Preliminary Joint Call for Evidence on video compression with capability beyond HEVC

(Ref: Report SG16-R1, Annex H; Approved 2017-01-27)

1 Introduction

ITU-T VCEG (Q6/16) and ISO/IEC MPEG (JTC 1/SC 29/WG 11) are jointly studying the potential need for standardization of video coding technology with a compression capability that significantly exceeds that of the current HEVC standard (including its extensions). Such future standardization could take the form of additional extension(s) of HEVC or an entirely new standard.

The scope of technology consideration performed by the Joint Video Exploration Team (JVET), an informal collaboration group of VCEG and MPEG, includes a broad variety of video source content, e.g. camera-view content, screen content, VR/360 video and high-dynamic-range video for such applications as broadcast (with live or pre-authored content), real-time video conferencing and video chat, on-demand viewing, storage-based media replay, consumer generated content, and surveillance with fixed or moving cameras [3][4].

It is now planned to enter a more rigorous evaluation phase, by issuing a Call for Evidence (CfE) on new superior-performance video compression technology, potentially followed by a formal Call for Proposals (CfP). This preliminary CfE will be updated in early April 2017, and responses to the CfE will be evaluated in July 2017, as further detailed below. The possible subsequent CfP would follow the evaluation of the responses to the CfE.

Companies and organizations who have developed compression technology that they believe to have better compression capability than the Main 10 Profile of the HEVC standard (Rec. ITU-T H.265 | ISO/IEC 23008-2) are kindly invited to bring such information to JVET in the context of this Call for Evidence. The main purpose of a possible subsequent Call for Proposals and standards development will be to increase compression performance, but contributions are also sought as well that better support newly emerging application areas of video coding.

1.1 Tentative timeline

– Final Call for Evidence: 2017-04-07

– Test sequences and HEVC Main 10 Profile anchors available: 2017-04-21

  NOTE 1 – a preliminary set of test sequences will be made available 2017-02-01

  NOTE 2 – a preliminary version of the HEVC anchors will be made available 2017-02-15

– Expression of interest to submit a response: 2017-06-16

– Submission of contributions (descriptive document): 2017-07-04

– Decoded sequences should be made available by the submitter for the 7th JVET meeting in Torino by 2017-07-13

– Evaluation of responses: July 2017 JVET meeting (expected to be attended by submitters)

– Depending on the outcome of the Call for Evidence, the parent bodies of the JVET collaboration intend to issue a Draft Call for Proposals by the end of the July meeting.

2 Test Cases

Submitters are encouraged (but not required) to submit results for all test cases. However, submitters are required to provide results for all sequences in a given test case.

Submissions must provide decoded YUV files from bitstreams encoding the complete test sequences as defined in the subsections below.

NOTE – In the final CfE, higher rate points than defined in the target bit rate tables below may be added for objective quality measurement purposes.

2.1 UHD 4:2:0 SDR

2.1.1 Sequence formats and frame rates

Table 1 – UHD SDR test sequence example pictures

 Tango	 CampfireParty	 DaylightRoad	 CatRobot
 ParkRunning1	 FoodMarket2	 BuildingHall	 Crosswalk

Table 2 – HD SDR test sequence example pictures

 BasketBallDrive	 BQTerrace	 Cactus
 Timelapse	 Ritual Dance	

Table 3 – SDR test sequences

Sequence ID	Sequence name	Resolution	Frame count	Frame rate	Chroma format	Bit depth
UHD1	Crosswalk	4096×2160	470	60	4:2:0	10
UHD2	Food Market2	4096×2160	720	60	4:2:0	10
UHD3	Tango	4096×2160	600	60	4:2:0	10
UHD4	CatRobot	3840×2160	600	60	4:2:0	10
UHD5	DaylightRoad	3840×2160	600	60	4:2:0	10
UHD6	BuildingHall	3840×2160	500	50	4:2:0	10
UHD7	ParkRunning1	3840×2160	500	50	4:2:0	10
UHD8	CampfireParty	3840×2160	300	30	4:2:0	10
HD1	BQTerrace	1920×1080	600	60	4:2:0	8
HD2	RitualDance	1920×1080	600	60	4:2:0	8
HD3	Timelapse	1920×1080	600	60	4:2:0	8
HD4	BasketballDrive	1920×1080	500	50	4:2:0	8
HD5	Cactus	1920×1080	500	50	4:2:0	8

