International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.7045
***Article Received**: 06 March 2023 **Revised**: 24 April 2023 **Accepted**: 10 May 2023*

Optimized Visual Internet of Things in Video Processing for Video Streaming

Bagade Shilpa[1], Budati Anil Kumar[2], L Koteswara Rao[3]
[1]Research Scholar , Department of ECE, KLEF,
Hyderabad, India
Shilpa.me1437@gmail.com
[2]Associate Professor, Department of ECE,
KG Reddy College of Engineering &Technology
Hyderabad, India
anilbudati@gmail.com
[3]Principal,KLEF,
Hyderabad, India
Kots.lkr@gmail.com

Abstract— The global expansion of the Visual Internet of Things (VIoT) has enabled various new applications during the last decade through the interconnection of a wide range of devices and sensors.Frame freezing and buffering are the major artefacts in broad area of multimedia networking applications occurring due to significant packet loss and network congestion. Numerous studies have been carried out in order to understand the impact of packet loss on QoE for a wide range of applications. This paper improves the video streaming quality by using the proposed framework Lossy Video Transmission (LVT) for simulating the effect of network congestion on the performance of encrypted static images sent over wireless sensor networks.The simulations are intended for analysing video quality and determining packet drop resilience during video conversations.The assessment of emerging trends in quality measurement, including picture preference, visual attention, and audio visual quality is checked. To appropriately quantify the video quality loss caused by the encoding system, various encoders compress video sequences at various data rates.Simulation results for different QoE metrics with respect to user developed videos have been demonstrated which outperforms the existing metrics.

Keywords-*:* Visual Internet of Things**,**Visual Sensor**,**Video streaming**,**video compression**,**5G networks**,**Packet loss

I. INTRODUCTION

Over the last two decades, the development of the Internet of Things (IoT) has demonstrated its tremendous conditions having the potential to be outfitted with "smart" things enriched with modern digitalization, allowing for flawless internet collaboration. This Internet of Things concept has progressed from modest and personal implementations to world wide scale integrated operations.RFID (Radio-frequency identification) devices were widely used to support IoT applications in the late 1990s [1].Video streaming has been so popular in recent years that it is now easily accessible on any and all OTT platforms such as Prime, and YouTube.Frame freezing and the resulting temporal jerkiness are common fragments in Internet video applications caused by both packet drops and latencies. There are two types of frame freezing that contribute to the improvement by the underlying media software programmes limitation, depending on the situation. In applications with security system designed to prevent requirements (Any frame that is not completely received by its display deadline (for example, communication or online coverage) is rendered unusable, and the recipient selects a specific error concealment strategy to recover the frame. Implementing its previously acquired proper frame is a common and popular error exposing approach. Even if the second frame is correctly received if it is predicatively coded using the preceding frame, it will have a decoding fault. Packet loss in wireless communications can be triggered by a wide range of events, including channel defects, crashes, and congested roads. Furthermore, these losses might increase in Wireless Visual Sensor Networks (WVSNs), when resource-constrained sensor networks attempt to transmit and relay data packets. Due to the delivery of numerous packets per obtained image, WVSNs with vision capabilities [2] provide distinct issues. While a scalar measurement (such as temperature or pressure) can be encoded in 2 or 3 bytes, the coding of an entire image requires hundreds of packets (depending on image size, codification type, and payload, among other considerations).Frequent packet filtering of data information may occur during data transfer, resulting in network congestion and video streaming delay.

The incompatibility of TCP fails to satisfy the current delay less video streaming .During transmission the data bits may be corrupted/packets may be

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.7045
***Article Received**: 06 March 2023 **Revised**: 24 April 2023 **Accepted**: 10 May 2023*

