INTERNATIONAL STANDARD

ISO/IEC 29170-2

First edition
2015-08-15

AMENDMENT 1
2020-03

Information technology — Advanced image coding and evaluation —

Part 2: Evaluation procedure for nearly lossless coding

AMENDMENT 1: Evaluation procedure parameters for nearly lossless coding of high dynamic range media and image sequences

Reference number
ISO/IEC 29170-2:2015/Amd.1:2020(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see http://patents.iec.ch).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

A list of all parts in the ISO/IEC 29170 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Information technology — Advanced image coding and evaluation —

Part 2: Evaluation procedure for nearly lossless coding

AMENDMENT 1: Evaluation procedure parameters for nearly lossless coding of high dynamic range media and image sequences

Introduction

Replace the text with:

This document normalizes evaluation and grading of a light coding system used for displays and display systems but is independent of the display technology. The procedure measures whether an observer can distinguish between an uncompressed reference and the reconstructed image or image sequence to a pre-determined, statistically meaningful level.

The procedure compares individual images or image sequences with two possible forced choice comparison test methods. The procedure relies on subjective evaluation methods designed to discern coding imperfections on electronic colour displays of any technology or size.

Selections for testing a specific coding system has bearing on the results this procedure will yield, but specific images or image sequences required for testing are not within scope, excluding an informative annex describing self-test certification. Content categories may vary between end-usage products. For example, content relevant to television manufacturers may or may not be relevant to computer display manufacturers. Due to the nature of this procedure as a visual psychophysical test, the observer's age is considered a meaningful parameter of the results.

Clause 3

After 3.18 add 3.19 to 3.23:

3.19
high dynamic range
image or image sequence format range conveying a larger range of perceptible shadow and highlight detail than in a standard dynamic range image or image sequence, with sufficient precision and sufficient separation of diffuse white and specular highlights

3.20
image sequence
plurality of images, either reference images or reconstructed test images, shown in progression

3.21
sensory feedback
audial, visual or haptic signal indicating the correctness of an observer's response

3.22
standard dynamic range
image or image sequence format conveying typical colour volume and rendering characteristics similar to those specified in Recommendations ITU-R BT.709 or ITU-R BT.1886 or IEC 61966-2-1 (sRGB)

3.23
wide colour gamut
image or image sequence format rendered a colour range larger than standard dynamic range systems, typically >75 % of the human visible spectrum

Clause 4

After AQL, add:

EOTF electro-optical transfer function

HDR high dynamic range

After RGB, add:

SDR standard dynamic range

WCG wide colour gamut

Subclause 5.1

Add a new row at the end of Table 1:

Annex H	forced choice paradigm with image sequences (no interleaving)	side-by-side, cropped image sequence comparison

Subclause 5.2

Add a new paragraph at the end of 5.2:

Media mastered in high dynamic range should be processed following guidance in Annex G. If an HDR display capable of rendering the media colour and brightness is available, experimenters should opt for the procedure in an HDR display hardware processed workflow, see G.5. If an appropriate HDR display is not available, this document provides an HDR software processed workflow that is display-independent, see G.4.

Subclause 5.3.3

Replace list item a) with:

a) Explain the use of the software to record image or image sequence assessments.

Replace list item e) with:

e) Explain if sensory feedback is provided (see 5.9), and which signal is correct or incorrect.

Subclause 5.3.4

Replace list item a) with:

a) Use the control images or sequences from the experiment as test images or image sequences.

Replace list item d) with:

d) Prompt the observer when a correct or incorrect response is entered. If incorrect, continue by repeating the test image or image sequence until a correct response is entered.

Subclause 5.4.1

Replace the subclause heading with:

5.4.1 Standard dynamic range lighting and display calibration

Replace list item c) with:

c) The surrounding walls and ceilings do not require a specific colour but shall not cause distracting reflections that may affect the vision of the observer. An appropriate viewing booth is optionally desirable.

