

# Comparative study of video coding solutions VVC, AV1, EVC versus HEVC

JVET –N0605

[Technicolor]

Geneva, March 2019

# Introduction

- Snapshot of the video coding landscape in terms of performance
- Existing solutions:
  - HEVC (HM-16.18)
  - AV1 (commit 52ea88f)
- Solutions in development:
  - VVC (VTM4)
  - EVC (ETM1)

# Algorithms comparison

| Category         | HEVC  | VVC  | AV1  | EVC (Main)  |
|------------------|---|--|--|---|
| Partition size   | <ul style="list-style-type: none"> <li>Coding Unit: 8x8 to 64x64</li> </ul>                                 | <ul style="list-style-type: none"> <li>Coding Unit: 4x4 to 128x128</li> <li>Pipeline size 64x64</li> </ul>   | <ul style="list-style-type: none"> <li>Superblock: 4x4 to 128x128</li> </ul>   | <ul style="list-style-type: none"> <li>Coding Unit: 4x4 to 128x128</li> <li>Pipeline size 128x128</li> </ul>                    |
| Partitioning     | <ul style="list-style-type: none"> <li>Transform: QT</li> <li>Prediction: QTBT</li> </ul>                   | <ul style="list-style-type: none"> <li>QTBT + TT (NS, QT, BTx2, TTx2= 6 partitions)</li> </ul>   | <ul style="list-style-type: none"> <li>Recursive tree (10 partitions)</li> </ul>   | <ul style="list-style-type: none"> <li>QTBT + TT (NS, QT, BTx2, TTx2= 6 partitions)</li> <li>Split Unit Coding Order</li> </ul> |
| Intra Prediction | <ul style="list-style-type: none"> <li>35 predictors</li> </ul>   | <ul style="list-style-type: none"> <li>81 modes: <ul style="list-style-type: none"> <li>65 angular + planar + DC + Sub</li> <li>6-MPM</li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>60 modes: <ul style="list-style-type: none"> <li>56 angles + 3 smooths + Paeth</li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>30 modes (angular, bilinear, DC, plane)</li> <li>2MPM + 2nd 8MPM</li> </ul>              |
|                  |   | <ul style="list-style-type: none"> <li>Chroma from Luma, PDPC, combined, Wide Angle, Sub-partitions</li> </ul>   | <ul style="list-style-type: none"> <li>Chroma from Luma</li> </ul>   | <ul style="list-style-type: none"> <li>Close to HEVC</li> </ul>   |
| Inter Prediction | <ul style="list-style-type: none"> <li>1/4 Pixel luma MV</li> <li>2 interpolation filters for MC</li> </ul> | <ul style="list-style-type: none"> <li>1/16 Pixel luma MV</li> <li>2 interpolation filters for MC</li> <li>DMVR</li> </ul>                                   | <ul style="list-style-type: none"> <li>1/8 Pixel luma MV</li> <li>5 interpolation filters for MC</li> </ul>                                  | <ul style="list-style-type: none"> <li>1/16 Pixel luma MV</li> <li>2 interpolation filters for MC</li> <li>DMVR</li> </ul>      |
|                  |   | <ul style="list-style-type: none"> <li>Sub-PU based MV Pred (ATMVP, Affine)</li> <li>GBi, BDOF, Triangular, Multi-hyp.</li> </ul>                            | <ul style="list-style-type: none"> <li>Extended reference frames (7 chosen per frame)</li> </ul>   | <ul style="list-style-type: none"> <li>Many similarities for inter prediction with VVC</li> </ul>                               |
|                  |   | <ul style="list-style-type: none"> <li>MMVD, HMVP, Pairwise Enhanced MV coding, AMVR, SMVD</li> </ul>  | <ul style="list-style-type: none"> <li>compound pred, wedge pred, OBMC, Global motion</li> </ul>   | <ul style="list-style-type: none"> <li>like VVC: AMVR, MMVD, Affine, HMVP</li> </ul>  |

# Simulations conditions

- JVET CTC for:
  - HM-16.18
  - VTM4
  - ETM1
- As close as possible to JVET CTC for AV1:
  - Two passes
  - No key frame every second
  - Constant quality (fixed QP)
  - Highest quality
- On « one intra period only » (  $\approx$  1 second)