lost due to delay and congestion. The assessment of packet loss is primary. for quantifying the quality of different types of channel impairments. The metrics define the error at equal intervals of time but fail to notice asymmetrical distribution of perceptible importance.The previous metrics developed so far were concentrating only on videos with defects such as blurring/blockingness. However, recent study has concentrated on the influence of packet delay on video performance measures. The authors of proposed No Reference video quality criteria to measure discontinuities in [3]. Traditional objective video quality evaluation metrics, such as peak signal to noise ratio (PSNR) and mean square error (MSE), take into account inaccuracy at all pixels identically and overlook the inequality in the distribution of visual value. The disadvantage of this nondiscrimination is exacerbated by packet losses, because data packets can have significant visual effects in different video segments. For example, one packet loss in the centre of an image is evident, although another in the background or corners is possible; one loss in a dynamic scene is noticeable, whereas another in a static place is less noticeable.As a result, assessing perceived visual quality as it changes due to packet loss is extremely challenging.In [4] authors proposed computational Visual attention Model (VAM) to predict the packet loss impaired videos. The reference and encrypted videos are first assessed by the blocking sensor as well as the perceptive model.Each of them continue to generate a rating of the the video quality. After decryption one of the values is then picked as the finished product rating for the video. The selection is unaffected by the expected level of deblurring detected in the decoded sequence. When there is noticeable blockiness in the demodulated video, the score of the blockiness analyzer is chosen. Otherwise, the value of the perceptual model is applied. This switching strategy ensures that a more precise noise detection method is always accessible. They employed a motion detection system with a high correlation that was analysed using stepwise linear regression. In [5] authors developed multi metric model for MPEG video. Subjective and objective scores were achieved. A combination of sophisticated wavelet transform analysis, gait analysis, and adaptive nonlinear averaging was used to evaluate the HVS features of perceptual blocking and nonlinear perception.

To begin, visual aberration is produced by fusing intensity difference with a spatio-temporal mixed tolerances mapping. Chromatic aberration in films is calculated by comparing the rf energy of the reference content to that of the distorted content and weighing it by the degree of neighbourhood motion. At the frame level, an augmented geometric mean model is employed to integrate these two components. A unique interpretation fusion model is presented in [6] for assessing video quality by using non-linear model was developed.Low-latency video distribution is a strenuous domain because the requisite for a low buffer in all components of the distribution chain severely limits the technological solutions available. Furthermore, in such a distribution scenario, there exists a variety of end consumers with varying device and connectivity characteristics. In terms of devices and connections, it is paramount that low-performing devices and connections do not negatively impact the service of their high-performing predecessors. Even for the identical end user who may be physically travelling over time, a considerable diversity in quality of service is probably inevitable, especially when such distribution occurs throughout all wireless volatile networks. The difference between spatial and temporal masking effects was carried out here. Several decades of research have been undertaken in order to increase the performance of QoE without packet loss during streaming. Classifying from the past research, the streaming can be addressed in two ways: The first, with protocols used by TCP like HTTP and second based on UDP like RTP [7]. Another typical application for RTP/UDP video streaming is the evaluation of network methods, such as transport media access control coding or path scheduling that provides the maximum multipath quality of experience (QoE).

II. VIOT IN VIDEO COMPRESSION

The initial technological challenge at the smart VIoT sensing phase is the development of VIoT-specific video data compression suitable for this sort of data. Despite research on compression algorithms has

Figure 1. Illustration of VIoT technology

already progressed with a set of global standards in MPEG and HEVC, due to fundamental differences in VIoT video streams, they may not be easily adopted for VIoT video compression. In general, while compressing entertainment films, high definition reproduction of virtually all images is essential for comfortable watching of video contents. The ultimate goal of VIoT applications is to preserve relevant information rather

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.7045
***Article Received**: 06 March 2023 **Revised**: 24 April 2023 **Accepted**: 10 May 2023*

than the pixel value. As a result, keeping the spatial contents and settings of VIoT data is the most crucial quality for VIoT video compression and sensor data processing. Novel video compression algorithms that are fundamentally different from current MPEG and HEVC standards are required to best keep the interpretation and background of the collected VIoT sensor data throughout the process of compression. The development of a new video coding standard for machine communications is one recent example. The new VIoT systems are producing enormous amounts of information for a wide range of applications. Because of the unique features of visual sensor data ,VIoT systems can reveal insights that regular IoT systems cannot. These unique features enable VIoT systems to reach a wide range of new application industries to add new aspects to existing IoT applications. The obtained visual sensor data can be pooled, analysed, and interpreted using modern techniques such as data modeling, machine learning, and supervised learning. The resulting knowledge, which includes the recognition of behaviors and trends, reveals new perspectives that have the potential to affect each aspect of our lives, from better congestion control to crime prevention, and from primary prevention to environmental protection.

2.1. LVT Model

LVT, or Lossy Video Transmission Simulator, is a framework for analysing the impact of network congestion on segmentation of video frames (at WVSNs) (on the decoder side). The main goal of WSNs in system simulators is scalability (i.e., their ability to grow) capability to manage large groups of nodes. Because LVT focuses solely on image quality assessments, it can manage enormous sets of simulations with multiple images, procedures, and loss patterns. Simply put, LVT simulation consists of five major components or stages.

