Vrije Universiteit Brussel

New procedures to evaluate visually lossless compression for display systems

Stolitzka, Dale F.; Bruylants, Tim; Schelkens, Peter

PROCEEDINGS OF SPIE

New procedures to evaluate visually lossless compression for display systems

Dale F. Stolitzka
Peter Schelkens
Tim Bruylants

SPIE.

New procedures to evaluate visually lossless compression for display systems

Dale F. Stolitzka[a], Peter Schelkens[b], Tim Bruylants[c]

[a] Samsung Electronics, Co. Ltd., 3655 N. 1st St., San Jose, CA, USA 95134;
[b] ETRO-Vrije Universiteit Brussel, Pleinlaan 2 - B-1050 Brussel, Belgium;
[c] IMEC Kapeldreef 75, B-3001 Leuven, Belgium

ABSTRACT

Visually lossless image coding in isochronous display streaming or plesiochronous networks reduces link complexity and power consumption and increases available link bandwidth. A new set of codecs developed within the last four years promise a new level of coding quality, but require new techniques that are sufficiently sensitive to the small artifacts or color variations induced by this new breed of codecs. This paper begins with a summary of the new ISO/IEC 29170-2, a procedure for evaluation of lossless coding and reports the new work by JPEG to extend the procedure in two important ways, for HDR content and for evaluating the differences between still images, panning images and image sequences.

ISO/IEC 29170-2 relies on processing test images through a well-defined process chain for subjective, forced-choice psychophysical experiments. The procedure sets an acceptable quality level equal to one just noticeable difference. Traditional image and video coding evaluation techniques, such as, those used for television evaluation have not proven sufficiently sensitive to the small artifacts that may be induced by this breed of codecs. In 2015, JPEG received new requirements to expand evaluation of visually lossless coding for high dynamic range images, slowly moving images, i.e., panning, and image sequences. These requirements are the basis for new amendments of the ISO/IEC 29170-2 procedures described in this paper. These amendments promise to be highly useful for the new content in television and cinema mezzanine networks.

The amendments passed the final ballot in April 2017 and are on track to be published in 2018.

Keywords: display electronic systems; data compression; advanced image coding and evaluation; subjective evaluation procedures; standardization; visual quality

1. INTRODUCTION

Light, visually lossless compression, also called display stream compression, promises a new level of visually lossless coding quality performed in real-time.[1] This class of codecs feature a unique combination of intra-frame coding with:

- moderate compression above four bits per pixel (bpp) with visually lossless image quality

- low latency measured in a number of lines and

- guaranteed real-time encoding and decoding

The Video Electronics Standards Association (VESA) found this combination was able to drive down implementation costs across many markets because consumer devices would benefit greatly from this new codec class.[2] The Joint Photographic Experts Group[3] (JPEG) developed requirements for real-time encoding either in FPGA hardware or with software running on a high-performance PC workstation. Their target is primarily compression in television and studio and cinematic production communication links.[4] However, the success of this codec class requires new techniques that are sufficiently sensitive to the small artifacts and subtle color variations. This paper summarizes the new ISO/IEC 29170-2 Evaluation procedure for nearly lossless coding[5], acknowledges the new codec requirements, discusses new work by JPEG to extend the existing procedures, reports test results of the procedure extensions and concludes with next steps in subjective visual quality evaluation.

Two types of full reference subjective evaluation of image or image sequence quality are in common use, comparison testing using a Lickert scale to obtain a mean opinion score or forced choice comparison. Both methods shows test material to a subject either side-by-side or sequentially in the case of full frame video for an "A" versus "B" comparison. In order to get a mean opinion score, the experimenter asks the subjects to quantify their opinion, "Grade the contents by which side is best and rate from 1 (worst) to 5 (best)". Results of the grading are tabulated across a range of test material and subjects to extract the mean opinion score. Forced choice comparisons either a binary choice, "Choose which side is better?" or "Choose which side is impaired?" or a ternary choice, "Choose either which side is better or select no difference."

The original evaluation procedure in ISO/IEC 29170-2 formalized the forced choice comparison method by enforcing rigor in all aspects of subjective testing. The subject selection is controlled with pre-screening for visual acuity, age, number and experimental instruction. The standard specifies the display quality, viewing environment, viewing distance test material presentation procedures, data collection automation techniques, data set size and statistical treatment of test data. Testing of the procedure underwent scrutiny for test repeatability both across differing populations and different test sites. VESA has made the subjective results and the test image data set available for download to facilitate data examination and verification of their codec, the Display Stream Compression (DSC) Standard.[6]

Subjective testing can be expensive and time consuming, however, comparison of the procedure's verification data shows relatively poor correlation with objective scoring by peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) and high dynamic range visual predictor (HDR-VDP2).[7] The need for exacting quality testing in order to attain visually lossless quality is demanding. Until such time that objective metrics catch up to fine artifact testing, the industry will need to rely on rigorous subjective testing coupled with a suitable acceptable quality level for products.

