https://doi.org/10.2352/EI.2023.35.3.MOBMU-351

Improving the Performance of Web-Streaming by Super-Resolution Upscaling Techniques

Yuriy Reznik and Nabajeet Barman; Brightcove, Inc.; Seattle, WA/USA

Abstract

In recent years, we have seen significant progress in advanced ML/AI-based image and video upscaling techniques, often commonly referred to as super-resolution (SR). Such algorithms are now available not only in the form of specialized software but also in drivers and SDKs supplied with modern graphics cards. Upscaling functions in NVIDIA Maxine SDK is one of the recent examples. However, to take advantage of this functionality in video streaming applications, one needs to (a) quantify the impact of super-resolution techniques on the perceived visual quality, (b) implement video rendering incorporating super-resolution upscaling techniques, and (c) implement new adaptation algorithms in streaming players, enabling such players to deliver better quality of experience or better efficiency (e.g., reduce bandwidth usage) or both.

Towards this end, we propose several techniques that may be helpful to the video streaming community. First, we offer a model quantifying the impacts of super-resolution upscaling on the perceived quality. Our model is based on a generalized version of the classic Westerink-Roufs model connecting the true resolution of images/videos to perceived quality. The generalized model provides several additional parameters, allowing its tuning to specific implementations of super-resolution techniques. We verify this model using several recent datasets, including MOS scores measured for several conventional upscaling and super-resolution algorithms. Then, we propose improved adaptation logic for video streaming players, considering encoded video resolutions, player size, and the upscaling method. This improved logic relies on our modified Westerink-Roufs model to predict perceived quality. The proposed algorithm suggests choices of renditions for SR upsampling algorithms that would deliver approximately the same quality as obtained using traditional upsampling algorithms while resulting in considerable bandwidth savings.

Introduction

Recent years have seen increasing growth and acceptance of streaming services with adaptive streaming, where the playback is adapted based on the changing network conditions being one of the fundamental technologies enabling a good user experience. In adaptive streaming of videos, the video is encoded in different representations, often called renditions. One of the most widely used adaptive streaming formats used by most over-the-top (OTT) service providers is HTTP-based Adaptive Streaming (HAS), where the streaming takes place over reliable transport protocols such as TCP.

In a typical HAS solution, such as HTTP Live Streaming (HLS) [6] and Dynamic Adaptive Streaming over HTTP (DASH) [7], the video is encoded in multiple resolution-bitrate pairs. The streaming client (player), depending on the available network throughput, buffer status, and player size, selects the appropriate rendition for playback [19]. The player at the end-user device typically upscale (or in some cases, downscales) the videos to fit the player/window resolution. In the special case of web streaming, as discussed by the authors in [19], player window size significantly impacts the selection of streams. In such systems where the videos are delivered embedded in web pages, the network bandwidth is no longer the only factor influencing the selection of streams. Many modern streaming clients also consider the player (window) size as one of the factors in their adaptation logic [18]. However, in most cases, the adaptation logic is very simplistic, e.g., limiting the upscale factor, selecting the nearest matching resolution in the ladder, etc. Such simple resolution adaptation algorithms do not necessarily account for viewing setup parameters such as pixel density and viewing angle. More importantly, to the best of the authors' knowledge, the existing resolution adaptation algorithms do not account for the effect of upsampling methods being used on the client side.

Advanced (AI-based) Upsampling Algorithms

So far, the traditional image/video scaling in web browsers has been limited to classical signal processing-based techniques such as *bicubic* interpolation or *sinc* and *lanczos* filters. More recently, however, there has been a growing interest and work in the field of AI/ML based upscaling, often termed as Super Resolution (SR) techniques [22, 10]. Such algorithms are primarily based on deep learning technologies such as Convolutional Neural Networks (CNNs) and, more recently, Generative Adversarial Networks (GANs). Such advanced SR algorithms are typically designed and used to perform $\times 2$ and $\times 4$ upsampling of lower-resolution images [9, 24, 4] and videos [20, 3, 15], with most of them outperforming existing traditional upsampling algorithms.

