INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2025, VOL. 71, NO. 4, PP. 1-7
Manuscript received May 21, 2025; revised September 2025. …….. doi: 10.24425/ijet.2025.155481

Subjective video quality assessment using Single and Double Stimulus methods

Stefan Brachmański, Janusz Klink, and Michał Łuczyński

***Abstract*—The article presents the results of video quality assessment using two methods recommended by the International Telecommunication Union. These are the Single Stimulus (SS) and the Double Stimulus Impairment Scale (DSIS) methods. The results obtained by both methods were compared. The studies were performed for two coding techniques, H.264 and H.265, and for spatial resolutions of 1280 × 720 and 1920 × 1080. The studies showed a very high correlation between the MOS values obtained by SS and DSIS methods for H.265 coding and both video resolutions. Regarding the H.264 coding, a very good consistency of the results was obtained for bitrates starting from 3000 kbps.**

***Keywords*—video quality; Single-Stimulus method; Double-Stimulus Impairment Scale method**

I. Introduction

THE development of technology in the IT industry has led to an increase in the demand for the provision of increasingly advanced multimedia services, including video content. A report published in January 2023 by Sandvine stated that video content accounted for almost 66% of the total online volume in the first half of 2022 [1].

Video transmission is carried out using various coding techniques. The International Telecommunication Union (ITU) recommends using the H.264 [2] and the newer H.265 [3] standards. The quality perceived by the viewer is influenced by, among other things, the coding bitrate and video resolution.

The aims of the presented work are as follows:

- assessment of video quality (VQ) using the selected Single-Stimulus (SS) method and the Double-Stimulus Impairment Scale (DSIS) method recommended by the International Telecommunication Union,
- examination of the impact of parameters such as coding technique (H.264 and H.265), resolution (1280 × 720 and 1920 × 1080), and bitrate on the assessment of video quality by the young end user,
- comparison of video quality scores obtained using the Single-Stimulus (SS) method and the Double-Stimulus Impairment Scale (DSIS) method.

Among the various subjective methods of video quality assessment [4–14], Single-Stimulus (SS) and Double-Stimulus Impairment Scale (DSIS) methods were used in the studies. These are methods recommended by the International Telecommunication Union [6] and based on subjective quality criteria. The subjective quality assessment proposed in this method is based on the opinion of a representative group of people called observers. Measurements are made under conditions specified by requirements, controlled and repeatable, and people taking part in these measurements are properly trained.

II. State of Art

The subjective assessment of the quality of services, including video services, plays a very important role because the results of this assessment best reflect the level of satisfaction of the user, who is the final recipient of these services and the verifier of their quality.

The topic of service quality discussed in the literature is very broad, ranging from standardization documents describing research methodology [7] through the tools used to numerous application cases. In addition to standardization documents, there are many scientific works devoted to testing and using subjective quality assessment methods in various conditions and using various types of video materials, both conventional and 3D video [6,15]. A good overview of subjective quality assessment methods can be found in [16]. Some of the work concerns the compression of still images or video compression and assessing the impact of this compression and different video codecs on the quality perceived by users [17]. Compression helps reduce memory size and the cost of transferring images and videos. However, compression may cause visual artefacts, depending on the compression level. Therefore, evaluating the performance of compression algorithms and coding efficiency [18,19] is a fundamental task necessary to visually reconstruct video with the least possible quality loss. The performance of compression algorithms is assessed using both subjective and objective video quality assessment methodologies. The paper [20] presents software for subjective and objective examination of image quality with various degrees of compression.

Because subjective quality assessment is very time-consuming and expensive, appropriate, objective methods are being sought that will speed up the quality assessment process and make it a cheaper alternative. However, the application of various objective metrics in real-world scenarios is limited by the lack of clear interpretations of how the metric values reflect the subjective video quality perceived by the user. Therefore, there is a need to look for relationships between individual objective measures and the quality perceived by the user. Much work was devoted to the construction of the so-called quality models that map objectively measurable service parameters expressed in specific scales to quality expressed on a 5-point MOS (Mean Opinion Score) scale [21,22], where the individual

All authors are with Wroclaw University of Science and Technology, Poland (e-mail: stefan.brachmanski@pwr.edu.pl, janusz.klink@pwr.edu.pl, michal.luczynski@pwr.edu.pl).

values mean respectively: 1 - bad, 2 - poor, 3 - fair, 4 - good and 5 – excellent. An example of such works is [23], where the authors propose a mechanism that uses data from public Video Quality Assessment (VQA) databases and allows the automatic creation of general rules for mapping the values of objective VQA metrics to the subjective MOS scale. There are also many references in the literature to subjective video quality databases containing human-annotated videos and quality scores that can be used to create tools for creating, testing, and comparing VQA algorithms [24]. These databases can be divided into general-purpose databases and special-purpose databases. They contain video materials described by many parameters, such as database type, number and video content, total number of test video sequences, video resolutions, video frame rate, video formats, distortion types, etc. Examples of such general-purpose databases can be found in [25,26,27], while special-purpose databases are described e.g. in [28,29].

