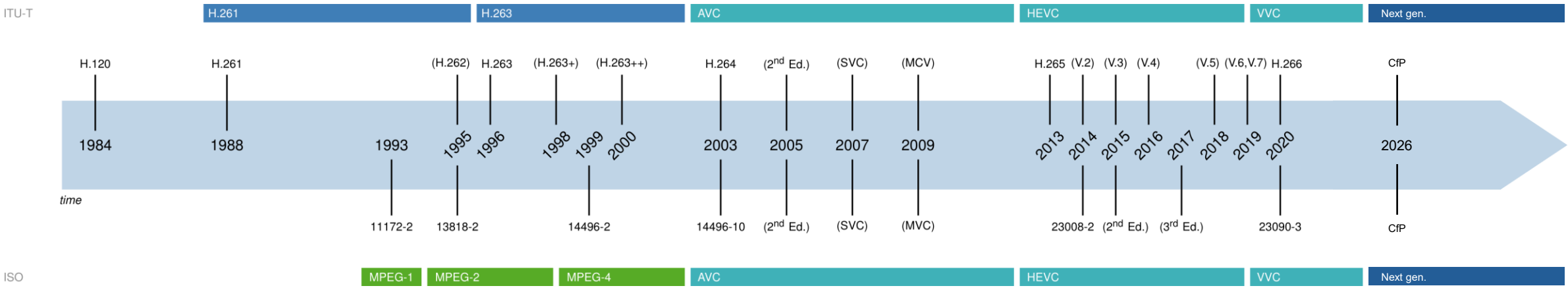
Next Generation Video Compression Standards

Mathias Wien and Jens-Rainer Ohm
RWTH Aachen University

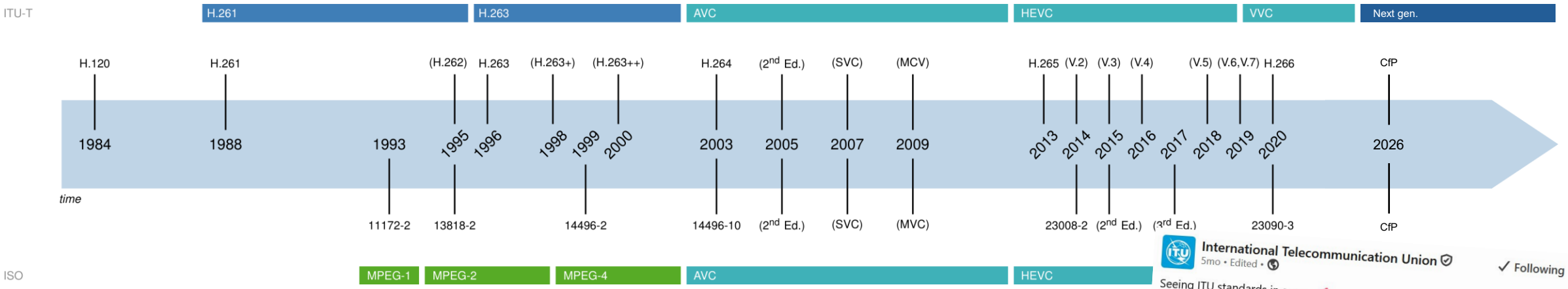
Joint Video Experts Team of ITU-T SG21 and ISO/IEC JTC 1/SC 29

