|  |  |
| --- | --- |
|  | Moving Picture, Audio and Data Coding by Artificial Intelligencewww.mpai.community |

**Public document**

|  |  |
| --- | --- |
| **N2266** | 2025/03/19 |
| **Source** | Requirements (EVC) |
| **Title** | Requirements (EVC) Report and Plans |
| **Target** | MPAI-54 |

The group has issued Call for Technologies [1] and is currently preparing the test sequences to be provided to the proponents.

The team coded the test sequences for the proponents (the proponents will use this sequence to demonstrate to MPAI the objective metrics specified in the CfT).

This task involves coding five sequences. The initial phase includes downscaling 4K sequences into HD and SD resolution and coding with various Quantization Parameters (QPs) - 22, 27, 32, 37, and 42. These sequences will be provided across three codecs: AVC, HEVC, and VVC, in both Random Access and Low Delay modes.

The complexity of this task is reflected in the numbers: 2 modes, 5 sequences, 5 QPs, 3 codecs, and 2 resolutions (SD and HD), resulting in a total of 300 test points.

Adding an element of confidentiality, an esteemed academic has selected 'secret' sequences for the MPAI test. The EVC members will utilize the proponents' software to encode these secret sequences, selecting a limited number of test points.

The list is as follows:

3 sequences from here: http://download.opencontent.netflix.com.s3.amazonaws.com/index.html?prefix=Netflix\_test\_conditions/

\* Chimera-ToddlerFountain\_3840x2160\_2997fps\_10bit\_420.yuv , first 300 frames (3840x2160, 10bit 29.97fps, 4:2:0) - Content: toddler playing in a fountain

\* Chimera-Aerial\_3840x2160\_2997fps\_10bit\_420.yuv, 300 frames starting at 120-th frame (3840x2160, 10bit 29.97fps, 4:2:0) - Content: mostly static view of people and pilot from a drone, that then takes off

\* Chimera-PierSeaside\_3840x2160\_2997fps\_10bit\_420.yuv, 300 frames (3840x2160, 10bit 29.97fps, 4:2:0) - Content: initially static view of a pier from distance, then pan on sea water.

Other 2 sequences 8bit from here:

\* https://www.cdvl.org/members-section/view-file/?videoid=2963 (NTIA traffic) first 600 frames, i.e., 10 sec (4096x2160, 8 bit, 59.94fps, 4:2:2) - Content: motorway traffic, different shots

\* https://www.cdvl.org/members-section/view-file/?videoid=2682 (NTIA violin) 600 frames starting at 300-th frame , i.e., 10 sec (4096x2160, 8 bit, 59.94fps, 4:2:2) - Content: woman playing violin, different shots (medium distance, far, close up)

(These 2 sequences require free registration on CDVL, but usage and licence terms it is suitable for our purposes).

Usage license details:

* first 3 sequences: CC by-nc-nd 4.0 (also licensed as CC-by 4.0 here, but in another format: http://download.opencontent.netflix.com.s3.amazonaws.com/Chimera/creative-commons-attribution-4-intl-public-license.txt )
* last 2 sequences (from CDVL): https://www.cdvl.org/license/ which says that all sequences in CDVL can be used by standards developing committees (SDO) including organizations that are peripherally involved (e.g., a university hired to do a subjective test for an SDO, etc.).
1. Reference

[1] https://mpai.community/standards/mpai-evc/ufv/v1-0/call-for-technologies/