**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG2013/N13943**

**November 2013, Geneva, CH**

|  |  |
| --- | --- |
| **Source:** | **Video** |
| **Title:** | **Conditions for visual comparison of VCB, IVC and WVC codecs** |
| **Status:** | **Draft** |

During the 106th MPEG meeting, the Video Subgroup decided to plan a comparative investigation of technologies that had been proposed in response to the first (N12204, July 2011) and the second (N13546, April 2013) Call for Proposals on Internet Video Coding, in order to identify merits and be able to plan further standardization actions. Part of this comparison is investigation of visual quality under comparable conditions of encoding settings as possible.

To enable comparison at approximately the same rate points, the following approach will be used:

* Produce for each of the codecs bit streams which are within +/-3% of the target rates for the sequences given below. To achieve this with a minimum change of quality, one change of the quantization parameters by one step is allowed per sequence. This change in quantization parameters shall persist from that point onward. (For example, in an AVC encoder, one may start encoding a video sequence with a quantization parameter QPx, and at frame N change the quantization parameter from QPx to QPx+1. From this point onward QPx+1 is used for encoding all remaining frames in an effort to achieve the target rate. The selection of the quantizer QPx can be achieved by finding the quantization parameter that results in a rate that would be higher than the target rate when coding the entire sequence with QPx, while, on the other hand, coding the entire sequence with QPx+1 would result in lower rate than the target rate).
* Rate points and sequences to be used in the visual comparison will be finally decided by the VCC Adhoc group meeting, seeking advice by the test manager. If one of the codecs is not able to meet a desired rate point, that point will not be used for comparison. It is planned to perform viewing with three of the four rate points for each sequence.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Rate 1 | Rate 2 | Rate 3 | Rate 4 |
| *1920x1080p* |  |  |  |  |
| Kimono 24fps | 1.6 Mbit/s | 2.5 Mbit/s | 4.0 Mbit/s | 6.0 Mbit/s |
| Park Scene 24fps | 1.6 Mbit/s | 2.5 Mbit/s | 4.0 Mbit/s | 6.0 Mbit/s |
| Cactus 50fps | 3.0 Mbit/s | 4.5 Mbit/s | 7.0 Mbit/s | 10.0 Mbit/s |
| BasketballDrive 50fps | 3.0 Mbit/s | 4.5 Mbit/s | 7.0 Mbit/s | 10.0 Mbit/s |
| *836x480p (WVGA)* |  |  |  |  |
| BasketballDrill 50fps | 512 kbit/s | 768 kbit/s | 1.2 Mbit/s | 2.0 Mbit/s |
| BQMall 60fps | 512 kbit/s | 768 kbit/s | 1.2 Mbit/s | 2.0 Mbit/s |
| PartyScene 50fps | 512 kbit/s | 768 kbit/s | 1.2 Mbit/s | 2.0 Mbit/s |
| RaceHorses 30fps | 512 kbit/s | 768 kbit/s | 1.2 Mbit/s | 2.0 Mbit/s |
| *1280x720p* |  |  |  |  |
| Johnny 60fps | 384 kbit/s | 512 kbit/s | 850 kbit/s | 1.5 Mbit/s |
| KristenAndSara 60fps | 384 kbit/s | 512 kbit/s | 850 kbit/s | 1.5 Mbit/s |
| FourPeople 60fps | 384 kbit/s | 512 kbit/s | 850 kbit/s | 1.5 Mbit/s |

The following constraints are to be followed by the encoders generating the bitstreams:

* Allow QP (or quantizer step size) variation within a sequence within a periodic pattern of frame types (where frame types are differentiated by syntax or by a reference picture handling mechanism) within a sequence. However, pictures that correspond to the same frame type shall be quantized with a constant step size. The difference in quantization quality for different frame types shall not exceed a ratio of 4 between the smallest and the largest quantizer step size (e.g. QPlarge=QPsmall+12 in the case of a logarithmic quantizer scale factor coding as defined in AVC, and QSlarge=QSsmall\*4 in the case of a linear quantizer scale factor). The encoder shall use the same quantizer ratio settings for all test cases.
* No per sequence adaptation of the pattern of frame types shall be used. (Note: The case of necessary alignment of GOP structures at the end of a sequence is exempted. Scene cuts in Kimono and FourPeople can also be handled like end of a sequence)
* No sequence specific tuning of coding parameters (such as enabling/disabling of special tools, certain modes, limitation of motion search range etc.) shall be used.
* No rate control shall be used.
* No pre-processing shall be used.
* No post-processing of the decoder output shall be used, unless it is part of a normative decoding process (such as filtering of reference pictures for subsequent predictions).