High Dynamic Range (HDR) technology has become increasingly common in recent years. HDR videos can represent a much wider range of brightness and colours than standard dynamic range (SDR), but on the other hand, they have more difficult capture, transmission and display requirements due to higher bit depth, wider colour gamut, etc. That is why there are works devoted to the specific challenges posed by HDR video, including the subjective assessment of its quality. Some publications present research on the impact of factors such as compression, aliasing, and ambient lighting on the quality of these videos as perceived by the user [30]. In parallel with the classic modelling of service quality, we observe the development of new methods for assessing and predicting the quality perceived by the user based on selected objective features of the video stream, using machine learning methods and artificial neural networks [31].

The list of topics and open problems regarding methods for assessing the quality of services, including video services, is very extensive and much broader than presented here. In the literature, we observe a large variety of described problems and quality assessment methods dedicated to many individual cases. The authors intended to check the substitutability of some methods with others. In particular, the work concerns the comparison of quality assessment results using double and single stimulus methods. Using the SS method should, in principle, shorten the time needed to assess video quality [7]. However, the question arises as to the possibility of replacing the DSIS method with the SS method and the possibility of obtaining comparable quality assessment results. The main aim of the work was to examine the correlation between the results of both methods and to check the conditions of their substitutability.

III. RESEARCH METHOD

A. Subjective measurement video quality using the Single Stimulus (SS) methods

The Single Stimulus (SS) [6] or Absolute Category Rating (ACR) [32] method involves observers who assess the quality of an image or sequence of images (video) on a five-point scale. In this method, test video sequences are presented without a reference signal. The test material may contain both test sequences and a corresponding reference sequence (standard). If the test material contains a reference sample, it is presented as a standalone stimulus for evaluation, just like any other test stimulus. The test sequences are separated by a grey screen. After each test sequence presented, during the presentation of a grey sequence lasting 5 to 10 seconds, the viewer assesses the quality of the video sequence viewed. The assessment is given in the Categorical Rating Scale (CRS), commonly used for evaluating video quality (Table I).

TABLE I
VIDEO QUALITY SCALE

Video quality	Rating
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

The average (final) rating is calculated for each video transmission condition tested as a result of averaging among observers. The rating is given as the MOS (Mean Opinion Score) index.

B. Subjective measurement video quality using the Double Stimulus Impairment Scale (DSIS) method

The Double Stimulus Impairment Scale (DSIS) [6] or the Degradation Category Rating (DCR) [32] method recommended by the International Telecommunication Union is an alternative to the SS method. In the DSIS method, the observer is presented with two video sequences. The first stimulus is the reference video, while the second is its distorted version. This method aims to compare the quality of the distorted video with that of the reference video. The observer reports the degree of deterioration of the second distorted video on a five-point Categorical Rating Scale (CRS) (Table II).

TABLE II
VIDEO QUALITY DETERIORATION SCALE

Video distortions	Rating
Imperceptible	5
Perceptible, but not annoying	4
Slightly annoying	3
Annoying	2
Very annoying	1

The average (final) score is calculated for each tested video transmission condition as a result of averaging across observers. The score is given as the MOS (Mean Opinion Score) indicator, however, to distinguish the results from those obtained using the SS method, the score is given as the DMOS (Degradation Mean Opinion Score).

C. Test material

The test reference material was a 20-second video sequence (without sound) with a resolution of 1920 × 1080 in avi format [8]. The video sequence contained dynamic scenes in the form of a fragment of the start of a horse race. This sequence was subjected to H.264 and H.265 encoding with different bitrates and different resolutions. In the DSIS measurements, three resolutions were considered: 640 × 360, 1280 × 720 and 1920 × 1080, while in the SS measurements, two resolutions 1280 × 720 and 1920 × 1080. For both the coding standards and each resolution, different transmission conditions were simulated with different bitrates:

- DSIS method: 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000 kbps;
- SS method: 100, 1000, 2000, 3000, 4000, 5000, 6000 kbps.

The reduced number of resolutions and bitrates in the SS measurements is the result of the analysis of the results obtained in earlier measurements performed using the DISS method [4].

Measurements were made in a laboratory adapted for the evaluation of video signals following the requirements of the recommendations of the International Telecommunication Union [6]. The test material was presented to the viewers on a 60-inch Full HD (1920×1080) TV screen. Lower-resolution videos have been resized to native resolution. Test videos representing different transmission conditions were presented to viewers in random order.