The first of two new amendments enhances the evaluation procedure for processing and viewing HDR test material. Additional rigor in a comfortable viewing distance viewing environment, display type and a software flow for pre encoding and post decoding images or image sequences prior to rendering for subject testing. The evaluation procedure for standard dynamic range (SDR) test materials stipulated a benign viewing environment in a darkened room and a display calibrated between 100 cd/m^2 and 120 cd/m^2. Professional monitors suitable for testing usually support Adobe RGB color and can store calibration data in non-volatile look-up-tables.[8]

However, HDR presents challenging display requirements:
- high brightness, > 540 cd/m^2 for organic light-emitting diode displays or > 1000 cd/m^2 for liquid crystal displays with a backlight
- high contrast > 10,000:1
- RGB 4:4:4, 30-bit color
- > 90% Digital Cinema Initiative's DCI-P3 color volume
- ability to turn off in-display video processing enhancements or sub-sampling when viewing with a television
- receive and interpret HDR static metadata[9] signaling in the video transport stream

A few professional HDR displays can meet the above specifications, such as, the Sony PVM-X550 55" 4K OLED monitor or the Dolby PRM-4200 42" cinema test monitor; however, both examples cost over $20,000, unavailable to many test labs that may want only occasional usage. Few TVs enabled for HDR can be used. TVs are designed for consumer entertainment where HDR settings usually include picture processing enhancements that do not support RGB 4:4:4 30-bit processing from the set input to the display panel.

The second amendment of ISO/IEC 29170-2 extending procedures to include image sequence evaluation. The amendment specifies sequence duration, size frame rate and pre-processing techniques to prepare test materials. A subset of the image sequence testing further recommends strict still image testing by panning an image through a small window in the subjects view over a short sequence of frames similar to slowly scrolling a picture across a screen.

The amendments described in this paper passed the final international standards ballot approval in April 2017, and are on track to be published in 2018.[10]

2. METHODS

2.1 Evaluation Procedure for nearly lossless coding

In 2015, the JPEG committee published the ISO/IEC 29170-2 Evaluation procedure for nearly lossless coding, called AIC-2 (Advanced Image Coding evaluation, Part 2). This international standard is becoming the basis for fine-tuned subjective image evaluation for any light compression codec that can approach visually lossless image quality.[11] VESA and the JPEG XS project use the procedure to produce sign-off image quality results for its codec developments.

The most difficult of two methods in the AIC-2 procedure calls for an subject to choose between two apparent images, one a reference (uncompressed) image and one an image sequence with the reference image interleaved with the test (reconstructed) image; see Figure 1. If the codec compression-reconstruction process altered the test image significantly, the interleaved test image sequence will flicker.
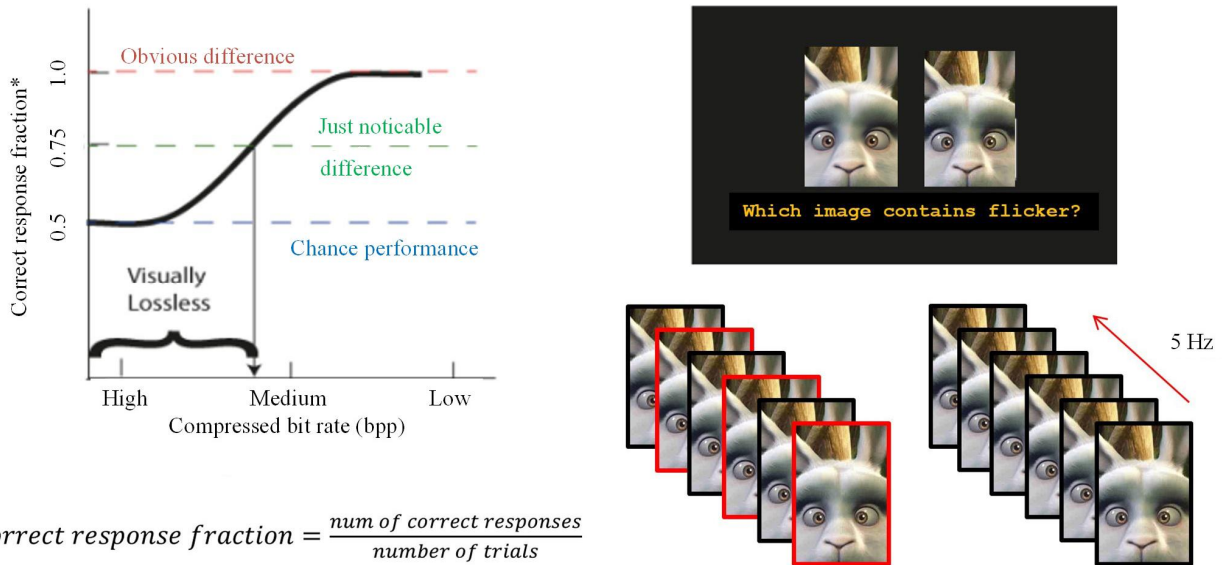

$$* \ correct \ response \ fraction = \frac{num \ of \ correct \ responses}{number \ of \ trials}$$

