

Overview of the JPEG XS core coding system subjective evaluations

Alexandre Willème^a, Saeed Mahmoudpour^{b,c}, Irene Viola^d, Karel Fliegel^e, Jakub Pospíšil^e,
Touradj Ebrahimi^d, Peter Schelkens^{b,c}, Antonin Descampe^f, and Benoit Macq^a

^aICTEAM, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

^bimec, Leuven, Belgium

^cVrije Universiteit Brussel, Brussels, Belgium

^dMultimedia Signal Processing Group (MMSPG), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

^eCzech Technical University in Prague, Prague, Czech Republic

^fintoPIX s.a., Belgium

ABSTRACT

The JPEG committee (Joint Photographic Experts Group, formally known as ISO/IEC SC29 WG1) is currently in the process of standardizing JPEG XS, a new interoperable solution for low-latency, lightweight and visually lossless compression of image and video. This codec is intended to be used in applications where content would usually be transmitted or stored in uncompressed form such as in live production, display links, virtual and augmented reality, self driving vehicles or frame buffers. It achieves bandwidth and power reduction for transparent and low latency coding for compression ratios ranging from 2:1 to 6:1. The subjective assessment of the impact of visually lossless compression poses particular challenges. This paper describes the subjective quality evaluation conducted on the JPEG XS core coding system. In particular, it details the test procedures and compares the results obtained by the different evaluation laboratories involved in the standardization effort.

Keywords: JPEG XS, lossless compression, near-lossless compression, image coding, video coding, subjective quality assessment

1. INTRODUCTION

Uncompressed video is transmitted and stored in many applications such as professional transport networks (SMPTE2022 5/6, IEEE 802.1AVB, etc.), display links (SDI, HDMI[®], and DisplayPort[™]), and frame buffers in screens. Recent improvements in video representation related to the emergence of Ultra-High Definition (UHD), High-Dynamic Range imaging (HDR), High-Frame Rate (HFR), and omnidirectional imaging have dramatically increased the bandwidth and memory space required to transmit or to store content. This growth is so significant that most currently deployed and emerging video processing infrastructures prove unable to provide the extra bandwidth required. To cope with this issue, low-latency, lightweight, and visually lossless compression is an attractive solution to reduce the size of video streams while preserving the advantages of uncompressed video. First, it extends the life of existing infrastructures, easing and reducing the cost of adoption of higher quality video formats. Second, it can be used to compress already supported content in order to reduce power consumption of video systems or multiplex several streams on the same link. Furthermore, in most uses cases of such a compression technology, devices from a variety of manufacturers are expected to communicate with each other. Therefore, interoperability is an important feature that must be taken into account during the definition of such schemes.

To this end, several standardization activities have been, or are currently, conducted.¹ For instance, the Society of Motion Picture & Television Engineers (SMPTE) released the SMPTE 2042 VC-2 standard, describing the BBC Research Dirac Pro technology.² Furthermore, the Video Electronics Standards Association (VESA) defined the Display Stream Compression (DSC) standard and, more recently, the VESA Display Compression-M (VDC-M) standard that both specifically target compression over display links.^{3,4} The latter is able to provide visually lossless compression at a 5:1 compression ratio. Moreover, the Joint Photographic Experts Group

(joint working group of ISO/IEC JTC1/SC29/WG1 and ITU-T SG16, better known as the JPEG committee) is currently working on the JPEG XS specification.^{5,6} This codec is designed so that it is a candidate technology wherever uncompressed video is used today including in video links and IP networks, real-time video storage, frame buffer,⁷ self driving car, drones and omnidirectional video capture and rendering in virtual and augmented reality. It only requires a low complexity implementation on CPU, GPU, and FPGA and provides transparent compression at ratios ranging from 2:1 to 6:1. The end-to-end latency it introduces is <32 lines. Additional information on the requirements of this new standard is available in the JPEG XS Call For Proposals (CfP) issued in early 2016.⁸

The JPEG XS core coding system has been defined based on different compression solutions proposed in response to the above CfP. The technology has been selected based on the outcome of a thorough study of the proposals by the JPEG experts and on the results of subjective and objective quality assessment conducted by independent evaluation laboratories.^{9,10} The resulting compression scheme is depicted in Figure 1. It consists of a Discrete Wavelet Transform (DWT), a quantizer and a predictive entropy coder. To ensure low-latency, the DWT is asymmetrical as it applies up to five horizontal decompositions but no more than two vertical decompositions. The entropy coder encodes the amplitude levels of groups of four neighboring wavelet coefficients belonging to the same sub-band; these levels are called bitplane counts. Then it removes the insignificant bits of the coefficients of each group before packing. The quantizer trims the least significant bits of the coefficients of a group according to a rate control strategy that ensures that the target bit rate is met. Coded bitplane counts, quantized coefficients and sign values are then packed together to form the final codestream. More details on the JPEG XS coding algorithm are given by Descampe et al.⁶

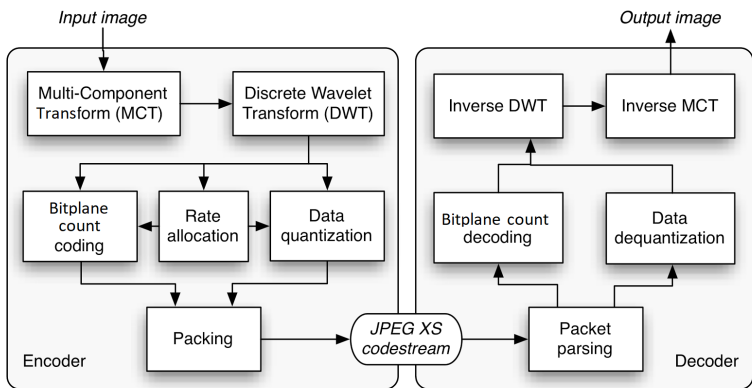


Figure 1: Block diagram of the JPEG XS codec.⁶

JPEG XS is a multi-standard specification. The first part, registered as ISO/IEC 21122-1,¹¹ normatively defines the core coding system able to generate JPEG XS-compliant codestreams. JPEG XS Part 2, registered as ISO/IEC 21122-2,¹² defines profiles, the buffer model, and identifies interoperability points that help manufacturers in the design of solutions that answer the specific needs of different use cases. JPEG XS Part 3, on transport and container, describes ways to transfer JPEG XS codestreams using well-known transport protocols. Finally, JPEG XS Part 4 (conformance testing) and Part 5 (reference software) provide guidelines to ensure compliance and a software implementation of the coding algorithm.