D. Observer Group

Two teams took part in the experiment, one in the SS method and the other in the DSIS method. In both types of measurement, the observer groups were students of the Wroclaw University of Science and Technology aged 20-21 years with normal visual acuity and correct colour discrimination. The International Telecommunication Union BT.500 recommendation specifies the number of observers, which should not be less than 15 people [6]. In the DSIS measurements, observers were divided into 2 groups. The size of each group was different; for the H.264 coding it was 45 people and for H.265 – 35 people. In the SS measurements, the observer group consisted of 80 people for both H.264 and H.265 coding techniques.

Before starting the measurements, the study participants were familiar with the evaluation method and had a training session. During the training session, viewers were familiarised with the method of presenting the test material and assessing the deterioration of the quality of the video signal. In the measurement session, after watching the original and encoded sequences, each participant in the study recorded his evaluation of the deterioration of quality in a special form. All evaluations were entered into a spreadsheet and statistically analysed according to the procedure described in the ITU-R BT.500 recommendation [6]. This allowed for the elimination of evaluations that exceeded the adopted 95% confidence interval.

IV. RESULTS AND DISCUSSION

A. Video quality assessment using the SS method

Estimates given by each of the 80 observers participating in the experiment were statistically analysed according to the procedure described in Recommendation ITU-R BT.500 [6]. This analysis allowed us to reject estimates that fell outside the assumed 95% confidence interval. The average results for the H.264 coding technique are presented in Table III, and for H.265 in Table IV. The tables provide the mean MOS value, as well as the standard deviation (S) and confidence interval coefficient (δ) values calculated according to ITU Recommendation BT.500 [6].

TABLE III
THE MEAN VALUE OF THE VQ ASSESSMENT (MOS) FOR H.264 CODEC

Bitrate [kbps]	1280×720			1920×1080		
	MOS	S	δ	MOS	S	δ
100	1.38	0.51	0.11	1.04	0.19	0.04
500	2.08	0.61	0.13	1.59	0.54	0.12
1000	3.03	0.80	0.17	2.51	0.53	0.12
2000	3.68	0.65	0.14	3.61	0.63	0.14
3000	4.16	0.60	0.13	4.18	0.61	0.13
4000	4.39	0.58	0.13	4.49	0.55	0.12
5000	4.66	0.50	0.11	4.68	0.47	0.10
6000	4.79	0.41	0.09	4.84	0.37	0.08

TABLE IV
THE MEAN VALUE OF THE VQ ASSESSMENT (MOS) FOR H.265 CODEC

Bitrate [kbps]	1280×720			1920×1080		
	MOS	S	δ	MOS	S	δ
100	1.50	0.50	0.11	1.31	0.47	0.10
500	2.26	0.61	0.13	2.25	0.65	0.14
1000	3.10	0.38	0.08	3.23	0.53	0.12
2000	3.84	0.37	0.08	3.94	0.63	0.14
3000	4.26	0.63	0.14	4.35	0.58	0.13
4000	4.41	0.52	0.11	4.56	0.50	0.11
5000	4.58	0.50	0.11	4.78	0.42	0.09
6000	4.70	0.46	0.10	4.85	0.36	0.08

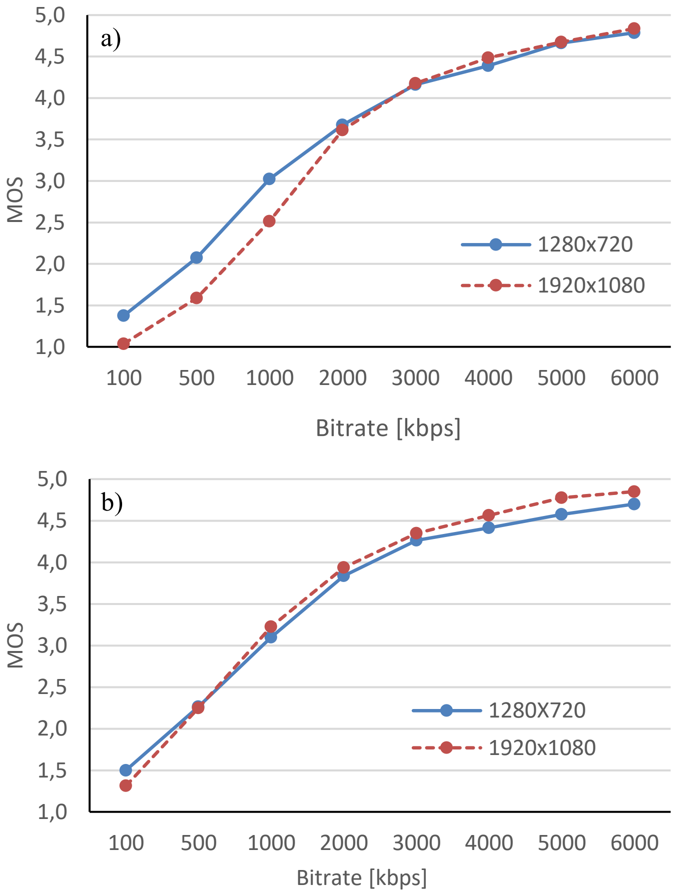

Fig. 1. Results of subjective quality assessment (MOS) for the H.264 (a) and H.265 (b) encoded video as a function of the coding bitrate for 2 resolutions