1) Forward Feature Extraction - The use of an image processing algorithm on an input image. This method yields a customized rendition of the original image.
2) Packetization - The use of a packetization system. This step connects processed image data to packets.
3) Packet loss Modeling - This stage simulates data loss during network transmission. The losses are created either randomly or by applying a loss pattern file as input.
4) Depacketization - The packetization scheme's inverse application.
5) Reverse Image Acquisition - The process of running the inverse image processing method. As a result, a version of the source image is created with some lost sequences.
6) Error Hiding - To fill the vacant gaps, an error concealment approach might be used.

A) LVT simulated model

1) proposed simulation model: The models used are fundamental.An incoming video frame I is a (L, B) matrix, I = {Ir,c},with r, c ϵ X $\wedge$ $0 \leq r < L$ $\wedge$ $0 \leq c < B$, each Ir,c pixel having its b bits in each pixel, b ϵ R+ Consider the communication system Γ́ where we transmit I in $\lfloor (L * B * b)/m \rfloor$ packets P, where m is the number of bits allocated for video data transfer in a packet. During communication, each packet pl has a chance of being lost.Various loss models can be employed to accomplish this. Because packet drops are expected to occur over a wireless channel with one or more intermediary nodes, the path characteristics should be irrelevant to the simulation (clearly, the number of nodes and customised communication protocols may alter the loss rate). Averaging the well-received neighbouring pixels yields an estimate of lost data for error concealing.

2) Encoding of Frames: As previously noted, the earliest form of LVT included frame coding error-resilient computation technique. In a typical single block-based transmission, the image is first divided into $\frac{L \, X \, B}{L_f \, X \, B_f}$ blocks Fi,j, where Lf and Bf are the length and breadth of the block frame, Fi,j = {Irf,cf}, where i. $L_f \leq r_f < L_f(i+1)$, j. $B_f \leq c_f < B_f(j+1)$.We allocate and deliver the ith block to the tth packet in a sequential transmission, with $t = \left\lceil \frac{i.b}{m} \right\rceil$. This sequence is disrupted by interleaving. It can be thought of as a bijective function: $\mathcal{V}$: I→ $\bar{\mathrm{I}}$, where $\bar{\mathrm{I}}$ is a new bitmap with all original blocks Fi,j put in a position $\bar{\mathrm{I}}\,\bar{\mathrm{J}}$). An improved model covers sequential processes on a low-resource network (requiring less memory and calculations).During the packetization process, semi pixel intensities are produced, but interleaving methods are utilised to select the data to put into the under construction packets.

B Video Quality Assessment

LVT's goal is to provide support for measuring the quality of produced image frames in WVSNs. Both subjective and objective assessment indicators for quality evaluation are done. Primary visualisation of the rebuilt frames provides subjective judgement.

$$\mathrm{PSNR(dB} = 10\ \log_{10}\left(\frac{(2^{n}-1)^{2}}{\mathrm{MSE}}\right) \qquad (2)$$

$$\mathrm{MAD\_measure} = \frac{\mathrm{MAD(input)}}{\mathrm{MAD(output)}} \qquad (1)$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.7045
Article Received: 06 March 2023 Revised: 24 April 2023 Accepted: 10 May 2023

1) Peak Signal to Noise Ratio

As shown below, the peak signal-to-noise ratio (PSNR) is employed to evaluate the restoration quality of suggested image restoration by SR

Where MSE denotes the mean squared error.

2) Mean absolute deviation (MAD)

Mean absolute deviation IQA model is formulated as a deviation of spatial regions from its average values as follows:

$$MAD = \sum_{i-1}^{n} \frac{|x_i - \bar{x}|}{n}$$

Where x_i=pixel values; $\bar{x}$=Mean value;
n = number of pixel values;

3) Structural similarity index (SSIM)

The structural similarity index (SSIM) is used to compare the resemblance of input Low Resolution and High Resolution images using orthogonal quantitative measures such asluminance (μ), contrast (σ) as follows:

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

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

Where C1 and C2 are constants, the picture structure is determined by normalising as illustrated in Equ. (4)

$$S = (I - \mu I)/\sigma I \quad (4)$$

And the measure the structural similarity is evaluated based on its correlations.

2.4 Encoding Standard

2.4.1 Video Coding Layer

The encoding standard used is H.265.It consists of video coding layer, used to encode the video and transmits it over the network. The video is broken up into three frames. I is the primary/reference frame that is unaffected by other frames. The P frame is interdependent on its previous but can be decoded independently. A B frame can be used to refer to another B frame.. All the frames in sequence can be defined as Group of Pictures (GoP).The concept of Macro blocks is used in H.265 . Typical size of each macro block can be different, 16 x 16 or 8 x 8 luma/chroma channels define the best size. Contemporary to macro block is Coding Tree Unit; both can be combined to form a slice. The slice can be decoded separately even if the same frame is not available[8-11].

