Keeping Video Quality Pristine throughout the Production Process: 4:2:2 10-bit AVC/H.264 encoding

Pierre Larbier
ATEME (Burbank, USA & Bievres, France)

Agenda

- Current States
- Talk about the **AVC/H.264 4:2:2 10-bit profile**
- Empirical comparisons against 4:2:0 8-bit & 4:2:2 8-bit
- Multi-Generation Encoding
- Production Cases

Production & Transmission Processes

- Production uses 4:2:2 10-bit over SDI
  - Used in video editing, tape decks, file storage, etc
  - Video needs to be transmitted by internal system or across networks
- Today's transmission mostly uses
  - 4:2:0 8-bit (MPEG-2 or AVC/H.264) for consumer applications
  - 4:2:2 8-bit(MPEG-2) for contribution applications
    - 20-60 Mbps with latency from 1 sec-250ms

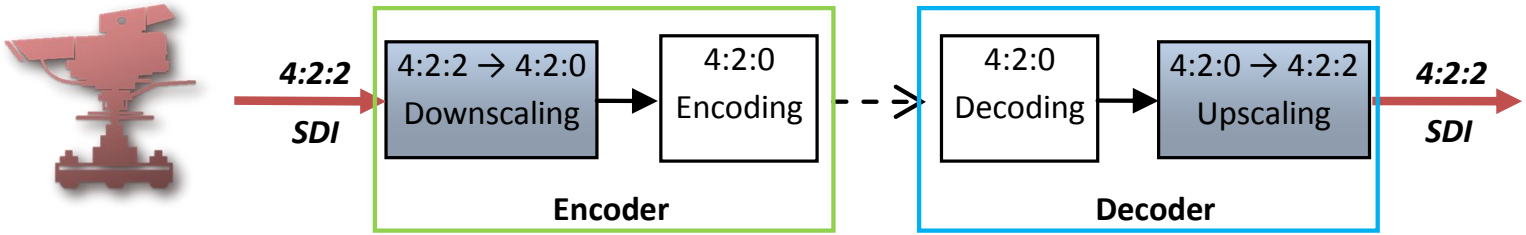
AVC/H.264 Studio Profile VS. Current Implementations

- Most AVC/H.264 products are currently geared towards consumer video market
  - 4:2:0 8-bit and <30Mbps encode rates @ High Profile
  - Using this equipment will necessitate resampling & incur mismatching if used for production transmission applications
- High 4:2:2 Profile
  - Can handle 4:2:2 processing @ 10-bit and is ideal for production transmission applications
  - Can handle upto 200 Mbps(HD) and 40 Mbps (SD)

ateme

About 4:2:0 Compression

- Professional video sources are 4:2:2
- 4:2:0 compression requires resampling filters



- Progressive and Interlace chroma location schemes are different
- Challenges for existing chroma resampling architecture:
  - Possible mismatch between the 2 resampling filters
  - Resampling filters quality not guaranteed
  - Choosing between Progressive and Interlace filters is not obvious

ateme

4:2:0 Chroma Artifacts

- Mainly caused by a mismatch between downscale/upscale filters
  - Color bleeding
  - Chroma detail loss
  - Interlace/Progressive artifacts
- Worsen with each generation
- Difficult to control in a Production chain

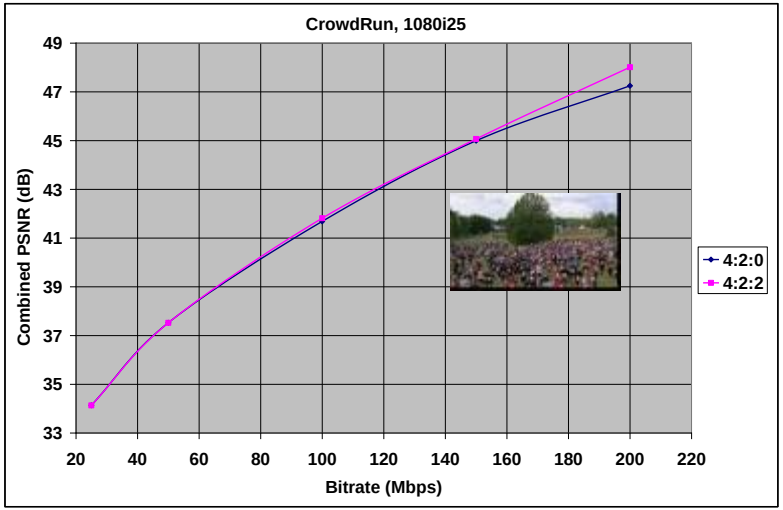
After 5 4:2:2 ↔ 4:2:0 conversions

Source picture

ateme

About 4:2:2 Compression

- At the 1st generation, with perfect resampling filters, 4:2:2 and 4:2:0 overall quality is roughly the same:



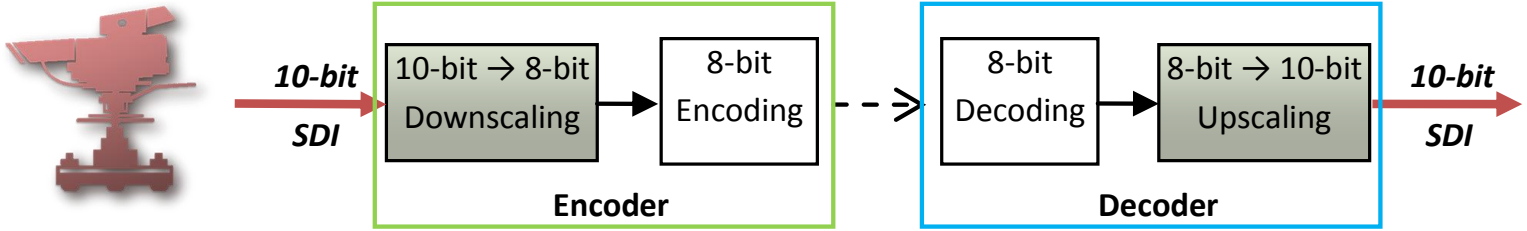
4:2:2 vs 4:2:0 with perfect resampling filters

- But in multi-generation and/or with mismatched sampling filters, 4:2:2 quality is much better than 4:2:0

ateme

Current 8-bit Compression

- Professional video sources are 10-bit
- 8-bit compression requires rounding before compression



- Motion compensation and loop-filtering performed using 8-bit samples
- Challenges for 8-bit architecture:
  - Scalers match and quality not guaranteed
  - Reduced accuracy in the encoding process reduces coding efficiency
  - Luma shift caused by lack of rounding control

ateme

10-bit compression: banding removal

- Processing video with 8-bit sample depth may create banding artifacts in shallow changing light scenes:
  - Blue skies
  - Underwater scenes
  - Sunsets
- These defects are not visible at 10-bit or more

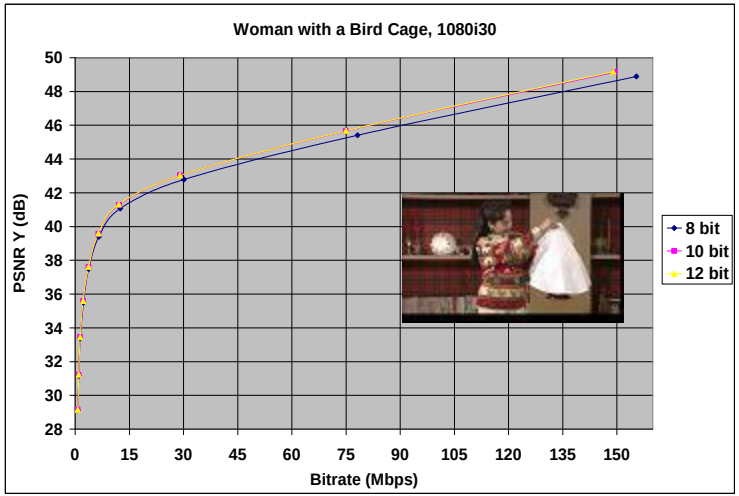
8-bit compression

10-bit compression

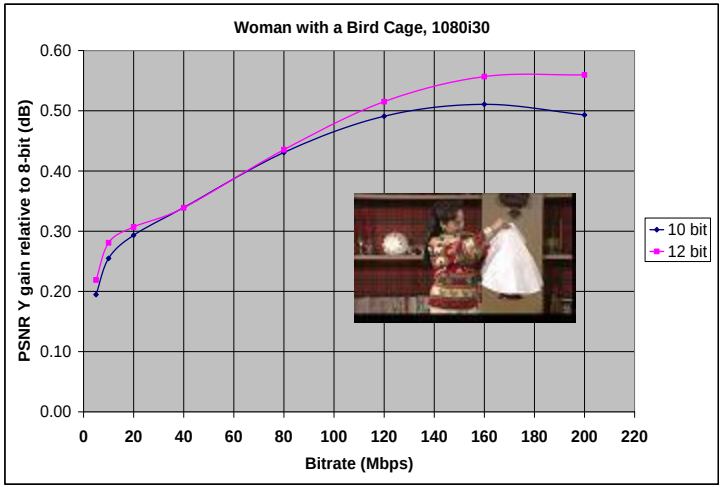
ateme

10-bit compression: increased coding efficiency

- Coding efficiency is increased: **less bit-rate for the same quality**
- Rate saving between 5% and 20% on most sources
- Most of the gain is provided with 10-bit coding. With 12-bit and 14-bit, coding efficiency gain is there but much smaller