The MOS values obtained for the H.264 coding technique presented in Tables III and IV are shown in Figure 1a and for

H.265 in Figure 1b. Comparing the curves in Figures 1a and 1b, we can see that the video quality is directly proportional to the encoding rate used. Moreover, using a higher video resolution allows for higher quality. However, it is visible that the positive effect of using higher video resolution is only visible at higher encoding speeds. At lower encoding rates, higher-resolution video may appear to be of lower quality. In the case of the H.264 codec, this effect was observed at codec rates below 2 kbps, while in the case of the H.265 codec, this effect occurred at codec rates below 1 kbps. For the lowest encoding speed, in the case of FHD resolution, the H.265 codec achieved MOS values approximately 25% higher than the HD codec (i.e. MOS = 1.31 and MOS = 1.04, respectively).

B. Video quality assessment using the DSIS method

The presented results were obtained as part of the first stage of research on the influence of selected coding techniques, resolution, and bitrate on video signal quality [4]. The averaged results of the video quality assessment obtained by the DSIS method in the first stage of the study are presented in Table V (H.264 coding) and Table VI (H.265 coding). The tables provide the mean MOS value, as well as the values of the standard deviation (S) and confidence interval coefficient (δ) values calculated according to ITU Recomm.BT.500 [6]. The MOS values obtained for the H.264 coding technique presented in Tables V and VI are shown in Figure 2a and for H.265 in Figure 2b.

TABLE V
THE MEAN VALUE OF THE VQ ASSESSMENT (DMOS) FOR H.264 CODEC

Bitrate [kbps]	1280x720			1920x1080		
	DMOS	S	δ	DMOS	S	δ
300	1.16	0.43	0.13	1.05	0.32	0.10
400	1.33	0.57	0.17	1.19	0.45	0.13
500	1.81	0.55	0.16	1.35	0.57	0.17
600	2.24	0.66	0.20	1.58	0.59	0.18
700	2.71	0.56	0.17	1.98	0.64	0.19
800	2.86	0.64	0.19	2.40	0.54	0.16
900	3.12	0.54	0.16	2.81	0.55	0.16
1000	3.26	0.54	0.16	3.16	0.57	0.17
1500	3.74	0.66	0.20	3.86	0.47	0.14
2000	4.07	0.7	0.21	4.26	0.49	0.15
2500	4.19	0.59	0.18	4.40	0.54	0.16
3000	4.28	0.55	0.16	4.49	0.51	0.15
3500	4.33	0.47	0.14	4.54	0.55	0.17
4000	4.40	0.49	0.15	4.63	0.49	0.15
4500	4.49	0.51	0.15	4.70	0.46	0.14
5000	4.56	0.50	0.15	4.79	0.41	0.12
5500	4.65	0.48	0.14	4.84	0.37	0.11
6000	4.72	0.45	0.14	4.91	0.29	0.09

The analysis of the curves presented in Figure 2, showing the influence of coding speed on the quality assessment perceived by users, determined by the double stimulus method, indicates MOS values very close to the results obtained for the single stimulus method (see Figure 1). At the same time, it can also be observed that the video quality increases with the increase in encoding speed, and the positive effect of using higher video resolution is observed at higher encoding speeds.

TABLE VI
THE MEAN VALUE OF THE VQ ASSESSMENT (DMOS) FOR H.265 CODEC

Bitrate [kbps]	1280×720			1920×1080		
	DMOS	S	δ	DMOS	S	δ
300	1.09	0.29	0.12	1.13	0.34	0.14
400	1.61	0.50	0.20	1.65	0.49	0.20
500	2.09	0.42	0.17	2.04	0.64	0.26
600	2.48	0.59	0.24	2.48	0.51	0.21
700	2.87	0.34	0.14	2.83	0.58	0.24
800	3.09	0.42	0.17	3.13	0.76	0.31
900	3.26	0.45	0.18	3.35	0.49	0.20
1000	3.43	0.51	0.21	3.48	0.59	0.24
1500	3.78	0.42	0.17	3.91	0.51	0.21
2000	4.09	0.67	0.27	4.22	0.52	0.21
2500	4.30	0.56	0.23	4.43	0.51	0.21
3000	4.43	0.59	0.24	4.57	0.51	0.21
3500	4.48	0.59	0.24	4.70	0.47	0.19
4000	4.57	0.51	0.21	4.78	0.42	0.17
4500	4.61	0.50	0.20	4.83	0.39	0.16
5000	4.70	0.47	0.19	4.87	0.34	0.14
5500	4.74	0.45	0.18	4.91	0.29	0.12
6000	4.78	0.42	0.17	4.91	0.29	0.12

In this situation, the question arises about the impact of the quality assessment method used on the results obtained. For this purpose, the next step will be to compare the results of studies conducted using single and double stimulus methods.

