Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Sigmedia
Signal Processing for Media Applications

4K HDR video with AV1 : A Reality Check
FOSDEM 2023

Vibhoothi*, Francois Pitie*, Angeliki Katsenou*, *Anil Kokaram**,
* *Trinity College Dublin,*
February 2023
vibhoothi@tcd.ie

https://people.videolan.org/~mindfreeze/fosdem2023.pdf

VideoLAN
ORGANIZATION

Who am I?

Vibhoothi

- PhD Student and Research Assistant@Trinity College Dublin (TCD), circa ~2020.
    - *Research on optimising video codecs for streaming and internet use-cases*.
- Involved in Open-source multimedia, circa ~2018.
    - *VideoLAN Association*, Xiph. org Foundation, and Alliance for Open-media (AOM).


What are we going to do today?


Main motivation is to talk about the technical challenges for AV1 HDR playback !!

1. HDR = Brighter Pixels

SDR, 100-200 nits
HDR, up to 10,000 nits

Image is tonemapped for representation

Sequence: sparks_aom_11198-11570, Luminance measured with JETI Spectravel 1511 Spectroradiometer on Sony BVM-X300 (1000 nits)

2. HDR = More Bits

SDR = typically 8-10 bits

HDR = typically 10+ bits

Visual representation of dynamic range

3. HDR = Different Transfer Function

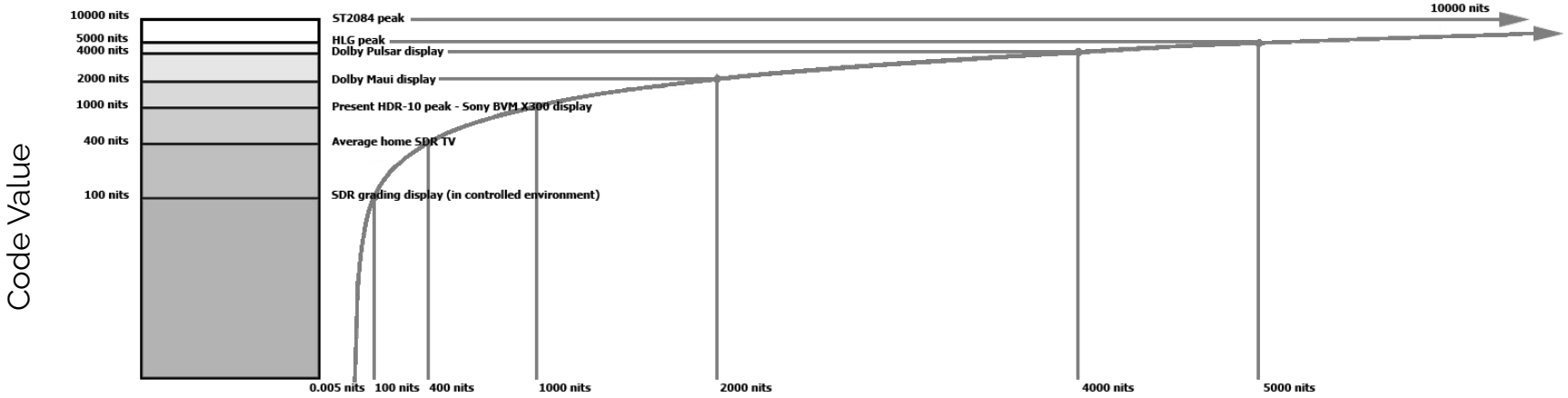
Now,

More Nits, More Bits…

Now,

Different mapping for Nits to Bits.

Black -> White, Modified Barten Model based on “banding” (“Perceptual Quantization” (PQ)),

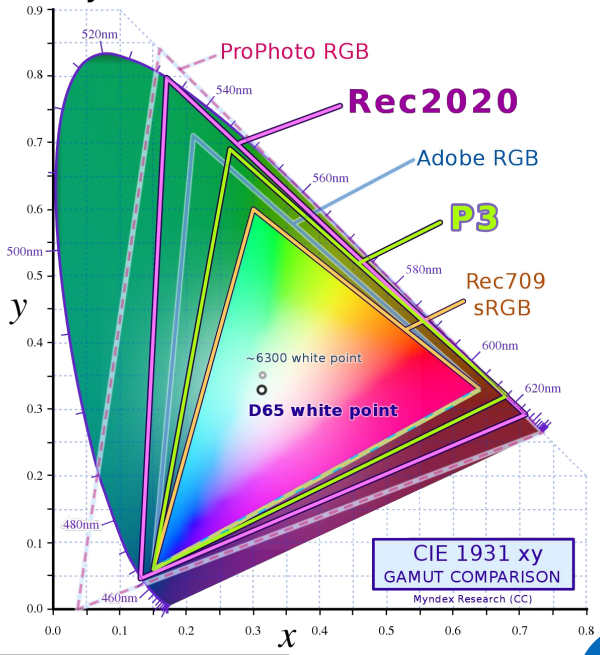


Source

4. HDR = Wider Gamut

The display technology have been improved since SDR standardisation (Rec 709). Rec.2020 proposed wider color gamut using primary colors closer to spectral locus.

BT.709 for SD TVs,
BT.2020 for UHD 4K, 8K HDR,
sRGB, *REDLog* for Post-production
CIE XYZ a perceptually uniform colour space



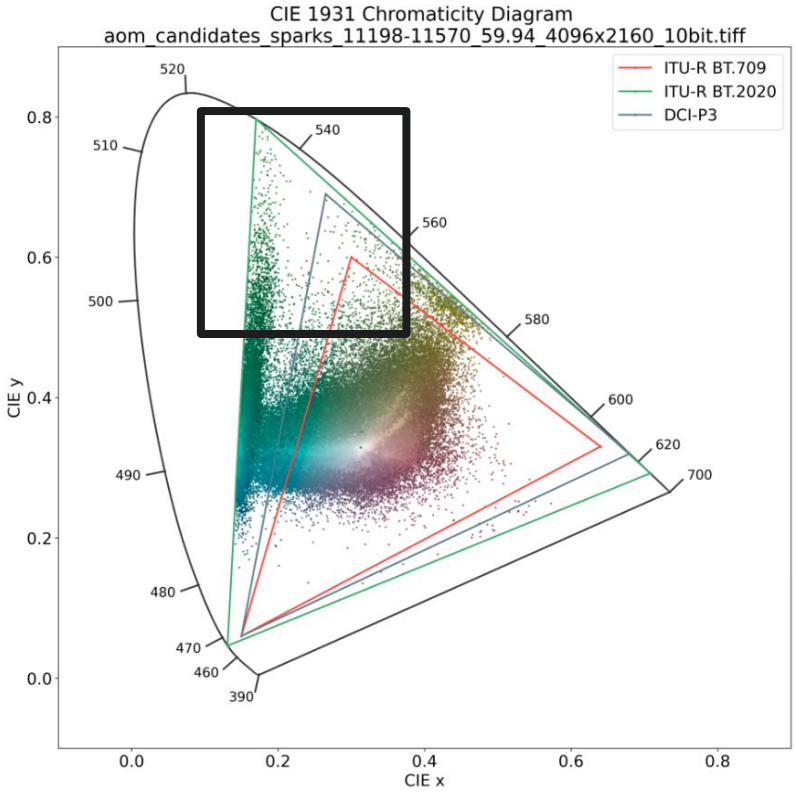
Wide Color Gamut (WCG.)

4. HDR = Wider Gamut

The displa…proved s… 709). Rec.2…gamut us… spectral locus.

