SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I1P101

Volume 11 Issue 1, 1-7, January 2024

Original Article

Mapping of Temporal Space Slicing for Video Quality Metrics Assessment

Shilpa Bagade[1], Budati Anil Kumar[2], L. Koteswara Rao[3]

[1,2,3]*Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad, India.*
[1]*Department of Electronics and Communication Engineering, Gokaraju Rangaraju Institute of Engineering and Technology, Telangana, India.*
[2]*Institute of Computer Science And Digital Innovation, UCSI University, Malaysia.*

[1]*Corresponding Author : shilpa.me1437@gmail.com*

Received: 12 October 2023 Revised: 14 November 2023 Accepted: 12 December 2023 Published: 20 January 2024

***Abstract** - Full-Reference (FR) video quality evaluation approach that combines frame-based Visual Quality Assessment (VQA) with analysis of space-time slices to provide an efficient video quality predictor is proposed. The sample and test video clips are first put into a temporal space slice form by the proposed method. Each reference-test video pair is subjected to the computation of several distortion-aware maps to define space-time distortions more thoroughly. Then, a standard visual quality model, such as Peak Signal to Noise Ratio (PSNR) or Structural Similarity Index (SSIM), is used to process these reference-distorted maps. A final video quality score is created by combining several VQA outputs using a straightforward, learnt pooling method. The method thoroughly evaluated the Temporal Space Slicing (TSS) algorithm using three publicly accessible video quality assessments and discovered that TSS-PSNR performed noticeably better than leading-edge video quality models.*

***Keywords** - HEVC, Packet loss, Video streaming, Video compression, Video Quality Metrics.*

1. Introduction

In recent years, video-based applications such as digital television, camera surveillance, and video teleconference have become increasingly popular in all spheres of society. Particularly with the advancement of wireless and video technology, individuals can capture recordings of their daily lives at any time they choose, using portable mobile devices and through social networks. Without a doubt, video traffic has increased most of the bandwidth requirements on the Internet. Videos, however, expire ahead of time and finally reach the processing phases, end users usually, and finally, human consumers. The majority of these phases attempt to increase the apparent video quality, while others attempt to decrease it. In order to provide a positive end-user experience, Video Quality Assessment (VQA) is a crucial step in many video-based applications [1].

Many practical uses for VQA exist, such as optimizing video systems perceptually and evaluating how well they work in terms of visual capture, shrinking, transmission, augmentation, and display. Video assessment can be accomplished through the use of either subjective or objective VQA. The most accurate technique for estimating perceived video quality is subjective VQA, which rates recipients' estimated video quality and uses arbitrary scores to get the overall video quality score. However, because subjective tests are cumbersome, expensive, and time-consuming, it is challenging to conduct the subjective study in real-time video-based systems.

Even though subjective VQA has so many flaws, it is nevertheless vital. The results of the objective VQA methodologies need to be compared to the "ground truth" provided by subjective VQA. Many researchers devote their time and effort to building subjective benchmarking databases [2].

The approach for fast subjective VQA based on hybrid data along with active learning built on the idea of interactive instruction for data inscription, HA-SVQA, is proposed [3] with iterative evaluations of the most beneficial or instructive movies, which are chosen using data from both the subject's prior choices and the objective quality forecasts. HA-SVQA can expedite the subjective VQA process by removing the unnecessary (or less important) films to be evaluated by creating a repository of video clips that may portray the range of videos studied subjectively, allowing the database to serve as a benchmark for video quality. For MPEG video, a multi-metric model that combines a blockiness detector and a perceptual model is suggested [4].

Because choices must be made locally, creating a statistic that can also be included in a video codec's rate-distortion optimization process can be more difficult. An approach that employs a fusion of local content elements to determine the optimal use of a number of state-of-the-art objective Video Quality Metrics at the coding block level [5] was developed.

A refined non-linear model that mimics the perception process of the Human Visual System (HVS) was suggested. It combined noticeable distortion and blurred vision aberrations adaptively. Noticeable distortion is determined by crossing absolute variances using the spatial and temporal tolerance maps that describe texture masking effects when the distorted video quality is similar to the original footage. This definition is essential for assessing quality. It has been discovered that metric performance can be further enhanced by characterizing blurring artefacts, which are estimated by computing high-frequency energy changes and weighted with motion speed [6].