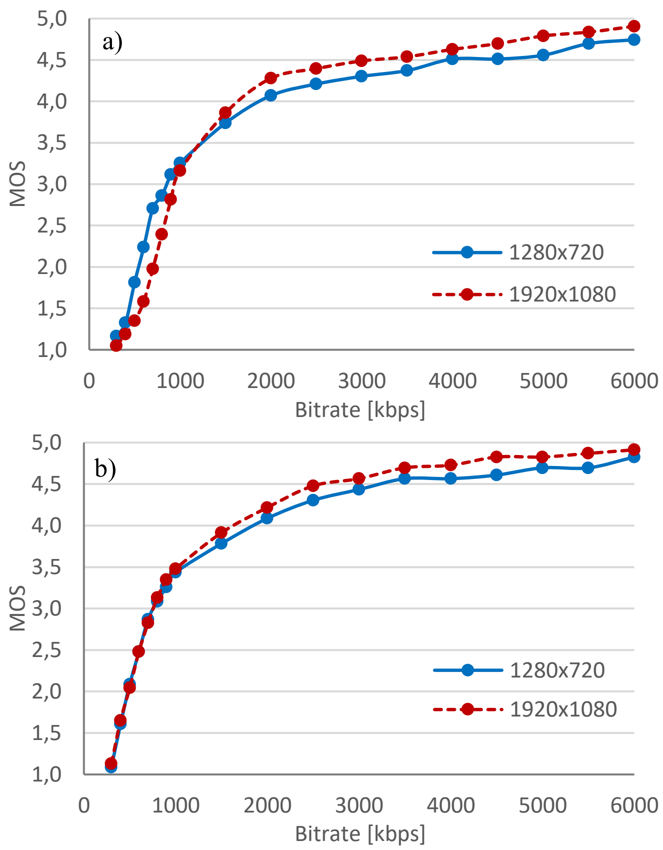

Fig. 2. Results of subjective quality assessment (DMOS) for the H.264 (a) and H.265 (b) encoded video as a function of coding bitrate for 2 resolutions

C. Video quality assessment using the SS vs DSIS method

The average results of video quality assessment using the SS and DSIS methods for the H.264 encoding technique are presented in Table VI and graphically in Figure 3.

TABLE VII
MOS (SS) VS DMOS (DSIS) FOR H.264

Bitrate [kbps]	1280×720		1920×1080	
	MOS	DMOS	MOS	DMOS
100	1.38	-	1.04	-
500	2.08	1.81	1.59	1.35
1000	3.03	3.26	2.51	3.16
2000	3.68	4.07	3.61	4.28
3000	4.16	4.28	4.18	4.49
4000	4.39	4.4	4.49	4.63
5000	4.66	4.56	4.68	4.79
6000	4.79	4.72	4.84	4.91

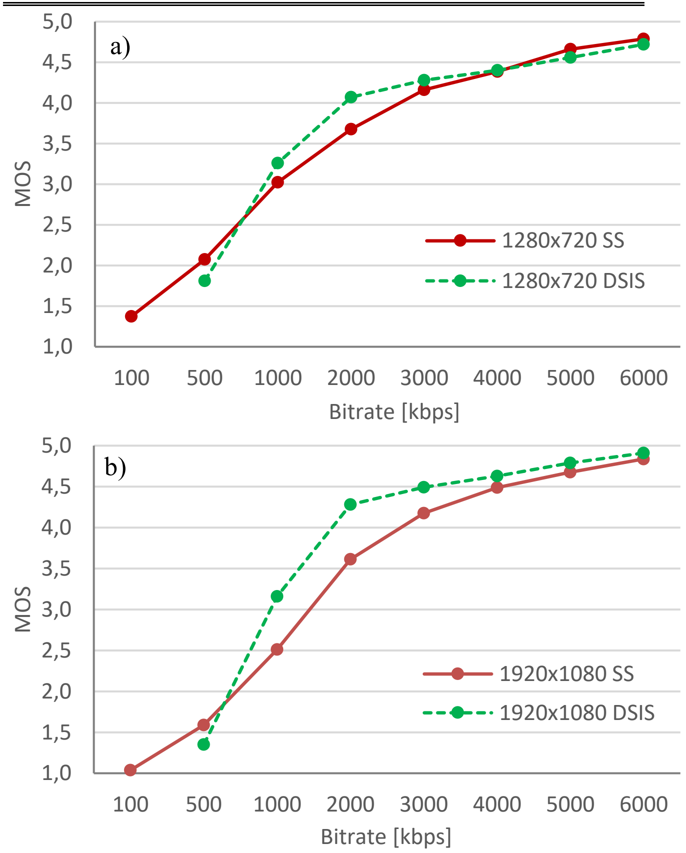

Fig. 3. Results of subjective quality assessment (DMOS and MOS) for the H.264 encoded video as a function of coding bitrate for resolutions 1280×720 (a) and 1920×1080.