Subclause 5.4.2

Replace Table 2 with:

Table 2 — Viewing distance versus display size and resolution

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

[a] The experiment requires a consistent display orientation to be maintained and a 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 H_{RES} 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 viewing distance may be used for SDR evaluation when the evaluator determines the display (television) is large enough to cause observer discomfort when at a close viewing distance based on 30 PPD.

Subclause 5.4.3

Add a new subclause after 5.4.3:

5.4.4 High dynamic range lighting and display calibration

Viewing conditions shall be consistent with ISO 3664 viewing conditions for images displayed on a high brightness display, such as an HDR-capable television or a wide colour gamut test monitor. Exceptions include:

a) The luminance of the peak brightness displayed on the monitor shall be >300 cd/m^2.

b) Displays that do not contain calibration tables should be avoided. However, if used, this document recommends televisions and monitors with colour, contrast and tint adjustment controls so that the light output can be characterized with a spectrometer and colour calibration completed using any of a number of procedures noted by the monitor or TV maker. Deviations shall be noted in the test report. The maximum luminance of the display used in the procedure shall be recorded by using a 2 % white centre on a black field.

c) The surrounding walls and ceilings do not need to be a specific colour but shall not cause distracting reflections that may affect the vision of the observer. An appropriate viewing booth is optionally desirable.

d) The display monitor shall render at least the number of bits per component present in the tested images.

Subclause 5.5

Replace the first paragraph with:

An observer shall view image paired stimuli on the display for no more than 4 s. An observer shall view image sequence paired stimuli for no more than 10 s.

Add the following paragraph at the end of 5.5:

The image viewing time may be extended to 10 s if the evaluator finds that the test display shows temporal dithering effects, which tend to distract observers from clearly identifying dithering from scintillations caused by image interleaving employed by the protocol in Annex B.

Subclause 5.8

Add a new subclause after 5.8:

5.9 Sensory feedback

An experiment may provide an observer with sensory feedback after a response that indicates a correct or incorrect response to the last trial. Feedback should be immediately recognizable but an otherwise subtle cue through an audial, haptic or visual signal.

The evaluator should take care to not mix sensory feedback and a retry method (see 5.6) where a retry can be initiated after sensory feedback. Use of both sensory feedback and retries is not allowed during an experiment unless the feedback mechanism disables the retry mechanism.

Subclause B.1.2

Add a new row at the end of Table B.1:

0,125	24	3	4

Subclause C.3

Add two new subclauses after C.3:

C.4 Image sequence search and cropping procedure

C.4.1 Image sequence processing

The process of determining the image sequences to be used in the experiment is performed by the evaluator. The process involves the steps from Figure C.1, which uses an image an example. The evaluator compresses the full sequence of images and inspects the results for artefacts.

The goal of this search is to identify regions within the test sequence that exhibit a bit-wise difference between the reference (original) and test image sequences and are thus potentially visually degraded. As part of the selection, the evaluator may review many sequences and select only a small fraction of the examples of artefacts for inclusion in the experiment to determine the point at which an image sequence may be visually impaired.

Once a sequence of artefacts is found, the evaluator crops a region centred about the artefact for inclusion in the test set. Typically testing should exhibit artefacts for a high percentage of the sequence duration, rather than only one frame, unless the objective of the testing is to identify whether an artefact type is visible even if the artefact is present for a short duration. The cropping of the region of interest will greatly reduce the time needed for observers to view the image sequence before making a response.

Guidance for content categories (see C.2), for session duration and for image selection (see C.3) is applicable to image sequence testing.

C.4.2 Stimulus orientation

Stimuli may be presented to the observer in either landscape or portrait orientation. The evaluator shall ensure the viewing position requirements for the observer are met, see subclause 5.4.3.