The High-Efficiency Video Coding (HEVC) standard is the most recent joint video initiative of the ISO/IEC Moving Picture Experts Group (MPEG) and the ITU-T Video Coding Experts Group (VCEG) standardization organizations, working together as the Joint Collective Team on Video Coding (JCT-VC) [7-8]. The first release of the HEVC standard is expected to be finished in January 2013, at which point ISO/IEC and ITU-T will publish an aligned document. Additional work is planned to enhance the standard to cover other application scenarios, including extended-range usage with optimum resolution and colour format support, scalable video coding, and 3-D/stereo/multiview video coding. The ISO/IEC 23008-2 MPEG-H Part 2 and the ITU-T MPEG-H standard will replace the HEVC standard [9].

Multiple versions of video coding standards are evaluated for their compression powers using the Peak Signal to Noise Ratio (PSNR) and qualitative testing results. The designs that are examined utilizing a consistent methodology include H.262/MPEG-2 Video, H.263, MPEG-4 Visual, H.264/MPEG-4 Advanced Video Coding (AVC), and High-Efficiency Video Coding (HEVC). Subjective experiments on WVGA and HD sequences demonstrate that HEVC encoders can attain the same perceptual replication quality as encoders compliant with H.264/MPEG-4 AVC, provided they are utilized at an average bit rate that is approximately 50% lower. It has been shown that the HEVC design offers significant advantages for low bit rates, precise video files, and short-delay communication applications. The measured subjective improvement is marginally more significant than the PSNR meter's measurement of improvement [10].

The most popular options for vehicle networks supporting different kinds of applications are in 5G mobile communication technology. In order to achieve more excellent video streaming quality, the double-buffer technique [11] is introduced to lessen the delay impact produced by a vehicle's frequent exchanges between 5G small cells and the irregular link effect caused by millimetre-wave propagation features. To evaluate the perceived quality of videos with packet loss, there are two types of measurement errors for full-reference saliency-based quality metrics [12].

The spatial variation in saliency values between the initial and deformed videos is measured, as well as a calculated average of pixel imperfections among the genuine and deformed films. The outcome of both measurements additionally makes use of the temporal fluctuation of the saliency map of the warped video. There must be an extensive, unbiased, publicly available database of distorted movies and subjective ratings [13]. A pertinent topic for the resilient design and adaption of multimedia infrastructures, services, and applications is the analysis of the effects of video content and transmission impairments on Quality of Experience (QoE) [14].

Large data sizes are produced when uncompressed video signals are transmitted. Higher fidelity, clarity, and resolution are also necessary when watching the video material. Utilizing videos exclusively for purposes like wearable cameras, remote home surveillance, and video chats at high resolutions like HD, UHD, and 4K increases video traffic and strains telecommunication networks and data preservation.

To address the expectations for such applications, 5G technology with low latency and high speed (100 times quicker than 4G) makes 5G more appealing. Packet loss via the internet is one of the many types of losses that frequently plague video transmission. It is possible to estimate the predicted average squared distortion resulting from compressed video packet loss by taking into account the significance of the dropped packet pattern, particularly the spike duration [15].

Vehicle communications are so important that they must be thoroughly tested and assessed. If analytical models can accurately simulate all the underlying factors that affect vehicular communications performance, they may offer a desirable and affordable method for this kind of assessment. So far, several analytical models relying on the IEEE 802.11p (or DSRC) standard have been presented to analyse vehicle communications. Nonetheless, current models typically simulate the Medium Access Control (MAC) in detail while oversimplifying the consequences of propagation and interference.

As a result, their usefulness as a substitute for measuring vehicular communications performance is diminished. To close this gap, the IEEE 802.11p standard-based novel analytical models that accurately represent the efficacy of vehicle-to-vehicle communications are presented in this study [16].

