

Studio requirements for next-generation consumer distribution video codecs (including “HExt”: HEVC HDR Extensions)

[M36276](#)
[JCTVC-U0046](#)

Summary

Deliver to end consumers:

- Full chrominance spatial resolution (4:4:4)
- Full gamut of human vision (beyond BT.2020)
- Sufficient bit-depth to render content transparently
 - Barten JND requires ~4000 distinct, perceptually uniform stepped sample code levels in smooth picture areas
 - **Main444-12bit PQ should be in HDR CfE / CfP with professional quality pre-processing and post-processing steps**
- A profile developed concurrently for high tier consumer distribution, rather lowest common denominator or single median spanning tiers. (home theater is in different tier than mobile, small screens)

Background

- MovieLabs is a joint venture of the six major Hollywood studios.
- next-gen requirements delivered at the July 2013 Vienna MPEG & VCEG meeting
- ST 2084 (PQ), ST 2085 (YDzDx), ST 2086 (Mastering Display Color Volume) now completed, and carried in HEVC metadata.
 - Content Light Level SEI added to HEVC v3 draft

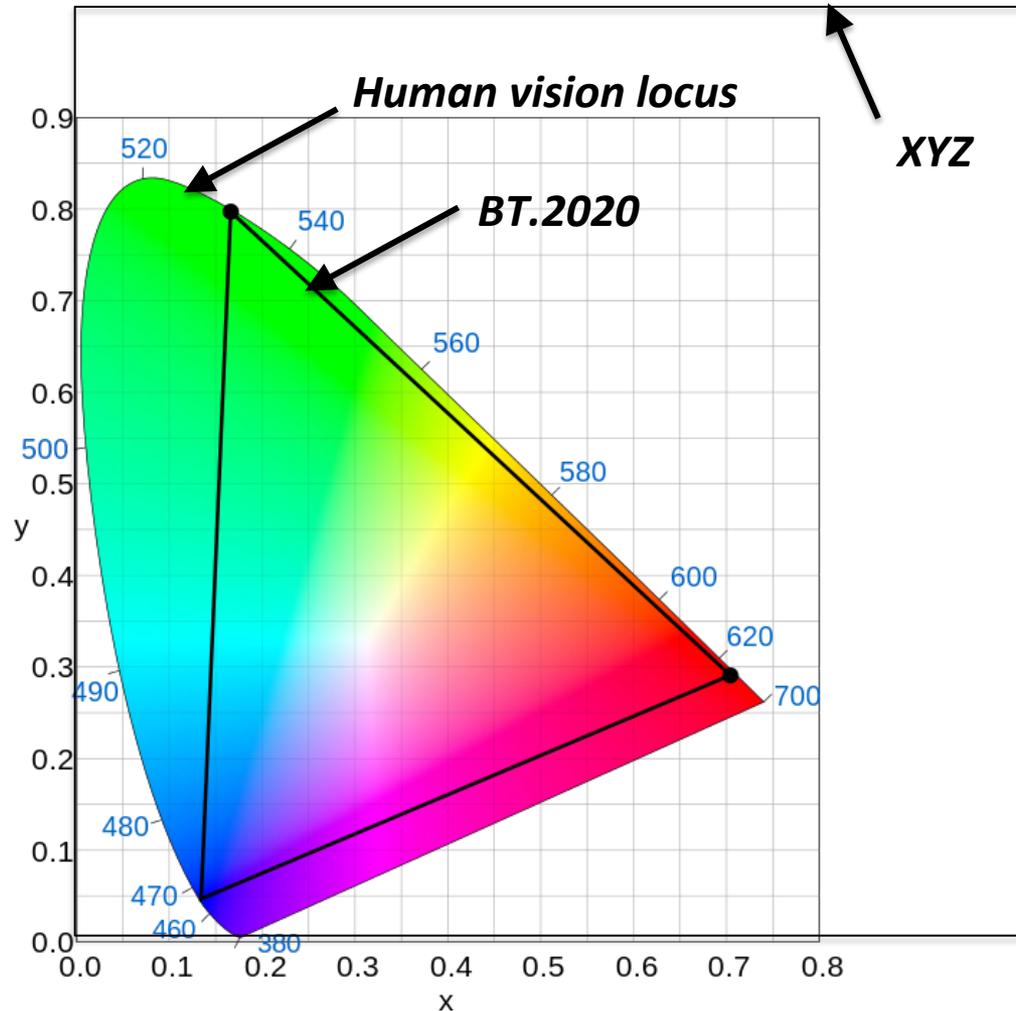
Current generation HDR/WCG

- Requires dithering and anti-contouring in display device to remove appearance of banding in 10-bit HDR video.
- HDR-10, coded with single layer HEVC Main10 as described in UltraHD Blu-Ray VUI parameters:
<http://www.blu-raydisc.com/en/Technical/TechnicalWhitePapers/>
- Parameters common to UHDA, UHDF, some OTT

Full chroma format (4:4:4)

- Down-sampling and up-sampling for 4:2:0 chooses code levels that were not in the master
 - Simple bi-linear example: $(a+b)/2 \neq a$ or b
- Chroma up-sampling (4:2:0 \rightarrow 4:4:4) may not sufficiently match down-sampling (4:4:4 \rightarrow 4:2:0)
- Current recommended practice: deliver 4:2:0 over HDMI, let HDR television perform better color up-sampling and conversion to RGB than set-top box.
- For 4:4:4: Cross component prediction tools, for example, permit sensitive picture areas (blocks) to be coded directly in primaries (RGB or XYZ), while more common blocks code color difference.