BT.709 for SD TVs,
BT.2020 for UHD 4K, 8K HDR,
sRGB, *REDLog* for Post-production
CIE XYZ a perceptually uniform colour space

Reds *and* ***Greens****, have wider range,*
Blues*, do not change much.*



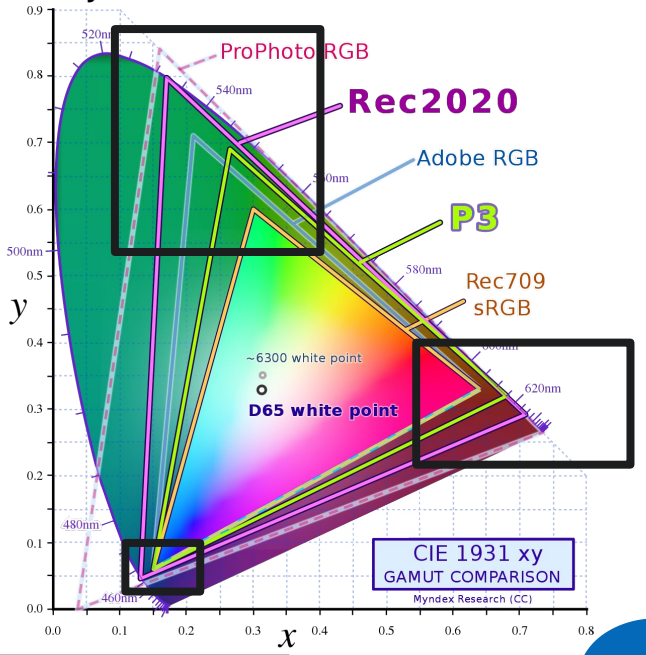
Wide … (Wcg)

4. HDR = Wider Gamut

The display technology have been improved since SDR standardisation (Rec 709). Rec.2020 proposed wider color gamut using primary colors closer to spectral locus.

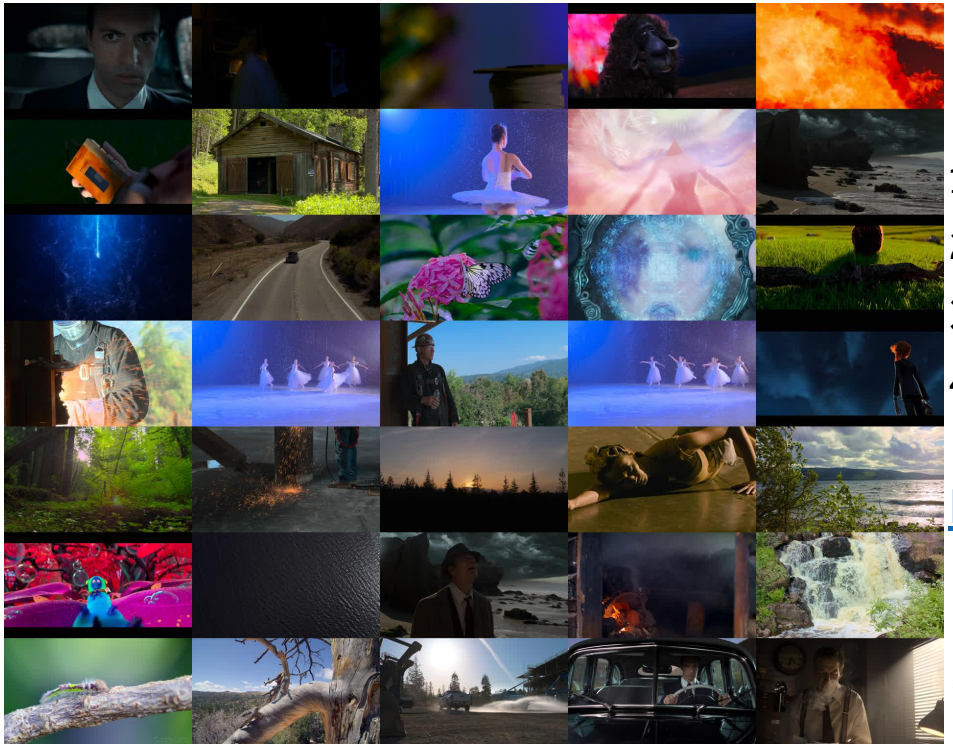
BT.709 for SD TVs,
BT.2020 for UHD 4K, 8K HDR,
sRGB, *REDLog* for Post-production
CIE XYZ a perceptually uniform colour space

Reds *and* ***Greens****, have wider range,*
Blues*, do not change much.*



Wide Color Gamut (WCG.)

Where to find HDR Sequences



1. *Netflix Open Content*
2. SVT Open Content
3. CableLabs 4k
4. Digital production from AWSF
   i) ASC StEM2 - Standard Evaluation Material 2

We currently used only (1, 2, 3).

Turns out it's a bit hard to get real HDR

Landscape of AV1 Playback and Decoding

Initially became popular and adopted with VideoLAN's dav1d, software decoder, with 100k+LoC of hand-written ASM which is fast.

Major browsers (except Safari*) support AV1 video playback,

Lately hardware decoders are on rise for AV1, from major vendors including latest Qualcomm, NVIDIA, AMD, Mediatek, Broadcom, Samsung etc.

*** Only Images (AVIF), latest TP 161 seems to be adding support(?)**

Landscape of AV1 Playback and Decoding

Initially became popular and adopted with VideoLAN's dav1d, software decoder, with 100k+LoC of hand-written ASM which is fast.

Major browsers (except Safari*) support AV1 video playback,

Lately hardware decoders are on rise for AV1, from major vendors including latest Qualcomm, NVIDIA, AMD, Mediatek, Broadcom, Samsung etc.

So what is the **problem**??

*** Only Images (AVIF), latest TP 161 seems to be adding support(?)**

Playback of AV1

Playback of AV1 streams with HDR signals natively is not working as expected

macOS,

Display and OS level support is available, but video playback capabilities is limited. Some players do *tonemapping*, others have ***limited support.***

Linux,

Wayland Protocols for signalling HDR is WIP, so at the moment ***No***

Windows,

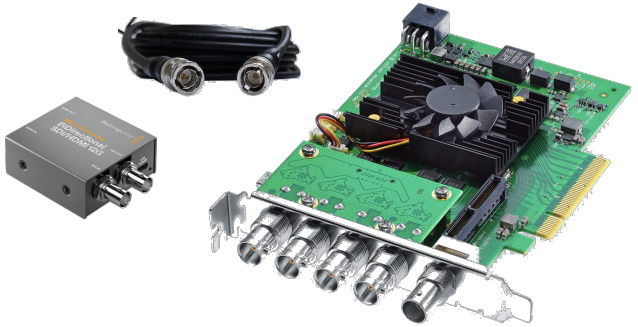
Display and OS level support available, video playback with DirectX is available, but playback is not the optimal due to ***display transition to HDR***.

Playback of AV1: A different approach

Playback cards, following paths of post-production industry

Blackmagic Decklink series, Decklink 8K Pro,

- Upto 8K60 DCI, 12bit 4:4:4, 12G SDI



FFmpeg and Gstreamer for driving playback

- Manual build with `decklink` support and blackmagic SDK

```
ffmpeg -max_frame_delay 200 -threads 5 -i $input -f decklink -pix_fmt v210 -r 60 -s 4096x2160
-an  'DeckLink 8K Pro (1)'
env GST_PLUGIN_SYSTEM_PATH='/usr/local/lib/' gst-play-1.0 --videosink="decklinkvideosink mode=
2160p60 sync=false" $input
```

Playback of AV1: Displays

Requirements for displays

- Need to display HDR content with **little-no changes**
- No TV level modifications of signals, ie. **no tone mapping**
- **Strictly calibrated** as per SMPTE 2084 with constant peak luminance (>= 1000 Nits).

Playback of AV1: Displays

Requirements for displays

- Need to display HDR content with **little-no changes** → ✅ Playback card
- No TV level modifications of signals, ie. **no tone mapping**
- **Strictly calibrated** as per SMPTE 2084 with constant peak luminance (>= 1000 Nits).

Playback of AV1: Displays

Requirements for displays

- Need to display HDR content with **little-no changes**
- No TV level modifications of signals, ie. **no tone mapping**
- **Strictly calibrated** as per SMPTE 2084 with constant peak luminance (>= 1000 Nits).

Reference monitor, Sony BVM-X300-V2 (32" OLED)

- Source of truth for the controlled HDR playback
- Force signal properties

Playback of AV1: Displays

Requirements for displays

- Need to display HDR content with **little-no changes**
- No TV level modifications of signals, ie. **no tone mapping**
- **Strictly calibrated** as per SMPTE 2084 with constant peak luminance (>= 1000 Nits).

Reference monitor, Sony BVM-X300-V2 (32" OLED)

- Source of truth for the controlled HDR playback
- Force signal properties

Once this link is established, can extend to play HDR videos on consumer TV

Playback of AV1: Scientific Testing

How to check if the HDR playback link respects REC.2100 standard?