C.5 Image panning in a sequence

C.5.1 Image panning processing

C.5.1.1 Panning setup, direction and sequence length

An alternative image sequence preparation method analyses a single image by panning the image either horizontally, vertically or diagonally within the test crop one pixel shift at a time. This method tests a coding system for sensitivity to the start and end of coding blocks. Use the side-by-side image sequence procedure in Annex H for the subjective task.

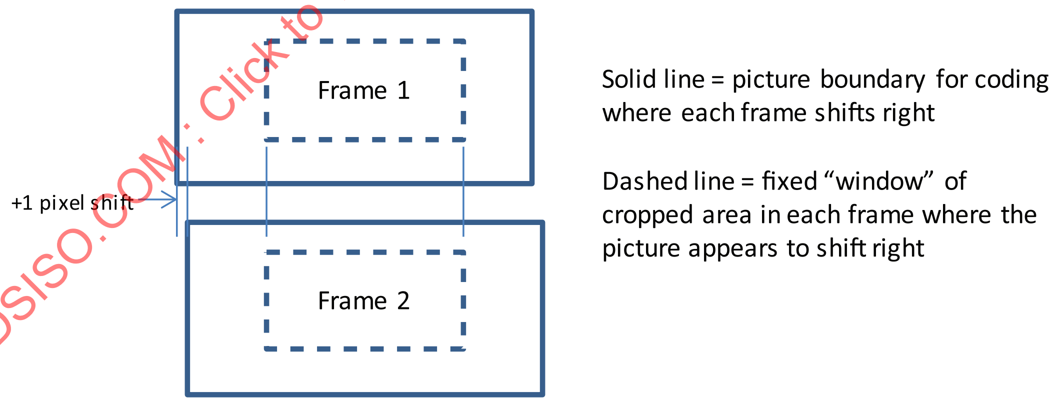

Figure C.3 — Example image panned horizontally

The evaluator prepares the image sequence as follows:

1) Select one image and a cropping area using the image processing techniques in subclause C.1.1 that will compose the first frame of the image sequence. Figure C.3 (top image) shows the crop area overlaying the coded and decoded image.

2) The crop area becomes a window at a fixed location on the display for viewing the reference and test image sequences as the image pans through the window for several frames.

3) The evaluator determines the direction, panning rate and number of panned frames.

   i) Figure C.3 shows an example panning horizontally and to the right. Panning may travel horizontally, vertically, or diagonally. The direction does not need to be the same for each trial in an experiment.

   ii) The evaluator selects a panning rate less than or equal to 30 Hz, otherwise motion silencing may render impairments unidentifiable. The panning rate equals the display frame rate divided by the number of still image frames shown per shift in the sequence:

$$panning\_rate = \frac{display\_frame\_rate}{still\_frames\_per\_shift}$$

See Table C.2 for examples of the panning rate versus display refresh rate.

Table C.2 — Panning rate versus display refresh rate

Panning rate (Hz)	Display refresh (Hz)	Show frame this number of times then pan
15	30	2
15	60	4
20	60	3
24	24	1
25	50	2
30	30	1
30	60	2

   iii) The number of images panned is k and shall be less than or equal to one half of the smallest dimension of the crop region determined in C.1 in order to keep the panning within a region of interest for the image under test. Usually panning includes no more than 90 shifted frames. Once the images have panned through the entire sequence, the frames plays in reverse to Frame 1. If the observer does not select a choice or has not reached the time limit, when frame 1 is shown, the panning plays again in the forward direction, and so on.