Table 4 – SDR test sequence md5sums

Sequence ID	MD5Sum
UHD1	978a5dea90fe9125f6bce42aade55b61
UHD2	a3cb399a7b92eb9c5ee0db340abc43e4
UHD3	To be provided in final CfE
UHD4	03a89792693fd9ecfd72ef2590025e97
UHD5	165c70e3008d37b9ff476e997297fc5e
UHD6	836a5a0558b24e8dde6b9a256e7aa468
UHD7	9de83b1bc2bca1afedb5342a2df572ba
UHD8	b676cf8de483c1b890379976323f92af
HD1	efde9ce4197dd0b3e777ad32b24959cc
HD2	a3cb399a7b92eb9c5ee0db340abc43e4
HD3	To be provided in final CfE
HD4	d38951ad478b34cf988d55f9f1bf60ee
HD5	3fddb71486f209f1eb8020a0880ddf82

Table 5 – SDR target bit rates

	Target bit rates [kbit/s]			
Sequences	**Rate 1**	**Rate 2**	**Rate 3**	**Rate 4**
UHD1, UHD2	1000	1500	2400	4000
UHD3, UHD4, UHD5	1500	2400	4000	7000
UHD6	800	1200	2000	3300
UHD7, UHD8	2000	3300	6000	10000
HD1	400	600	1000	1700
HD2	900	1500	2600	4300
HD3	600	900	1500	2600
HD4	800	1200	2000	3500
HD5	500	800	1200	2000

2.1.2 Coding conditions for HEVC anchors

In this test case, a Random Access scenario is used for evaluation, following the JVET common test conditions and software reference configurations [7]. The intra refresh period is dependent on the frame rate of the source and the GOP size in use: a value 32 shall be used for sequences with a frame rate equal to 24 fps, 25 fps or 30 fps; 48 for 50 fps; 64 for 60 fps; and 96 for 100 fps.

Configuration files are provided in the cfg/ folder of version HM16.14 of the reference software package (available at https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.14/cfg/). A static quantization parameter (QP) setting is applied for generation of the anchors. A one-time change of the quantization parameter from value QP to value QP+1 may be applied in order to meet the defined target bit rates. The quantization parameter settings applied for the anchors will be reported.

2.1.3 Coding conditions for submissions

Submissions to the CfE shall obey the following rules:

– Be encoded to within ±2% of the target bit rates defined above;

– Allow for random access at intervals not larger than the intra refresh period of the respective anchor;

– Quantization settings should be kept static. When change of quantization is used it shall be described;

– A one-time change of the quantization settings to meet the target bit rate is admitted and must be documented;

– Preprocessing is not used;

– Use of a postfilter is allowed only if it is part of the decoding process.

2.1.4 Coding conditions for JEM anchors

The JVET maintains a Joint Exploration Test Model (JEM) software package embracing coding features that are under coordinated test model study [5]. It is intended to provide JEM bitstreams at the target bit rates following the conditions defined above in the final CfE.

2.2 HDR

2.2.1 Sequence formats and frame rates

Table 6 – HDR test sequence example pictures

 Market3	 ShowGirl2	 Hurdles	 Starting
 CosmosTreeTrunk			

Table 7 – HDR test sequences

Sequence ID	Sequence name	Resolution	Frame count	Frame rate	Chroma format	Bit depth
HDR1	Market3	1920×1080	400	50	4:2:0	10
HDR2	ShowGirl2	1920×1080	339	25	4:2:0	10
HDR3	Hurdles	1920×1080	500	100	4:2:0	10
HDR4	Starting	1920×1080	500	100	4:2:0	10
HDR5	Cosmos_TreeTrunk	1920×856	240	24	4:2:0	10

Table 8 – HDR test sequence md5sums

Sequence ID	MD5Sum
HDR1	c97abe47455fd12f6d6436cecfad7c7d
HDR2	44f1974d68f7799c71eea29fb72b245b
HDR3	bc3cba849d6f4ee74d39056600722aa5
HDR4	1cbc416696cb0dfcf4da9886eeb6a4a2
HDR5	To be provided in final CfE

Table 9 – HDR target bit rates

	Target bit rates [kbit/s]			
Sequences	**Rate 1**	**Rate 2**	**Rate 3**	**Rate 4**
HDR1	1200	1500	2400	4700
HDR2	550	900	1500	3000
HDR3	450	700	1500	2400
HDR4	1000	1700	2700	6000
HDR5	500	900	1500	3000

2.2.2 Coding conditions for HDR anchors and submissions

In this test case, the Random Access scenario as described in §2.1.2 is used for evaluation. The constraints defined in §2.1.3 for submissions apply.

The QP adaptation scheme will be detailed in the final CfE.

2.3 360° video

2.3.1 Test sequence formats and frame rates

Table 10 – 360 Test sequences

 SkateBoardInLot	 Chairlift	 KiteFlite
 Harbor	 GasLamp	 Trolley

The test sequences are defined in the table below. All input sequences are in the equirectangular projection (ERP) format, and represent omnidirectional 360º × 180º degree video.