Conform using multiple methods,

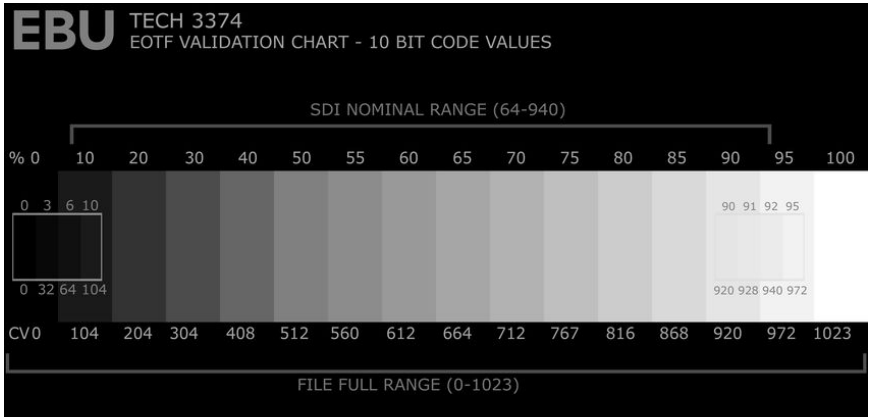
Playback of AV1: Scientific Testing

How to check if the HDR playback link respects REC.2100 standard?

Brightness

Conform using multiple methods,

- **Use a PQ EOTF Chart from EBU**



*This can help to **find the peak brightness**.*

Playback of AV1: Scientific Testing

How to check if the HDR playback link respects REC.2100 standard?

Brightness

Conform using multiple methods,

- **Use Test Patterns from EBU**

*This can help to **find maximum area** in your screen which can show **peak brightness.***

Playback of AV1: Scientific Testing

How to check if the HDR playback link respects REC.2100 standard?

Brightness

Signal

Conform using multiple methods,

- **Use HDR cross-converter monitor (Atomos Shogun 7) for pass-through of video to the TV**

*This can help to check **existence** of the **signal.***

Playback of AV1: Scientific Testing

How to check if the HDR playback link respects REC.2100 standard?

Brightness

Signal

COLOR

Conform using multiple methods,

- **Turn on Gamut Marker on Reference Monitors.**

*This can show pixels which is **beyond SDR colorspace** (BT.709) in Reference monitor.*

Playback of AV1: Scientific Testing

How to check if the HDR playback link respects REC.2100 standard?

Brightness

Signal

COLOR

Conform using multiple methods,

+ **Use a Spectroradiometer**

*This can help to **measure color volume** (Color-space, brightness) of patch in the screen.*

Playback of AV1: Scientific Testing

How to check if the HDR playback link respects REC.2100 standard?

Brightness

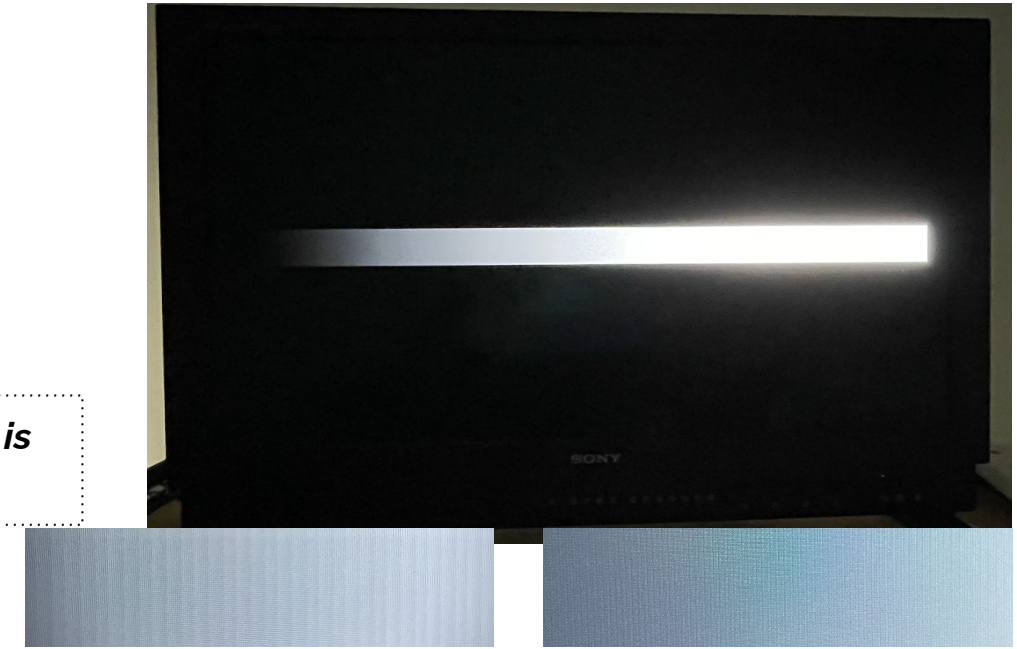
Signal

COLOR

BIT DEPTH

Conform using multiple methods,

+ **Use 10 bit gray ramp**



*This can help to validate if your **full pipeline is 10 bits** or any decimation happening.*

https://people.videolan.org/~mindfreeze/grayRamp.tiff
https://people.videolan.org/~mindfreeze/grayRampWithNoise.tiff

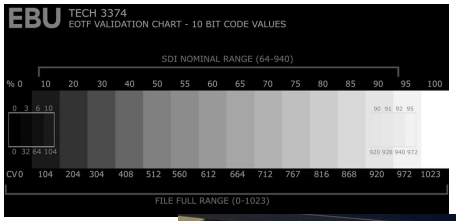
Playback of AV1: Scientific Testing

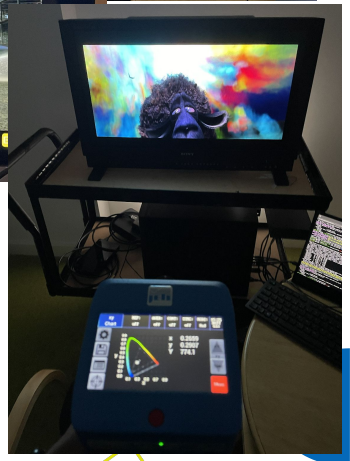
How to check if the HDR playback link respects REC.2100 standard?

Brightness
Signal
COLOR
BIT DEPTH

Conform using multiple methods,

- Use a PQ EOTF Chart from EBU
- Use Test Patterns from EBU
- Use HDR cross-converter monitor (Atomos Shogun 7) for pass-through of video to the TV
- Turn on Gamut Marker on Reference Monitors.
- Use a Spectroradiometer
- Use 10 bit gray ramp





Can we extend to Consumer TVs?

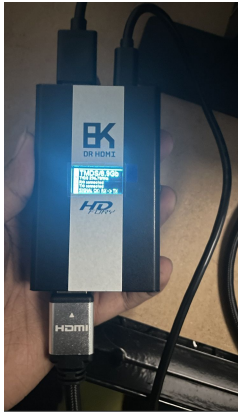
Sony OLED Critical Reference monitor

Sony 4K OLED Consumer TV (A80J)

Yes,

How?,

- Using SDI->HDMI converter to send HDR signals to TV
- Force the HDR metadata from the setting, set them correctly,
- Can use Dr. HDMI to signal Metadata



Setting up scientific testing environment

Conforming video playback is not only enough for HDR videos,

The **viewing environment** has a big impact on perception of colors

Setting up scientific testing environment

Conforming video playback is not only enough for HDR videos,

The **viewing environment** has a big impact on perception of colors

Depends on

- *Display panel technology* (Peak luminance, and color temperature)
- ***Surrounding lights*** and light reflection from screen
- *Video materials*
- *Perception* of compression *artifacts*

Setting up scientific testing environment

Conforming video playback is not only enough for HDR videos,

The **viewing environment** has a big impact on perception of colors

Depends on

- *Display panel technology* (Peak luminance, and color temperature)
- ***Surrounding lights*** and light reflection from screen
- *Video materials*
- *Perception* of compression *artifacts*

Based on the viewing environment, individuals can experience ***fatigue and dizziness*** on prolonged viewing.

Setting up scientific testing environment

Conforming video playback is not only enough for HDR videos,

The **viewing environment** has a big impact on perception of colors

Depends on

- *Display panel technology* (Peak luminance, and color temperature)
- ***Surrounding lights*** and light reflection from screen
- *Video materials*
- *Perception* of compression *artifacts*

Based on the viewing environment, individuals can experience ***fatigue and dizziness*** on prolonged viewing.