C.5.1.2 Panning with trimming images

4) A full-size reference image sequence of k images in length to be cropped for the subjective task in a later step is created by using the original image chosen in step 1 and trimming k pixels from the edge or edges of the original image at the end of the pan as shown in Figure C.4. The sub-steps (i) through (iv) create a series of trimmed reference images.

   i) The original image is h pixels horizontally and v pixels vertically.

   ii) Frame 1 is the original image, reduced by at least $(k - 1)$ in each direction of panning, $(h - k + 1)$ horizontally, if panning horizontally and $(v - k + 1)$ pixels vertically, if panning vertically. If panning diagonally, both dimensions are reduced.

   iii) Frame 2 trims the original image by one pixel on each image edge from where the panning moves and by $(k - 1)$ pixels from edges where the panning approaches to yield an image $(h - k + 1)$ horizontally if panning horizontally and at least $(v - k + 1)$ vertically, if panning vertically.

   iv) Frame 3 through Frame k repeats the process in list item 4)iii).

5) The created frames 1 through k form the reference image sequence prior to cropping for the subjective task.

6) Compress and then decompress frames 1 through k with the coding system under test to form the test image sequence frames 1′ through Frame k′.

7) Crop the reference frame sequence and the test frame sequence to the chosen size of the crop viewing window as described in C.1.1.