Due to their improved performance over existing traditional upsampling algorithms, such advanced algorithms are getting increasingly popular, with many companies offering such AI-based upscaling solutions[1,2]. Another more popular and widely used example is the NVIDIA Maxine which is made available as part of NVIDIA's graphic card drivers and is used in many applications, especially in gaming. However, such a rapidly evolving field brings a new range of questions that need to be answered before such technologies can be adopted into mainstream video streaming applications, a few of which are discussed next.

Open Questions

1. *What are the advantages of SR over traditional scaling?* Many of the proposed models have been evaluated on a

[1] https://www.topazlabs.com/topaz-video-ai
[2] https://www.avclabs.com/

small number of datasets. Their performance evaluation for real-world video streaming applications considering adaptive streaming applications and their benefits compared to traditional upscaling algorithms remains an open question.

2. *How to model/quantify super-resolution scaling capability?* Most of the performance metrics used to quantify existing SR techniques have been limited to PSNR and, in some cases, SSIM [21]. However, such metrics are often limited in terms of their correlation with human-perceived quality and, oftentimes, unsuitable for measuring the quality of AI-based algorithms [25].

3. *How to use SR for improved image/video delivery?* Advanced AI-based upsampling algorithms and newer, alternative contents such as HDR result in increased quality saturation at lower angular resolutions. Understanding the impacts of the encoded video resolution and scaling algorithms on the perceived quality is essential for applications to allow them to select the optimal renditions [19, 16]. Such intelligent adaptation algorithms, considering the effect of upsampling algorithms, can result in significant bandwidth and storage savings.

Contributions

Towards this end, this paper presents several contributions which can help understand and quantify the impacts of more advanced, AI-based upscaling algorithms with a focus on optimal rendition selection by streaming players/clients. We first present and discuss a generalized model of the well-known Westerink and Roufs model, which uses angular parameters, viewing angle, and angular resolution to predict the perceived picture quality. The generalized model provides additional parameters which allow it to be tuned to adapt to the differences in perceived picture quality due to the use of different upsampling algorithms. Using the generalized model, we present an improved adaptation logic for video streaming clients, considering the player size and the upsampling algorithm used at the client for upscaling the received video. We conclude the paper with some results for a sample case, demonstrating the utility of the proposed model for selecting an optimal set of renditions, resulting in significant bandwidth savings.

Understanding the Impact of Scaling on Perceived Quality

Angular Metrics

Table 1 presents a list of the main parameters of the video, player, and characteristics of the viewing setup. Figure 1 illustrates a typical video reproduction chain explaining the relationship between the various parameters considering a typical case where an encoded video sequence of size $W \times H$ [pixels] is scaled to fit a player/display of size $W_p \times H_p$ [pixels]. Here d [inches] is the viewing distance between the observer and the display. Of prime interest to us are the two angular metrics: (a) viewing angle (ϕ), which is the angular span of the video frame, as visible on screen, and (b) angular resolution (μ), which is the maximum spatial frequency that can be reproduced by the display. These two angular metrics are discussed next.

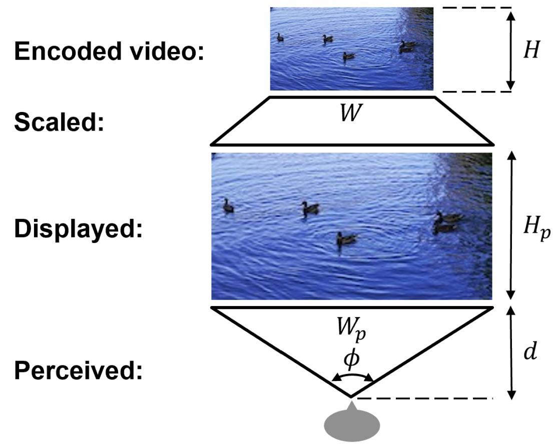


Figure 1: A typical video reproduction setting illustrating an example case where an encoded video of dimension $W \times H$ pixels is scaled and displayed on a screen of dimension $W_p \times H_p$ pixels. The distance between the observer and the screen is d inches, and ϕ is the angular size called as picture angle or viewing angle.