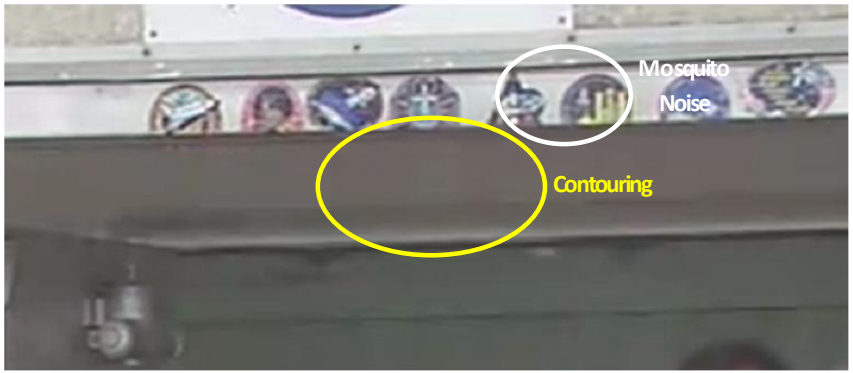
RD curve at 8,10 and 12-bit



Distortion gain vs 8-bit compression

10-bit compression: artifacts reduction

- 10-bit compression reduces 3 kind of artifacts:
  - Contouring
  - Smearing
  - Mosquito noise
- As a side-effect, 10-bit compression also reduces random distortion of shallow textured objects (clouds for instance)



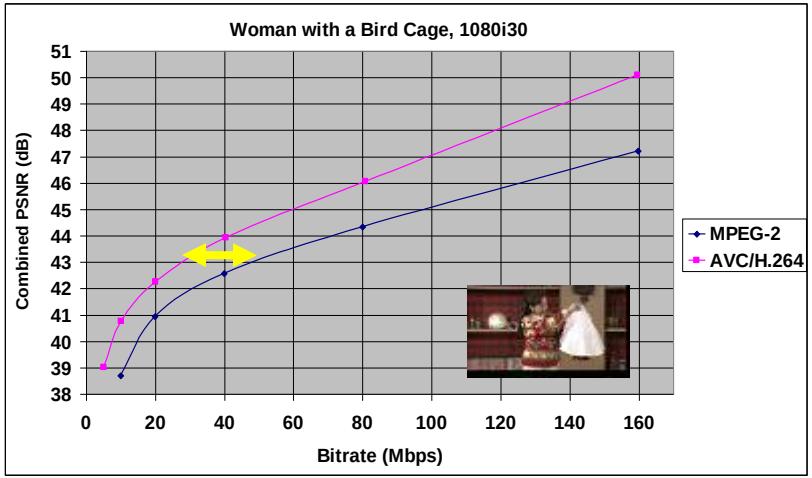
8-bit compression

10-bit compression

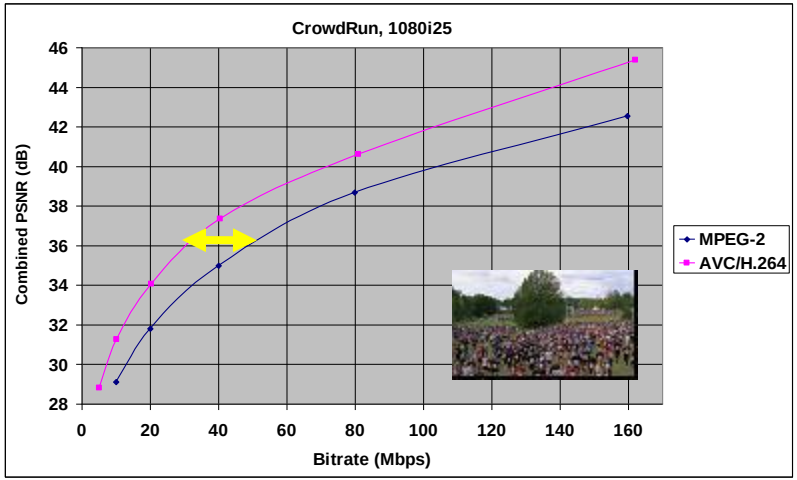
ateme

AVC *H422P* vs MPEG-2 *422P*

- H.264 *H422P* outperforms MPEG-2 *422P* in Contribution applications:
  - 10-bit support
  - More than **20Mbps** saved at Contribution rates!