Figure 1. Evaluation procedure that measures the ability to discern induced flicker.[12]

The procedure asks the subject whether the left or right image is not flickering. Over the course of the experiment, the subject should view the same test image coded with one algorithm 30 times. If the subject cannot see a difference, the result is the same as random guessing, 0.5 correct response rate, while if the difference is obvious, the subject should answer correctly for every view, 1.0 correct response rate. Half way in between is 0.75 response fraction, also known as 1 JND (just noticeable difference). AIC-2 defaults to one just noticeable difference (JND) as the threshold for visually lossless quality, a strict metric.

The AIC-2 procedure also allows a less strict method where the reference image and test image are presented side by side. The subject still must choose one of two images compared to a third reference image, also a forced choice procedure. The same statistical analysis applies to report the correct response fraction. The side-by-side presentation AIC-2 differs from subjective assessment specified in Recommendation ITU-R BT.500, which asks the subject to rate the test image on a Lickert scale[13] from 1 to 5, unacceptable to excellent image quality.

AIC-2 is a forced choice comparison that automatically removes subject biases present in Lickert scaling. AIC-2 is stricter, too. The flicker method is the stricter of the two methods in AIC-2 which is a proxy for identifying both spatial and temporal image artifacts that could be discernable in a mobile display or television.
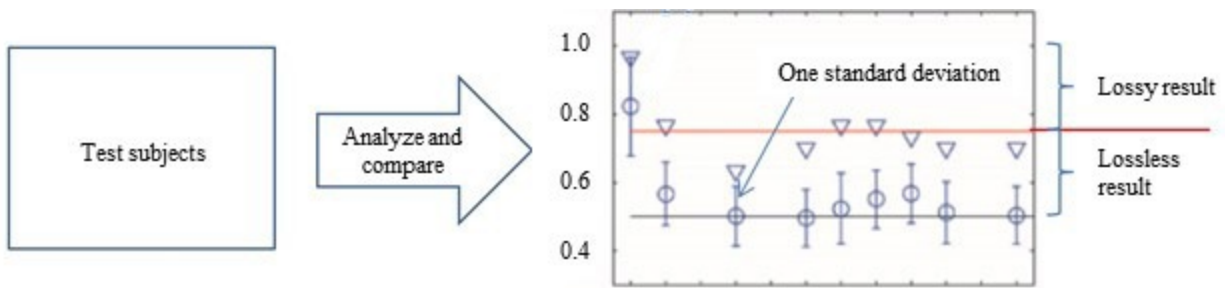
Figure 2. Report format for ISO/IEC 29170-2: ○:= mean, ▽:= maximum response, bars:= 1σ

2.2 Treatments for high dynamic range contents amendment

AIC-2 has been used to test contents by using a calibrated, 30-bit color SDR display that can be driven by professional graphics cards installed in a PC. The amendment will add a software flow that relies on high brightness SDR mode to bypasses the unwanted TV HDR processing. The software flow and a high brightness SDR monitor or TV can support testing of a large portion of the HDR image's wide color gamut when a professional HDR graphics card and monitor setup is not available or too expensive. Figure 3 shows the software flow for converting an HDR image processed from a movie graded to 1000 cd/m^2
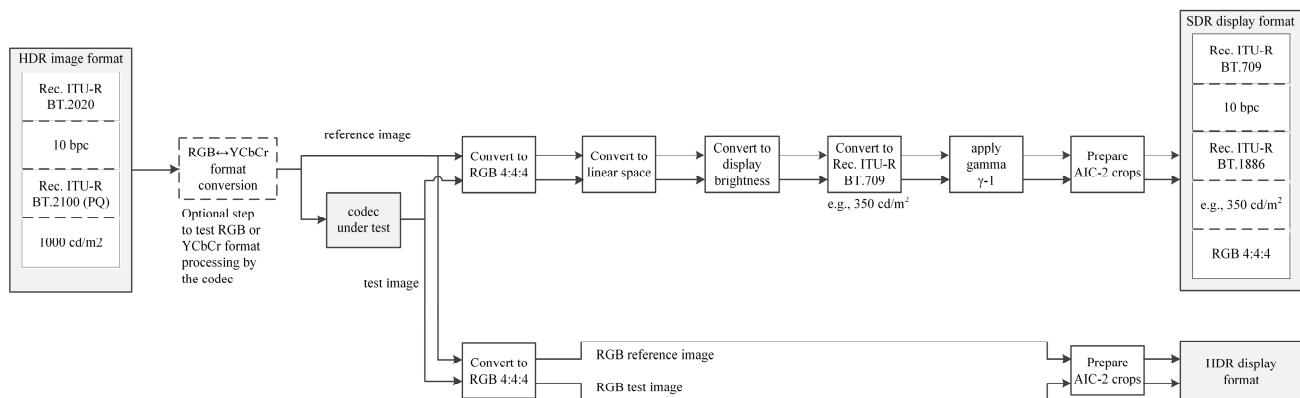

Figure 3. Image preparation software flow for HDR image and image sequence testing. Following one of these flows, the contents should be sent to the full reference testing using either flicker (strict) or side-by-side comparison.