Evolution of video coding standards

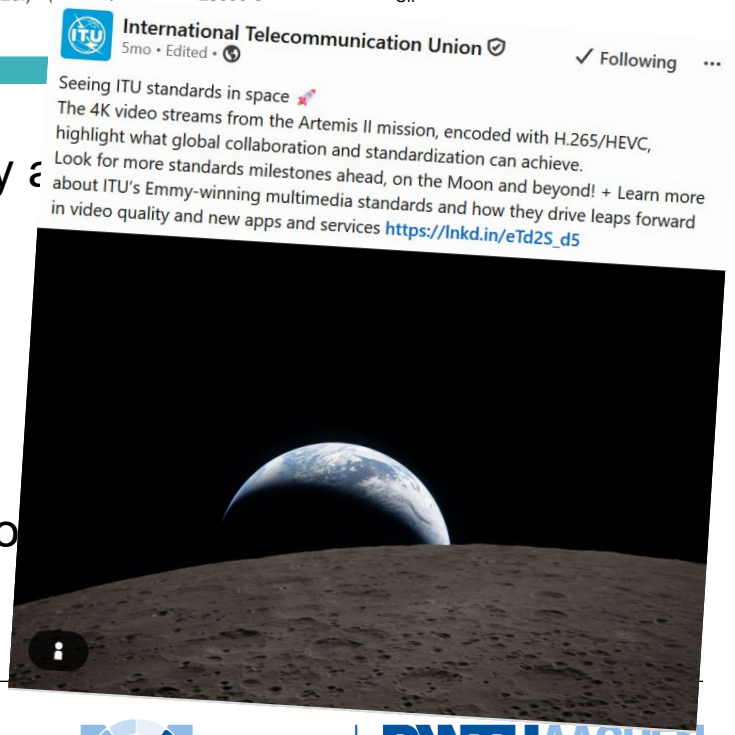


- Series of video coding standards delivered by ISO/IEC and ITU-T, continuously as joint work since H.264 | AVC
- H.265 | HEVC (High Efficiency Video Coding, 1st version in 2013)
  - Approximately 50% bitrate saving relative to AVC
  - Main (8-bit) and Main10 (10-bit) profiles most widely used
  - Central role in deployment of HDR video, such as HDR10 (PQ), HLG10
  - Profiles for contribution, screen content, scalable coding, and multiview video

Evolution of video coding standards



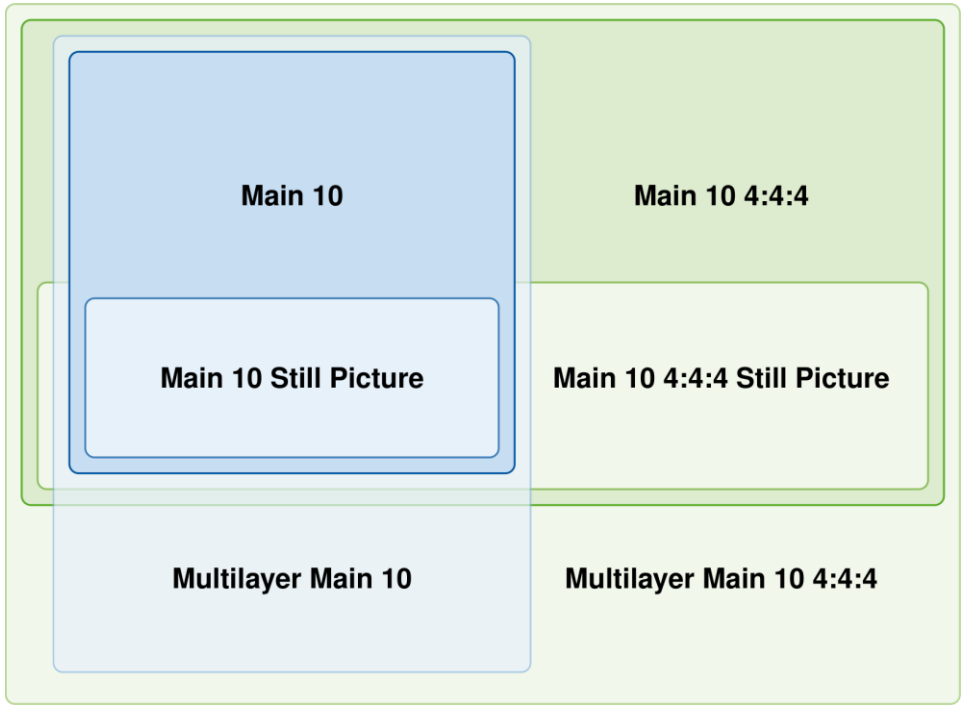
- Series of video coding standards delivered by ISO/IEC and ITU-T, continuously a
- H.265 | HEVC (High Efficiency Video Coding, 1st version in 2013)
  - Approximately 50% bitrate saving relative to AVC
  - Main (8-bit) and Main10 (10-bit) profiles most widely used
  - Central role in deployment of HDR video, such as HDR10 (PQ), HLG10
  - Profiles for contribution, screen content, scalable coding, and multiview video