ITU BT.500-14: Methodologies for the subjective assessment of the quality of television images, explains the testing environment

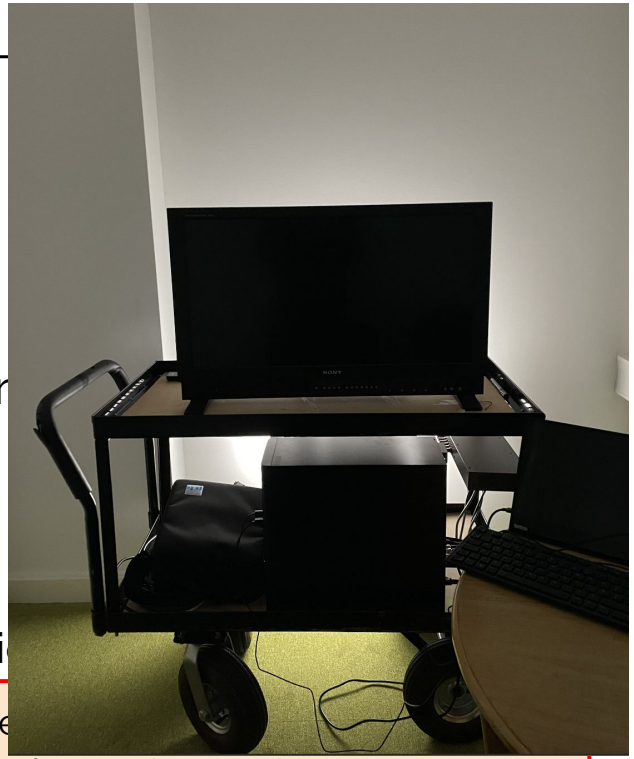
Setting up scientific testing environment

Conforming video playback is not only enough for H

The **viewing environment** has a big impact on

Depends on

- *Display panel technology* (Peak luminance, and
- ***Surrounding lights*** and light reflection from scr
- *Video materials*
- *Perception* of compression *artifacts*



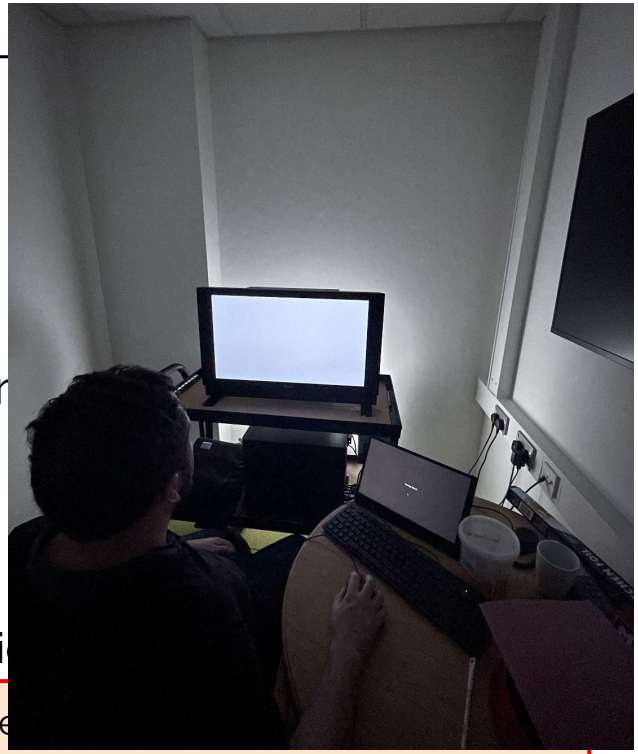
Based on the viewing environment, individuals can experi
prolonged viewing.

ITU BT.500-14: Methodologies for the
quality of television images, explains the testing environment

Setting up scientific testing environment

Conforming video playback is not only enough for H

The **viewing environment** has a big impact on



Depends on

- *Display panel technology* (Peak luminance, and
- ***Surrounding lights*** and light reflection from scr
- *Video materials*
- *Perception* of compression *artifacts*

Based on the viewing environment, individuals can experi
prolonged viewing.

ITU BT.500-14: Methodologies for the
quality of television images, explains the testing environment

Closing Remarks

- Signalling Metadata is **secondary** aspect of HDR, primarily it is,
  - Wide range of **brightness** due to different quantization scheme (**PQ**).
  - **WCG** can enhance viewing experience with more **colors**.

- Setting up playback pipeline of HDR content of AV1 is **non-trivial** accompanied by **high costs** despite the fact HDR was standardized in **2012** and AV1 in **2018.**

- Subjective evaluation of HDR videos is significantly influenced by the **viewing environment**.

References

[1]: `ffmpeg -i input.y4m -vf zscale=tin=smpte2084:min=bt2020nc:pin=bt2020:rin=tv:t=smpte2084:m=bt2020nc:p=bt2020:r=tv,zscale=t=linear:npl=100, format=gbrpf32le,zscale=p=bt709,tonemap=tonemap=hable:desat=0,zscale=t=bt709:m=bt709:r=tv,format=yuv420p image.png`

[2]: https://gitlab.freedesktop.org/pq/color-and-hdr

[3]:High Dynamic Range Video, from Acquisition to Display and Applications https://www.sciencedirect.com/book/9780081004128/high-dynamic-range-video

[4]: https://github.com/Netflix/sureal

[5]; https://aomedia.org/docs/CWG-B075o_AV2_CTC_v2.pdf, AOM-CTC

[6]: https://2019.acmmmsys.org/program/slides/walt-husak.pdf, 2019 HDR presentation from Dolby

[7]: "Direct optimisation of λ for HDR content adaptive transcoding in AV1." In. SPIE, 2022.

[8]: https://www.lightillusion.com/guides.html,

[9]: https://www.colour-science.org/,

Thanks to Sigmedia.tv, AOMedia, YouTube Media & Algorithms Team, and other Open-Source members for helping and supporting the Research and Development.

THE END

For questions, please email to vibhoothi@tcd.ie/anil.kokaram@tcd.ie.

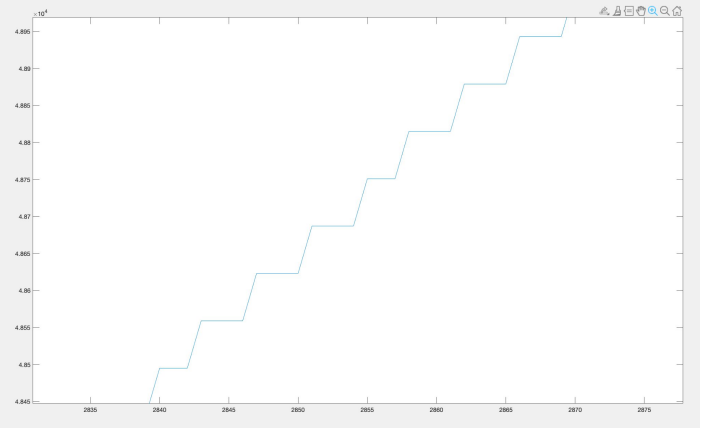
This project is funded by Enterprise Ireland under Disruptive Technology Innovation Fund (DTIF.), ADAPT-SFI Science Research Center, Ireland. Grant No DT-2019-0068.

Special mention to John Squires from TCD, and other various FFmpeg devs;)

Additional Resources

GrayBar.m

```
rows = 2160;
cols = 3840; % 4K Resolution
grayRamp = uint16(linspace(0,65535,1024)); %1024 Bars
grayRamp = repmat(grayRamp, [rows, 1]);     %Make it 2160 Wide
grayRamp = imresize(grayRamp,[2160 3840],"nearest"); % Reshape
to approximately 3840 using NearestNeighbours for step size
grayRamp = uint16(round(256*randn(size(grayRamp))) +
double(grayRamp)); % Do this only for adding noise
grayRamp(1 : 1000, :) = 0; % Remove first 1000 pixels from top
grayRamp(1200 : end , :) = 0; % Remove pixels from 1200,
making only 200 pixels with data
imshow(grayRamp); % Show image
imwrite(grayRamp,"grayRampWithNoise.tiff"); % Save as
grayRampWithNoise
```



Encoding Process

Encoding process is **same** as any other videos in the **current** implementation, only difference is, for correct *playback*, we have to signal,

- *Color primaries*: BT.2020 (Color space)
- *Transfer characteristics*: SMTPE2084 (PQ)
- *Matrix coefficients*: BT.2020NCL (Non-constant Luminance)
- *Chroma sample position*: Colocated

Libaom have *optional* flag (disabled by default) which can assign *different chroma quantizers* when these options are signalled.