# Results

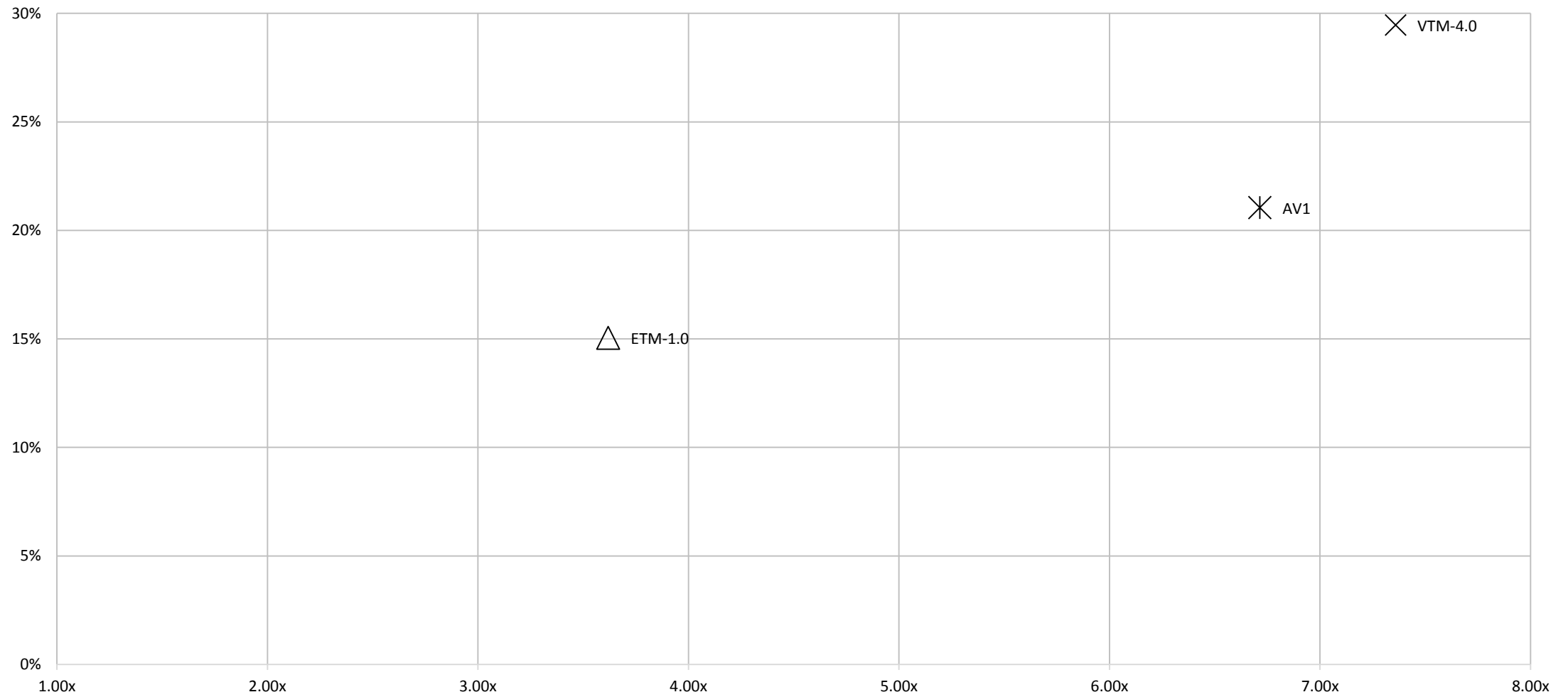
| BD-Rate<br>PSNR YUV* | Reference |          |         |       |         |
|----------------------|-----------|----------|---------|-------|---------|
| Test                 |           | HM-16.18 | VTM-4.0 | AV1   | ETM-1.0 |
|                      | HM-16.18  |          | 42.7%   | 27.6% | 18.5%   |
|                      | VTM-4.0   | -29.5%   |         | -8.8% | -17.1%  |
|                      | AV1       | -21.0%   | 10.4%   |       | -7.1%   |
|                      | ETM-1.0   | -15.0%   | 21.0%   | 9.3%  |         |

| BD-Rate<br>PSNR<br>YUV*    | Reference (UHD + HD only) |          |         |        |         |
|----------------------------|---------------------------|----------|---------|--------|---------|
| Test<br>(UHD + HD<br>only) |                           | HM-16.18 | VTM-4.0 | AV1    | ETM-1.0 |
|                            | HM-16.18                  |          | 50.8%   | 29.5%  | 25.1%   |
|                            | VTM-4.0                   | -33.5%   |         | -11.4% | -17.4%  |
|                            | AV1                       | -22.4%   | 13.5%   |        | -4.0%   |
|                            | ETM-1.0                   | -19.8%   | 21.3%   | 5.4%   |         |