The aim of this paper is to describe the subjective quality assessment conducted on the JPEG XS core coding system after the 77th JPEG meeting in Macau, in October 2017. The goal of this assessment phase was to make sure that the JPEG XS core coding system meets the visually lossless requirement. Moreover, it is also intended to compare different configurations for different application categories with different decoder complexity. The remaining of this paper is organized as follows. Section 2 describes the methodologies adopted in other image and video codec standardization initiatives and compares them with the procedure defined for JPEG XS. Section 3 details the test methodology, the test content, the JPEG XS configurations under test, and the evaluation setup. It also provides an analysis of the results of the assessment. Finally Section 4 draws a conclusion.

2. RELATED WORK

To overcome the limitations objective metrics face to properly characterize the response to a visual stimulus, subjective quality assessment methods are defined in ITU Recommendations to reliably evaluate the visual quality of multimedia contents. In particular, the ITU-T Recommendation P.910¹³ and P.913¹⁴ define the methodology that should be followed to assess video, audio or audiovisual quality in 'multimedia services', and of 'Internet video and distribution quality television', respectively, whereas the ITU-T Recommendation BT.500-13¹⁵ defines the methodologies for subjective quality assessment of television pictures.

In past standardization efforts, different methodologies have been adopted to evaluate the performance of the upcoming standards in terms of visual quality. The JPEG XR image compression standard¹⁶ was evaluated using the Double Stimulus Continuous Quality Scale (DSCQS) method, aptly modified for the evaluation of still images. Specifically, a side-by-side comparison with implicit reference was implemented. To reduce the complexity of the test, subjects were asked to identify on which side the compressed content was displayed, and to rate its quality on a scale from 0 to 100, with 5 distinct quality levels.

For the JPEG XT HDR compression standard,¹⁷ the Double Stimulus Impairment Scale (DSIS) Variant I methodology was selected to assess the level of impairment caused by compression artifacts. As a side-by-side presentation was selected, the test images were cropped and displayed on the same HDR monitor. The subjects were informed on which side the reference would be displayed, and were asked to rate the level of impairment of the test image on a discrete scale from 1 (Very annoying) to 5 (Imperceptible).

Similarly, for the evaluation campaign of the ISO/MPEG and ITU-T/VCEG HEVC/H.265 video compression standard,¹⁸ the DSIS method, Variant II with a continuous rating scale was selected. The reference and test video were sequentially presented twice, with a two-second interval between each video sequence. The subjects were asked to rate the level of impairment of the test sequence with respect to the reference, on a scale from 0 to 100, with 5 distinct impairment levels.

More recently, the evaluation for the JPEG Pleno Call for Proposals on Light Field Coding has adopted a Double Stimulus Comparison Scale (DSCS) methodology.¹⁹ The light field test contents were shown as video sequences alongside the reference, in a side-by-side fashion. Subjects were informed on which side the reference was displayed, and were asked to rate the quality of the test content on a scale from -3 (much worse) to +3 (much better).

Despite the rich literature available on methodologies for subjective quality assessment of multimedia contents in a variety of scenarios, the aforementioned methods are seldom suitable for the evaluation of visually lossless compression algorithms and standards. Due to the nature of the compression artifacts under test, displaying the stimuli in a temporal sequence or in a side-by-side fashion, as proposed in the DSIS Variant I and II, respectively, is not practical, as it would create unnecessary strain and fatigue for the subjects without guaranteeing statistically relevant results. To facilitate the assessment of visually lossless codecs in a reliable and reproducible way, the JPEG committee has issued an international standard for evaluation procedures of nearly lossless coding.²⁰ In particular, the ISO/IEC 29170-2 (AIC Part-2) Draft Amendment 2 specifies two alternative test methodologies (Annex A and B). In the first, two test images and one reference image are presented, and the subject must choose which of the test images is closer to the reference. A temporal window of at least 4 seconds is given for the subject to perform their selection. In the second, only one test image is presented along with the reference, in a side-by-side fashion. The test image is interleaved with the uncompressed reference at a given temporal frequency. With this method, any visual differences between the original and the compressed content would create visual discrepancies in the form of "flickering". The reference is displayed on either left or right side of the screen, and subjects are not made aware of where the reference is located in the screen. They are obliged to choose the image they consider to be the test image on the side where such a flickering is perceived.

The methodology defined in ISO/IEC 29170-2 (AIC Part-2) has been successfully used for several subjective evaluation campaigns for visually lossless codecs, including the JPEG XS Call for Proposals¹⁰ and the VESA DSC standard.²¹

3. SUBJECTIVE QUALITY ASSESSMENT OF JPEG XS CORE CODING SYSTEM

This section describes the subjective quality assessment conducted in order to evaluate the efficiency of the JPEG XS core coding system. This process is mainly intended at validating that JPEG XS compression is visually transparent. The evaluation was conducted in parallel by three test laboratories affiliated to, respectively, Czech Technical University in Prague (CTU), Ecole Polytechnique Fédérale de Lausanne (EPFL), and Vrije Universiteit Brussel-imec (VUB-imec). This section details the test methodology, the test content selection, and the JPEG XS configurations that have been assessed. It also describes the evaluation setup defined for these experiments. Finally, the results provided by the test laboratories are discussed and compared.