Chair of Imaging and Computer Vision | RWTH AACHEN UNIVERSITY

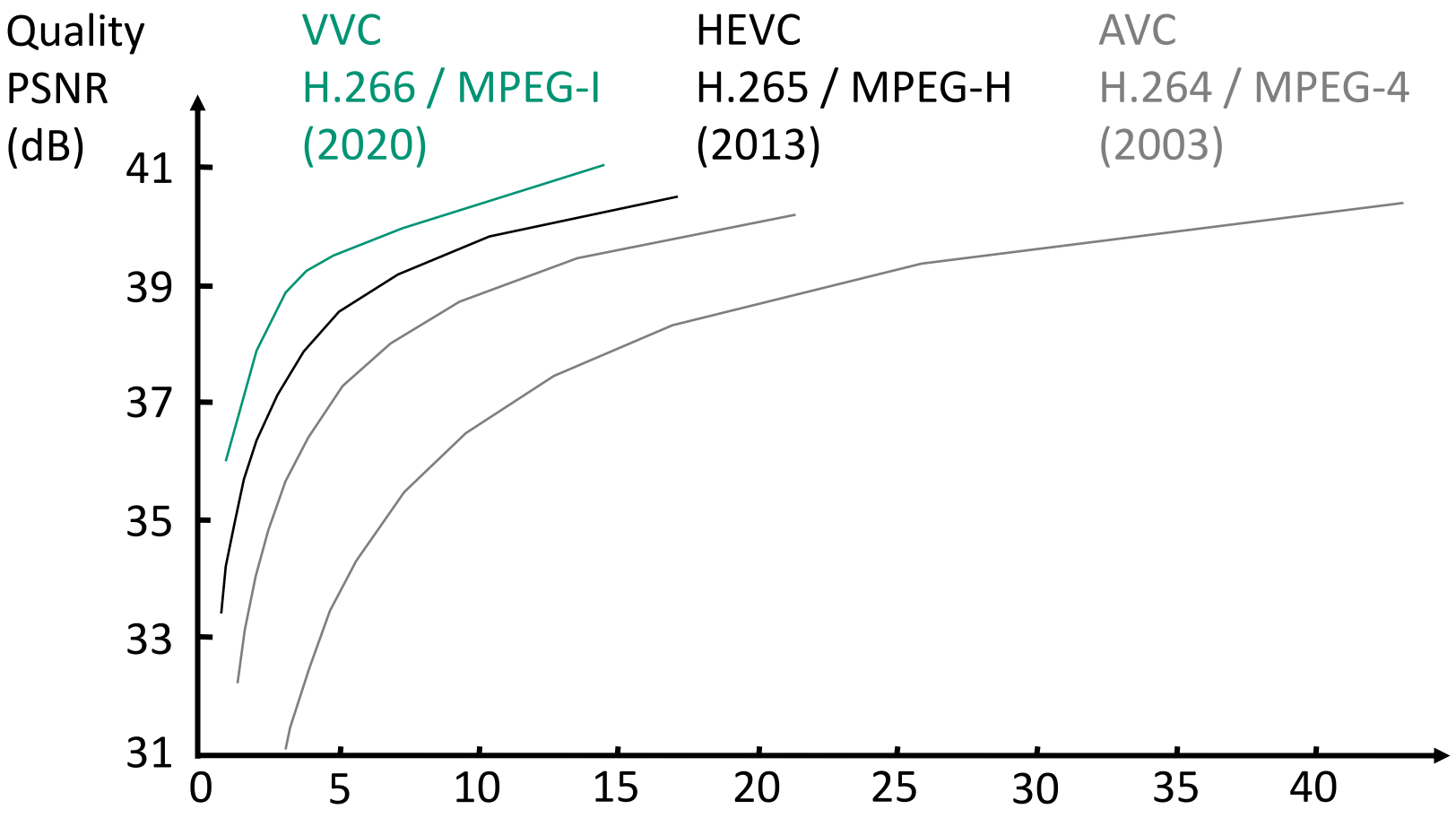
Versatile Video Coding (VVC)