2.4.2 Network Abstraction Layer

The generated encoded video from the VCL is converted into bits by NAL which further improves the efficiency of video transmission.eg: RTP can be further converted into mp3, mp4 for video storage. NAL can be further classified into two types-VCL NAL units and Non-VCL NAL units. The former carries the
data about video requirements and the latter carries additional overhead information. H.265 defines different types of NAL standards [12-13].

III. METHODOLOGY

To realize the packet loss and to improve video streaming, a simple architecture is developed with 3 modules 1) Transmitter 2) A packet failure model 3) A Receiver.The architecture consists of compressed video data using modified H.265 protocol [14]. The compressed video data is transmitted using 802.11p protocol. The streamed data utilizing the virtual network interface is sent to the receiver. The several forms of errors that occur after encoding H.265 data are examined in terms of packet loss.Finally the video is stored in the form of RTP packets as distorted video. The evaluation is done by following considerations and different encoding parameters [15-17]. As a preliminary example, we investigate the performance evaluation of new network coding techniques. One proposed evaluation methodology for this is to stream videos over an emulated network with packet delivery ratio even without network coding and then once with network coding then compare the QoE of both by calculating quantitative QoE metrics [18]. In this paper second method is used i.e.with packets are transferred to the network with network coding.

3.1. Transmitter

libx265 based H.265 model used as reference for the encoder. The primary parameters used while sending the video is listed in Table 1.Group of Pictures: In general any GoP structure for the evaluation of packet loss can be used but the structure used here is I-B-B-B-P-B-B-B-I with GoP size of 8.

3.2. Packet failure module

To determine the probability of packets sensed and transmitted, errors need to be taken into account which occurs during packet transmission. The four types of packet errors Propagation error, Sensing error, Busy Receiver error and Collision error were analyzed in [19-22] helped us to determine the improvement of throughput efficiency. This model used Hidden Markov Model to calculate interference and propagation losses.

3.3. Receiver

The major goal of this design is to ensure that data is sent without packet delays. In order to achieve high level of accuracy several constraints are analyzed. The encoding is

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.7045
***Article Received**: 06 March 2023 **Revised**: 24 April 2023 **Accepted**: 10 May 2023*

done once at a time to avoid the impact on QoE metrics, which reduces the necessity of encoding several times when multiple streams are needed. The second criterion is the loss sequence should match the parameter set e.g. If there is 10% increase in the PLR then same 10% increase must be seen with packet loss also. This requirement is satisfied by using random number generator of the packet loss model with number of seeds fixed. Next by keeping payloads fixed, RTP packets are transmitted by H.265, the loss sequence should not depend on cross traffic from any other application.

Table 1. List of Encoding Parameters

Parameter	**Value**
Video User Profile	Video Data of the user
Compressor, Transmission	Compressor, Transmission
Simulation Software	FFMPEG version 4.2.2
VMAF	Version 1.8.1
Quality	480⇥ 1080
2Frame Rate	Frame frequency of the compressed video
Video data rate	64M bit/s
Strategy for error coverage	None
libx265 Encoder Parameter	
Compressor Profile	Main
Compressor Level	4
Group of Pictures Structure	I-B-B-P-B-B-P-I
Packet Measurement Mode	binary NAL element
Scalable B frame Mode	Deactivated
JCTVC HEVC Reference Encoder Parameter	
Compressor Profile	Main
GOP Structure	I-B-B-B-P-B-B-B-I
Scalable B frame Mode	Active

IV. RESULTS

In this section results are discussed in both packet loss environment is discussed in first sub section and assessment of metrics like PSNR, SSIM and MAD is compared with existing H.264 & H.265 codecs done in separate section.