3.1 Test Methodology

The test methodology adopted in this subjective assessment follows the recommendations defined in ISO/IEC 29170-2 (AIC Part-2) Draft Amendment 2 that provides guidelines to perform the subjective evaluation of near lossless image coding schemes.²⁰ In particular, the assessment of the JPEG XS core coding system broadly conforms to the test methodology described in Annex B of AIC Part-2 and introduced above. According to this test, the screen is split in two parts. On one side a crop of the original still image is displayed and on the other side, the same crop of the encoded-decoded image is displayed, interleaved with the crop from the original image. Both the encoded-decoded image and its original counterpart appear four times per second on the interleaved side. As a consequence, visual discrepancies between the original image and the encoded-decoded image appear as "flickering". The subjects are required to guess on which side of the screen the interleaved content is displayed. In the JPEG XS subjective assessment, the forced binary choice defined in AIC Part-2 is replaced by a ternary voting system where the subjects have the option to cast a *no decision* vote. As a consequence, each vote could fall into three categories: (i) *correct*, when interleaved side is correctly identified, (ii) *wrong*, when interleaved side is wrongly identified, (iii) *no decision*, when the subject could not decide which side is interleaved. Offering a *no decision* vote is intended to reduce subjects stress and fatigue and was deemed beneficial for the reliability of the test.¹⁰

During a session, each stimulus (one image encoded by one JPEG XS configuration at one given bit rate) was shown 8 times at randomized points (instead of 30 times as defined in AIC Part-2) during a maximum of 5 seconds. In order to improve the reliability of the assessment, both *training stimuli* and *control stimuli* were inserted in the test material. Training stimuli use encoded-decoded images that present obvious compression artifacts. They were shown to the subject at the beginning of the session, with an instructor beside, to help identify visual discrepancies. Control stimuli also present visible compression artifacts and were inserted at random points during the test session. Subjects who were unable to make *correct* decisions on these control stimuli were regarded as unreliable and their results were discarded during the statistical analysis.

The encoded-decoded images were undergone seven successive encoding-decoding cycles. This allows one to assess the *multi-generation robustness* of JPEG XS. The bit rates used for encoding images varied from one image to another and also depended on the JPEG XS configuration (as detailed in Table 1).

3.2 Test content selection

The content selected for this assessment consisted of eight still images. As mentioned above, no video sequences were used in the tests. The images are depicted in Figure 2. They cover a wide range of formats, various content types (i.e. camera-captured, CGI, and screen-content) and present several coding challenges (e.g. complex textures or sharp edges). Two images displayed on the left-end of the first row of Figure 2 (namely *Tools* and *HintergrundMusik*) were only used to generate the training stimuli. One must note that these images have been cropped so as to help the viewer to focus on some specific areas of the pictures.

3.3 JPEG XS configurations under test

Four different configurations of JPEG XS were assessed during this subjective assessment campaign. These configurations are based on four of the profiles defined in the Draft International Standard of JPEG XS Part 2, registered as ISO/IEC 21122-2,¹² as of July 2018*. These profiles are intended to offer various trade-offs between

*Some discrepancies exist between the actual profiles defined in ISO/IEC 21122-2 and the configurations described here.

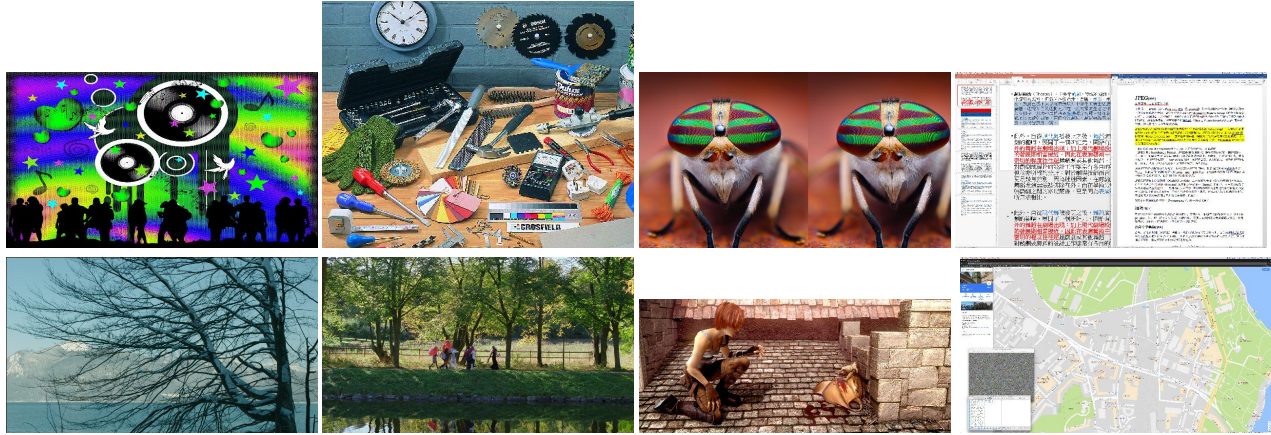


Figure 2: Still images used for the subjective quality assessment.

decoder complexity, buffering needs, and coding efficiency. Therefore, they provide a range of solutions that answer the needs of specific applications while meeting the requirements defined in the JPEG XS Cfp.⁸ Even though a detailed description of the profiles is out of the scope of this paper, it is worth summarizing each for the purpose of completeness:

- **Main** configuration (related to Main profile) provides an intermediate trade-off between complexity, buffering needs, and coding efficiency. It enables five horizontal DWT decompositions using a 5-3 wavelet structure but only one single vertical DWT decomposition. It allows for three different prediction modes for the bitplane counts. It targets use cases such as broadcast production, Pro-AV, frame buffer compression (FBC) in digital TV (DTV), and compression over display links.
- **High** configuration (related to High profile) aims at providing the best coding efficiency (and visual quality preservation) at the cost of increased complexity and more buffering. Compared to the Main configuration, it enables two vertical DWT decompositions and uses additional coding tools within its entropy coder. Targeted applications include high-end Pro-AV, broadcast contribution, cinema remote production, FBC and compression over display links in high-end DTV.
- **Low Logic** configuration (related to Light profile) models applications that can only afford the lowest complexity implementations at the cost of a slight coding efficiency drop. Compared to the Main configuration, only two prediction modes for the bitplane counts are available and the quantizer is simpler. It is intended to be used in use cases such as broadcast production, industrial cameras, and in-camera compression (for compression ratios up to 4:1).
- **Low Memory** configuration (related to Light-Subline profile) is intended at reducing significantly the buffering needs at the cost of a coding efficiency drop. To do so, it does not perform any vertical DWT decomposition. It also splits the frames into two columns that are encoded independently. Targeted use cases include in-camera compression (for compression ratios up to 2:1) and many cost-sensitive applications.

3.4 Evaluation setup

The subjective evaluation setup used by the test laboratories followed the relevant recommendations described in Annex B of ISO/IEC 29170-2 (AIC Part-2).²⁰ Selected important setup characteristics and requirements¹⁰ specific to JPEG XS Core Experiments are summarized below, including notes on condition differences between different test laboratories.

- Monitor with 4K UHD resolution (3840×2160 pixels) and 10-bit color depth processing (Eizo CG318-4k[†] recommended and used by the test laboratories).

[†]Eizo CG318-4k. <https://www.eizoglobal.com/products/coloredge/cg318-4k/index.html>. Accessed: 2018-07-30

- Monitor settings were applied in order to minimize additional stimuli processing, e.g., no picture expansion (dot by dot native resolution), full signal range (10 bpc input signal processing allowed). The color representation mode was set to ITU-R BT.709-6.²²
- Observers were seated at the viewing distance of 60 ppd (pixels per degree) which approximately corresponds to 58.6 cm for Eizo CG318-4k (native resolution 3840×2160 pixels, viewable image size 698.0×368.1 mm). **Note** – At the EPFL and VUB-imec laboratories, the observers were seated approximately at 50 cm viewing distance. At the CTU laboratory, the observers were seated at 59 cm and the viewing distance was verified before each session.
- Head/chin-rest (recommended in ISO/IEC 29170-2) was not used, only one single observer was viewing the monitor at a time. **Note** – At the EPFL and VUB-imec laboratories the observers were allowed to get closer to the screen, whereas, at the CTU laboratory, the observers were instructed to keep the viewing distance. However, as the head/chin-rest was not used, the viewing distance could not be kept exact to the specification during the entire test session.
- Viewing room with controlled lighting similar or compliant with ITU-R BT.500-13.¹⁵ Background luminance was low in the range between 15 and 20 cd/m². **Note** – The background luminance was approximately 15, 18, and 20 cd/m² for the CTU, EPFL, and VUB-imec laboratories respectively.
- Target peak screen luminance was set to 120 cd/m² with allowed luminance range of 100-140 cd/m². **Note** – The target peak luminance was set to 120 cd/m² for hardware calibration of Eizo CG318-4k monitor at all the three laboratories. The luminance at the CTU laboratory measured using Photo Research SpectraScan PR740[‡] spectroradiometer was approximately 116.5 cd/m², which is well within the tolerances and similar results for the EPFL and VUB-imec laboratories can be expected.
- Minimum specification for the video server was Intel[®] i7 with PCIe or M2 SSD and GPU with 10 bpc support. The resolution was set to 3840×2160, refresh rate 24 Hz and 10 bpc output activated. **Note** – The CTU laboratory used the refresh rate of 24 Hz, whereas 30 Hz was used at EPFL, and 60 Hz at VUB. Detailed technical inspection of the CTU laboratory setup, after finished subjective experiments, revealed slight banding in 10 bpc smooth gradient test pattern, which indicates questionable 10 bpc transparency of this particular configuration.
- Open source, and cross-platform media player MPV[§] with the proper setting was capable of correctly displaying stimuli with 10 bpc color content.
- All the test laboratories used the same testing software package including scripts for evaluation sessions and tools for automatic consolidation of obtained results.

3.5 Test validation and results analysis

A parsing tool including a set of Python scripts and Excel worksheets was provided to extract required data from subjective test reports and to generate analytic graphs.

3.5.1 Voting method and validation metric

For a given test image, three types of decisions are defined: correct, wrong and no decision. Based on the ternary voting procedure,¹⁰ an empirical metric is defined to obtain a score for each test image:

$$\text{Score} = 2 \times \left(1 - \frac{OK + 0.5 \times ND}{N} \right) \quad (1)$$