2. Space-Time Slice Mappings

Space-time slice mappings are a technique used in video quality assessment to analyze and evaluate the quality of a video sequence. This approach combines spatial and temporal information to assess the perceptual quality of a video. Here is an overview of how it works:

1. Space-time slices: in video quality assessment, a video sequence is divided into a series of space-time slices. These slices can be thought of as small chunks of video data where both spatial and temporal information is considered.
2. Spatial information: for each space-time slice, the spatial information is analyzed. This involves examining the quality of the individual frames or frames within a small time interval. Spatial information often includes the sharpness, contrast, colour accuracy, and other visual attributes of the frames.
3. Temporal information: the temporal aspect considers the change in video content over time. It evaluates how smoothly the video plays and how well it maintains a consistent frame rate. Temporal information may also include the presence of artefacts like frame drops, stuttering, or motion blur.
4. Feature extraction: various characteristics are retrieved from the spatial and temporal information of each space-time slice. These features can be both objective (quantitative) and subjective (perceptual). Objective features might include measures like Mean Squared Error (MSE), Peak Signal to Noise Ratio (PSNR), or Structural Similarity Index (SSIM). Subjective features could involve human assessments or models based on human perception.
5. Quality prediction: after extracting features from the space-time slices, a model or algorithm is used to assess the overall video quality. Machine learning techniques, like support vector machines or deep learning models, can be employed to make these predictions based on the extracted features.
6. Quality score: the output of the model is typically a quality score that reflects how good the video is thought to be. Higher scores indicate better video quality, while lower scores indicate poorer quality.

Space-time slice mappings provide a more comprehensive approach to video quality assessment by considering both spatial and temporal aspects of the video. This is important because video quality issues can arise from various sources, such as compression artefacts, temporal inconsistencies, or spatial distortions, and this method helps in capturing these issues accurately.

3. Quality Metrics

The most recent video encoding method, known as High-Efficiency Video Coding (HEVC), delivers far higher compression efficiency than previous coding guidelines. Because satellite broadcasts are lengthy in duration for each trip (RTT), using packet acknowledgements is challenging. Strict latency limits in satellite networks prevent such packet confirmations for applications that stream media in true time. Using UDP at the transport layer, combining the usage of Network Coding (NC) and Turbo Coding (TC) approaches to improve video quality across the noisy satellite links. [17].

When distributing live video via networks that are prone to errors, like wireless networks, there are two issues from the standpoint of video coding: random access and packet loss correction techniques that do not impact consumers or equilibrium consumption with dependable connections are hard to come, valuable information is given regarding the provision of error recovery and low-latency fast channel switching capabilities with minimal impact on quality [18]. The Full Reference Video Quality Assessment (FRVQA) methods may typically produce an acceptable performance because they have complete access to the reference data. Given that structural data has been shown to be crucial for Image Quality Assessment (IQA), it should also be helpful for Video Quality Assessment (VQA). Videos feature a third dimension over the period axis compared to photos. Therefore, for VQA, tracking data should also be essential.

Furthermore, learning more about the Human Visual System (HVS) is particularly beneficial for creating an FR-VQA approach that corresponds well with human perception, dividing the FR-VQA approaches into three groups, namely, methods that are directed by structural information, methods that are tuned for motion, and perceptually hybrid methods that HVS inspires. When it comes to video coding, random access and packet loss repair provide the two biggest obstacles to live video distribution across networks that are prone to errors, including wireless networks. The suggested approach lessens this effect by adding a companion stream made up entirely of keyframes to an efficient video stream for compression. In addition, comprehensive quantification sheds light on how to offer error recovery and rapid, minimal latency channel conversion with little effect on average video quality [19].

3.1. Video Quality Assessment

Retransmission is usually not employed in exceptionally interactive multimedia applications like telepresence, teleoperation, or video conferencing because of the application's tight deadline. In these situations, the misplaced or inaccurate data needs to be hidden. There is no set standard by which to evaluate the perceived quality of the several mistake-concealing strategies that are available. The performance of currently used metrics for assessing image and video quality (such as PSNR, SSIM, VQM, etc.) in assessing error-concealed video quality is examined in this work. Packet loss occurs to the encoded video, and several error concealment strategies are used to mask the loss. We demonstrate that the visual appeal of the frame with errors hidden alone does not always indicate the subjective quality of the video. Next, we utilize the metrics on the error-veiled photos and videos, and the goal of HEVC is to facilitate evaluating the output image frame quality in WVSNs [20]. For quality evaluation, both subjective and objective assessment markers are used. Subjective evaluation is provided by immediate display of the reconstructed frames.