Table 1: List of commonly used parameters of player, video, and characteristics of viewing set.

Parameter	Description	Units
W_p	display/player window width	pixels
H_p	display/player window height	pixels
W	horizontal image resolution	pixels
H	vertical image resolution	pixels
d	viewing distance	inches
ρ	display pixel density	pixels per inch

Viewing Angle

Viewing angle, ϕ describes the horizontal angular size of the video as visible to the user. Considering a video player window of size $W_p \times H_p$, display pixel density ρ, and viewing distance d, the viewing angle ϕ can be computed as:

$$\phi = 2 \arctan\left(\frac{W_p}{2d\rho}\right). \tag{1}$$

Angular Resolution

The angular resolution effectively describes the Nyquist frequency of the video, presented in angular units. Using the same parameters as defined above and considering that the resolution of the video being played is $W \times H$ [pixels], the angular resolution, μ, of the video at that resolution can be computed as:

$$\mu = \left(2 \arctan\left(\frac{W_p}{W d\rho}\right)\right)^{-1}. \tag{2}$$

Encoding Ladders

A typical adaptive bitrate streaming application encodes the media sequence into multiple resolution-bitrate pairs. Multiple versions of a particular media sequence are often called renditions (sometimes also as streams) of the adaptive bitrate ladder. Considering such an application, let $H_1, H_2, ...H_n$ be the height (in pixels) of the different renditions available for a particular video stream. Here n refers to the number of renditions that can vary from video to video. In this work, we will assume a fixed aspect ratio of all the renditions in an ABR ladder for a particular video for simplicity. This simplification will allow for the specification of a single resolution parameter, e.g., height H_i, to derive the other.

Video Player Sizes

In a given display of resolution $W \times H$, the player window size can vary as per user choice. Let H_p be the player window height, where H_p can vary, e.g., 240p up to the maximum display height, H. For simplicity, similar to renditions, it is considered that the aspect ratio is fixed for a given video streaming session. Hence, knowledge of one parameter among width and height of the player window size is sufficient.

Quality model based on viewing setup parameters

A very well-known basic model of perceived quality based on the parameters of viewing setup is the Westerink and Rouf (WR) model [23]. The authors found that at a constant viewing distance, the subjective image quality was influenced independently by both the viewing angle of the projected image as well as the angular resolution of the projected picture on display. The model since then has been validated by others and used in many works [11, 1, 17] and more recently in [13].

In this work, we use the generalized version of the WR model proposed in [13]:

$$Q(\phi, \mu) = \log\left(\alpha + \beta\left(1 + \left(\frac{\phi}{\phi_s}\right)^{-k}\right)^{-\frac{\gamma}{k}}\left(1 + \left(\frac{\mu}{\mu_s}\right)^{-l}\right)^{-\frac{\delta}{l}}\right) \quad (3)$$

where γ, δ, k, l, ϕ_s and μ_s are model parameters controlling the behavior with respect to viewing angle and angular resolution. As we will see later, this allows the model to be tuned to consider differences in HDR and SDR content and upsampling algorithms.

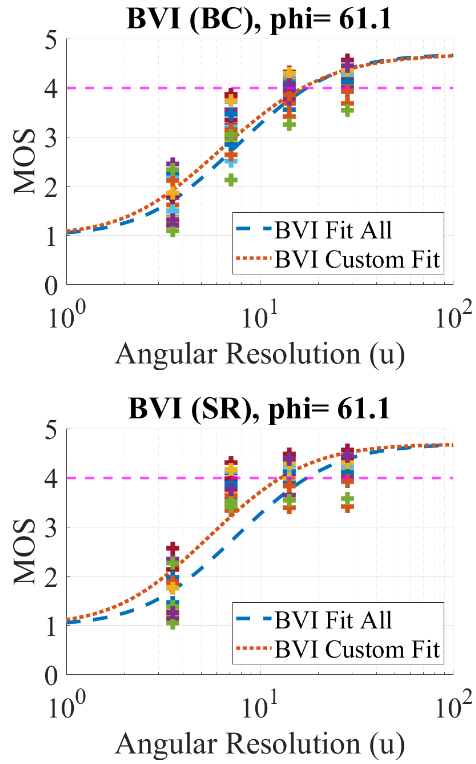
By fitting the Generalized WR (GWR) model to the six modern datasets considering different viewing setups (UHD TV to smartphones and tablets) and resolutions (QCIF to 4k/UHD) combined, the model parameters obtained are α= 2.72, β=145.69, γ = 1.55, δ=2.12, k=6.01, l=2.11, ϕ_s=35.0, and μ_s=16.93. The proposed model outperforms the original WR model with authors in [14] using this along with other distortion metrics to propose parametric quality models for multi-screen systems.