A. Analysis of video metrics in packet loss Environment

The purpose of this section is to investigate the behaviour of prospective metrics in the presence of packet delay in various IoT scenarios. To model packet losses in these error-prone conditions, we use a Markov model and the methods outlined in [18]. The Markov model is well known for its ability to describe bursty behaviour, ease of design, short execution times, and wide applicability. They are suitable for accelerating the QoE review process for video distribution applications on IoT scenarios while providing equivalent results to those obtained through simulation or real-world testbeds. The Markov model is used to build a packet loss model for VANET that faithfully reproduces video content. Figure 2&3 depicts the objective quality value in the classic PSNR scale for two distinct compression levels (low and high) amid a significant packet loss surge.During this big burst, the spectator sees a frozen frame with varying degrees of quality depending on the compression level. According to the PSNR statistic, quality reduces dramatically with the initial burst-affected frame and continues to drop as the imbalance between the previous frame and the current frame grows. An additional reduction in quality may be seen near the middle of the burst. It relates to a scenario transition that uses the H.265 codec with the error resilience settings set to the values indicated in [19], allowing the processor to rebuild even massive sequences. H.265 is configured to generate one I frame every 29 P frames, with no B frames, and to break each frame into seven slices, which we put into their own packets and encapsulate in RTP packets. In addition, as explained in [20-23], we force 1/3 of the macroblocks in each frame to be randomly encoded in intramode. This procedure replicates packet losses in adhoc circumstances, resulting in a distorted bitstream being sent to the decoder.

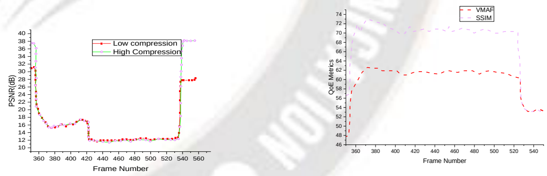

Figure 2&3. PSNR, SSIM, VMAF frame values packet loss burst (at different bitrates from frame 350 to 550).

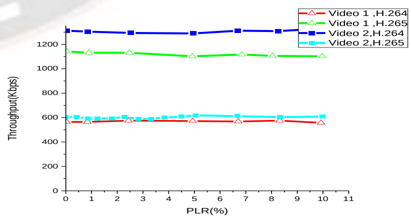

Figure 4. Throughput (kbps) vs PLR(%)

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.7045
***Article Received**: 06 March 2023 **Revised**: 24 April 2023 **Accepted**: 10 May 2023*

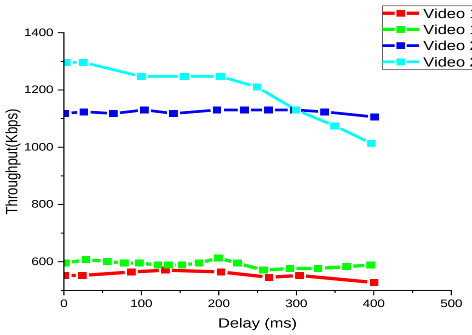

Figure 5. Throughput (kbps) vs Delay (ms)

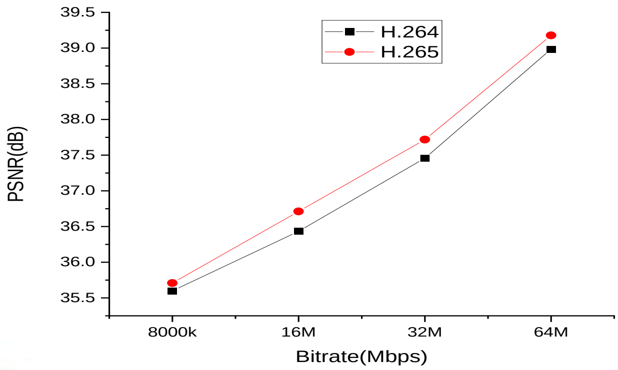

Figure 6. PSNR vs Bitrate

To begin, the emphasis is on determining how the introduced impairments, namely PLR and delay, impacted on a fundamental level.Throughput is one such metric. Thus, Fig. 4 depicts the evolution of throughput measured for both video-sources (598 Kbps for video #1; 1158 Kbps for video #2) as the intensity of the aforementioned impairments is varied. Both video clips were sent using the two CODECs .With increasing PLR and delay, the throughput value appears to decrease slightly. In the event of longer delays. Figure 5 shows how throughput with H.264 decreases as delay increases, whereas H.265 produces a more consistent result.This behaviour is due to H.265 LVT famework adaptive feature when network degradation is detected, the transmitter reduces the throughput rate to allow the receiver to recover. Nonetheless, when transmitting the heavier video, H.265 using LVT framework achieves a higher throughput.

B. Full Reference Metric Calculation

The video quality assessment is the common way to assess and verify the quality used by humans, with the availability of reference data, the metrics can be divided into Full reference, reduced Reference, No-reference. As per the research carried out by the authors Full reference method is used in this paper. Mean squared error (MSE) and Peak signal-to-noise ratio (PSNR) were widely been employed as authenticity metrics in the video processing discipline (PSNR is simply a nonlinear version of MSE). The prominence of these two measurements can be attributed to a variety of factors. The formulas for computing them are as simple to grasp and apply as they are to compute [24-25]. From a logical perspective, minimising MSE is also fairly understood. PSNR has become so known to video researchers over the years that they can easily analyse the values. PSNR is likely the most widely accepted statistic, owing to a lack of other criteria. The improvement of video streaming can be measured by the metrics like PSNR, SSIM, MAD for the codecs H.265 with LVT framework using the user defined videos.