The average results of video quality assessment using the SS and DSIS methods for the H.265 encoding technique are presented in Table VIII and graphically in Figure 4. A preliminary evaluation of the results presented in Table VII and Table VIII, and Figure 3, and Figure 4 indicates slight differences in the estimates obtained using the SS and DSIS methods. This observation applies to both tested resolutions and codecs used. To check whether the observed small differences in the results obtained were statistically significant, we used the nonparametric Wilcoxon-Mann-Whitney test. The test result, with the assumed significance level of 5%, showed that the hypothesis about the equivalence of both methods, i.e. SS and DSS, should not be rejected. The results presented here correspond to the conclusions of the works of other authors [12,33], where no statistically significant differences were observed between the results of subjective tests using single and double stimulus methods.

TABLE VIII
MOS (SS) VS DMOS (DSIS) FOR H.265

Bitrate [kbps]	1280×720		1920×1080	
	MOS	DMOS	MOS	DMOS
100	1.50	1.09	1.31	1.13
500	2.26	2.09	2.25	2.04
1000	3.10	3.43	3.23	3.48
2000	3.84	4.09	3.94	4.22
3000	4.26	4.43	4.35	4.57
4000	4.41	4.57	4.56	4.78
5000	4.58	4.7	4.78	4.87
6000	4.70	4.78	4.85	4.91

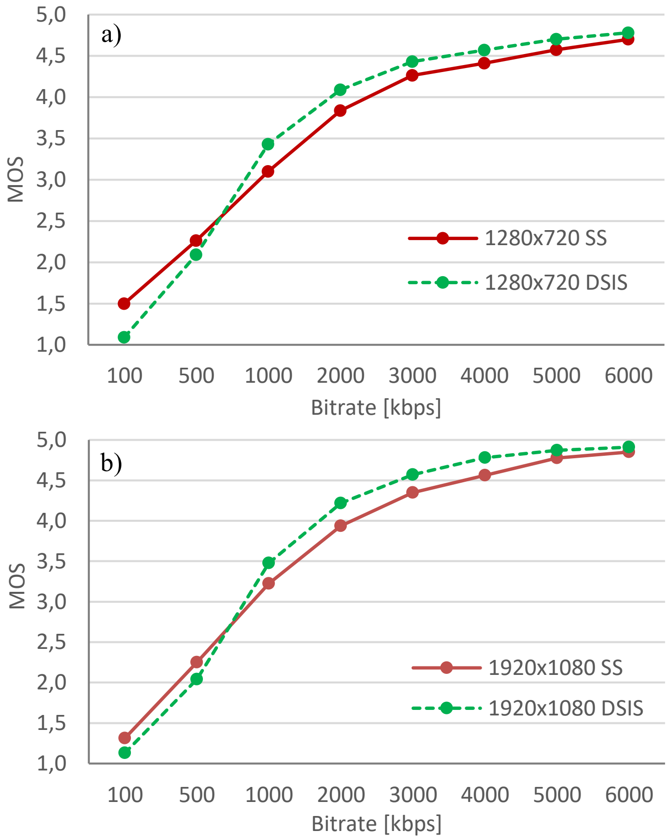

Fig. 4. Results of subjective quality assessment (DMOS and MOS) for the H.265-encoded video as a function of the coding bitrate for resolutions 1280×720 (a) and 1920×1080.

CONCLUSION

This article presents the results of research on video quality assessment using two subjective methods, i.e. single and double stimulus. Video samples encoded using various codecs, spatial resolutions, and encoding speeds were analyzed.-The obtained results indicate the expected superiority of the H.265 codec over the H.264 codec [4, 37], which in the tested range of coding speeds is reflected in better average image quality for a given resolution (see Table VII and Table VIII). These observations are consistent with the results presented in other publications [34] and with the results of our previous studies [4,9]. At the same time, these results show that the use of higher video resolution allows for higher quality perceived by the user only from a certain coding speed threshold upwards. For low encoding rates (i.e. high compression), higher video resolution may result in lower quality perceived by users. The result of the comparison of the tested quality assessment methods deserves special attention. The conducted video quality tests show a high convergence of the results obtained with both the single and double stimulus methods, which was demonstrated in their statistical analysis. This happens regardless of the codec used and the encoding bitrate. Understanding the basic differences and applications of both types of methods, i.e. single and double stimulus, it seems that the single stimulus method can be used as an equivalent of the double stimulus method where there is no access to a reference video or where it is important to shorten the time of performing subjective tests.–Although the single-stimulus method requires less time, by eliminating the need to compare the evaluated video with a reference sample, it produces results that are statistically indistinguishable from those obtained with the double-stimulus method. The compatibility of the DSIS and SS methods presented here fully corresponds to the research results presented in the literature [35,36], where a very high correlation of research results carried out using the above-mentioned methods was also obtained (from R=0.97 to R=0.99 depending on the resolution of the video encoded in the H.264 format). The authors of the above-mentioned work also drew attention to the advantage of the SS method, which was characterized by a shorter research assessment time than the DSIS method.