- First version released in 2020
- Approximately 50% bit-rate savings at less than twice the decoder complexity compared to HEVC
- Versatility as inherent feature, supported from VVC version 1 on
  - HDR and WCG support
  - Resolution-adaptive multi-rate streaming
  - Support of ultra-low latency coding
  - 360-degree immersive video
  - Multilayer coding (multiview, scalability), alpha/depth maps
  - Enhanced concepts of tiles, slices, and wavefronts
  - …
- Main profiles in version 1
- Addition of operations extensions in 2022
  (12-bit, 16-bit versions of Main and Still Picture profiles)
- VSEI as separate specification (Versatile Supplemental Enhancement Information, H.274 | ISO/IEC 23001-7)



Profiles in VVC version 1

Performance history of video coding standard generations since the 1990's



BasketballDrive (1920 x 1080, 50Hz)

Slide by courtesy of J.-R. Ohm

Explorations beyond VVC

- **Exploration experiment 1: Neural network-based video coding**
  - Systematic exploration for better understanding (started in late 2020, terminated in 2026)
    - benefits of NN-based technology in terms of compression
    - methods to assess implementation complexity, both on CPU and GPU based processors
    - impact of training (materials, strategies, parameters, etc.) on the performance of NN-based algorithms
  - Exploration software freely available: NNVC [JVET-AQ2019]
    - Extending the VVC reference software VTM [JVET-AO2002]
    - NN-based loop/post filters, intra/inter coding tools, super-resolution up-sampling (post processing)
    - Evaluation of various operation points: Impact of complexity
      - High Operation Point (HOP), Low Operation Point (LOP), Very Low Operation Point (VLOP)
      - HOP: Gains relative to VTM of up to 13.5% in all-intra (AI), up to 15% in random-access (RA) configurations, based on PSNR improvement, CPU decoding runtime more than 1000x, encoder 2.5x
      - LOP: lower runtimes, still 8.7% in AI and 8.2% in RA configurations.

JVET document repository: https://jvet-experts.org

Explorations beyond VVC (2/2)

- **Exploration experiment 2: Enhanced coding tools**
  - Systematic exploration for better understanding (started in 2021, terminated in 2026)
    - Goal: explore potential improvements, no target for real applications
    - Assess the potential of new or extended conventional coding tools
    - Tools based on signal processing, entropy coding, error minimization, and related techniques
  - Exploration software freely available: ECM [JVET-AP2025]
    - Extending the VVC reference software VTM [JVET-AO2002]
    - Numerous tools
      - Intra and inter prediction, transforms, loop filtering, quantization
      - Parameter coding, enhanced entropy coding
      - Lower-complexity elements from NNVC
    - Bit-rate reductions of about 17%, 28%, and 24% for AI, RA, and low-delay configurations compared to VTM
    - CPU encoding and decoding runtimes increase by more than 10x

JVET document repository: https://jvet-experts.org

JVET Call for Evidence (CfE) on video compression with capability beyond VVC (JVET-AM2026)