Setting up scientific testing environment

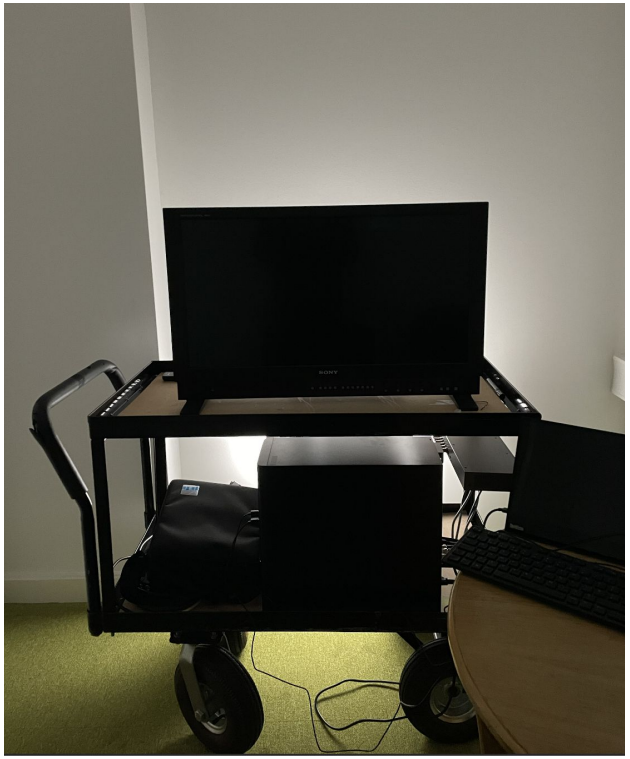
The current ITU recommendations is focused for SDR videos viewing conditions,

- Adopted the “Laboratory environment” condition,
- Grey picture brightness chosen carefully based on the viewing environment and video sequences,
  - Grey image of brightness of **14.9** nits (cd/m^2) for reducing viewing discomfort (***#555555***)
  - Background luminance is a controlled studio light to be **2.62** nits (cd/m^2).

Setting up scientific testing environment

The current ITU recommendations is focused for SDR videos viewing conditions,

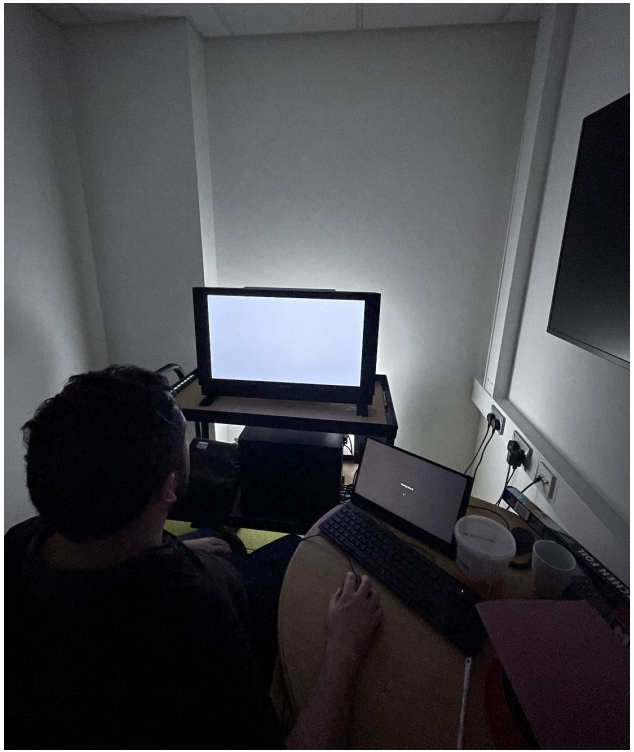
- Adopted the “Laboratory environment” condition,
- Grey picture brightness chosen carefully based on the viewing environment and video sequences,
  - Grey image of brightness of **14.9** nits (cd/m^2) for reducing viewing discomfort (***#555555***)
  - Background luminance is a controlled studio light to be **2.62** nits (cd/m^2).



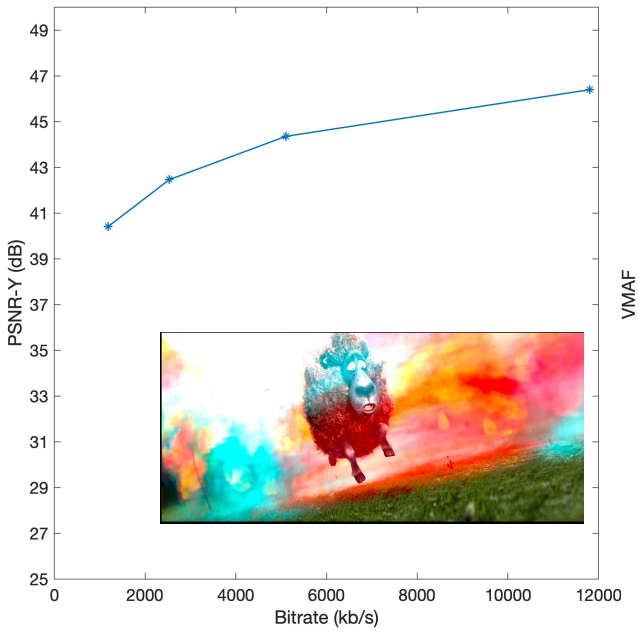
Setting up scientific testing environment

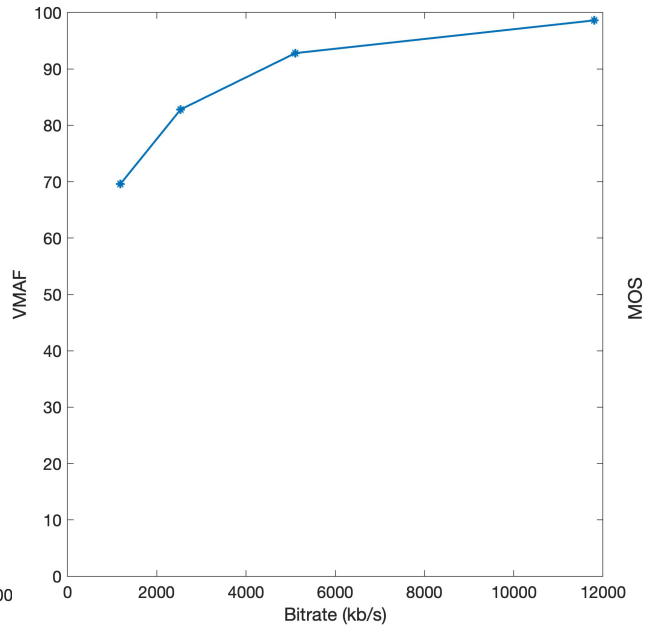
The current ITU recommendations is focused for SDR videos viewing conditions,

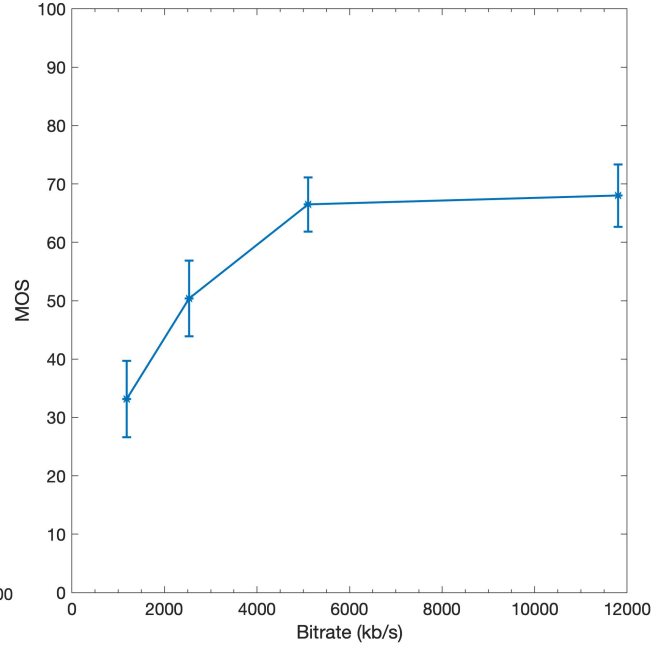
- Adopted the “Laboratory environment” condition,
- Grey picture brightness chosen carefully based on the viewing environment and video sequences,
  - Grey image of brightness of **14.9** nits (cd/m^2) for reducing viewing discomfort (***#555555***)
  - Background luminance is a controlled studio light to be **2.62** nits (cd/m^2).