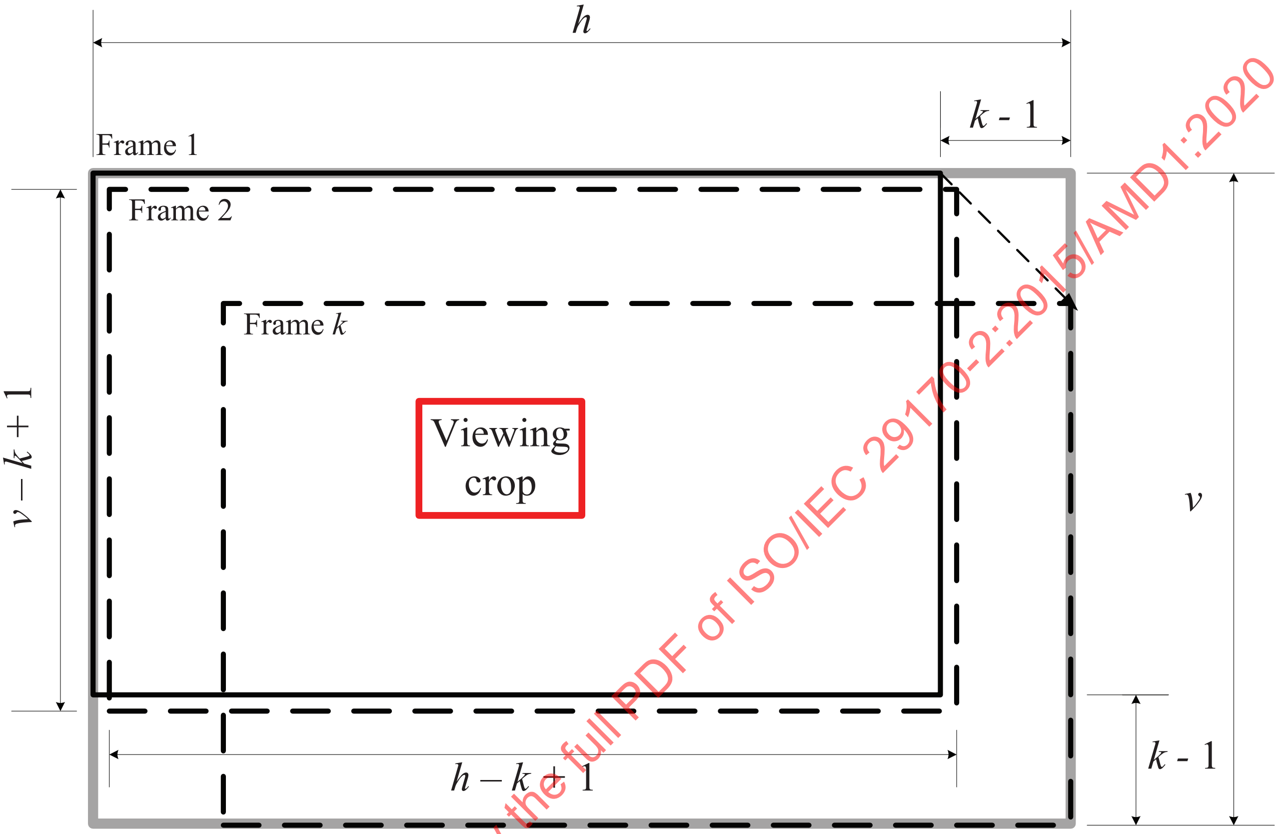

Figure C.4 — Image sequence k frames long showing a pan diagonal right and down

The above procedure has formed a cropped reference image sequence and a cropped test image sequence for the subjective task that will use the side-by-side image sequence procedure in Annex H.

C.5.1.3 Panning by circular shifting images

If the experimenter must use images but without trimming as performed in subclause C.5.1.2, an image sequence may be prepared by circular shifting the image horizontally, vertically or both, using the following steps and illustrated in Figure C.5:

1) Frame 1 is the original image selected in subclause C.5.1.1, step 1.

2) Frame 2 is a circular shift of image 1 by shifting one pixel horizontally, vertically or both at a frame repetition rate using a value from Table C.2. The pixel column (horizontal pan) on the far side of the image shift is moved to the opposite side of the image or the pixel row (vertical pan) on the far side of the image shift is moved to the opposite side of the image. If panning diagonally, both the column and row move are needed.

3) Repeat step 2) for the remainder of the sequence.

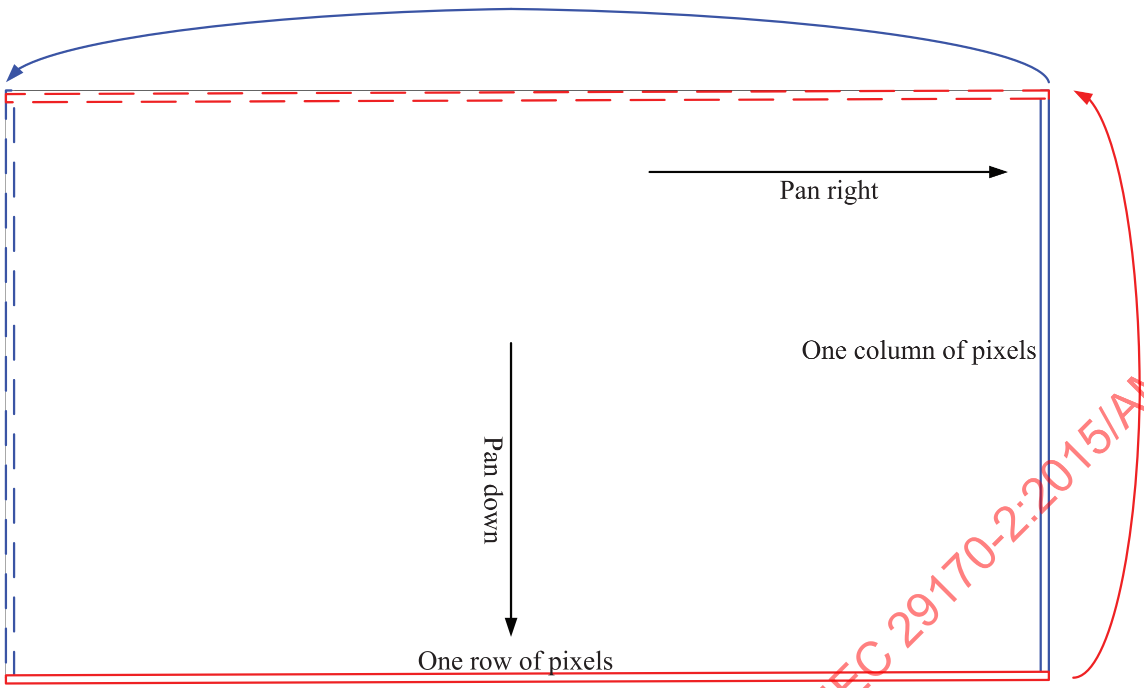

Figure C.5 — Panning image sequence preparation by circular shift

C.5.2 Subjective task guidance

This image processing induces flicker by highlighting position-dependent coding differences. Observers should be prompted to look for the presence or absence of flicker even though the presentation uses a side-by-side test protocol (see Annex H.3).