HDR images may be tested by the software flow using either an SDR display (top path) or an HDR display (bottom path). The HDR testing amendment provides two flows because HDR test displays are rare or expensive or both. Televisions may be useful for consumer satisfaction testing.

The software flow is designed to provide a reliable and repeatable base line for testing HDR contents. The flow in Figure 3 duplicates the image processing pipeline to simulate delivery from a set-top box or PC to a television where a known inverse electro-optical transfer function (EOTF), such as Recommendation ITU-R BT.2100 (PQ) is applied by the graphics hardware prior to the codec. A lower than maximum brightness avoids brightness regions where a consumer television product must apply tone-mapping to ensure a pleasing, unclipped presentation.

Researchers at Samsung found subjects were uncomfortable when sitting closely to a large screen television.[14] A 65" UHD television has 3840 x 2160 active pixels where the required 30 pixel per degree (PPD) angular viewing distance positioned the subject at 64 cm, or 0.8×H, where H is the screen height. Subjects moved back to 60 PPD (1.4×H) were more comfortable and at the minimum recommended viewing distance for UHD televisions by *AVS Forum*.[15]

The 60 PPD viewing distance found in Table 1 were employed during the verification phase when a using television.

Table 1. HDR amendment changes to viewing distance

Condition	PPD[a]	D[b] cm
Viewing distance for SDR evaluation	30	D equals the larger of the value in the following equation or 12cm[c] $$D = \dfrac{W}{H_{RES} \times \tan(\dfrac{1}{PPD})}$$
Viewing distance for HDR evaluation	60[d]	

[a] The experiment requires a consistent display orientation to be maintained and mobile display may have a different width and pixel resolution in landscape versus portrait orientation. PPD is calculated for each orientation. Detailed work on computer displays and mobile devices tends to be closer than for general entertainment, e.g., television, and requires evaluation with a more aggressive PPD than would be the case for Snellen acuity (30 cycles/degree or PPD = 60)

[b] W is the screen width (cm) and HRES is the number of pixels across the display horizontally as viewed by the observer.

[c] The minimum focusing distance for normal vision is predetermined as 12 cm by this document.

[d] Snellen distance viewing distance may be used for SDR evaluation when the evaluator determines the display (television) is large enough to cause observer discomfort.

2.3 Image sequence testing amendment

The second new AIC-2 amendment is an evaluation procedure of video using image sequences or image stacks. The procedure is not useful for video that used a temporal codec, only for measuring artifacts between full frames. The existing still image procedure is rigorous for spatial compression artifacts and is a good proxy of temporal artifacts. The second new amendment enhances the existing evaluation procedure by supporting video within image sequences to ensure that temporal artifacts do not escape the analysis even when using a good proxy for temporal artifacts.

An additional process in the second new amendment is designed to simulate a panning image on a display. Images pan diagonally, horizontally or vertically by one pixel per frame within a fixed window. The reference and test images incrementally pan through a fixed window (Figure 4). If the codec has a spatial dependency during the encoding process, artifacts can be seen that may have escaped.
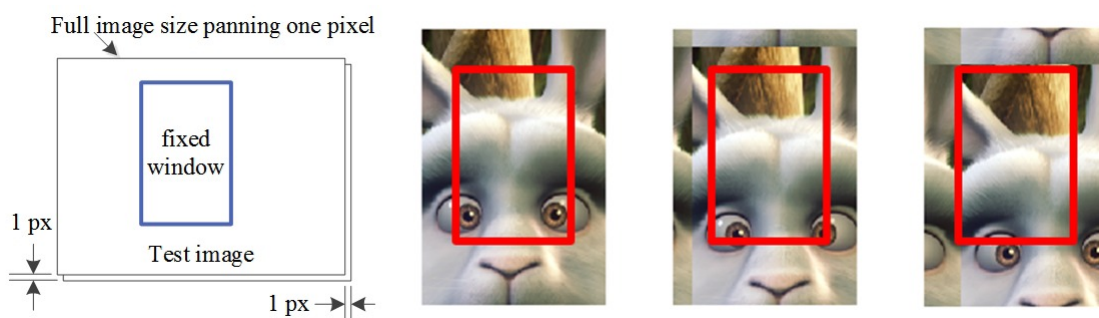

Figure 4. Image panning by a one pixel shift, in this case diagonally through the fixed window.