3.1.1. Peak Signal to Noise Ratio

Peak Signal to Noise Ratio (PSNR) compares the greatest strength of an image to the amount of noise that is deteriorating to determine how well it can be reproduced. Using the most feasible power, an image must be compared to an optimum clear image to measure its PSNR. This is how PSNR is defined:

$$PSNR = 10\log_{10}\left(\frac{(L-1)^2}{MSE}\right) = 20\log_{10}\left(\frac{L-1}{RMSE}\right) \quad (1)$$

In this case, L stands for the total number of possible intensity levels in a picture where the minimum intensity level is 0. The following is a description of the average squared error or MSE:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(O(i,j) - D(i,j)^2) \quad (2)$$

The matrix data of the original and degraded images are denoted by the letters O and D, respectively. I is the index of the row inside the picture, where m stands for the number of rows of pixels. The term Mean Squared Error (MSE) refers to the relationship between the number of pixels (n) and the index (j) of each column inside the image.

3.1.2. Structural Similarity Index (SSIM)

The Structural Similarity Index (SSIM), which uses orthogonal quantitative metrics like brightness and contrast, is used to assess how comparable input images in both high- and low-resolution are,

$$C_L(I.I_0) = \frac{2\mu I\mu I_0 + C1}{\mu I^2 + \mu I_0^2 + C1} \quad (3)$$

$$C_c(I.I_0) = \frac{2\sigma I\sigma I_0 + C2}{\sigma I^2 + \sigma I_0^2 + C2} \quad (4)$$

Since C1 and C2 are constants, normalizing yields the image structure, as shown in Equation (5) and its correlations are used to evaluate the structural similarity measure.

$$(I, I_0) = \frac{2\sigma II_0 + C3}{\sigma I\sigma I_0 + C3} \quad (5)$$

Where

$$\sigma II_0 = \frac{1}{N-1}\sum_{i=1}^{N}(I_i - \mu I)(I_0 - \mu I_0)$$

3.1.3. RR-IQA

The reference image in RR-IQA is not entirely available. Instead, the reference image is used to extract several features. These characteristics are used by the quality evaluation approach as auxiliary data to measure the test image quality. RR-IQA techniques can be used for a variety of purposes.

They could be used to monitor how much the image's aesthetic quality and video data being transferred via real-time visual communication networks is degrading. The frequency at which the side information is encoded is a crucial design factor for RR-IQA systems. If a high data rate is available, more information about the reference image can be included, allowing for more precise quality predictions.

3.1.4. VMAF

A video quality metric built on machine learning is known as Video Multi-method Assessment Fusion (VMAF). Many have been drawn to it as a substitute metric to assess perceptual quality because it has been experimentally demonstrated to have a greater affinity with the visual system of the human system than traditional metrics like Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM) in a variety of scenarios. Netflix created VMAF with the explicit intent of having a significant correlation with MOS ratings.

A quality estimation model was trained using methods for machine learning using a large sample of MOS scores as the ground truth. It is a full-reference, perceptual video quality metric designed to mimic how people perceive video quality. The quality deterioration brought on by compression and rescaling is the main emphasis of this statistic. Through the computation of scores from several algorithms to determine quality and their fusion using a Support Vector Machine (SVM), VMAF calculates the perceived quality score.

3.1.5. VQM

The Video Quality Metrics (VQM) is an objective statistic that compares the amount of distortion people notice when assessing video quality. It helps assess video flaws such as jerkiness, block distortion, global noise, and colour distortion. The Video Quality Experts Group (VQEG) developed the VQM algorithm, tested it, and compared the results to subjective metrics. The results showed that the algorithms values correlated with subjective viewer ratings up to 0.9 (90%).