Encoded bitstreams are to be provided for the following two constraint cases:

* Constraint set 1 (CS1): structural delay of processing units not larger than an 8-picture "group of pictures (GOPs)" and random access intervals of 1.1 seconds or less.
* Constraint set 2 (CS2): no structural delay of processing units, with essentially no picture reordering between decoder processing and output. Furthermore, no sequence level multi-pass encoding shall be allowed (Note: The procedure described above for approaching a target rate by one QP change per sequence is exempted from this limitation). Further analysis will be applied to measure bit rate fluctuation characteristics.

In CS1, the distance of random access intervals should be as close as possible (e.g. as multiple of the used GOP size) to the value given above, but not supersede it.

Testing methodology:

The evaluation will be made by means of an “Expert Viewing” evaluation Protocol (EVP), roughly based on the protocol of the DSIS test method from ITU-R Rec. BT.500-13.

Three experts will be seated in front of a 30” (or greater) high quality monitor (High quality TV set is also allowed), at 3 times the screen height, while taking care that the widest viewing angle does not exceed 60° from the center axis of the screen.

It is desirable (where possible) that more than three experts participate in the viewing sessions.

The Basic Test Cell (BTC) of the EVP will include the viewing and the evaluation of samples from all the coding schemes under consideration, using the same video clip at the same rate point.

The timing the BTC will be as follows:

* “BTC n” (message on the screen) 1 second
* Reference 10 seconds
* Codec 1 10 seconds
* Codec 2 10 seconds
* “Vote 1 and 2” (message on the screen) 6 seconds
* Reference 10 seconds
* Codec 3 10 seconds
* Codec 4 10 seconds
* “Vote 3 and 4” (message on the screen) 6 seconds

Beyond the three codecs IVC, WVC and VCB, the visual tests will include AVC High Profile (HP) anchors. Encoding of those anchors will be performed under same conditions as for the other codecs.

The order of presentation of the coded clips (including the “anchor”) inside the BTC will be randomly changed (in order to avoid any bias in the judgment) in a way hidden to the viewers.

A 5 grade quality scale will be used to assess the visual quality of the coded video clips.

An example scoring sheet is shown below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | BTC 1 | | | | | | | | | |  | BTC 2 | | | | | | | | | |  |
|  |  | 1 |  | 2 |  |  | 3 |  | 4 |  |  |  | 5 |  | 6 |  |  | 7 |  | 8 |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | … |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

The viewers will write a number (from 1 to 5) in the corresponding box labeled with the number of the BTC and the presentation inside each BTC.

Based on the raw data, further analysis of the results will be made.

Work plan:

2013-12-01: Availability of Excel sheet to fill in results (to be further refined by reflector discussion)

2013-12-15: Report for each codec whether it is possible to meet the rate points

2014-01-05: Sequences to be made available to the Test Manager (Vittorio Baroncini) who prepares a “dry run” of the test before the meeting. Responsibilities:

* AVC HP anchors: RWTH (J. Ohm/M. Wien; R. Sjöberg provides configuration files by 2013-12-15)
* IVC: Peking University (R. Wang)
* WVC: Apple (K. Kolarov)
* VCB: Google (M. Raad)

2014-01-07: Document upload deadline, bit rates and PSNR values of each codec to be provided in the Excel sheet.

2014-01-11: Adhoc meeting, final selection of rate points and sequences, and further planning of test.

2014-01-12: Start of subjective viewing, with the intent to have the first set of results presented to the MPEG plenary on 2014-01-15.