[‡]Spectrascan PR-740. <http://www.photoresearch.com/>. Accessed: 2018-07-30

[§]MPV. <https://mpv.io>. Accessed: 2018-07-17

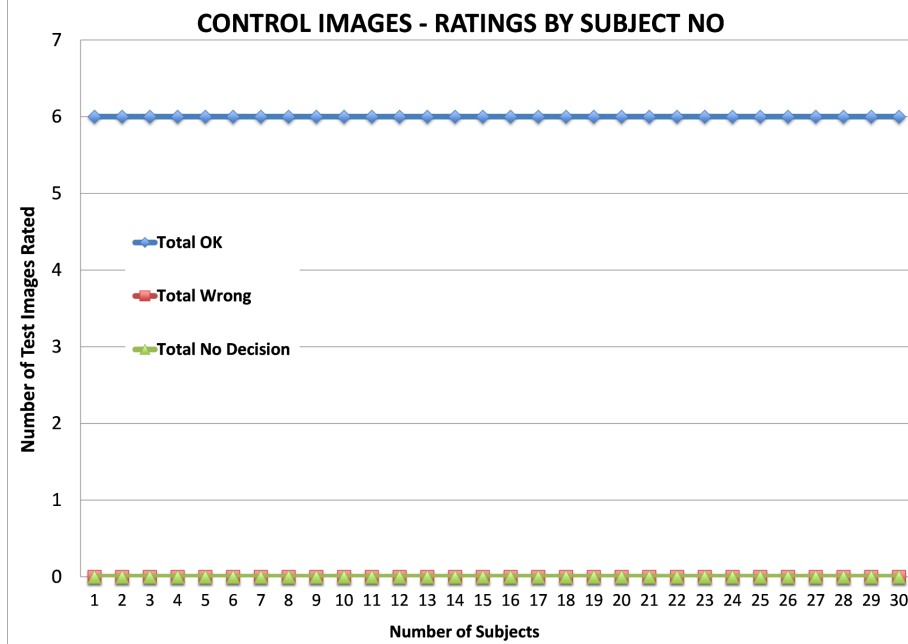


Figure 3: Ratings of control images by subject number at the VUB-imec lab.

where N is the total number of votes for a given stimulus, OK the number of times the flickering is correctly identified and ND the number of times that a subject decided to cast a no decision vote. The score is bounded to $[0, 2]$, where 0 means all subjective votes are correct; score 1 is achieved when all the votes are casted as no decision, implying that no subject could identify flickering (transparent coding). The upper bound score 2 is obtained if all votes are wrong. This case is avoided by using the outlier removal procedure described next.

3.5.2 Outlier removal and result analysis

To remove outlier subjects, a two-stage removal procedure was deployed:

Step 1 – Subjects who either misunderstood the test procedure or that appeared to have delivered insufficient effort in spite of the given instructions were removed. Six control stimuli with obvious rating choices were randomly inserted in the test sessions. It was expected that all subjects provide correct ratings (OK) for six images as the flickering was quite visible on images. Figure 3 illustrates the ratings of all 30 subjects for the control images at the VUB-imec test laboratory. As shown in the figure, all ratings were correct (OK) implying that subjects spent sufficient attention during the test. Identical results were achieved at EPFL and CTU labs.

Step 2 – To remove unreliable votes, the following threshold was applied to exclude subjects with too many wrong decisions:

$$\text{Threshold} = \mu_{Wrong} + 2 \times \sigma_{Wrong} \quad (2)$$

where μ_{Wrong} and σ_{Wrong} are the mean and standard deviation of the number of 'Wrong' decisions per subject. Subjects with a number of 'Wrong' decisions larger than the threshold were excluded from the test.

Figure 4 summarizes the overall vote cast by all subjects and the reliability threshold (black lines) at the VUB-imec test laboratory. According to the figure, the numbers of 'Wrong' decision votes cast by subjects 15, 19, and 23 were over the threshold. The reliability thresholds were similarly computed for EPFL and CTU data, where two outliers were identified in each lab.

Figure 5 compares the overall vote casts per image in the three labs. The results obtained from the three participating laboratories, namely CTU, EPFL, and VUB-imec, show some differences. The differences are more prominent for CTU vs. EPFL and CTU vs. VUB-imec results. The VUB-imec and EPFL results are more consistent. The results from the CTU laboratory showed a significant reduction in artifact sensitivity when

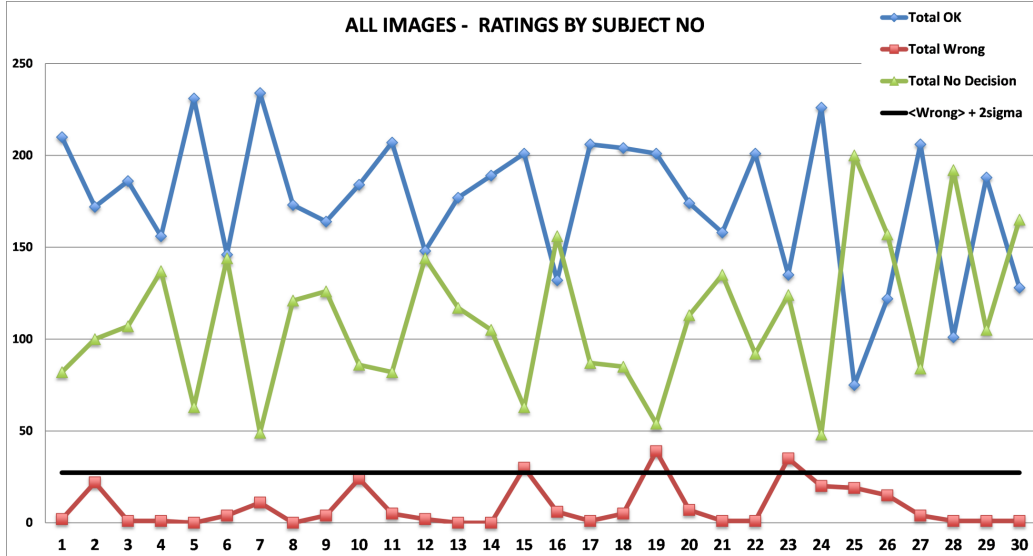


Figure 4: Number of ratings per subject at the VUB-imec lab.

compared to EPFL and VUB-imec. The following major inconsistencies and hypothetical causes of differences in obtained results could be identified in the setup and procedures used by the laboratories.