4. Results and Discussion

The choice of video quality metric depends on factors like the specific application, the nature of the content, and the available resources. It is common to use a combination of metrics, both objective and subjective, to get a comprehensive understanding of video quality. The evaluation of video quality is a field of active investigation, and new metrics and models continue to be developed to improve the accuracy of quality evaluation. The objective quality value in the traditional PSNR scale is shown in Figure 1 for three different compression settings (low, medium, and high) during a notable packet loss surge. Depending on the degree of compression, the viewer perceives a rigid frame with distinct characteristics throughout this large burst.

According to the PSNR data, quality begins to deteriorate significantly with the first frame affected by a burst. It continues to do so as the difference between the prior and current frames gets larger. There can be a further decline in quality around the middle of the burst.

It has to do with a scene transition where the processor reconstructs even big sequences by using the H.265 codec and setting the error resilience parameters to the numbers. H.265 is configured to generate one I frame for every 29 P frames without any B frames and seven in total slices of each structure, that we pack into respective packets and encase in RTP packets.

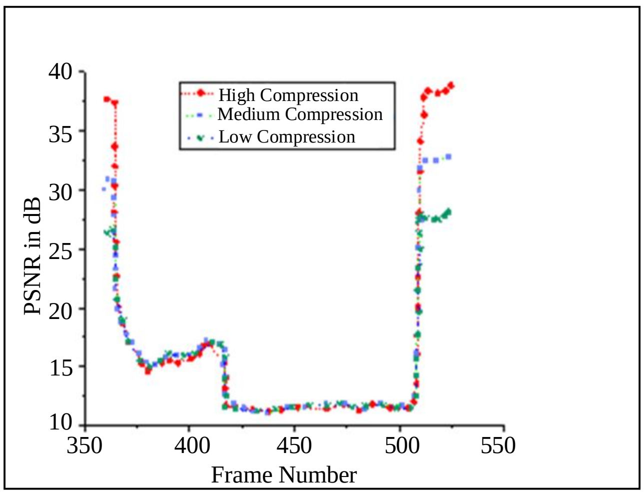

Fig. 1 PSNR for different compression levels

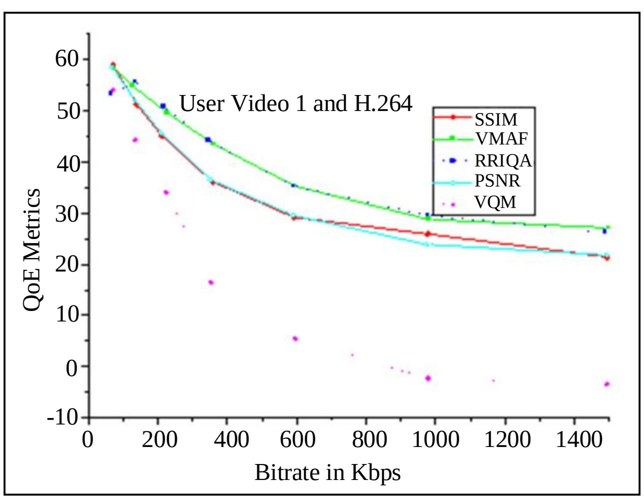

Fig. 2 QoE metrics vs. Bitrate (using H.264 codec)

Furthermore, as explained below, one-third of the macroblocks in every frame are encoded intramode-randomly. This procedure simulates packet losses in ad hoc scenarios by sending a corrupted bitstream to the decoder.

Consequently, the QoE parameters acquired for viewer video source #1 are shown evolving when the bit rate intensity is changed in Figure 2. The expected QoE grading pattern for a file or sequence at various bit rates is as follows. a) If frame rate levels are beneath the saturation threshold, it should offer a decreasing degree of accuracy as the bit rate decreases. b) When bit rate levels are above the saturation threshold, a person's perception of quality degree should be roughly the same. Video footage from users was transmitted using H.264. The measures have an exponential tendency to decline as the bit rate increases.

Figure 3 shows how the QoE metrics vary with H.265, whereas H.265 produces a more consistent result. We can observe that regardless of the encoder, all QoE at high bitrates (high quality) captures the perceptual saturation.