Full color gamut

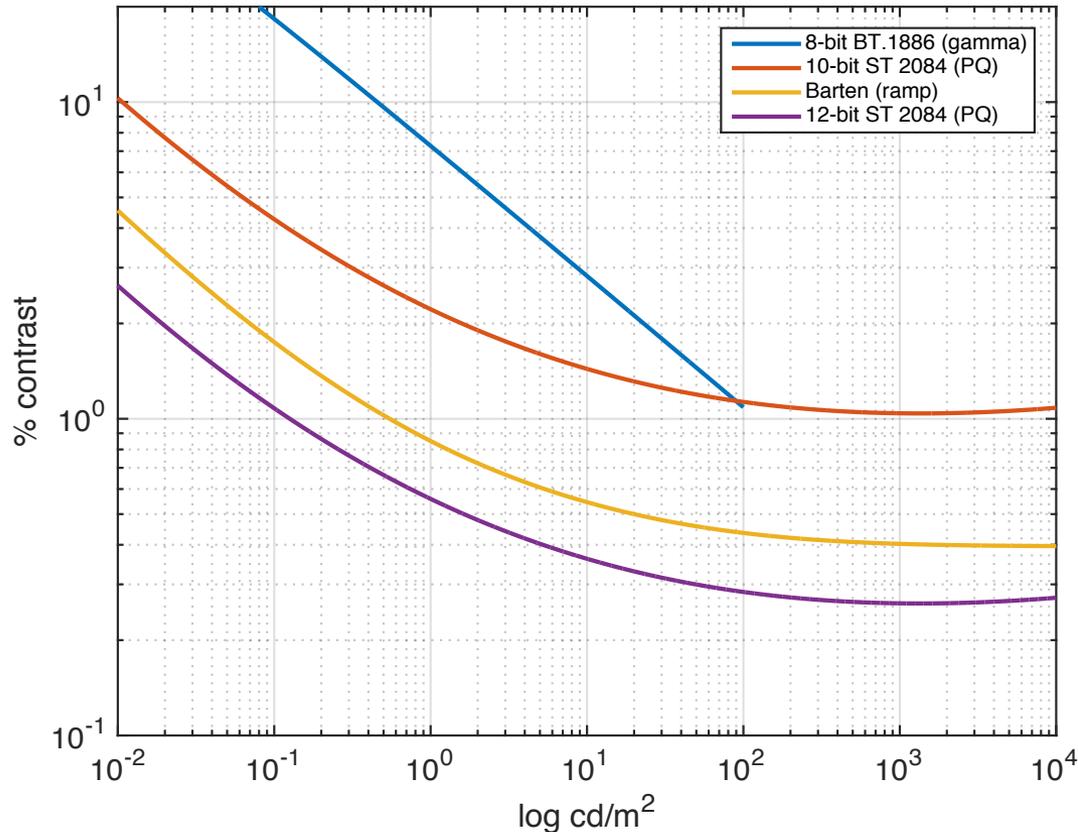


Reduce number of masters

Future proof:
create now, render
or display (possibly
without tone map)
later.

CGI can go beyond
current camera
spectra today

How many unique code levels needed?



Barten model predicts ~4000 unique code levels to transparently deliver content, but more codes may be necessary to provide system and codec margin

12-bit min. codec
input / output

Possible 14-bit IBDI
(but be prepared for
16-bit I/O samples)

Linear light equivalent
requires at least 16-bit
half-float

Other potential requirements on the horizon

Engineering buzz for:

- Decoder-side Output Device Transform (ODT)?
- Coding of uniform (linear) intensity stepped samples?
- Multiple frame rate deliveries? (with different shutter speeds / motion blur at each rate)
 - .. or: *universal* 120+ fps delivery to consumer?
 - From there, “easy” frame rate conversion or rendering to 60, 50, 30 Hz, etc.
- High-frame rate (HFR) beyond 120 fps ?
- Variable frame-rate (VFR) ?
 - some 120 fps camera captured scenes coded at 120 fps with one shutter / motion blur, while other scenes conveyed at 24 fps with usual 24 Hz motion blur... all within same bitstream / CLVS, rather than assume hierarchy of nested, fixed-rate temporal scalable layers.
 - Instantaneous frame rate changes easy to signal in codec, but requires changes to wire protocol timing models (HDMI)
 - All current systems set HRD fixed_pic_rate_flag today
- convey to real observers the colors as captured and mastered.

It may be 5+ years before these advanced requirements reach consensus.

References

- MovieLabs Next Generation Video (NGV)
 - <http://www.movielabs.com/ngvideo>
- ACES high precision workflow:
 - <http://www.oscars.org/science-technology/sci-tech-projects/aces>
- HDR Intermediates:
 - <https://www.smpte.org/webcasts/HighDynamicRange>