Example 1

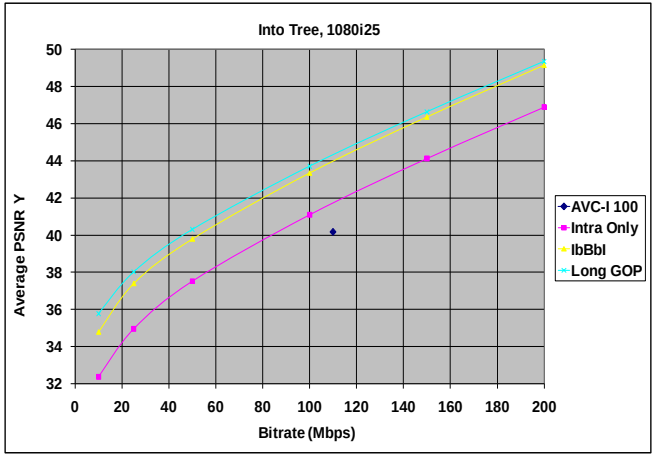


Example 2

GOP Structure Types for Contribution

- GOP structure needs to be carefully considered
- Possible Contribution GOP Structures
  - Intra-only, AVC-I, Short GOP, Adaptive Long GOP

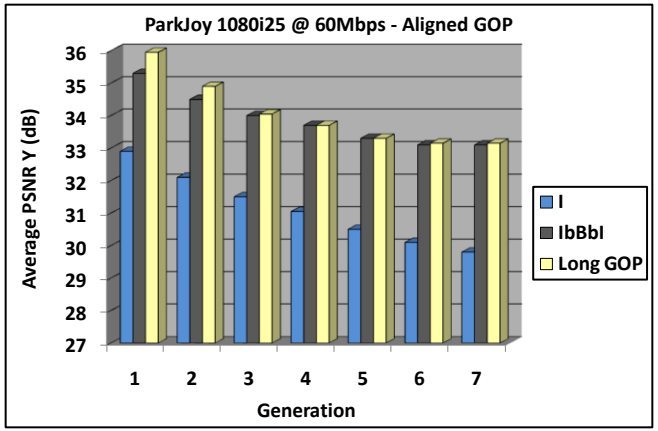
- Adaptive Long GOP and Short GOP give better performance at same bit rate
- Guide: let encoder make picture type decision wherever possible



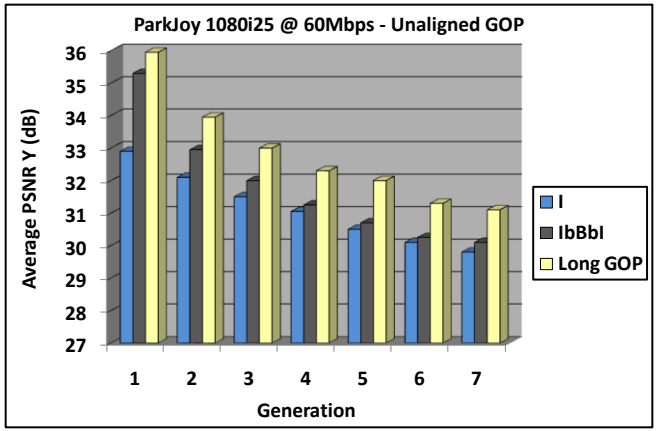
Importance of GOP structure on coding efficiency

Role of GOP in Multi-Generation Encoding

- 3-4 Re-encodes can happen in a production transmission workflow path
- Aligning Encodes at GOP Boundaries can retain quality
  - PSNR Drop 3dB (Aligned) Vs 5dB (non-aligned) @ 7th GEN



Multiple generations with aligned GOPS



Multiple generations without aligned GOPS

ateme

Production Uses

- Remote Truck Uplinks
  - Transmit in AVC 422 10-bit and at lower bit rates than MPEG-2 422P
- Playout to Tape Infrastructures
  - Easily integrate transmission back into an SDI infrastructure (international applications)
- Remote Collaboration
  - Realtime 422 10-bit at lower bit rates
- Multiple Generations Re-encodes
  - Typical Streams can re-encode 3-4 times before it gets to consumer, so transmission in 422 10-bit aligned GOP will preserve quality

Advantages

- Significantly **lower transmission costs**
- **Improve video quality** on existing transmission links

ateme

Summary

Using 4:2:2 10-bit allows:

- Process original source video in as close to an original form
- Enables most demanding application in terms of quality & rate
- Offers significant gain in quality and/or rate over existing solutions
- Reduces banding, smearing, and mosquito noise
- Retains quality in production processes
- Integrates into an SDI infrastructure at highest allowable quality

Thank-You

p.larbier@ateme.com