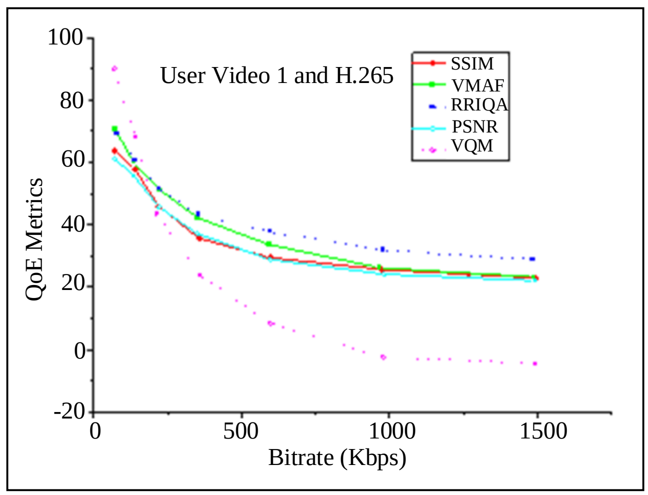

Fig. 3 QoE metrics vs. Bitrate (using H.265 codec)

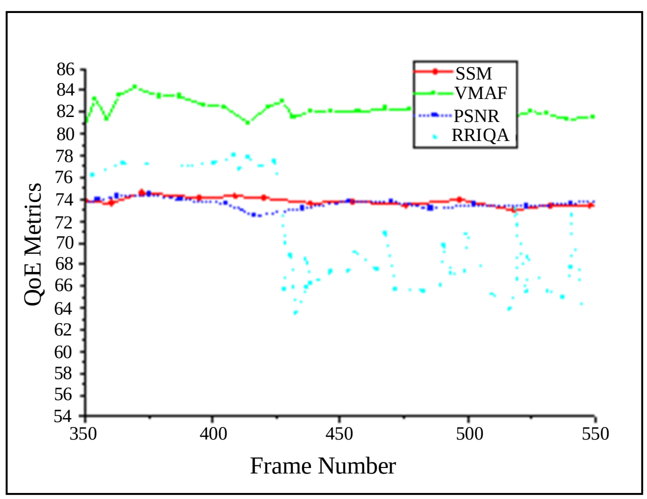

Fig. 4 QoE metrics

As said in the above two cases, identical holds for all encoders and sequences. The mapping function is anticipated to exhibit monotonicity. Likewise, the expected behaviour should be flat, meaning that as bitrates drop, the metrics should signal lower quality values. However, if we focus on the two lowest bitrates in Figure 3, we see that the quality score provided by the RRIQA and H.264 metrics rises as the bitrate value drops. That is different from how a QoE should act according to expectations.

The score of RRIQA for H.265 decreases with an increased bit rate. Using the H.265 codec, Figure 4 examines how the candidate metrics performed during a notable spike in QoE metrics. If one or more frames are lost, the quality must drastically decrease from a perceptual standpoint and stay there until the data flow is restored. Whether a scene shift occurs before or after the massive explosion should not matter.

Despite the compression level, both the VIF and MSSIM measures have virtually reached their “bad quality” threshold at the time of the burst, where the scene changes; therefore, the reported quality has not significantly changed. The quality dropping to the bare minimum at the start of the burst indicates that entire frames were lost.

5. Conclusion

The new model performs substantially better than previous models, reaching its peak performance STS-PSNR. A significant portion of the predictive ability of the STS concept is probably derived from the combined analysis of temporal and spatial information as it functions in a domain different from where standard principles are applicable. Furthermore, the video is broadcast over RTP for both H.264 and H.265 compression methods, and it evaluates packet loss across Full Reference QoE metrics such as PSNR, SSIM, VMAF, and RRIQA. To get the simulation results using the FFMPEG reference software, the authors employed the setting of the libx265 encoder parameters. HEVC reference software in various resolutions was used to replicate the measurements. The measurements were examined to show that compressed raw data provides a higher-quality streaming experience. When the video is compressed at low and high compression rates, the Full Reference measure displays nondeterministic behaviour with packet losses, making it challenging to detect and quantify this impact. The quality metrics should be used as a compromise between a high-quality measurement technique (like human visual perception) and communication complexity, notwithstanding minor variations in the packet drop strategy.