Modeling the Effects of Super Resolution

For AI/ML based upsampling algorithms such as Super Resolution (SR) when compared to the traditional algorithms such as Nearest Neighbour (NN) or BiCubic (BC) interpolation algorithms, it is observed that MOS scores often reach saturation at lower angular resolutions. In this section, we discuss how the proposed model can be re-tuned to consider the differences in subjective perception due to different upsampling algorithms.

Dataset

The open-source *BVI* dataset [12] provides subjective scores considering three different upsampling algorithms. The dataset consists of 24 10-bit five-second source reference videos sequences of 60 fps encoded in three different resolutions (1920×1080, 960×540 and 480×270), which were then upsampled to the source resolution of 3840x2160p using three different upsampling algorithms: BiCubic (BC), Nearest Neighbour (NN) and Super Resolution (SR) [8]. The upscaled videos were then displayed to test subjects on a display measuring 65.4 × 36.8 cm of BT.2020 color space (full range) at a viewing distance of 1.5H.



(b) Super Resolution Upsampling.

Figure 2: MOS vs Angular resolution (cpd) plot considering two 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 the full BVI dataset) are shown. The colors of the markers in each plot represent a particular video sequence.

Model Refit to HDR content

Since the new dataset is of a different content type (10-bit HDR content) than the GWR model was designed for, we first perform a refit of the GWR model using all MOS scores from the BVI dataset. Allowing for different values of α and β, as well as adding a common scale factor, ε for γ and δ, we obtain the new parameter values as: $\alpha = 2.72, \beta = 106.91, \varepsilon = 1.08, \gamma = 1.55\varepsilon, \delta = 2.12\varepsilon$. Figure 2 shows the plot of MOS vs angular resolution (μ) for two of the upsampling algorithms (BiCubic and Super Resolution) from the BVI dataset considering the generic fit ("BVI Fit All") to all of the BVI dataset using default GWR model parameters. One can observe that while the fit captures the behavior when considering the traditional upsampling algorithm, BiCubic, the fit is not that great when considering the Super Resolution upsampling. Hence, to take into account the differences

Table 2: 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
Bi-Cubic (BC)	13.93	1.76
Super Resolution (SR)	12.24	2.06

due to upsampling algorithm, one needs to perform a re-tuning to each specific algorithm subset of the dataset, as discussed next.

Model Re-tuning

To account for the differences in upsampling algorithms, we now fit the GWR model to all three upsampling algorithms (BC, SR, and NN) subsets of the BVI dataset corresponding to the upsampling method used. To adapt to the differences, we allow only parameters μ_s and l to vary (since they control the model behavior wrt μ). Rest all parameter values are set to default (as obtained from the refit to the BVI dataset). The new fit values of μ_s and l for BC and SR upsampling algorithms are summarized in Table 2. The customized fit to each upsampling algorithm with new fitting parameters (μ_s) and l) ("BVI Custom Fit") is also shown in Figure 2. It can be noted that the re-tuning of the model by adapting model parameters, μ_s, and l helps take into account the differences in the upsampling algorithm, resulting in a much better fit.

In Figure 2, the pink dashed line MOS = 4 demonstrates an example where one can establish a "baseline" MOS that could be used to compare relative savings. For example, suppose one wants to provide a service where the customers get a "good" MOS score (here assumed to be MOS=4). In that case, one can see that if SR upsampling algorithm is used on the client side, this can be achieved at a lower angular resolution of 12.8. In contrast, a similar MOS score for BC upsampling is achieved at higher angular resolutions of 16.6. Service providers can then use such information to compute relative savings in encoding resolutions or optimal encoding ladder generation.