REFERENCES

[1] Sandvine. "Video Permeates, Streaming Dominates," *Phenomena. Global Internet Phenomena Report 2023,* Online: https://www.sandvine.com/hubfs/Sandvine_Redesign_2019/Downloads/2023/reports/Sandvine GIPR 2023.pdf (accessed on 13 August 2024).

[2] ITU-T Rec. H.264, "Audiovisual and multimedia systems: Infrastructure of audiovisual services-Coding of moving video, Advanced video coding for generic audiovisual services," 2021.

[3] ITU-T Rec. H.265, "High efficiency video coding," 2023.

[4] S. Brachmański, J. Klink, "Subjective Assessment of the Quality of Video Sequences by the Young Viewers," In *30th Int. Conf. on Software, Telecomm. and Computer Networks (SoftCOM 2022)*, Split, pp. 1–6, 2022, https://doi.org/10.23919/SoftCOM55329.2022.9911375

[5] Y. Chen, K. Wu, Q. Zhang. "From QoS to QoE: A tutorial on video quality assessment," *IEEE Communications Surveys & Tutorials*, vol. 17, no 2, pp. 1126-1165, 2014. https://doi.org/10.1109/COMST.2014.2363139

[6] ITU-R BT 500-14, "Methodologies for the subjective assessment of the quality of television images," 2020.

[7] ITU-T Rec. P.910, "Subjective video quality assessment methods for multimedia applications," 2021.

[8] ITU-R BT 1788, "Methodology for the subjective assessment of video quality in multimedia applications," 2007.

[9] J. Klink, S. Brachmański, M. Łuczyński, "Assessment of the Quality of Video Sequences Performed by Viewers at Home and in the Laboratory," Appl. Sci., vol. 13, no. 8, 2023, https://doi.org/10.3390/app13085025

[10] F. M. Moss et al., "On the optimal presentation duration for subjective video quality assessment," *IEEE Trans. on Circuits and Systems for Video Tech.*, *26*(11), 1977-1987, 2015, https://doi.org/10.1109/TCSVT.2015.2461971

[11] S. Péchard, R. Pépion, R., P. Le Callet, "Suitable methodology in subject-tive video quality assessment: a resolution dependent paradigm," In *Int. Workshop on Image Media Quality and its Appl., IMQA2008* , p. 6, 2008, https://hal.science/hal-00300182/ (accessed on 13 August 2024).

[12] M. Pinson, S. Wolf., "*Comparing subjective video quality testing methodologies"*, Visual Communications and Image Processing 2003. SPIE,. pp. 573-582, 2003, https://doi.org/10.1117/12.509908

[13] F. P. Pongsapan, T. Hendrawan, "Evaluation of HEVC vs H.264/A VC video compression transm. on LTE network," *11th Int. Conf. on Tel. Syst. Serv. and Appl. (TSSA)*, pp 1-4, 2017, https://doi.org/10.1109/TSSA.2017.8272949

[14] S. Winkler, "Video quality measurement standards – current status and trends", *7th Int. Conf. on Information, Comm. and Signal Processing (ICICS). IEEE*, pp. 1-5, 2009. https://doi.org/10.1109/ICICS.2009.5397585

[15] J. Gutierrez, P. Perez, F. Jaureguizar, J. Cabrera, "Validation of a novel approach to subjective quality evaluation of conventional and 3D broadcasted video services," In *4th Int. Workshop QoMEX*, 2012, pp. 230–235, https://doi.org/10.1109/QoMEX.2012.6263870

[16] X. Min, H. Duan, W. Sun, Y. Zhu, G. Zhai, "Perceptual Video Quality Assessment: A Survey," 2024, [Online]. http://arxiv.org/abs/2402.03413 (accessed on 15 October 2024)

[17] A. V. Katsenou, F. Zhang, M. Afonso, D. R. Bull, "A Subjective Comparison of AV1 and HEVC for Adaptive Video Streaming," in *Proc. Int. Conf. on Image Proc. (ICIP)*, 2019. https://doi.org/10.1109/ICIP.2019.8803523.

[18] P. Hanhart, T. Ebrahimi, "Calculation of average coding efficiency based on subjective quality scores," *J. of Visual Com. and Image Repr.*, vol. 25, no. 3, pp. 555–564, Apr. 2014, https://doi.org/10.1016/j.jvcir.2013.11.008.