A quick glance over mean opinion scores





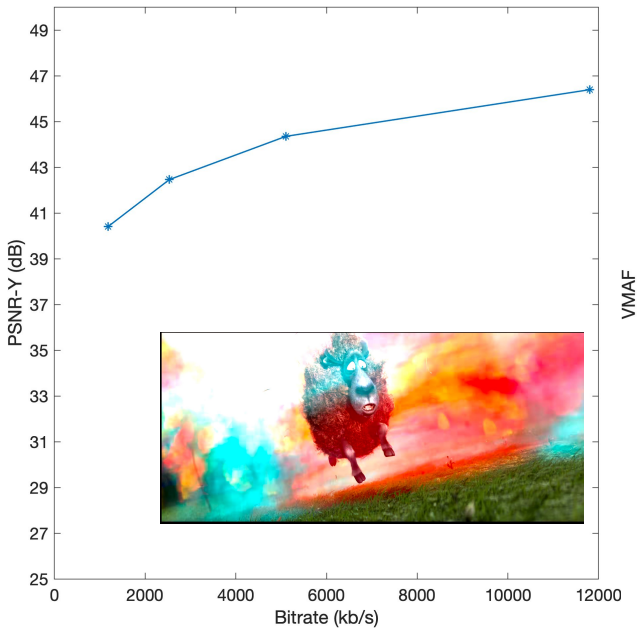


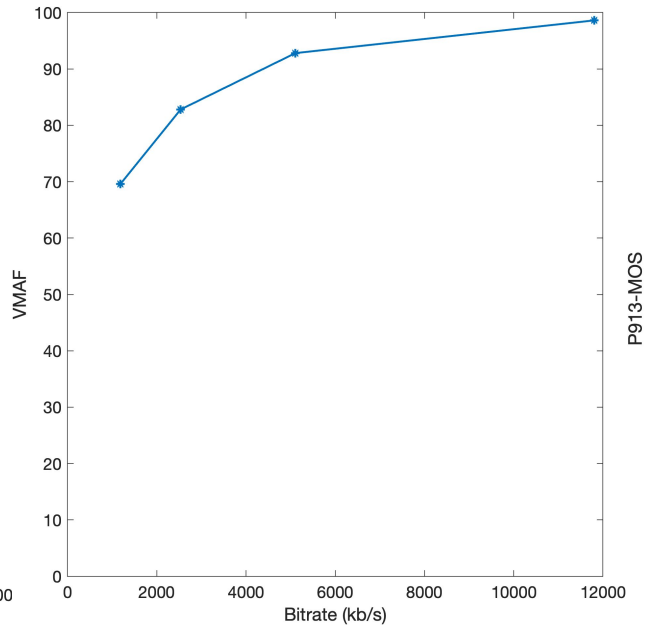
Comparison of PSNR, VMAF, and MOS score,

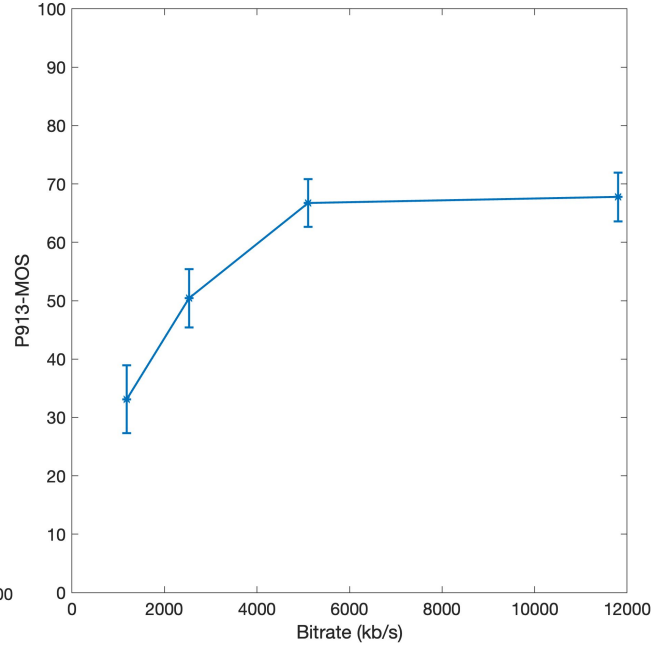
Comparison of MOS score with Objective metrics from 42 Subjects with Age 22-55, (30M, 12F) with 11 Experts and 31 Non-Experts.

CosmosSheep Encoded with libaom-av1@cpu-used 2 with 4 QPs 27, 39, 49, 59, MOS recovery carried out with Netflix Surreal Toolkit; Subjective study carried out with DSCQS protocol@TCD.

A quick glance over mean opinion scores



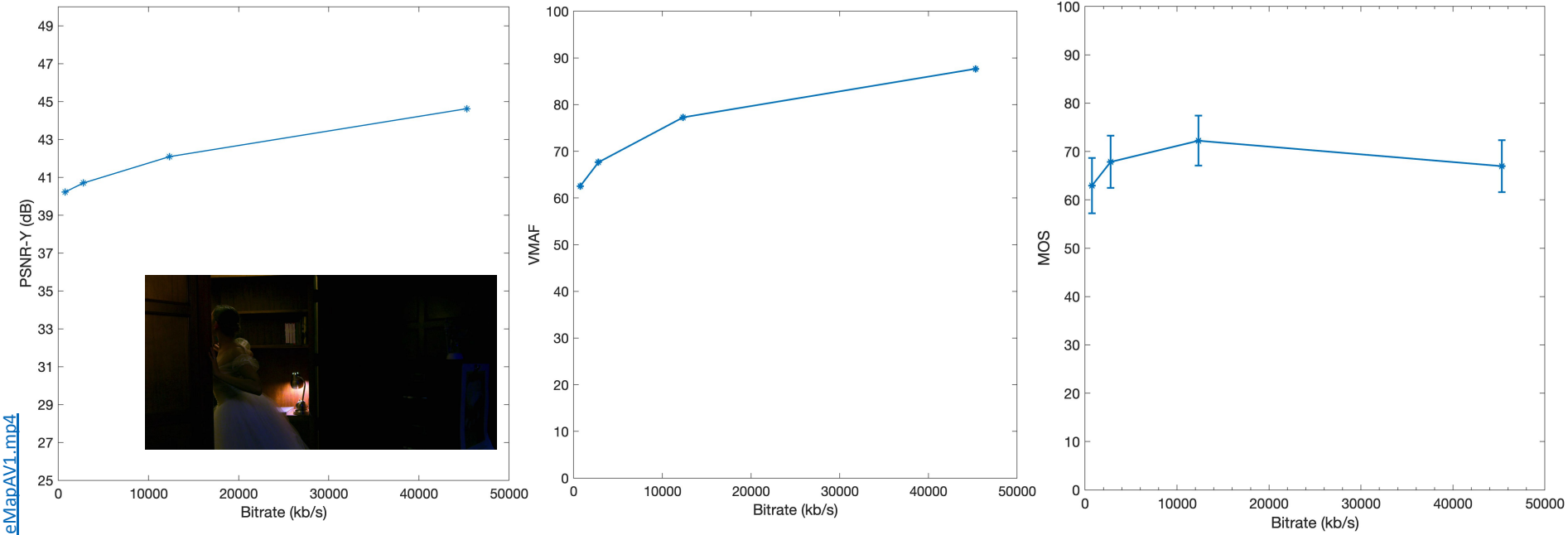




Comparison of PSNR, VMAF, and MOS score after ITU-P.913 Score recovery.,

Comparison of MOS score with Objective metrics from 42 Subjects with Age 22-55, (30M, 12F) with 11 Experts and 31 Non-Experts.