D.1.3

Replace the subclause heading with:

D.1.3 Aggregate statistics for a binary choice experiment

Add a new subclause after D.1.3:

D.1.4 Aggregate statistics for a ternary choice experiment

Repeat an analysis for each observer who participated in the study.

Upon completion of the analysis, there are a total of $B \times N$ response fractions corresponding to a total of

N test images that were used in the testing and B observers. The experimenter shall assign a response fraction for a “no difference” response equal to 0,5 rather than discard the “no difference” responses. For each of the N test images, calculate the mean and standard deviation (1σ) of the response fractions for all qualifying observers.

Record the minimum and maximum response fraction for the group of observers for each of the N test images.

Annex F

Add two new annexes after Annex F:

Annex G
(informative)

HDR image preparation

G.1 Colour space

The following guidance is offered for HDR image or image sequence preparation:

1. Content should be mastered to the same or lower maximum luminance of the display to be used in order to minimize effects and modifications of colours by display tone mapping functions.
2. Determine a suitable target average brightness range and stabilize salient image content to the target brightness range.
3. Document in the test report and map content into the HDR colour space to be testing with a coding system for the duration of any one test session. Known and useful colour spaces are provided in Table G.1.

Table G.1 — Colour space identification

Rec. ITU-R BT.709
IEC 61966-2-1
IEC 61966-2-2 (scRGB)
IEC 61966-2-4 (xvYCC)
Adobe RGB
SMPTE RP 431-2 (P3)
Rec. ITU-R BT.2020
Linear
User specified

G.2 Electro-optical transfer function

1) Ensure test contents are prepared for use with a display that uses a specific electro-optical transfer function. Known and useful EOTFs are shown in Table G.2.

Table G.2 — EOTF identification

Rec. ITU-R BT.1886
Rec. ITU-R BT.2100 PQ
Linear
User specified

2) Document the colour space of the contents and map decompressed content to the display colour space, if different than used for input to the coding system. Refer to Table G.1 for known colour spaces.

3) Document the EOTF in use and ensure contents uses an electro-optical transfer function compatible for the rendering display. Refer to Table G.2 to choose an EOTF for transmission to the display.

4) Ensure that average light level does not undergo dramatic change within a test sequence. Trial duration must be longer than display's temporal response window.

5) During between-trial intervals, use a screen with similar average luminance as test materials, recommended at <25 % grey level.

6) Characterize any local backlighting of the display to avoid or mask significant display non-uniformity, such as, halos.

7) Specify static metadata for each image in a test sequence to properly trigger display responsiveness in its HDR mode, if used.

G.3 Other HDR display characteristics

An alternative HDR processing signal chain following Recommendation ITR-R BT.2100 HLG may be used rather than the EOTF system noted in G.2. If such a system is to be tested, the experimenter should provide documentation of the appropriate signal processing for test contents in the test report.

G.4 HDR software processed workflow — option A

This workflow is available for testing when the experimenter either has little control over or cannot sufficiently disable HDR tone mapping and other hardware processing in the display unit that may interfere with codec evaluation.

The option A workflow bypasses HDR processing in the display and relies entirely on the experimenter's image or image sequence preparation prior to and after coding. Figure G.1 depicts the recommended workflow to process HDR images and image sequence for coding system evaluation.