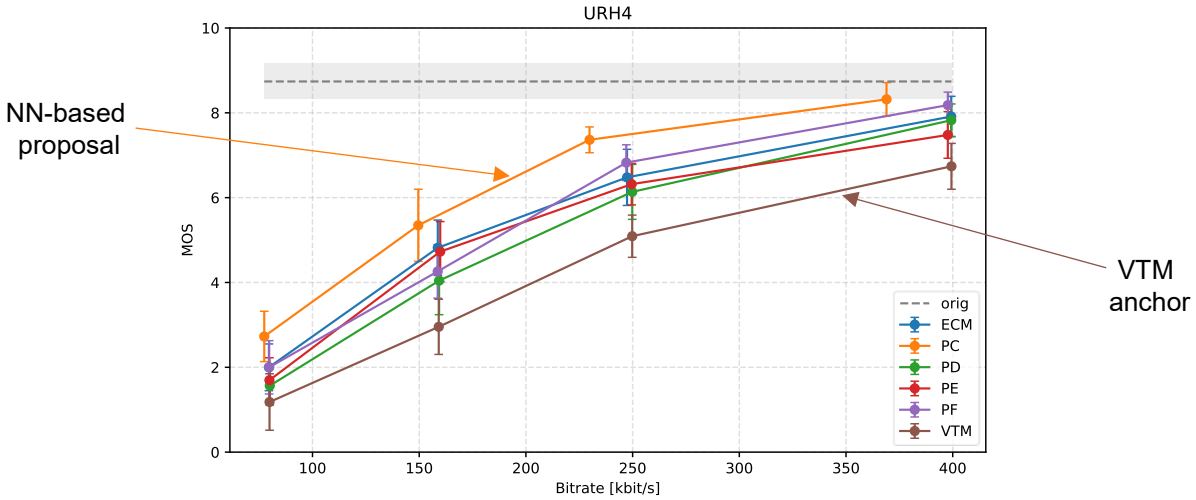
- **Conditions**
  - **30 test sequences** including camera-captured, computer-generated, user-generated, and high dynamic range content
    Resolutions: HD, UHD, 8K (cropped to UHD)
  - **4 target bitrates** (R1..R4)
  - **2 encoder configurations**
    (selected based on content type)
    - Random Access (RA): distribution/streaming
    - Low Delay (LD): conversational/gaming
  - **Compression efficiency** test case
  - **Reduced runtime** test case
- Conditions similar to a Call for Proposals

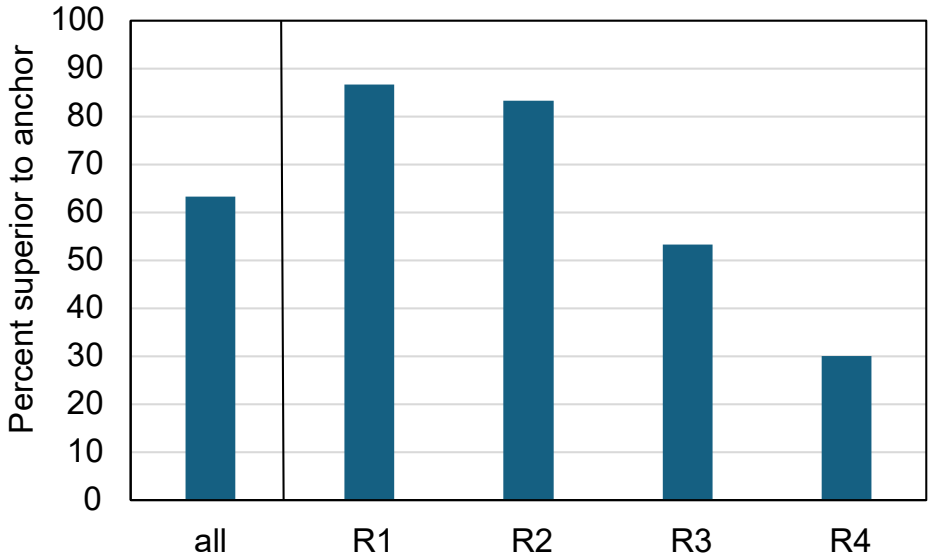
- **Submissions**
  - **Compression efficiency test case (CP)**
    - **ECM + 4 submissions** + anchor
    - All but one submissions in all test categories
  - **Runtime test case (RT)**
    - **ECM_2x + 15 submission variants** + anchor
    - All submissions responded to all test categories
    - Evaluated on subset of test sequences
  - **Functionality test (FC)** (error concealment)
    - One submission with subset of test sequences
    - Broadcast and unicast scenarios

JVET document repository: https://jvet-experts.org

JVET CfE on video compression with capability beyond VVC (JVET-AM2026)

- Summary of formal subjective expert viewing tests
  - Study across test categories and sequences
    - At least one proposal significantly better than VTM in roughly 2/3 of the test points, higher numbers at low bitrates
    - Reduced runtime configurations provide similar numbers (!)
  - Significant benefits observed for NN-based technology (higher than indicated by objective measurements)