Note: The GWR model was retuned considering MOS scores of all three upsampling algorithms (BC, NN, and SR) of the BVI dataset. However, for brevity and ease of comparison between traditional and AI-based algorithms, Figure 2 and Table 2 present the results for only two upsampling algorithms, one traditional (BiCubic) and one AI-based (Super Resolution).

Optimal SR-aware Adaptation Algorithms

As discussed earlier, in a modern-era adaptive streaming system delivering videos embedded in web pages, the stream selection logic is jointly influenced by both available network bandwidth and output video size (player size). While adaptation to network bandwidth has been widely studied [2], stream adaptation based on the resolution of available renditions and the player size has received little attention [16]. In this section, we present a discussion about improving the ABR algorithms considering the second aspect, i.e., adaptation to resolution, while also taking into account the effect of client-side upsampling algorithms (traditional vs super resolution based upsampling).

Adaptation Algorithms

We present two optimal rendition resolution selection algorithms based on the player size and the upsampling algorithm.

Algorithm 1: Optimal Rendition Resolution Selection Based on Player Size and BiCubic Upsampling Algorithm

Data:
Viewing angle ϕ
Angular resolution μ
Available video rendition heights, $H_{renditions} = H_1, ...H_n$, such that $H_1 \leq .. \leq H_n$
Player Window Height H_p
Distance from the display d
Effective pixel density of the screen, ρ
Model fit parameter values, $\alpha = 2.72$, $\beta = 106.91$, $\varepsilon = 1.08$, $\gamma = 1.55\varepsilon$, $\delta = 2.12\varepsilon$.

Result: Best rendition height (BiCubic Upsampling), $H_{best_{BC}}$, and Best MOS (BiCubic Upsampling), $MOS_{best_{BC}}$

$best_{mos} = 0$;
$best_{rendition-index} = 1$;
for $i \leftarrow 1$ **to** n **do**
  Calculate Viewing angle ϕ
  Calculate Angular resolution μ
  $\mu_s = 13.93$; $l = 1.76$; `/* BiCubic Upsampling, Table 2 */`
  Calculate MOS, $Q(\phi, \mu)$; `/* Using Eqn 3 */`
  if $MOS\ is \geq best\_mos$ **then**
    best_mos = MOS ;
    $best_{rendition-index} = i$;
  end
end
$H_{best_{BC}} = H_{renditions}(best_{rendition-index})$
$MOS_{best_{BC}} = best_{mos}$

The first algorithm, Algorithm 1, finds the best set of renditions from a list of available renditions, which delivers the best possible quality for a given viewing setup (viewing distance, player size, and device type) when considering the BiCubic upsampling algorithm. The perceived quality is estimated using the retuned GWR model (Eqn 3), with saturation point $\mu_s = 13.93$, and slope, l =1.76 (corresponding to BiCubic upsampling, cf. Table 2).

The second algorithm, Algorithm 2, in addition, takes as input the quality of the "optimal" renditions corresponding to BiCubic upsampling obtained using Algorithm 1. Considering the retuned GWR model (Eqn 3) saturation point (μ_s) and slope (l) for angular resolution parameters for Super Resolution (SR) upsampling (μ_s = 13.93 and l=1.76), this algorithm select renditions which are approximately of the same quality as obtained using Algorithm 1.

It should be noted that the presented algorithms are simplistic in nature and take into account only available rendition resolution and upsampling algorithms into account. It is assumed for simplicity that the available renditions in the bitrate ladder generated are optimal in nature (of the highest possible quality, thus minimizing any compression or scaling-related artifacts). The algorithms assume that the player dimensions ($W_p \times H_p$) cannot exceed the dimension of the display, which is a logical assumption. As mentioned before, we assume that the display aspect ratio

Algorithm 2: Optimal Rendition Resolution Selection Based on Player Size and **Super Resolution** Upsampling Algorithm