- At the EPFL and VUB-imec laboratories, the observers were allowed to get closer to the screen than initial 50 cm, whereas the observers at CTU were instructed to keep the viewing distance of 59 cm.
- The CTU laboratory reported questionable 10 bpc transparency of their setup, which appeared in the detailed technical inspection of the setup after the subjective campaign was finished, see Section 3.4 for more details.

Especially the viewing distance control seems to be the main cause of the differences. The viewers participating in the experiment at CTU often commented that the perceived differences are hardly visible at the nominal viewing distance, which made the evaluation frustrating, and that the artifact detection could be less cumbersome for shorter viewing distance than specified. However, this hypothesis requires further investigation. The second issue of questionable 10 bpc transparency revealed high 10 bpc configuration functionality dependence, especially due to software version combinations, e.g., OpenGL vs. GPU drivers, vs. MPV player version.

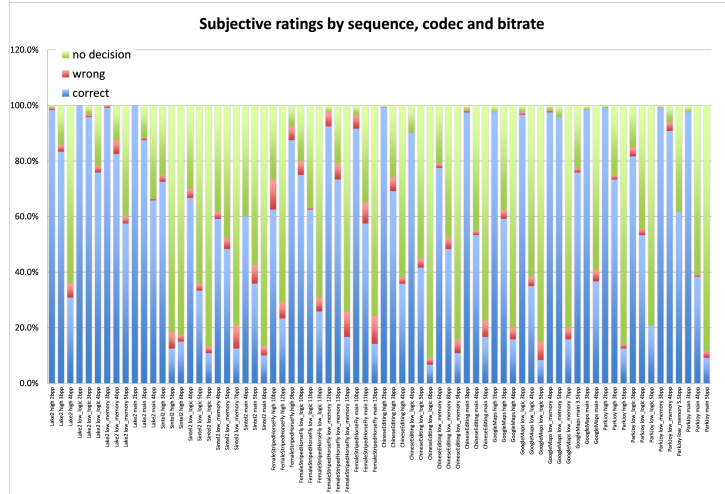
Figure 6 compiles the data collected by all three labs. The rating scores are reported using equation (1) for each codec configuration and target bitrate. For each image, the ratings are ordered by target bitrates and arranged by codec configuration. A higher score implies less flickering identification and better coding performance.

Table 1 summarizes the scores shown in Figure 6. The ratings higher than 50 % are considered as accepted (passed) coding performance (marked as green).

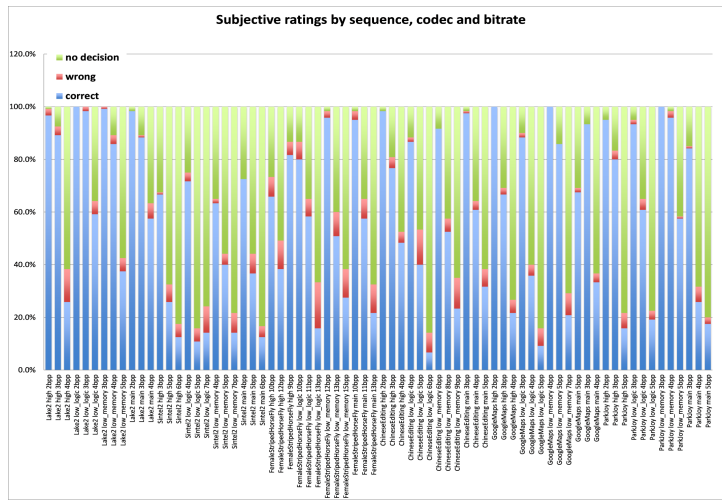
4. CONCLUSIONS

Basic characteristics of JPEG XS, low-latency, lightweight and visually lossless compression were reviewed and more importantly, the JPEG XS core coding system subjective study was described and obtained results were presented in this paper. This assessment aimed to make sure that the JPEG XS core coding system meets the requirements for visually lossless compression for different configurations, application categories, and decoder complexity.

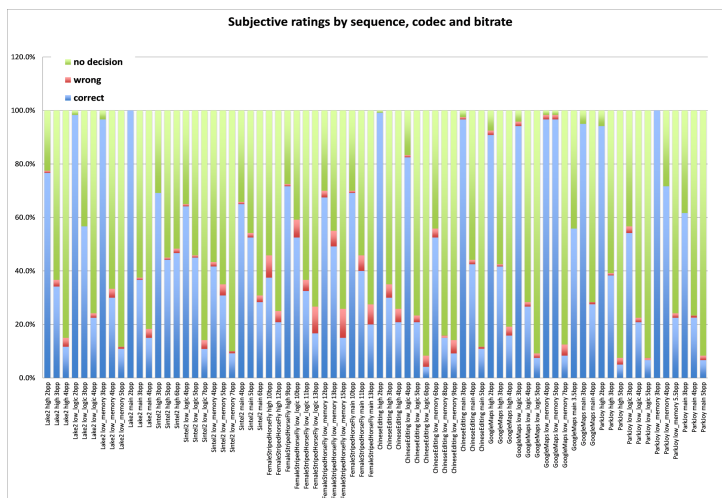
The methodology described in Annex B of ISO/IES 29170-2 (AIC Part-2) with flickering detection by the subjects was used in this experiment. Three test laboratories participated in the subjective campaign, six contents were used with four different codec configurations, and three bitrates for each content and configuration. The obtained results allow for analysis of important visually lossless property or transparency of the JPEG XS core