The new panning method minimizes changes for the subject; the presentation of two images remains the same, the time for responding remains the same, only the reference and test images appear to move smoothly through the field of view.

For image panning, the image under test and reference do not interleave, only the reconstructed test image is shown. The test will nevertheless show scintillations or flicker as the image pans through a small window.

3. EXPERIMENTAL VERIFICATION

3.1 Experimental Setup

Both amendments underwent verification testing by several universities and companies. Shared software tools assisted the process that was combined with codec verification for preliminary testing of a 4:1 codec by VESA and the JPEG XS proposal down selection process.

The HDR amendment testing verified the software flow and verified the viewing distance changes. Materials from The Blender Foundation provided appropriate cinematic content from *Big Buck Bunny*[16]*, Sintel*[17] and *Tears of Steel*[18]. Frames from Tears were directly usable in HDR processing, *Big Buck Bunny* and *Sintel* were converted from wide color gamut original materials. The picture in Figure 5 shows the full image and the crop region to be viewed.

Figure 5. Full screen cropped to the region of interest for subjective evaluation

Equipment for HDR testing of the software flow used a Samsung JS9500 65" diagonal television driven to a maximum 350 cd/m^2 brightness with a PC and discrete graphics card. This experimental value was found to be far away from any tone mapping by the television video card, therefore, colors and artifacts due to compression would be reproduced.

Subjects can be sensitive to small timing differences that could result in display flicker not related to the compression artifacts. Sometimes the display has dithering, there is little the experiment can do to avoid this flickering other than not use a dithered display, which is part of the original standard's cautionary notes on displays. Sometimes the flicker is induced by poorly replicating presentation timing where an image may be buffered before presentation. There are two techniques used to avoid uncertain presentation timing.

The first way used by Samsung and York built scripts with OpenGL that write directly to a graphics card that controls display timing. The open source Matlab toolbox Psychtoolbox uses this technique.[19] Equivalent scripts using Python can do the same thing. Both Matlab and Python support subject automated subject feedback recording so that data can be collected efficiently and without error. The second way used by Vrije Universiteit Brussel (VUB) builds two videos of the stacked images for playback. Using MPV, a precise-timing video player, and controlled through a Lau script that also recorded the subject response input.[20]

Experimenters should always test the bit-depth support of any third party software to ensure 30 bit per pixel is enabled and rendered correctly when testing wide color gamut or high dynamic range imagery.

Equipment for testing video sequences and panning sequences used both the JS9500 and an Eizo Coloredge® 24" professional monitor with calibrated color.

All labs performed visual acuity screening and restricted age to the extent allowed under societal norms of the locations.

3.2 Stimuli

Images from several images sets contained materials of different types. Figure 6 shows images and image crops used from a few of the experiments. A set of experiments used full images rather than crops.[21]
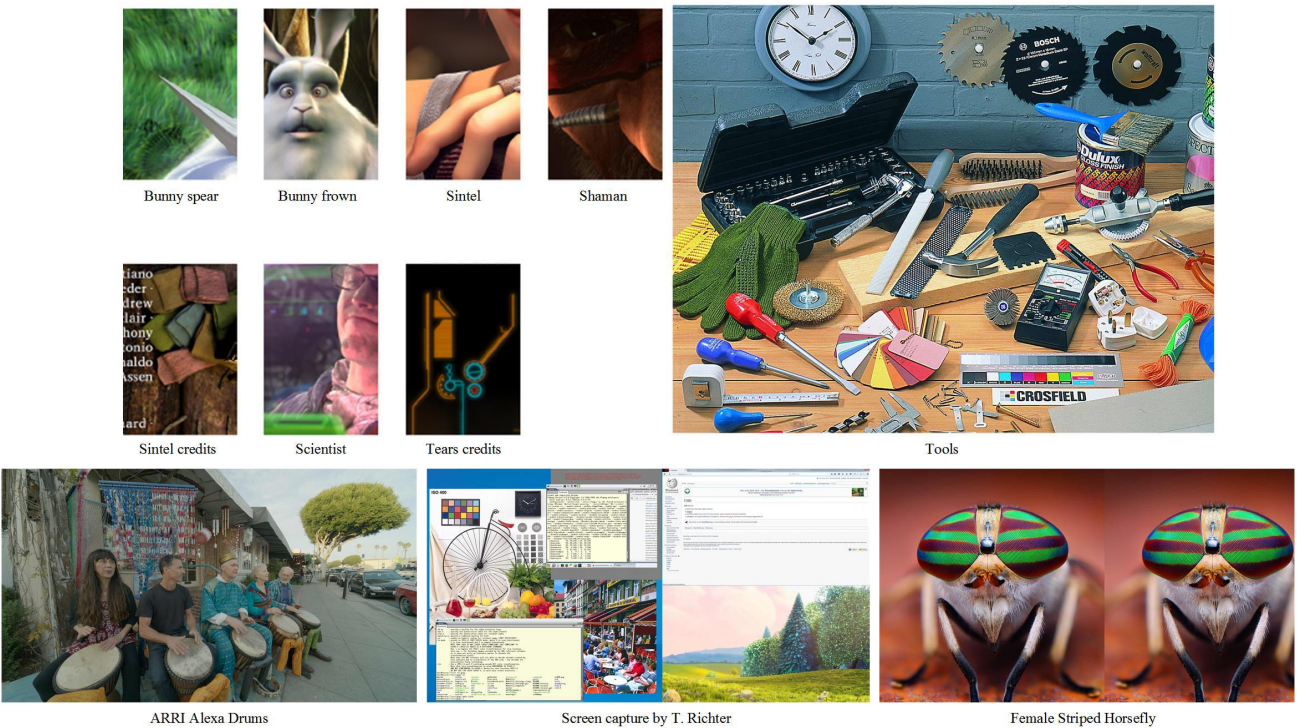