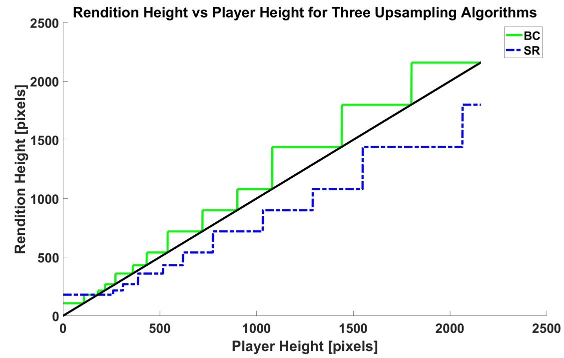
Data:

Viewing angle ϕ
Angular resolution μ
Available video rendition heights, $H_{renditions} = H_1, ...H_n$, such that $H_1 \leq .. \leq H_n$
Player Window Height H_p
Distance from the display d
Effective pixel density of the screen, ρ
Model fit parameter values, $\alpha = 2.72$, $\beta = 106.91$, $\varepsilon = 1.08$, $\gamma = 1.55\varepsilon$, $\delta = 2.12\varepsilon$.
MOS value corresponding to BiCubic Upsampling for the same device, $MOS_{best_{BC}}$

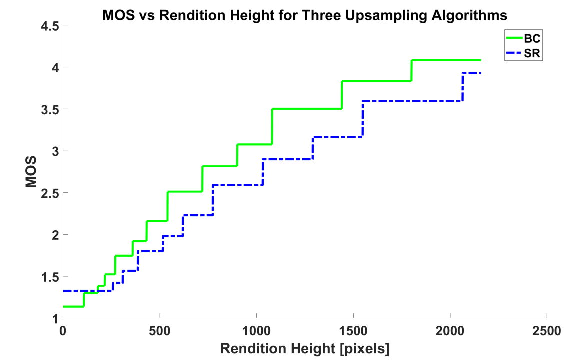
Result: Best rendition height (Super Resolution Upsampling), $H_{best_{SR}}$, and Best MOS (Super Resolution Upsampling), $MOS_{best_{SR}}$

$best_{mos} = MOS_{best_{BC}}$; `/* MOS from Algorithm ?? */`
$best_{rendition-index} = 1$;
for $i \leftarrow 1$ **to** n **do**
  Calculate Viewing angle ϕ
  Calculate Angular resolution μ
  $\mu_s = 12.24$; $l = 2.06$; `/* SR Upsampling, Table 2 */`
  Calculate MOS, $Q(\phi, \mu)$; `/* Using Eqn 3 */`
  if MOS *is* $\geq$ *best_mos* **then**
    best_mos = MOS ;
    $best_{rendition-index} = i$;
    break; ; `/* Minimum Rendition found, exit */`
  end
end
$H_{best_{SR}} = H_{renditions}(best_{rendition-index})$
$MOS_{best_{SR}} = best_{mos}$

(DAR) is fixed, which in our case is assumed to be 16:9 (one of the most commonly used DAR values). As discussed earlier, the parameters (μ_s and l) are adapted depending on the upsampling algorithm used at the client end.



(a) Rendition Height (pixels) vs Player Height (pixels) for BiCubic (BC) and Super Resolution (SR) Upsampling.



(b) MOS vs Rendition Height (pixels) for BiCubic (BC) and Super Resolution (SR) Upsampling.

Figure 3: MOS and Best Rendition Selection for both SR and BC algorithms.

Table 3: List of available renditions considered in this work.

Horizontal (Width)	Vertical (Height)
3840	2160
3200	1800
2560	1440
1920	1080
1600	900
1280	720
960	540
768	432
640	360
480	270
384	216
320	180
192	108

Results

We present next the results of the selected renditions and MOS scores obtained for the two different upsampling algorithms for one sample test case considering the BVI dataset [12]. The device parameters used in the design of the original BVI dataset are as follows: display size of 3840x2160, ppi = 202.57, and viewing distance = 21.75 inches. We consider that 13 different renditions are available irrespective of the device type. We also assume that the renditions are proper in that the resolutions are non-decreasing such that $0 < H_1 \leq ... \leq H_n$ for all renditions in the ladder. The resolutions of the considered renditions are the recommended resolution values in the DVB Bluebook A168 [5] as shown in Table 3. These values are typical of any streaming solution varying from very low resolution (192×108) to UHD (3840×2160) and hence will allow us to obtain realistic performance figures.