[19] P. Akyazi, T. Ebrahimi, "Comparison of Compression Efficiency between HEVC/H.265, VP9 and AV1 based on Subjective Quality Assessments," in *2018 Tenth Int. Conf. on Quality of Multimedia Experience (QoMEX), IEEE*, May 2018, pp. 1–6. https://doi.org/10.1109/QoMEX.2018.8463294.

[20] A. Afnan et al., "Subjective Assessment of Objective Image Quality Metrics Range Guaranteeing Visually Lossless Compression," *Sensors*, vol. 23, no. 3, p. 1297, 2023, https://doi.org/10.3390/s23031297

[21] N. Staelens et al., "Measuring Video Quality in the Network : From Quality of Service To User Experience," In *9th Int. Workshop on Video Processing and Consumer Electronics (VPQM 2015)*, pp. 1–6, 2015.

[22] Y. Chen, K. Wu, and Q. Zhang, "From QoS to QoE: A Tutorial on Video Quality Assessment," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 2, pp. 1126–1165, 2015, https://doi.org/10.1109/COMST.2014.2363139.

[23] A. N. Moldovan et al., "VQAMap: A Novel Mechanism for Mapping Objective Video Quality Metrics to Subjective MOS Scale," *IEEE Trans. Broadcast*., vol. 62, no. 3, 2016, https://doi.org/10.1109/TBC.2016.2570002

[24] A. Saha et al., "Perceptual video quality assessment: the journey continues!," *Frontiers in Signal Processing*, vol. 3, June 2023, pp. 1–22, https://doi.org/10.3389/frsip.2023.1193523

[25] M. Nuutinen et al. "CVD2014 - A Database for Evaluating No-Reference Video Quality Assessment Algorithms," *IEEE Trans. Image Process*., vol. 25, no. 7, pp. 3073–3086, 2016, https://doi.org/10.1109/TIP.2016.2562513

[26] C. Keimel, A. Redl, K. Diepold, "The TUM High Definition Video Datasets," In *2012 Fourth Int. Work. on Quality of Multim. Experience*, Australia, 2012, pp. 97–102, https://doi.org/10.1109/QoMEX.2012.6263865

[27] H. Wang et al., "MCL-JCV: A JND-Based H.264 / AVC Video Quality Assessment Dataset," In *2016 IEEE Int. Conf. on Image Proc. (ICIP)*, USA, September 2016, pp. 1509-1513. https://doi.org/10.1109/ICIP.2016.7532610

[28] C. Chen et al. "Modeling the time-Varying subjective quality of HTTP video streams with rate adaptations," *IEEE Trans. on Image Processing*, vol. 23, no. 5, pp. 2206–2221, 2014, https://doi.org/10.1109/TIP.2014.2312613

[29] P. C. Madhusudana et al., "Subjective and Objective Quality Assessment of High Frame Rate Videos," *IEEE Access*, vol. 9, 2021, pp. 108069–108082, https://doi.org/10.1109/ACCESS.2021.3100462

[30] Z. Shang, J. P. Ebenezer, A. C. Bovik, Y. Wu, H. Wei, S. Sethuraman, "Subjective Assessment of High Dynamic Range Videos Under Different Ambient Conditions," In *2022 IEEE Int. Conf. on Image Proc. (ICIP)*, France, 2022, pp. 786–790, https://doi.org/10.1109/ICIP46576.2022.9897940

[31] M. Giannopoulos et al., "Convolutional Neural Networks for Video Quality Assessment," 2018, http://arxiv.org/abs/1809.10117 (accessed on 17 October 2024)

[32] ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality," Geneva, Switzerland, 1996

[33] L. Polak et al.,, "A comparison of single and double stimulus subjective assessment of full HD video sequences,"*24th Int. Conf. Radioelektronika*, Slovakia, 2014, pp. 1-4, https://doi.org/10.1109/Radioelek.2014.6828436

[34] L. Guo, J. De Cock, and A. Aaron, "Compression Performance Comparison of x264, x265, libvpx and aomenc for On-Demand Adaptive

Streaming Applications," in *2018 Picture Coding Symposium*, PCS 2018 - Proceedings, 2018, https://doi.org/10.1109/PCS.2018.8456302

[35] T. Tominaga et al., "Performance comparisons of subjective quality assessment methods for mobile video," Sec. Int. Work. on Quality of Multimedia Experience (QoMEX),, Norway, 2010, pp. 82-87, https://doi.org/10.1109/QOMEX.2010.5517948

[36] T. Tominaga et al., "Performance compar. of subjective quality assessment methods for video." *IEICE Trans. on Com. v*ol. 97. No 1, 2014, pp. 66-75.

[37] N. Barman, M. G. Martini, "H.264/MPEG-AVC, H.265/MPEG-HEVC and VP9 codec comparison for live gaming video streaming" in 2017 9th Int. Conf. on Quality of Multimedia Experience, QoMEX 2017, 2017. https://doi.org/10.1109/QoMEX.2017.7965686