Figure 6. Examples of stimuli used by laboratories for amendment verification.[22 23 24 25]

4. RESULTS

This paper reports experiments from three sources, JPEG XS core experiment #1 at VUB, Samsung in San Jose, CA and York University. Table 2 summarizes testing at each site.

Table 2. Summary of testing by test site

	JPEG XS core exp #1	Samsung, San Jose	York University
Subjects	6	10	130
Repetitions per image	4	20	30
HDR	No	Yes	No
Still image flicker	Yes	Yes	Yes
Image panning	Yes	Yes	Yes
Image sequence (video)	Yes	Yes	No
Interlaced sequences	Yes	No	No
Codec vehicle	JPEG 2000 (restricted tiles), VC-2 HQ, and six JPEG XS candidates	DSC	DSC, VC-2 HQ and JPEG 2000
Bit rate for 24-bit RGB	4 bpp to 12 bpp	8, 12 bpp	4, 6, 8, 10, 12 bpp

This report used only RGB 4:4:4 (no sub-sampling) for its amendment verification but several modes may have been tested by each lab, for example Allison also tested YCbCr 4:2:2 and YCbCr 4:2:0 versions of the same test images.

4.1 HDR testing

The HDR testing followed the software flow shown in Figure 3 for low brightness monitors because no source of controlled HDR metadata through a PC was available at the time of the testing. For all HDR viewing, the JS 9500 television rendered images with viewers positioned at 60 PPD.

Results are reported by Hoffman, Wang and Stolitzka at the International Display Workshop in December 2016.[26] Testing with the software flow technique were found to be consistent with results reported by Hoffman and Stolitzka (2015). In this regard, artifacts have been preserved and the 60 PPD viewing distance verified as suitable for large screen testing.

4.2 Image sequence testing

Image sequence testing had three comparison points, below, and the summary in this section comment on an example result for each case.

1. Panning versus still image flicker testing
2. Panning versus image sequence as a video test
3. Side-by-side image sequence versus flickered image sequence both as a video test

The first case is analyzed by results in Figure 7, which is data from Hoffman that demonstrates two cases where the panning procedure identified artifacts at higher response fraction rate than the AIC-2 static image flicker method. The selected images from *Sintel*[27] and *Tears of Steel*[28] are relatively low in luminance and fine line details. In the *Sintel* crop, subjects often obtained a clue from Sintel's hair, barely visible at the top of the image but revealed with flicker during the panning. This image also included a false clue. The shoulder strap fabric will scintillate in both the reference and test images from a moiré effect.
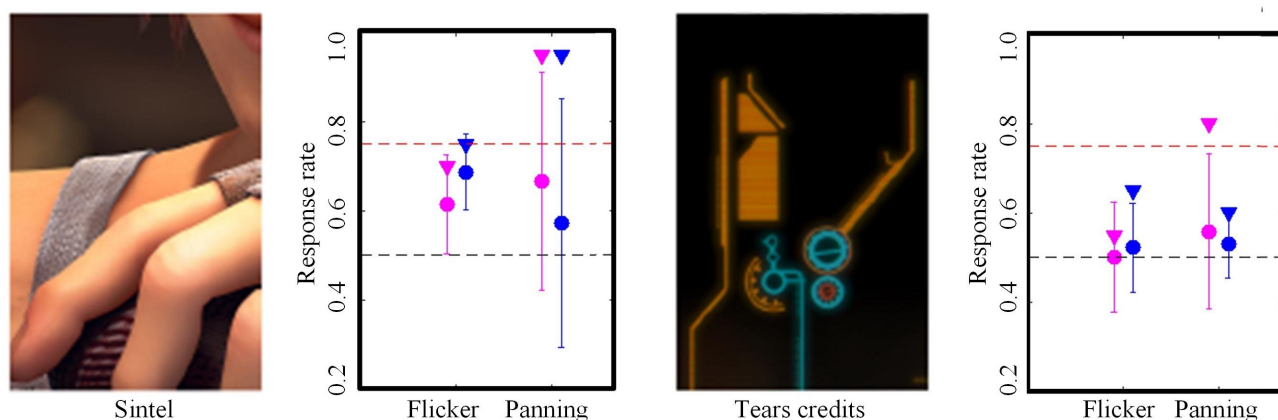

Figure 7. Panning test reports that indicate slight sensitivity increase in detection for some scenes in *Sintel* and *Tears of Steel*.