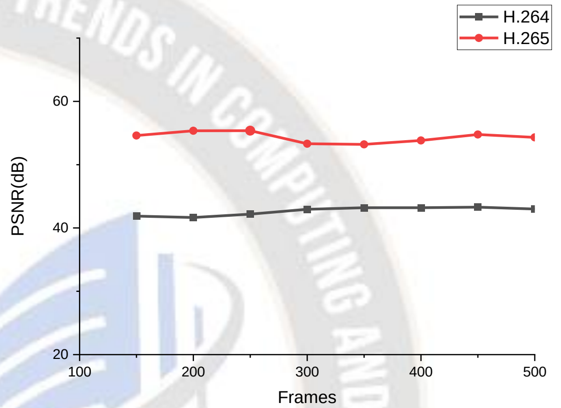

Figure 7. PSNR Vs No. of Frames

Fig. 6 ,7 shows the PSNR curves for the user developed videos for number of frames and bit rate .As can be seen from the figures that higher the resolution and bit rates the PSNR is high.

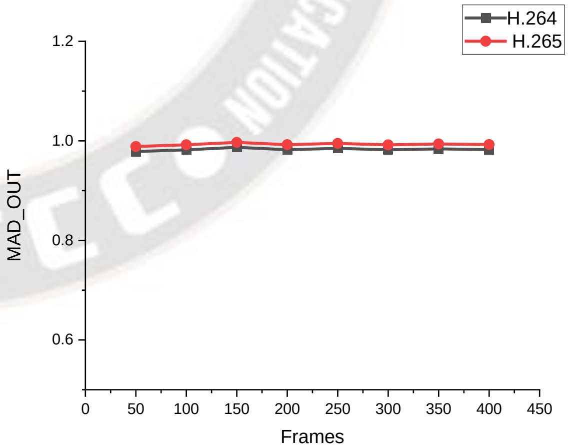

Figure 8. MAD Vs No. of Frames

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.7045
***Article Received**: 06 March 2023 **Revised**: 24 April 2023 **Accepted**: 10 May 2023*

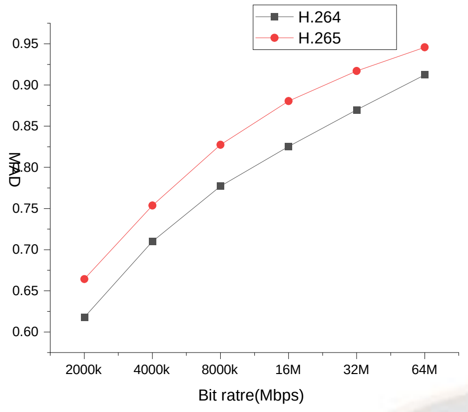

Figure 9. MAD Vs Bit rate

Fig 8 & 9 shows MAD from number of frames and different bit rates .The maximum value of MAD is 1.For higher bit rates and frame number, the deviation is also high.

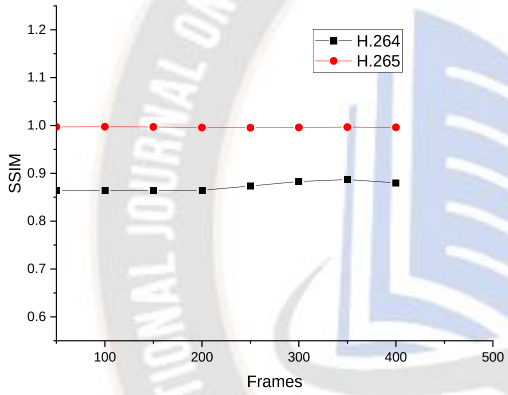

Figure 10. SSIM vs No.of Frames

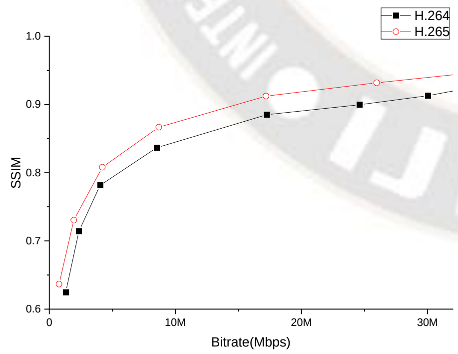

Figure 11. SSIM vs Bit rate

Fig 10 & 11 shows the metric SSIM wrt. Number of Frames and Bit rate .The range of SSIM is from 0 to 1.Higher the bit rate higher is the SSIM