A quick glance over mean opinion scores



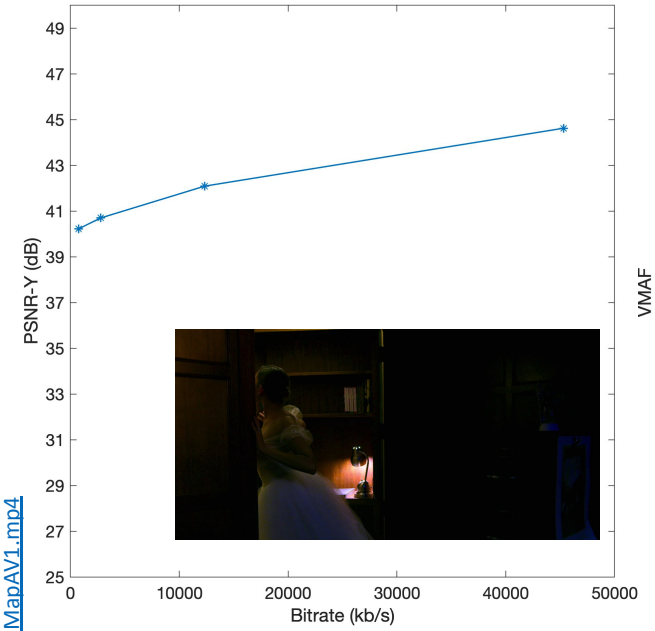
Comparison of PSNR, VMAF, and MOS score,

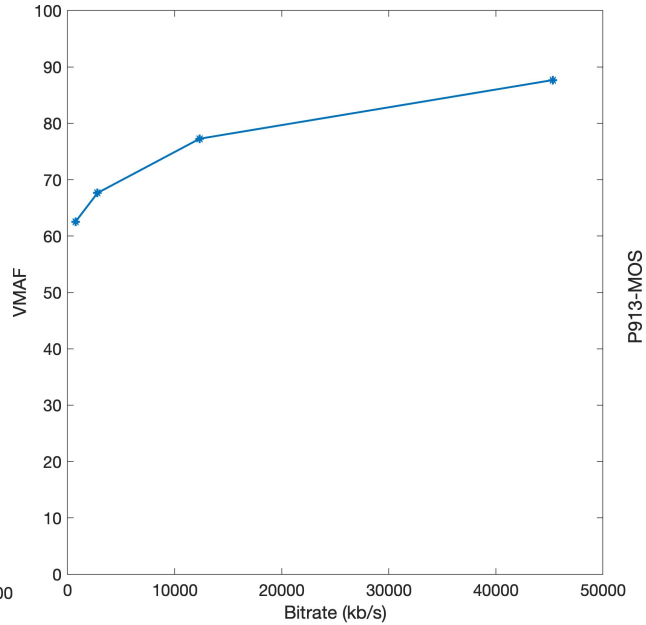
- Do not expect MOS to be "**monotonic**" like PSNR, VMAF (Objective Metrics), as perception of quality across people varies
  - You may not be able to able to compare like objective metrics using BD-rate (%)

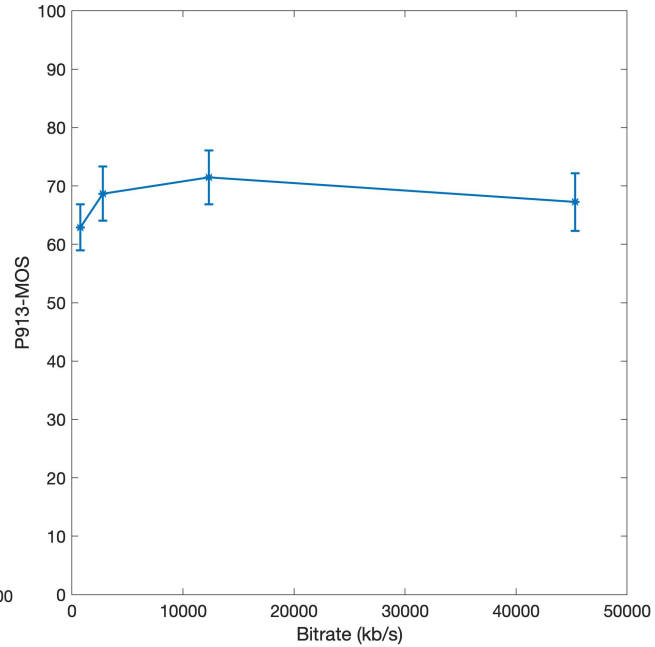
NocturneRoomToneMapAV1.mp4

NocturneRoom Encoded with libaom-av1@cpu-used 2 with 4 QPs 27, 39, 49, 59, MOS recovery carried out with Netflix Surreal Toolkit; Subjective study carried out with DSCQS protocol@TCD.

A quick glance over mean opinion scores







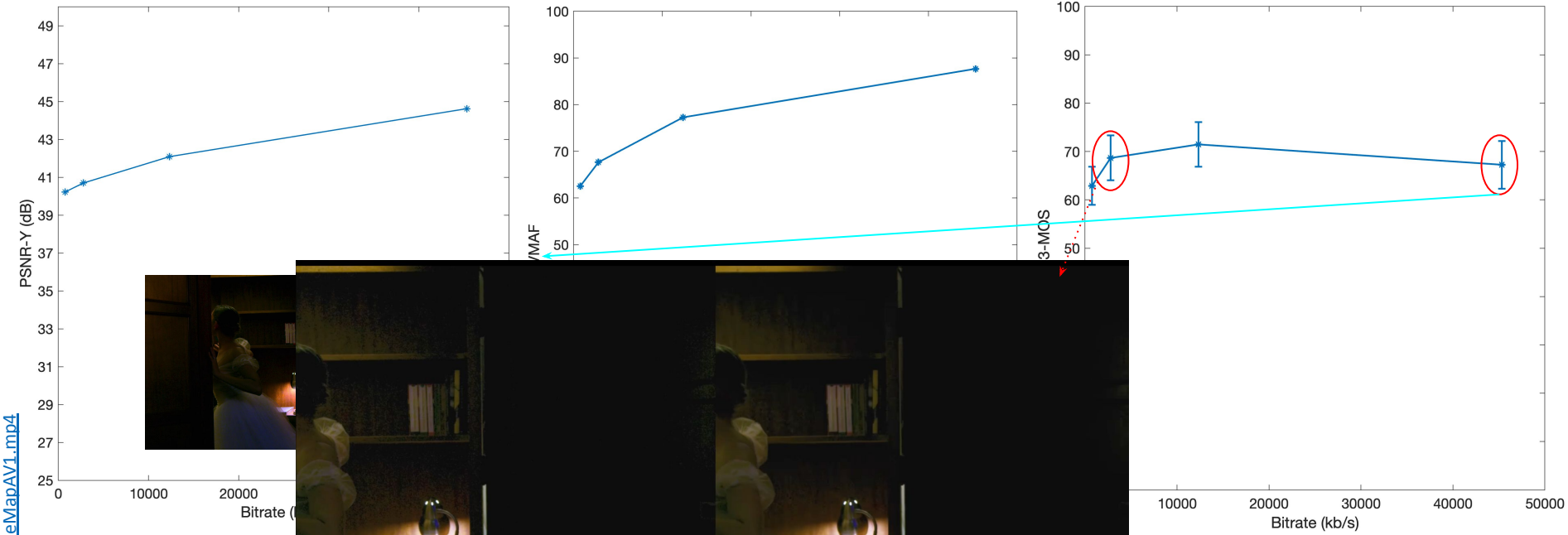
Comparison of PSNR, VMAF, and MOS score after ITU-P913 recovery

- Do not expect MOS to be "**monotonic**" like PSNR, VMAF (Objective Metrics), as perception of quality across people varies
  - You may not be able to able to compare like objective metrics using BD-rate (%)

NocturneRoomToneMapAV1.mp4

NocturneRoom Encoded with libaom-av1@cpu-used 2 with 4 QPs 27, 39, 49, 59, MOS recovery carried out with Netflix Surreal Toolkit; Subjective study carried out with DSCQS protocol@TCD.

