanced upscaling algorithms such as super-resolution, implemented at the client side. Such a metric that can take into account client post-processing capabilities such as more advanced upsampling on perceived quality, can potentially allow for the design of more network-friendly clients (requiring less bandwidth) or for the system to deliver a better overall quality of experience.

VI. Conclusions and Future Work

In this paper, we presented and validated the Westerink and Roufs model [8] using six recent open-source datasets. We observed that the original WR model performs reasonably well despite its limitations. We also proposed an extension of the original WR model allowing for graceful saturations at the extreme values and resulting in a better fit on the datasets. We also showed that the proposed model allows tuning to HDR and different and more advanced upsampling algorithms. We also discussed the possible uses of the proposed model. They include optimal rendition selection algorithms in ABR streaming, selection of optimal encoding resolution in adaptive video conferencing applications, and implementation of adaptive-resolution encoders.

In the future, we plan to extend the study by considering additional datasets, and using different datasets as basis for training and validation.

Additionally, we plan to use the proposed generalized model jointly with measures of codec-introduced noise (e.g., PSNR, SSIM, or VIF)to arrive at a more complete parametric quality model for QoE assessment in multiscreen systems. Some preliminary results towards development of such a hybrid metric have already been reported in [3], [19] and [24].

References

[1] "HTTP live streaming, RFC 8216," https://tools.ietf.org/html/rfc8216, 2019, [Online: accessed 19-March-2022].

[2] "ISO/IEC 23009-1:2019 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats," https://www.iso.org/standard/79329.html, 2019, [Online: accessed 17-March-2022].

[3] Y. A. Reznik, K. O. Lillevold, and R. Vanam, "Perceptually Optimized ABR Ladder Generation for Web Streaming," *Electronic Imaging*, vol. 2021, no. 3, pp. 75–1–75–11, 2021.

[4] Y. A. Reznik, K. O. Lillevold, A. Jagannath, and X. Li, "Towards Understanding of the Behavior of Web Streaming," in *2021 Picture Coding Symposium (PCS)*, 2021, pp. 1–5.

[5] N. Barman and M. G. Martini, "QoE modeling for HTTP Adaptive Video Streaming– A Survey and Open Challenges," *IEEE Access*, vol. 7, pp. 30 831–30 859, 2019.

[6] B. Belmudez and S. Möller, "An approach for modeling the effects of video resolution and size on the perceived visual quality," in *2011 IEEE International Symposium on Multimedia*, 2011, pp. 464–469.

[7] ITU-T Recommendation P.1203, *P.1203 : Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport*. Geneva: International Telecommunication Union, Oct 2017.

[8] J. H. D. M. Westerink and J. A. J. Roufs, "Subjective Image Quality as a Function of Viewing Distance, Resolution, and Picture Size," *SMPTE Journal*, vol. 98, no. 2, pp. 113–119, 1989.

[9] M. Lombard, T. B. Ditton, M. E. Grabe, and R. D. Reich, "The role of screen size in viewer responses to television fare," *Communication Reports*, vol. 10, no. 1, pp. 95–106, 1997. [Online]. Available: https://doi.org/10.1080/08934219709367663

[10] P. Barten, "The effects of picture size and definition on perceived image quality," *IEEE Transactions on Electron Devices*, vol. 36, no. 9, pp. 1865–1869, 1989.

[11] ——, *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*, ser. Press Monographs. SPIE Optical Engineering Press, 1999. [Online]. Available: https://doi.org/10.1117/3.353254

[12] S. Winkler, "Digital Video Quality: Vision Models and Metrics," *John Wiley & Sons, Ltd*, 2005, 200 pages.

[13] R. R. R. Rao, S. Göring, W. Robitza, B. Feiten, and A. Raake, "AVT-VQDB-UHD-1: A Large Scale Video Quality Database for UHD-1," in *2019 IEEE ISM*, Dec 2019, pp. 1–8.

[14] Netflix, "VMAF - Video Multi-Method Assessment Fusion," https://github.com/Netflix/vmaf, [Online: accessed 12-Feb-2022].

[15] N. Barman, S. Zadtootaghaj, S. Schmidt, M. G. Martini, and S. Möller, "GamingVideoSET: A Dataset for Gaming Video Streaming Applications," in *2018 16th Annual Workshop on Network and Systems Support for Games (NetGames)*, 2018, pp. 1–6.

[16] ITU-T SG 12 (Study Period 2017) Temporary Document 1612-GEN, *Output - Draft New Recommendation J.op-tr "Methods for Optimizing Bitrates and Transmission Resolution by Considering Display Characteristics and Available*, Std., Oct 2021, https://www.itu.int/md/T17-SG12-211012-TD-GEN-1612.

[17] T. Teixeira, B. Zhang, and Y. Reznik, "Adaptive Streaming Playback Statistics Dataset," in *Proceedings of the 12th ACM Multimedia Systems Conference*, ser. MMSys '21, 2021, p. 248–254. [Online]. Available: https://doi.org/10.1145/3458305.3478444

[18] Y. Bababekova, M. Rosenfield, J. E. Hue, and R. R. Huang, "Font Size and Viewing Distance of Handheld Smart Phones," *Optometry and Vision Science*, vol. 88, no. 7, pp. 795–797, 2011.

[19] Y. A. Reznik, "Average Performance of Adaptive Streaming," in *2021 Data Compression Conference (DCC)*, 2021, pp. 263–272.

[20] A. Mackin, M. Afonso, F. Zhang, and D. Bull, "A Study of Subjective Video Quality at Various Spatial Resolutions," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 2830–2834.

[21] Nabajeet Barman and Yuriy Reznik and Rahul Vanam , "Optimal Rendition Resolution Selection Algorithm for Web Streaming Players," in *SPIE Optical Engineering + Applications*, vol. OP22O, no. OP332-50, 2022, pp. 1–10, accepted.

[22] ITU-T Recommendation H.266, *H.266 : Versatile Video Coding*. Geneva: International Telecommunication Union, 2020.

[23] ITU-R Recommendation BT.500-14, *Methodologies for the Subjective Assessment of the Quality of Television Images*. Geneva: International Telecommunication Union, 2019.

[24] Yuriy Reznik and Nabajeet Barman and Rahul Vanam, "Parametric Quality Models for Multiscreen Video Systems," in *10th European Workshop on Visual Information Processing (EUVIP)*, Lisbon, Portugal, Sept 2022, pp. 1–6, submitted.