V. CONCLUSION

The new Internet of Video Things refers to an emerging type of IoT system incorporating wireless visual sensors at the front end (VoIT) is introduced. H.265 coder with LVT programmable framework for simulating video communications in Wireless Visual Sensor Networks utilising error-tolerant methodologies such as block connecting has been experimented .In addition,the video for both compression techniques H.264 and H.265 is streamed over RTP measures both packet loss over the Full Reference QoE metrics like PSNR, SSIM and MAD. The authors used libx265 encoder parameters configuration to achieve the simulation results by using FFMPEG reference software. The metrics were simulated by using Lossy Video Simulation (LVT) with different resolutions (SD and HD).The novelty in this is that unlike the other transmission authors did not use uncompressed data and the video frames that were captured by wireless visual sensor transmitted by 802.11p protocol standards and successfully recovered through the decoder. The metrics were analyzed in such a way they owe a better experience of Quality during streaming compared to uncompressed raw data.In the presence of packet losses, the Full Reference measure exhibits nondeterministic behaviour, making it difficult to identify and quantify this effect when video is encoded at low and high compression rates. When it comes to the other measurements, SSIM, MAD and PSNR behave consistently.In conclusion, despite slight differences in the packet drop approach, we believe that the SSIM metric should be employed as a balance between a high quality measuring technique (equivalent to human visual perception) and communication complexity.

REFERENCES

[1] C. W. Chen, 'Internet of Video Things: Next-Generation IoT With Visual Sensors', IEEE Internet Things J., vol. 7, no. 8, pp. 6676–6685, Aug. 2020, doi: 10.1109/JIOT.2020.3005727.

[2] S. Soro and W. Heinzelman, 'A Survey of Visual Sensor Networks', Advances in Multimedia, vol. 2009, pp. 1–21, 2009, doi: 10.1155/2009/640386.

[3] X. Feng, T. Liu, D. Yang, and Y. Wang, 'Saliency Inspired Full-Reference Quality Metrics for Packet-Loss-Impaired Video', IEEE Trans. on Broadcast., vol. 57, no. 1, pp. 81–88, Mar. 2011, doi: 10.1109/TBC.2010.2092150.

[4] K. T. Tan and M. Ghanbari, 'A multi-metric objective picture-quality measurement model for MPEG video', IEEE Trans. Circuits Syst. Video Technol., vol. 10, no. 7, pp. 1208–1213, Oct. 2000, doi: 10.1109/76.875525.

[5] F. Zhang and D. R. Bull, 'A Perception-Based Hybrid Model for Video Quality Assessment', IEEE Trans. Circuits Syst. Video Technol., vol. 26, no. 6, pp. 1017–1028, Jun. 2016, doi: 10.1109/TCSVT.2015.2428551.

[6] J. You and J. Korhonen, 'Deep Neural Networks for No-Reference Video Quality Assessment', in 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, Sep. 2019, pp. 2349–2353. doi: 10.1109/ICIP.2019.8803395.

[7] S. Winkler and P. Mohandas, 'The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics', IEEE Trans. on Broadcast., vol. 54, no. 3, pp. 660–668, Sep. 2008, doi: 10.1109/TBC.2008.2000733.

[8] D. Laniewski, B. Schutz, and N. Aschenbruck, 'On the Impact of Burst Loss for QoE-Based Performance Evaluations for Video Streaming', in 2020 IEEE 45th LCN Symposium on Emerging Topics in Networking (LCN Symposium), Sydney, Australia, Nov. 2020, pp. 78–87. doi: 10.1109/LCNSymposium50271.2020.9363272.

[9] J. Qiao, Y. He, and X. S. Shen, 'Improving Video Streaming Quality in 5G Enabled Vehicular Networks', IEEE Wireless Commun., vol. 25, no. 2, pp. 133–139, Apr. 2018, doi: 10.1109/MWC.2018.1700173.

[10] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, 'Study of Subjective and Objective Quality Assessment of Video', IEEE Trans. on Image Process., vol. 19, no. 6, pp. 1427–1441, Jun. 2010, doi: 10.1109/TIP.2010.2042111.