A quick glance over mean opinion scores

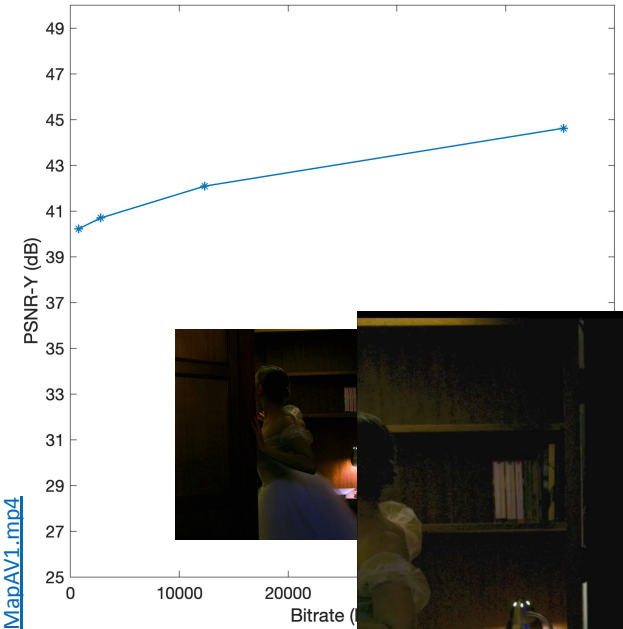


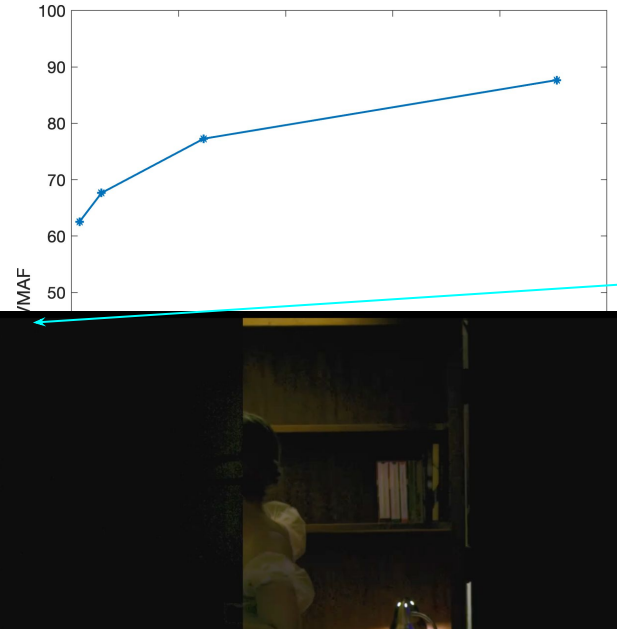
- Do not exp… …etrics), as perception of quality acr…
  - You n… …sing BD-rate (%)

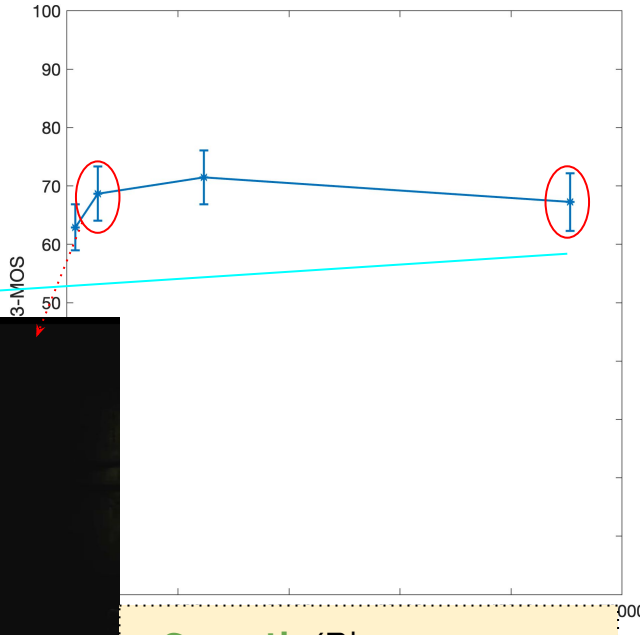
NocturneRoomToneMapAV1.mp4

NocturneRoom Encoded with libaom-av1@cpu-used 2 with 4 QPs 27, 39, 49, 59, MOS recovery carried out with Netflix Surreal Toolkit; Subjective study carried out with DSCQS protocol@TCD.

A quick glance over mean opinion scores







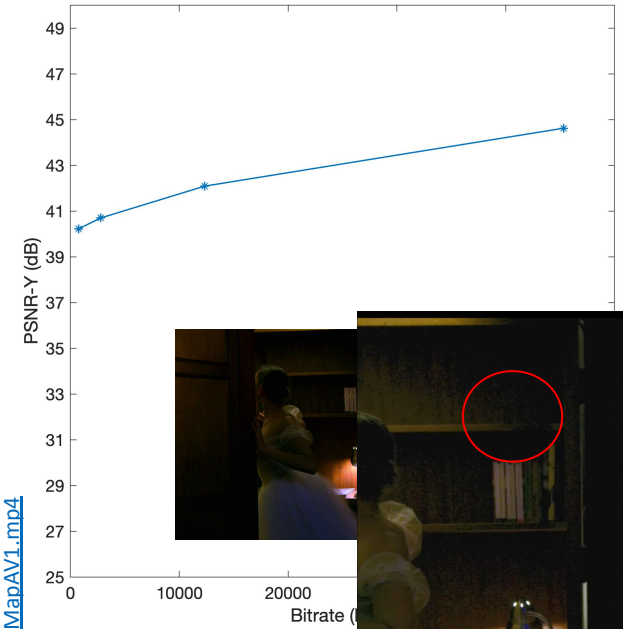
+ Do not exp
quality acr
  + You n

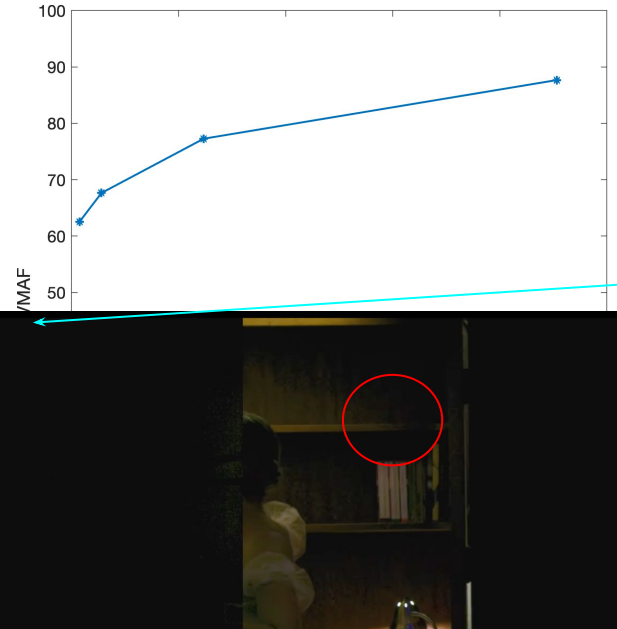
+ **Smooth**/Blurry

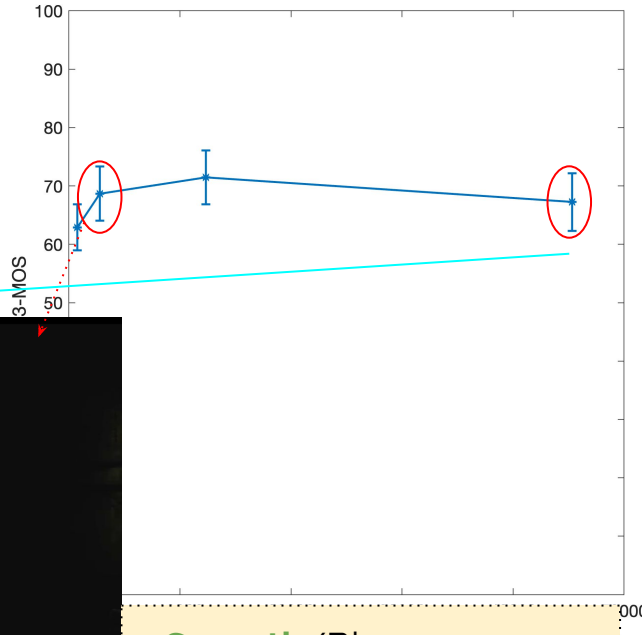
NocturneRoomToneMapAV1.mp4

NocturneRoom Encoded with libaom-av1@cpu-used 2 with 4 QPs 27, 39, 49, 59, MOS recovery carried out with Netflix Surreal Toolkit; Subjective study carried out with DSCQS protocol@TCD.

A quick glance over mean opinion scores







- Do not exp
  quality acr
  - You n

+ **Smooth**/Blurry
+ **Different** noise pattern, perceptually same for subjects

NocturneRoomToneMapAV1.mp4

NocturneRoom Encoded with libaom-av1@cpu-used 2 with 4 QPs 27, 39, 49, 59, MOS recovery carried out with Netflix Surreal Toolkit; Subjective study carried out with DSCQS protocol@TCD.

FIN

For Questions, reach out to vibhoothi@tcd.ie/anil.kokaram@tcd.ie