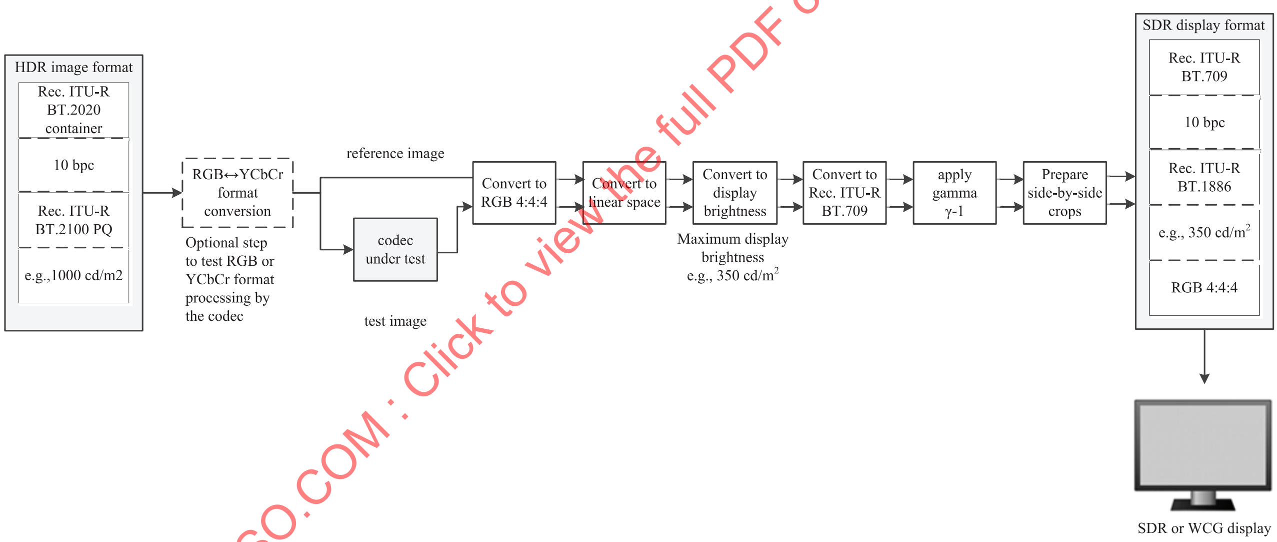

Figure G.1 — Display-independent HDR workflow

G.5 HDR display hardware processed workflow — option B

This workflow is available for testing when the experimenter has access to a display with transparent HDR processing, well-controlled and known with no tone mapping and other hardware processing in the display unit that may interfere with codec evaluation.

The option B workflow relies on the HDR processing in the display, correct EOTF application and relies entirely on the experimenter's image or image sequence with correct formatting that may include static metadata. Figure G.2 depicts the recommended workflow to process HDR images and image sequence for coding system evaluation.

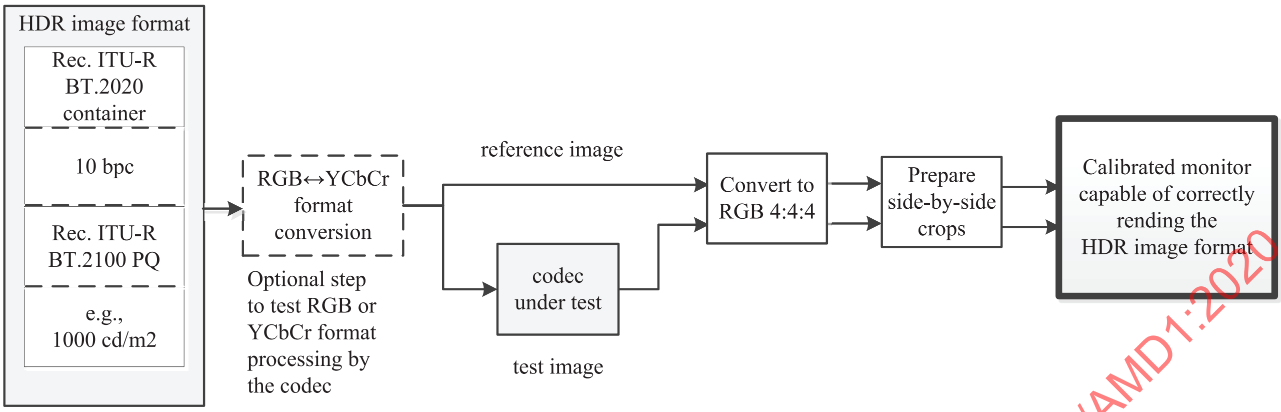

Figure G.2 — Production test monitor HDR workflow

Annex H
(normative)