Figure 3a presents the plot for Rendition Height (pixels) vs Player Height (pixels) for both BiCubic (BC) and Super Resolution (SR) upsampling algorithms. Here, Algorithm 1 is used to select the rendition delivering the best possible quality by considering standard BiCubic upscaling at the client end. On the other hand, the chosen renditions for SR upsampling are based on Algorithm 2, which tries to find the best rendition matching the level of quality achievable with algorithm 1. Based on Figure 3a, one can observe that SR-based upsampling enables much more conservative choices of rendition resolutions. In this example case considering the TV screen, we see about a 16% reduction in frame height or 30% in pixel count in the high-resolution regime. This indicates that the use of SR techniques can enable significant savings in network bandwidth. Figure 3b presents the MOS vs Rendition Height (pixels) for both BiCubic (BC) and Super Resolution (SR) upsampling. We can observe that the difference in quality (MOS) between the two upsampling algorithms is quite low, however, resulting in significant bandwidth savings as was observed in Figure 3a.

Conclusions and Future Work

In this paper, we presented a model based on parameters of viewing setup, viewing angle, and angular resolution. We demonstrated how the model can be re-tuned to account for differences in HDR content and upsampling algorithms. We also presented two algorithms to select optimal rendition based on the upsampling algorithm implemented on the client side. Using a sample case study of a TV device, we demonstrated how such upsampling algorithms aware adaptation algorithms can result in significant bandwidth savings. Such approaches, when used with existing Context Aware Encoding (CAE) solutions, can lead to even more bandwidth and storage savings, influence encoding decisions for manifest ladder generation, and improve adaptation algorithms.

This work was a first step towards demonstrating the utility and potential in terms of the benefits of using SR-based upsampling algorithms. However, for more widespread adoption by the industry, there is a need for clearly defined APIs universally supported by all browsers and platforms. At the same time, there is a need for understanding and ensuring the utilities/benefits of different SR implementations across different devices (TVs, PCs, Mobiles), codecs, etc. Our future work in this direction will address some of these open questions.

References

[1] P. Barten. "The effects of picture size and definition on perceived image quality". In: *IEEE Transactions on Electron Devices* 36.9 (1989), pp. 1865–1869.

[2] A. Bentaleb et al. "A Survey on Bitrate Adaptation Schemes for Streaming Media Over HTTP". In: *IEEE Communications Surveys & Tutorials* 21.1 (2019), pp. 562–585.

[3] K. C. Chan et al. "BasicVSR: The Search for Essential Components in Video Super-Resolution and Beyond". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2021.

[4] Y. Chen, S. Liu, and X. Wang. "Learning continuous image representation with local implicit image function". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021, pp. 8628–8638.

[5] *Digital Video Broadcasting (DVB) BlueBook A168 (Draft ETSI TS 103 285 V1.3.1)*. https://dvb.org/wp-content/uploads/2019/12/a168_dvb_mpeg-dash_oct_2019.pdf. [Online: Accessed 8-Dec-2022]. Oct 2019.

[6] *HTTP live streaming, RFC 8216*. https://tools.ietf.org/html/rfc8216. [Online: accessed 19-Sept-2022]. 2019.

[7] *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. [Online: accessed 17-Sept-2022]. 2019.

[8] J. Kim, J. K. Lee, and K. M. Lee. "Accurate Image Super-Resolution Using Very Deep Convolutional Networks". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 1646–1654.

[9] C. Ledig et al. "Photo-realistic single image super-resolution using a generative adversarial network". In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2016.

[10] H. Liu et al. "Video Super Resolution Based on Deep Learning: A comprehensive survey". In: *CoRR* abs/2007.12928 (2020). arXiv: 2007.12928.