In the scatter plots to the right of each image in Figure 7, both *Sintel* and *Tears* appear visually lossless by the flicker test. The panning test allowed a few subjects to obtain sufficient clues to reliably identify the reference from the test; however, the average subject response is visually lossless in both cases. The average remained below the 1 JND limit and the average flicker and panning outcomes are within one standard deviation.

In the second comparison to verify panning, the frame rate of the panning played a critical role. If the image moves too quickly, Allison found that motion silencing came into effect and rendered otherwise visible artifacts less susceptible to detection. However, Figure 4 illustrates using a slower panning of one pixel either horizontally, vertically or diagonally at a 30 Hz advance speed is very effective.[29]

Allison tested several panning sequences at 60 fps and 30 fps. His results show higher defect impairment rate at 30 Hz than 60 Hz (Figure 8). Further the defect detection rate at 60 Hz fell below the defect rate found with still image flicker.
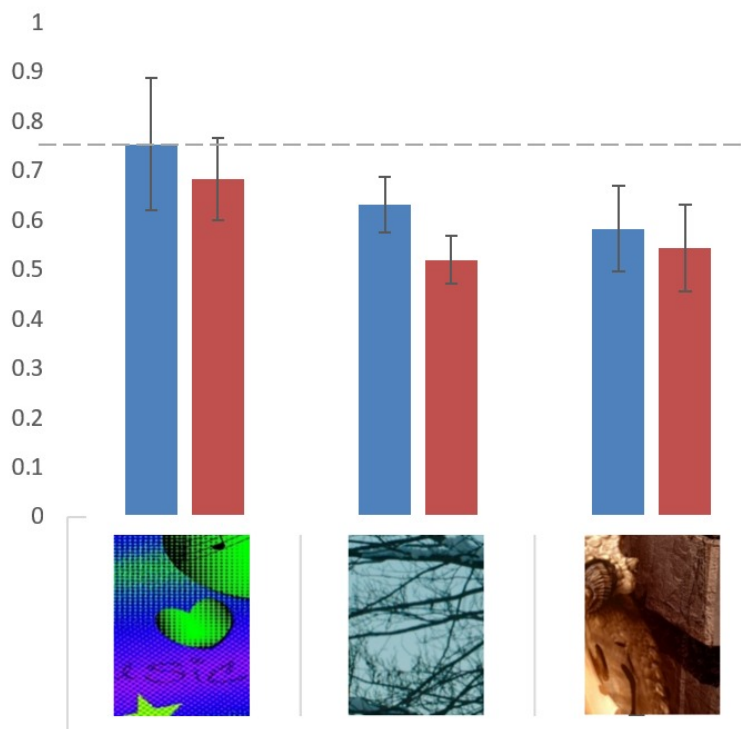

Figure 8. Effect of panning at 30 Hz (blue) versus 60 Hz (red)[30][31][32]

The third comparison case studied whether defect visibility could be improved in side-by-side "video" comparisons, which is common for high compression codec testing. First, the experiment tried side-by-side image sequence comparison, where the reference is one side, the reconstructed image sequence is the other side. Usually results are visually lossless due to the motion silencing found earlier, except in highly impaired image sequences. The results were compared by playing a reference unimpaired video on one side and comparing to a flicker rate of 1/8 second of reference and reconstructed images in the sequence. At 24 Hz frame rate, 1/8 second equals three frames.

The summary conclusion at VUB found no detectable improvement over side-by-side comparisons. The JPEG committee accepted this result and dropped the test case from the amendment.

Table 2, shows several test sites performed overlapping tests and the corresponding reports showed good correlation and agreeable results. The reader is encouraged to seek out these cited published data and data in JPEG committee records.

5. CONCLUSION

Results from three test sites verified the procedures that have been included in AIC-2 final amendments for HDR testing and image sequence testing. Work by Allison, at York, Hoffman, at Samsung and by Bruylants, at VUB, form a sound basis supporting the amended procedures.

It is worth sharing that the AIC-2 procedures are rigorous and a good stressing of the visually lossless codec class. Depending on the test material, nearly any codec can be induced into some flickering which allows for more discrimination between codecs rather than rely very mild differences or statistically insignificant measures.

The following points summarize the results conclusions for the new AIC-2 evaluation procedures:

1. Still image flicker is a highly effective test method.

2. 30 Hz image panning is effective and slightly more rigorous than still image flicker testing in a few cases.