The number of active coded luma sample is lower than the resolution of the input sequence, to attenuate the implicit advantage of ERP over other projection formats, as ERP is used in the inputs.

The coded active luma sample count should be met within ±3%.

Table 11 – 360° video test sequences

Sequence name	Input resolution	Anchor resolution	Coded active luma sample count	Frame count	Frame rate	Chroma format	Bit depth
SkateBoardInLot	8192×4096	4096×2048	8388608	300	30	420	10
Chairlift	8192×4096	4096×2048	8388608	300	30	420	10
KiteFlite	8192×4096	4096×2048	8388608	300	30	420	8
Harbor	8192×4096	4096×2048	8388608	300	30	420	8
GasLamp	8192×4096	4096×2048	8388608	300	30	420	8
Trolley	8192×4096	4096×2048	8388608	300	30	420	8

Table 12 – 360° video test sequence md5sums

Sequence	MD5Sum
SkateboardInLot	e8eae04c43e959060f641fec4892fced
ChairliftRide	9126f753bb216a73ec7573ecc4a280c3
KiteFlite	18c0ea199b143a2952cf5433e8199248
Harbor	aa827fdd01a58d26904d1dbdbd91a105
GasLamp	25c1082d1e572421da2b16530718156d
Trolley	3417d0b862ffb0fd34f65c3bc810d25c

Table 13 – Target bit rates for 360° video test sequences

	Target bit rates [kbit/s]			
Sequences	**Rate 1**	**Rate 2**	**Rate 3**	**Rate 4**
SkateboardInLot	2000	3300	6000	10000
Chairlift	1500	2400	4000	7000
KiteFlite	1500	2400	4000	7000
Harbor	1200	2000	3300	6000

NOTE – The set of test sequences and the definition of rate points may be updated in the final CfE.

2.3.2 Coding conditions for HEVC anchors and submissions

In this test case, the Random Access scenario as described in §2.1.2 is used for evaluation. The constraints defined in §2.1.3 for submissions apply.

Submissions to the CfE may use any type of projection mapping, but are restricted to match the same number of active luma samples as used in the anchors. The projection mapping may change dynamically within the sequence, if an automatic selection algorithm is used and described. The same projection mapping algorithm, including downsampling, shall be used for all sequences and bit rate points. The applied projection mapping algorithm should be documented in the input contribution. Additional information may optionally be provided using a different number of active luma samples. If global rotation or other multi-pass projection mapping is used, it should be described.

3 Evaluation methodology

Evaluation of the submissions in response to the Call for Evidence will be performed at the July 2017 JVET meeting in Torino, Italy.

The evaluation of the submissions to the CfE will be done using the Expert Viewing Protocol, as defined in Rec. ITU-R BT.2095-0 [10], with JVET participants serving as expert viewers.

In addition, proponents are required to submit an input contribution with documentation of PSNR values (at least average of frame PSNR for each test sequence and encoding point, separate for luma and chroma components, as well as Bjøntegaard Delta-Rate and Delta-PSNR [1][2] compared to the anchors) and, if possible, documentation of the compression technology.

NOTE – Excel templates for reporting of the results for each test case will be released one week after release of the anchor bitstreams.

3.1 HDR Video evaluation

In addition to the evaluation methodology described in §3, proponents are required to submit an input contribution with documentation of weighted PSNR values (at least average of frame wPSNR for each sequence and encoding point, separate for luma and chroma components), tPSNR-Y, deltaE100 and PSNR-L100. Metric definitions are provided in the JVET common test conditions and evaluation procedures for HDR/WCG video [8]. It is requested to also provide the Bjøntegaard Delta-Rate for each metric.

3.2 360º Video evaluation

For subjective evaluation, 2D rectilinear viewports will be extracted from the 360º × 180º omnidirectional video, using bi-linear interpolation, similar to the default viewport extraction used in the 360Lib software [6]. The 2D rectilinear viewports will be viewed on ordinary monitors, following the method described above for SDR content.

Dynamic rectilinear viewports are expected to be used, in which the yaw and pitch angles may change for each frame in the sequence. The particular dynamic viewports used for evaluation of each sequence will be selected after the submission of YUV files. If the projection and packing format used in a submission is not supported in the 360Lib software, either 8K ERP format YUV files shall be provided for use in generating viewports, or a binary decoder shall be provided which has the capability to generate a dynamic rectilinear viewport using the same metadata input file format used in the 360Lib software [6].