[11] P. Paudyal, F. Battisti, and M. Carli, 'Impact of video content and transmission impairments on quality of experience', Multimed Tools Appl, vol. 75, no. 23, pp. 16461–16485, Dec. 2016, doi: 10.1007/s11042-015-3214-0.

[12] E. Pakulova, K. Miller, and A. Wolisz, 'Adaptive low-delay video streaming in heterogeneous wireless networks using MPRTP', in 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC), Valencia, Spain, Jun. 2017, pp. 14–19. doi: 10.1109/IWCMC.2017.7986255.

[13] B. Shilpa, A. K. Budati, L. K. Rao, and S. B. Goyal, 'Deep learning based optimised data transmission over 5G networks with Lagrangian encoder', Computers and Electrical Engineering, vol. 102, p. 108164, Sep. 2022, doi: 10.1016/j.compeleceng.2022.108164.

[14] B. Schutz and N. Aschenbruck, 'Packet-Preserving Network Coding Schemes for Padding Overhead Reduction', in 2019 IEEE 44th Conference on Local Computer Networks (LCN), Osnabrueck, Germany, Oct. 2019, pp. 447–454. doi: 10.1109/LCN44214.2019.8990879.

[15] Y.-K. Wang, Y. Sanchez, T. Schierl, S. Wenger, and M. M. Hannuksela, 'RTP Payload Format for High Efficiency Video Coding (HEVC)', RFC Editor, RFC7798, Mar. 2016. doi: 10.17487/RFC7798.

[16] Y. J. Liang, J. G. Apostolopoulos, and B. Girod, 'Analysis of Packet Loss for Compressed Video: Effect of Burst Losses and Correlation Between Error Frames', IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 7, pp. 861–874, Jul. 2008, doi: 10.1109/TCSVT.2008.923139.

[17] S. Winkler and P. Mohandas, 'The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics', IEEE Trans. on Broadcast., vol. 54, no. 3, pp. 660–668, Sep. 2008, doi: 10.1109/TBC.2008.2000733.

[18] M. Sepulcre, M. Gonzalez-Martin, J. Gozalvez, R. Molina-Masegosa, and B. Coll-Perales, 'Analytical Models of the Performance of IEEE 802.11p Vehicle to Vehicle Communications', IEEE Trans. Veh. Technol., vol. 71, no. 1, pp. 713–724, Jan. 2022, doi: 10.1109/TVT.2021.3124708.

[19] S. Li, F. Zhang, L. Ma, and K. N. Ngan, 'Image Quality Assessment by Separately Evaluating Detail Losses and Additive Impairments', IEEE Trans. Multimedia, vol. 13, no. 5, pp. 935–949, Oct. 2011, doi: 10.1109/TMM.2011.2152382.

[20] C. T. Calafate, P. Manzoni, and M. P. Malumbres, 'Speeding up the evaluation of multimedia streaming applications in MANETs using HMMs', in Proceedings of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems - MSWiM '04, Venice, Italy, 2004, p. 315. doi: 10.1145/1023663.1023719.

[21] C. T. Calafate, M. P. Malumbres, and P. Manzoni, 'Performance of H.264 compressed video streams over 802.11b based MANETs', in 24th International Conference on Distributed Computing Systems Workshops, 2004. Proceedings., Hachioji, Tokyo, Japan, 2004, pp. 776–781. doi: 10.1109/ICDCSW.2004.1284121.

[22] G. Van Wallendael, H. Mareen, J. Vounckx, and P. Lambert, 'Keyframe Insertion: Enabling Low-Latency Random Access and Packet Loss Repair', Electronics, vol. 10, no. 6, p. 748, Mar. 2021, doi: 10.3390/electronics10060748.

[23] H. Hamdoun, S. Nazir, J. A. Alzubi, P. Laskot, and O. A. Alzubi, 'Performance Benefits of Network Coding for HEVC Video Communications in Satellite Networks', IJEEE, vol. 17, no. 3, Sep. 2021, doi: 10.22068/IJEEE.17.3.1956.

[24] M. Ghosh and C. Singhal, 'MO-QoE: Video QoE using multi-feature fusion based Optimized Learning Models', Signal Processing: Image Communication, vol. 107, p. 116766, Sep. 2022, doi: 10.1016/j.image.2022.116766.

[25] N.-N. Dao, A.-T. Tran, N. H. Tu, T. T. Thanh, V. N. Q. Bao, and S. Cho, 'A Contemporary Survey on Live Video Streaming from a Computation-Driven Perspective', ACM Comput. Surv., vol. 54, no. 10s, pp. 1–38, Jan. 2022, doi: 10.1145/3519552