References

[1] Stefan Winkler, and Praveen Mohandas, “The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics,” *IEEE Transactions on Broadcasting*, vol. 54, no. 3, pp. 660-668, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[2] Zhengzhong Tu et al., “UGC-VQA: Benchmarking Blind Video Quality Assessment for User Generated Content,” *IEEE Transactions on Image Processing*, vol. 30, pp. 4449-4464, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[3] Xiaochen Liu et al., “Speeding Up Subjective Video Quality Assessment via Hybrid Active Learning,” *IEEE Transactions on Broadcasting*, vol. 69, no. 1, pp. 165-178, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] K.T. Tan, and M. Ghanbari, “A Multi-Metric Objective Picture-Quality Measurement Model for MPEG Video,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 7, pp. 1208-1213, 2000. [CrossRef] [Google Scholar] [Publisher Link]

[5] Miltiadis Alexios Papadopoulos et al., “A Multi-Metric Approach for Block-Level Video Quality Assessment,” *Signal Processing: Image Communication*, vol. 78, pp. 152-158, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[6] Fan Zhang, and David R. Bull, “A Perception-Based Hybrid Model for Video Quality Assessment,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 6, pp. 1017-1028, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[7] Gary J. Sullivan et al., “Overview of the High Efficiency Video Coding (HEVC) Standard,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[8] T. Wiegand et al., “Overview of the H.264/AVC Video Coding Standard,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560-576, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[9] Frank Bossen et al., “HEVC Complexity and Implementation Analysis,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1685-1696, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[10] Jens-Rainer Ohm et al., “Comparison of the Coding Efficiency of Video Coding Standards-Including High Efficiency Video Coding (HEVC),” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1669-1684, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[11] Jian Qiao, Yejun He, and Xuemin Sherman Shen, “Improving Video Streaming Quality in 5G Enabled Vehicular Networks,” *IEEE Wireless Communications*, vol. 25, no. 2, pp. 133-139, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[12] Xin Feng et al., “Saliency Inspired Full-Reference Quality Metrics for Packet-Loss-Impaired Video,” *IEEE Transactions on Broadcasting*, vol. 57, no. 1, pp. 81-88, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[13] Kalpana Seshadrinathan et al., “Study of Subjective and Objective Quality Assessment of Video,” *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1427-1441, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[14] Pradip Paudyal, Federica Battisti, and Marco Carli, “Impact of Video Content and Transmission Impairments on Quality of Experience,” *Multimedia Tools and Applications*, vol. 75, pp. 16461-16485, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[15] Yi J. Liang, John G. Apostolopoulos, and Bernd Girod, "Analysis of Packet Loss for Compressed Video: Effect of Burst Losses and Correlation between Error Frames," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 7, pp. 861-874, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[16] Songnan Li et al., "Image Quality Assessment by Separately Evaluating Detail Losses and Additive Impairments," *IEEE Transactions on Multimedia*, vol. 13, no. 5, pp. 935-949, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[17] Monalisa Ghosh, and Chetna Singhal, "MO-QoE: Video QoE Using Multi-Feature Fusion Based Optimized Learning Models," *Signal Processing: Image Communication*, vol. 107, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] H. Hamdoun et al., "Performance Benefits of Network Coding for HEVC Video Communications in Satellite Networks," *Iranian Journal of Electrical and Electronic Engineering*, vol. 17, no. 3, pp. 1-10, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[19] Glenn Van Wallendael et al., "Keyframe Insertion: Enabling Low-Latency Random Access and Packet Loss Repair," *Electronics*, vol. 10, no. 6, pp. 1-17, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[20] Mohammad Kazemi, Mohammad Ghanbari, and Shervin Shirmohammadi, "The Performance of Quality Metrics in Assessing Error-Concealed Video Quality," *IEEE Transactions on Image Processing*, vol. 29, pp. 5937-5952, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[21] Lixiong Liu et al., "Video Quality Assessment Using Space-Time Slice Mappings," *Signal Processing: Image Communication*, vol. 82, 2020. [CrossRef] [Google Scholar] [Publisher Link]