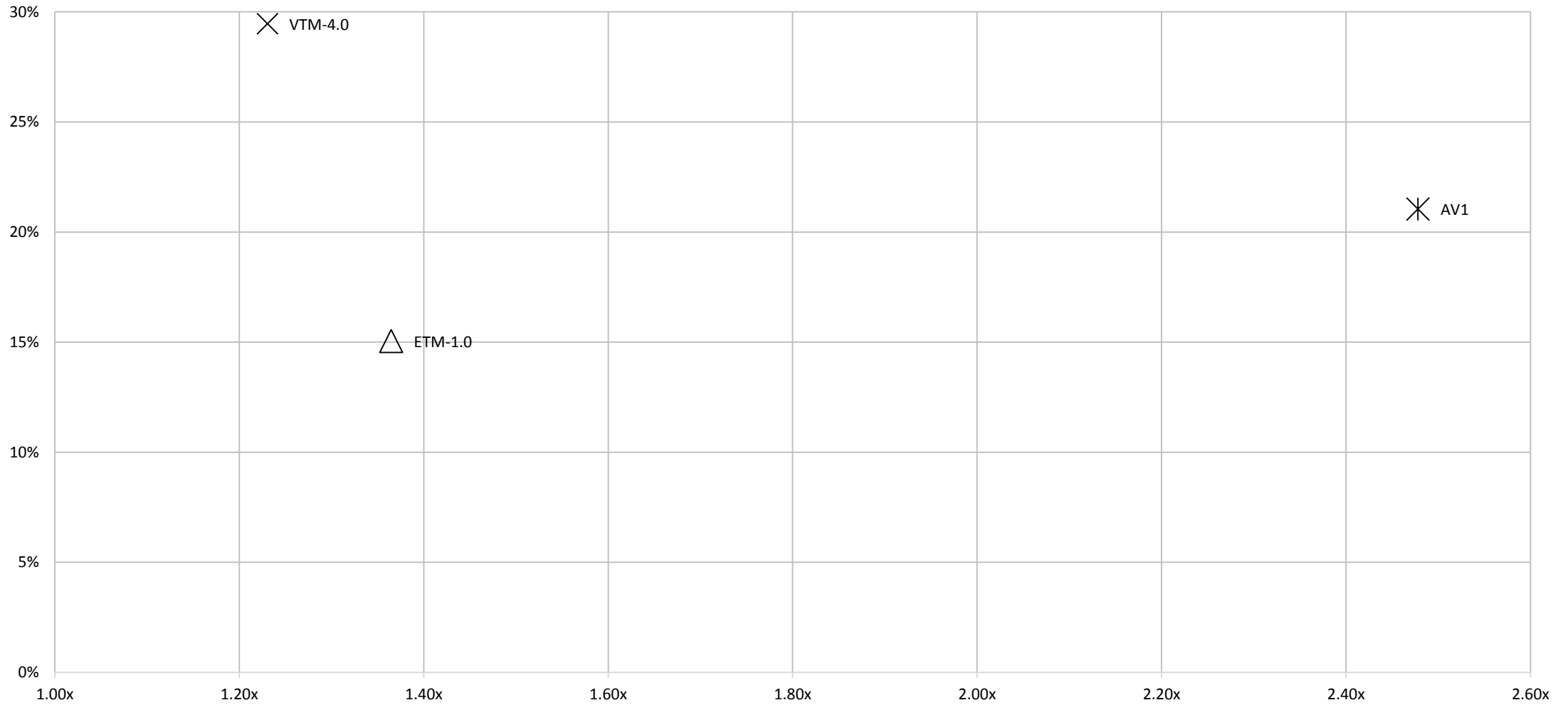
PSNR YUV =

$(6 \cdot \text{PSNRY} + \text{PSNRU} + \text{PSNRV}) / 8$

BD-Rate PSNRYUV = f(EncTime)  
over HM-16.18  
Overall



BD-Rate PSNRYUV = f(DecTime)  
over HM-16.18  
Overall



# Conclusion

- VTM4 (current version of VVC) is better than AV1 and EVC
- But, is the gap sufficient enough for VVC commercial success?
- What are the degrees of liberty to increase the gap?