3. 60 Hz image panning often will not show defects shown at 30 Hz panning or with static flicker testing

4. 60 PPD is a comfortable and effective viewing distance with UHD resolution large displays

5. The HDR test flow is effective at preserving defects, but a full HDR test monitor is preferred if available

6. Audio feedback is anecdotally helpful to maintain subject interest and attention

The authors encourage future work in this area to establish standard image sets that broadly represent specific application types, such as television or cinema production or gaming.

REFERENCES

[1] Stolitzka, D. "Developing Requirements for a Visually Lossless Display Stream Coding System Open Standard," SMPTE Motion Imaging Journal 124(3), 59-65 (2015).

[2] VESA, "VESA Issues Call for Technology: Advanced Display Stream Compression," <https://www.vesa.org/news/vesa-issues-call-for-technology-advanced-display-stream-compression/> (16 January 2015).

[3] JPEG, "ISO/IEC JTC 1/SC29/WG1," <https://jpeg.org/about.html> (1 July 2017).

[4] JPEG, "JPEG Initiates Standardization of Low-latency Lightweight Coding System - JPEG XS," <https://jpeg.org/items/20160311_cfp_xs.html> (26 February 2016).

[5] ISO/IEC 29170-2, [Information Technology – Advanced image coding and evaluation – Part 2: Evaluation procedure for nearly lossless coding], ISO/IEC, Geneva (2015).

[6] VESA DSC, "Purchase Standards,"<https://www.vesa.org/vesa-standards/> (18 January 2017).

[7] Hoffman, D.M., Stolitzka, D. "A new standard method of subjective assessment of barely visible image artifacts and a new public database," J Soc Info Display 22 (12), 631-643 (2015).

[8] Eizo, "ColorEdge® Color Management Monitors," <http://www.eizo.com/products/coloredge/> (1 July 2017).

[9] CTA-861.3, [HDR Static metadata Extension], Consumer Technology Association, Washington, D.C. (2015).

[10] ISO/IEC JTC 1/SC29, "Programme of work," <https://www.itscj.ipsj.or.jp/sc29/29w42901.htm> (1 July 2017).

[11] Federal Agencies Digital Guidelines Initiative, "Term Compression, visually lossless," http://www.digitizationguidelines.gov/term.php?term=compressionvisuallylossless> (1 July 2017).

[12] "Big Buck Bunny," (CC) The Blender Foundation | <https://peach.blender.org/>.

[13] Trochim, W.M.K., "Lickert Scaling," <http://www.socialresearchmethods.net/kb/scallik.php> (20 October 2006).

[14] Hoffman, D.M., Stolitzka, D., Wang, W., "Verification of visually lossless image quality for display stream compression in consumer devices," International Display Workshop, Fukuoka, Japan (Dec 2016).

[15] M2N Limited, "Optimal viewing distance by the size of the television and the resolution," <https://www.avforums.com/image.php?imageparameters=editorial/products/38c56aa3af5093a6cc7282bb9bf5eee8_0.jpg> (1 July 2017).

[16] op. cit., "Big Buck Bunny."

[17] "Sintel," (CC) The Blender Foundation | <https://durian.blender.org/>.

[18] "Tears of Steel," (CC) The Blender Foundation | <https://mango.blender.org/>.

[19] Kleiner, M., Psychtoolbox-3, <http://psychtoolbox.org> (11 June 2017).

[20] MPV, <https://mpv.io/> (1 July 2017).

[21] JPEG, "JPEG XS announced participation in core experiments by ISO delegates," <https://jpeg.org/items/20170405_press.html> (5 April 2017).

[22] Shahan, T., "Female Striped Horse Fly (Tabanus lineola)," (CC) Attribution 2.0 Generic.

[23] ARRI – Arnold and Richter Cine Technik GmbH, "Alexa Drums," <https://www.arri.com/DE/camera/alexa/learn/alexa_sample_footage/> permission given for technical research.

[24] Richter, T. "Richter Screen Content," (CC) 4.0 BY-SA.

[25] Clark, R., "Tools," no copyright claim.

[26] op. cit., D.M. Hoffman, D. Stolitzka, W. Wang.

[27] op. cit., "Sintel."

[28] op. cit., "Tears of Steel."

[29] Allison, R.S., Wilcox, L.M., Wang, W., Hoffman, D.M., Hou, Y., Goel, J., Deas, L., Stolitzka, D. "Large-scale subjective evaluation of display stream compression," SID Symp Dig Tech Papers 48(1), 1101-1104 (2017).

[30] Sauermaul, S., "Background Music 203", <http://www.publicdomainpictures.net/> public domain dedication.

[31] ARRI – Arnold and Richter Cine Technik GmbH, "Public University," (CC) Attribution 3.0.

[32] op. cit., "Sintel."