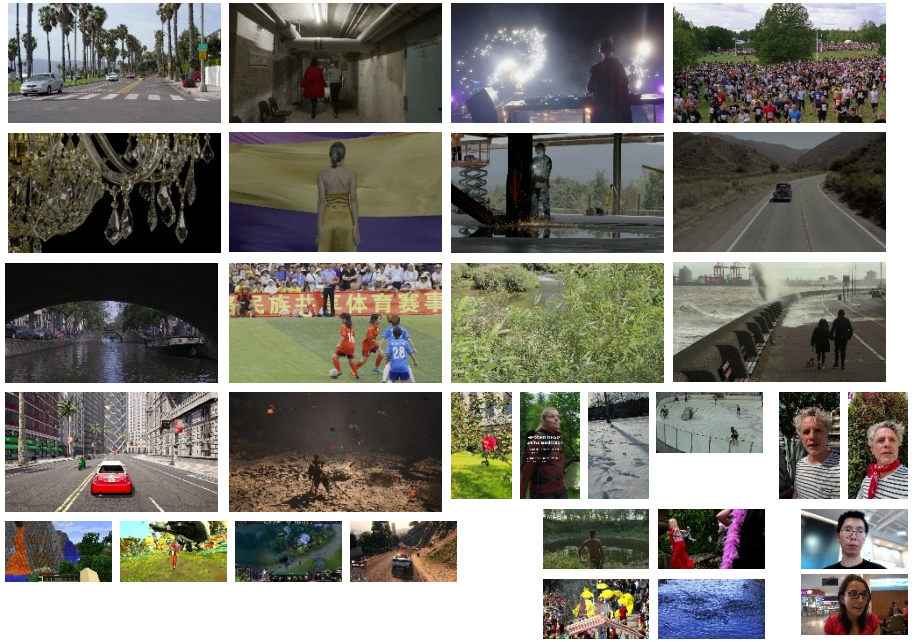




JVET document repository: https://jvet-experts.org

JVET Call for proposals (CfP) Overview – JVET-AQ2021

- **Timeline**
  - Call issued in July 2026
  - Submission of responses in October 2026
  - Evaluation in January 2027, first test model in 2027
  - Completion targeted before 2030



- **Test content**: Seven test categories, 30 test sequences (HD, UHD)
  - Standard dynamic range, HDR-PQ, HDR-HLG
  - Resolutions: HD, UHD, 8K (cropped to UHD)
    - Random Access (RA): distribution/streaming
    - Low Delay (LD): conversational/gaming
  - Content: Professional camera-captured, gaming, user-generated
  - Both landscape and portrait orientation for conversational and user-generated content

JVET document repository: https://jvet-experts.org

JVET CfP Overview – JVET-AQ2021

- **Complexity in focus**
  - Major take-away from a joint workshop of ITU and ISO/IEC JTC 1 SC 29 on “Future video coding – advanced signal processing, AI and standards”, Jan. 2025
- **Definition of Test Cases**
  - **Improved compression efficiency** (“classical call”) (COMP)
    - No constraints on encoder or decoder, but reporting of runtimes, memory usage etc.
  - Runtime-constrained encoding: **5 times (500%)** the encoder run time of the VTM anchor (RT1)
  - Runtime-constrained encoding: **Same (100%)** as the encoder run time of the VTM anchor (RT2)
  - Runtime-constrained encoding: **0.2 times (20%)** the encoder run time of the VTM anchor (RT3)
  - **Additional functionality**: open for further technical aspects
- **Evaluation dimensions**: subjective quality assessment, objective quality metrics, complexity figures
  - Formal subjective tests of COMP and RT1..3 by a group of international test laboratories
  - “Hidden Set”: evaluation of supplemental set of unknown/new test sequences

Summary

- **Evolution** of ISO/IEC and ITU-T standards

  JVET explorations on potential new technology
  - Learning-based tools
  - Conventional tools
  - Encoder and decoder complexity

- **Call for Evidence**
  - Assessment of expectable compression gains

- Call for Proposals on video compression with capability beyond VVC

  → **Development towards a new joint ISO/IEC and ITU-T standard**
  - Enhanced compression efficiency
  - Provisions for lean encoder design
  - Consideration of NN-based technology

Thanks for your attention
Questions welcome!

apl. Prof. Dr.-Ing. Mathias Wien
Visual Media Communication Group
Chair of Imaging and Computer Vision
RWTH Aachen University
wien@lfb.rwth-aachen.de