(a) VUB-imec

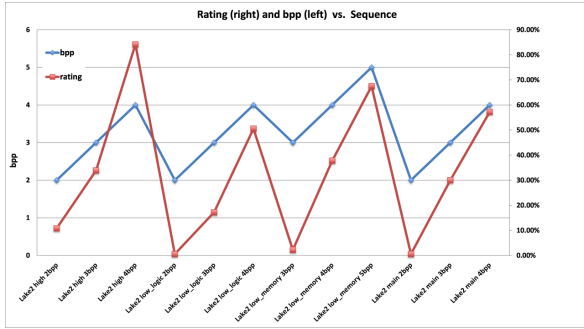


(b) EPFL

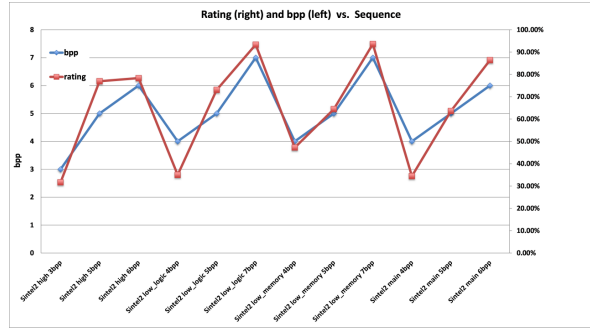


(c) CTU

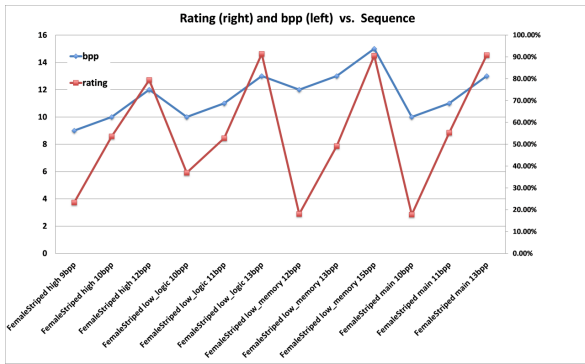
Figure 5: The stacked histogram of all votes by image, codec configuration, and bitrate.



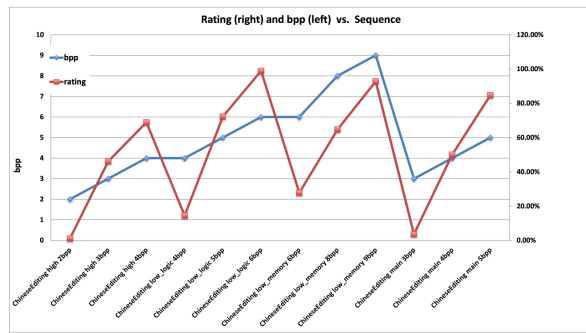
(a) Lake2



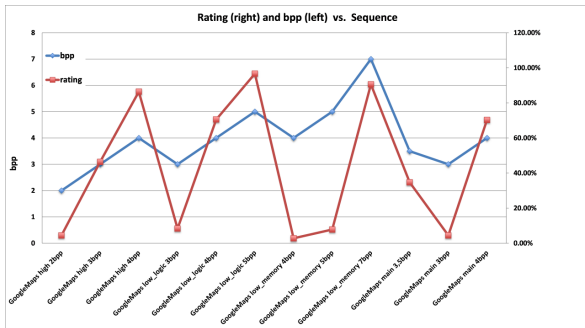
(b) Sintel2



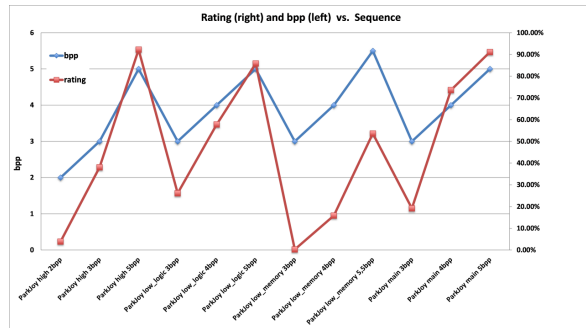
(c) FemaleStripedHorseFly



(d) ChineseEditing



(e) GoogleMaps



(f) ParkJoy

Figure 6: Subjective scores for each test stimulus ordered by target bitrates and codec configurations (All labs) for test images (a) Lake2, (b) Sintel2, (c) FemaleStripedHorseFly, (d) ChineseEditing, (e) GoogleMaps, (f) ParkJoy.

	Main		Low memory		Low logic		High	
	bitrate	score	bitrate	score	bitrate	score	bitrate	score
Lake2	2	0.56%	3	2.22%	2	0.56%	2	10.83%
	3	30.00%	4	37.78%	3	17.22%	3	33.89%
	4	57.22%	5	67.50%	4	50.56%	4	84.17%
Sintel2	4	34.44%	4	47.22%	4	35.00%	3	31.67%
	5	63.61%	5	64.44%	5	73.06%	5	76.94%
	6	86.39%	7	93.61%	7	93.33%	6	78.33%
FemaleStripedHorseFly	10	17.78%	12	18.06%	10	36.94%	9	23.33%
	11	55.28%	13	49.17%	11	52.78%	10	53.61%
	13	90.83%	15	90.56%	13	91.39%	12	79.44%
ChineseEditing	3	3.61%	6	27.78%	4	14.44%	2	1.11%
	4	50.00%	8	64.72%	5	72.22%	3	46.11%
	5	84.72%	9	92.78%	6	98.89%	4	68.89%
GoogleMaps	3	4.44%	4	2.78%	3	8.33%	2	4.44%
	3.5	34.72%	5	7.78%	4	70.56%	3	46.39%
	4	70.28%	7	90.56%	5	96.67%	4	86.39%
ParkJoy	3	19.17%	3	0.28%	3	26.11%	2	3.89%
	4	73.61%	4	15.83%	4	57.78%	3	38.06%
	5	91.11%	5.5	53.61%	5	85.83%	5	92.22%

Table 1: Pass-fail summary per image per bitrate at four codec configurations. The scores marked as red and green denote fail and pass conditions, respectively. (All labs)

coding system under test in the given context. At the chosen ratings higher than 50%, the tested codec proved to have required coding performance, which translates to compression ratios ranging from 2:1 to 6:1 for transparent compression.