Proponents are required to submit an input contribution with documentation of multiple objective metrics, including E2E WS-PSNR, E2E CPP-PSNR, E2E S-PSNR-I, E2E S-PSNR-NN, WS-PSNR, as described in JVET-E1003 [6]. Reporting CPP-PSNR, S-PSNR-I, and S-PSNR-NN is encouraged.

NOTE – In the final CfE, the set of measures may change.

4 Logistics

Prospective contributors of responses to the Call for Evidence should contact the following people:

Gary Sullivan (JVET co-chair)
Microsoft Corp.
1 Microsoft Way
Redmond, WA 98052 USA
Tel.: +1 425 703 5308, e-mail: garysull@microsoft.com

Jens-Rainer Ohm (JVET co-chair)
RWTH Aachen University, Institute of Communications Engineering
Melatener Str. 23, 52074 Aachen, Germany
Tel.: +49-241-8027671, e-mail: ohm@ient.rwth-aachen.de

Vittorio Baroncini (JVET test coordinator)
Technical Director
GBTech
Viale Castello della Magliana, 38, 00148 – Rome – Italy
Tel.: +39-3335474643 , e-mail: baroncini@gmx.com

Expressions of interest are requested by sending an e-mail to the contact persons above by **2017-06-16 at the latest**. Interested parties are kindly invited to express their intent of participation to the JVET chairs as early as possible.

Submitters should upload their contributions as regular input documents to JVET. The JVET chairs will provide assistance to submitters from outside JVET and will also assist in enabling them to attend the JVET meeting.

Test sequences and anchors (including configuration files used for encoding) will be made available by request by contacting one of the contact persons above.

References

[1] Gisle Bjøntegaard, "Calculation of Average PSNR Differences between RD curves", ITU-T SG16/Q6, 13th VCEG Meeting, Austin, Texas, USA, April 2001, Doc. VCEG-M33.

[2] Gisle Bjøntegaard, "Improvements of the BD-PSNR model", ITU-T SG16/Q6, 35th VCEG Meeting, Berlin, Germany, 16–18 July 2008, Doc. VCEG-AI11.

[3] "Requirements for a Future Video Coding Standard v4", ISO/IEC JTC1/SC29/WG11 MPEG, 115th Meeting, Geneva, June 2016, Doc. N16359.

[4] "Requirements for Future Video Coding (H.FVC)", Report SG16-R1 Annex I, ITU-T SG 16, Geneva, 16-27 January 2017. Also available here.

[5] "Algorithm Description of Joint Exploration Test Model 5", Joint Video Exploration Team (JVET) of ITU-T VCEG (Q6/16) and ISO/IEC MPEG (JTC 1/SC 29/WG 11), 5th Meeting, Geneva, January 2017, Doc. JVET-E1001.

[6] "Algorithm descriptions of projection format conversion and video quality metrics in 360Lib", Joint Video Exploration Team (JVET) of ITU-T VCEG (Q6/16) and ISO/IEC MPEG (JTC 1/SC 29/WG 11), 5th Meeting, Geneva, January 2017, Doc. JVET-E1003.

[7] "JVET common test conditions and software reference configurations", Joint Video Exploration Team (JVET) of ITU-T VCEG (Q6/16) and ISO/IEC MPEG (JTC 1/SC 29/WG 11), 2nd Meeting, San Diego, February 2016, Doc. JVET-B1010.

[8] "JVET common test conditions and evaluation procedures for HDR/WCG video", Joint Video Exploration Team (JVET) of ITU-T VCEG (Q6/16) and ISO/IEC MPEG (JTC 1/SC 29/WG 11), 5th Meeting, Geneva, January 2017, Doc. JVET-E1020.

[9] "JVET common test conditions and evaluation procedures for 360 video", Joint Video Exploration Team (JVET) of ITU-T VCEG (Q6/16) and ISO/IEC MPEG (JTC 1/SC 29/WG 11), 5th Meeting, Geneva, January 2017, Doc JVET-E1030.

[10] "Subjective assessment of video quality using Expert Viewing Protocol", ITU-R Recommendation BT.2095-0, Geneva, April 2016.

Glossary

CfE	Call for Evidence
CfP	Call for Proposals
ERP	Equirectangular projection
fps	Frames per second
HDR	High dynamic range
HEVC	High efficiency video coding (Rec. ITU-T H.265 | ISO/IEC 23008-2)
JVET	Joint Video Exploration Team
MPEG	Moving Picture Experts Group, Working Group 11 of ISO/IEC JTC 1/SC 29
PSNR	Peak signal-to-noise ratio
QP	Quantization parameter
VCEG	Video Coding Experts Group of ITU-T Question 6/16
VR	Virtual reality
WCG	Wide colour gamut