Forced choice paradigm using side-by-side image sequence test protocol

H.1 Stimuli generation

H.1.1 On-screen presentation

Each test sequence shall be presented side-by-side with the reference sequence. These sequences shall be separated by a 1,0° ± 0,1° gap separating them based on the PPD. The side on which the test and reference sequences are presented shall be randomly assigned on each trial.

Sequences may be arranged in either portrait or landscape orientation.

All content shall be shown with 1:1 mapping to the display pixels without anti-aliasing.

Figure H.1 shows a display with locations where the reference image sequence and test image sequence may be placed. The observer prompt is below the displayed content.

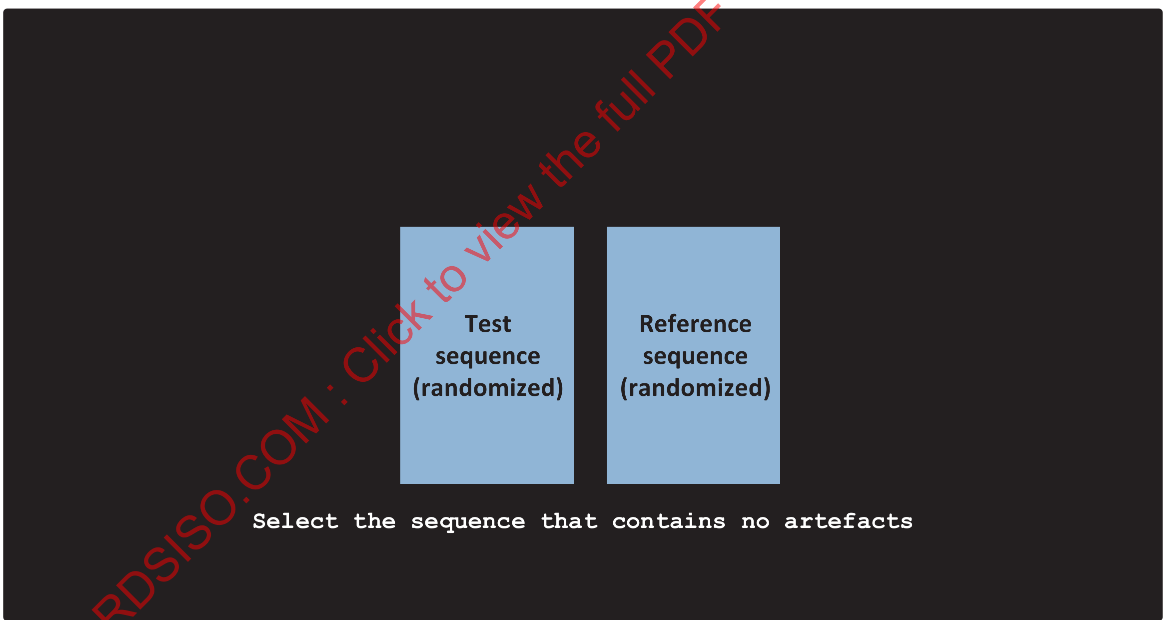

Figure H.1 — Example display showing sequence locations for the image sequence test protocol

H.1.2 Image trial ordering

Two consecutive trials shall present a different algorithm used to code the test sequence, even if the reference is the same.

During a session, the presentation order of the set of test sequences shall be randomized.