The presented results from the three test laboratories show some differences, while one of the laboratories provided results with a significant reduction in artifact sensitivity. As a possible cause of this discrepancy, two setup issues were identified, namely, the viewing distance and questionable transparency of the 10 bpc signal pipeline. To confirm or reject these hypotheses it would be necessary to perform an additional subjective experiment. However, as the results of the two remaining laboratories are highly correlated, and do not show statistically significant differences, it is possible to state that the tested codecs proved to have the required coding performance.

ACKNOWLEDGMENTS

This work was partially supported by the following: Project No. GA17-05840S “Multicriteria optimization of shift-variant imaging system models” of the Czech Science Foundation; Walloon Region (Belgium) “WALinnov TrustEye”; ImmersiaTV project under the European Unions Horizon 2020 research and innovation program (grant agreement no. 688619) and funded by Swiss State Secretariat for Education, Research and Innovation SERI; Swiss National Foundation for Scientific Research (FN 200021 159575) project Light field Image and Video coding and Evaluation (LIVE).

REFERENCES

- [1] Willème, A., Descampe, A., Lugan, S., and Macq, B., “Quality and Error Robustness Assessment of Low-Latency Lightweight Intra-Frame Codecs for Screen Content Compression,” *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* **6**(4), 471–483 (2016).
- [2] SMPTE, “VC-2 Video Compression, document SMPTE 2042-1:2012,” (August 2012).
- [3] VESA, “Display Stream Compression (DSC) Standard v1.2,” (January 2016).

- [4] VESA, “Display Compression-M (VDC-M) Standard v1.1,” (May 2018).
- [5] “Overview of JPEG XS.” <https://jpeg.org/jpegxs/index.html>. Accessed: 2018-07-09.
- [6] Descampe, A., Keinert, J., Richter, T., Föbel, S., and Rouvroy, G., “JPEG XS, a new standard for visually lossless low-latency lightweight image compression,” in [*Applications of Digital Image Processing XL*], **10396**, 103960M, International Society for Optics and Photonics (2017).
- [7] Willème, A., Descampe, A., Rouvroy, G., Pellegrin, P., and Macq, B., “JPEG XS-based frame buffer compression inside HEVC for power-aware video compression,” in [*Applications of Digital Image Processing XL*], **10396**, 103960C, International Society for Optics and Photonics (2017).
- [8] WG1, “Call for Proposals for a low - latency lightweight image coding system .” https://jpeg.org/downloads/jpegxs/wg1n71031-REQ-JPEG_XS_Call_for_proposals.pdf. Accessed: 2018-07-10.
- [9] Willème, A., Richter, T., Rosewarne, C., and Macq, B., “Overview of the JPEG XS objective evaluation procedures,” in [*Applications of Digital Image Processing XL*], **10396**, 103960N, International Society for Optics and Photonics (2017).
- [10] McNally, D., Bruylants, T., Willème, A., Ebrahimi, T., Schelkens, P., and Macq, B., “JPEG XS call for proposals subjective evaluations,” in [*Applications of Digital Image Processing XL*], **10396**, 103960P, International Society for Optics and Photonics (2017).
- [11] “ISO/IEC DIS 21122-1.” <https://www.iso.org/standard/74535.html>. Accessed: 2018-07-09.
- [12] “ISO/IEC DIS 21122-2.” <https://www.iso.org/standard/74536.html>. Accessed: 2018-07-09.
- [13] “Rec. ITU-R P.910 Subjective video quality assessment methods for multimedia applications.” International Telecommunication Union (2008).
- [14] “Rec. ITU-R P.913 Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment.” International Telecommunication Union (2016).
- [15] “Rec. ITU-R BT.500-13 Methodology for the subjective assessment of the quality of television pictures.” International Telecommunication Union (2012).
- [16] De Simone, F., Goldmann, L., Baroncini, V., and Ebrahimi, T., “Subjective evaluation of JPEG XR image compression,” in [*Applications of Digital Image Processing XXXII*], **7443**, 74430L, International Society for Optics and Photonics (2009).
- [17] Artusi, A., Mantiuk, R. K., Richter, T., Hanhart, P., Korshunov, P., Agostinelli, M., Ten, A., and Ebrahimi, T., “Overview and evaluation of the JPEG XT HDR image compression standard,” *Journal of Real-Time Image Processing*, 1–16 (2015).
- [18] Hanhart, P., Rerabek, M., De Simone, F., and Ebrahimi, T., “Subjective quality evaluation of the upcoming HEVC video compression standard,” in [*Applications of Digital Image Processing XXXV*], **8499**, 84990V, International Society for Optics and Photonics (2012).
- [19] Viola, I. and Ebrahimi, T., “Quality assessment of compression solutions for ICIP 2017 Grand Challenge on light field image coding,” in [*2018 International Conference on Multimedia and Expo Workshops*], (CONF) (2018).
- [20] “ISO/IEC 29170-2:2015 Information technology – Advanced image coding and evaluation – Part 2: Evaluation procedure for nearly lossless coding.” <https://www.iso.org/standard/66094.html>. Accessed: 2018-07-16.
- [21] Allison, R. S., Wilcox, L. M., Wang, W., Hoffman, D. M., Hou, Y., Goel, J., Deas, L., and Stoltzka, D., “75-2: Invited Paper: Large Scale Subjective Evaluation of Display Stream Compression,” in [*SID Symposium Digest of Technical Papers*], **48**(1), 1101–1104, Wiley Online Library (2017).
- [22] “Rec. ITU-R BT.709-6 Parameter values for the HDTV standards for production and international programme exchange.” International Telecommunication Union (2015).