[11] M. Lombard et al. "The role of screen size in viewer responses to television fare". In: *Communication Reports* 10.1 (1997), pp. 95–106. eprint: https://doi.org/10.1080/08934219709367663.

[12] A. Mackin et al. "A Study of Subjective Video Quality at Various Spatial Resolutions". In: *2018 25th IEEE International Conference on Image Processing (ICIP)*. 2018, pp. 2830–2834.

[13] Nabajeet Barman, Rahul Vanam and Yuriy Reznik. "Generalized Westerink-Roufs Model for Predicting Quality of Scaled Video". In: *The 14th International Conference on Quality of Multimedia Experience (QoMEX)*. Lippstadt, Germany, 2022, pp. 1–6.

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

[15] Z. Qiu et al. "Learning Spatiotemporal Frequency-Transformer for Compressed Video Super-Resolution". In: *ECCV*. 2022.

[16] Y. Reznik, N. Barman, and R. Vanam. "Optimal rendition resolution selection algorithm for web streaming players". In: *Applications of Digital Image Processing XLV*. Vol. 12226. International Society for Optics and Photonics. SPIE, 2022, 122260F.

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

[18] Y. A. Reznik, K. O. Lillevold, and R. Vanam. "Perceptually Optimized ABR Ladder Generation for Web Streaming". In: *Electronic Imaging* 2021.3 (2021), pp. 75–1–75–11.

[19] Y. A. Reznik et al. "Towards Understanding of the Behavior of Web Streaming". In: *2021 Picture Coding Symposium (PCS)*. 2021, pp. 1–5.

[20] X. Wang et al. "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks". In: *Computer Vision – ECCV 2018 Workshops*. Ed. by L. Leal-Taixé and S. Roth. Cham: Springer International Publishing, 2019, pp. 63–79.

[21] Z. Wang et al. "Image Quality Assessment: From Error Visibility to Structural Similarity". In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612.

[22] Z. Wang, J. Chen, and S. C. H. Hoi. "Deep Learning for Image Super-resolution: A Survey". In: *CoRR* abs/1902.06068 (2019). arXiv: 1902.06068.

[23] J. H. D. M. Westerink and J. A. J. Roufs. "Subjective Image Quality as a Function of Viewing Distance, Resolution, and Picture Size". In: *SMPTE Journal* 98.2 (1989), pp. 113–119.

[24] F. Yang et al. "Learning texture transformer network for image super-resolution". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 5791–5800.

[25] R. Zhang et al. "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric". In: *2018 IEEE/CVF Conference on Com-*

puter Vision and Pattern Recognition (CVPR). Los Alamitos, CA, USA: IEEE Computer Society, 2018, pp. 586–595.

Author Biography

Yuriy A. Reznik is a technology fellow and head of research department at Brightcove, Inc., Seattle, WA, USA. Previously he worked at RealNetworks, Seattle, WA (1998-2005), Qualcomm, San Diego, CA, (2005-2011), and InterDigital, San Diego, CA (2011-2016), and contributed to several wellknown standards (ITU-T H.264/MPEG-4 AVC, ITU-T H.265/MPEG HEVC, MPEG DASH, etc.) as well as products in the domains of internet streaming and wireless multimedia. He received the Ph.D. degree in computer science from Kyiv University, Kyiv, Ukraine, in 2005. In 2008, he was also a visiting scholar at Stanford University, Stanford, CA. He has coauthored more than 100 papers and co-invented more than 60 granted U.S. patents related to his studies and work.

Nabajeet Barman is a Principal Video Systems Engineer, Research at Brightcove, London, UK, and an Adjunct Lecturer (Assistant Professor) in Applied Computer Science at Kingston University London, United Kingdom. He obtained his PhD and MBA from Kingston University (2019), Masters in IT from Universität Stuttgart (2014), Germany, and a Bachelor of Technology (2011) in Electronics Engineering from the National Institute of Technology, Surat, India. Previously, he was a full-time lecturer and Marie-Curie fellow at Kingston University and a research intern at Bell Labs, Stuttgart. He is also a Fellow of the Higher Education Academy, UK, and is a board member of the Video Quality Expert Group (VQEG). He is also involved in other standardization activities.