Joint Call for Evidence on Video Compression with Capability beyond HEVC

(Ref: JVET-F1002; Approved 2017-04-07)

1 Introduction

ITU-T VCEG (Q6/16) and ISO/IEC MPEG (JTC 1/SC 29/WG 11) are jointly studying the potential for standardization of video coding technology with a compression capability that significantly exceeds that of the HEVC standard (Rec. ITU-T H.265 | ISO/IEC 23008-2) and its current extensions. Such future standardization could take the form of additional extension(s) of HEVC or an entirely new standard.

To better coordinate this study, VCEG and MPEG created the Joint Video Exploration Team (JVET) as an informal collaboration. The scope of the JVET activity includes consideration of a variety of video sources and video applications. Example sources include camera-view content, screen content, consumer generated content, virtual reality/360º content, and high dynamic range content, while example applications include broadcast (with live or pre-authored content), real-time video conferencing, video chat, on-demand viewing, storage-based media replay, and surveillance with fixed or moving cameras [3][4].

This Call for Evidence (CfE) has been issued jointly as part of this study. The CfE requests information regarding video compression technology that has compression performance beyond that of HEVC. Responses to the CfE will be evaluated in July 2017, as further described below. Depending on the result of the evaluation, a formal Call for Proposals (CfP) is likely to be issued in preparation for starting a formal standardization project.

Companies and organizations who have developed compression technology that they believe to have compression capability better than that of the Main 10 Profile of the HEVC standard are kindly invited to bring such information to the JVET in response to this Call for Evidence. Additionally, contributions are also welcome regarding technology that better supports newly emerging application areas of video coding.

1.1 Timeline

- Test sequences available: 2017-04-21
- HEVC SDR and HDR anchors available: 2017-04-21
- HEVC 360º anchors available: 2017-04-28
- JEM anchors available: 2017-06-10
- Expression of interest to submit a response: 2017-06-16
- Submission of contributions (descriptive document): 2017-07-05
- Submission of bitstreams and binary decoders via FTP: 2017-07-05
- 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

Test cases for SDR, HDR, and 360º content are defined in the three subsections below. 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.

2.1 SDR

2.1.1 Sequence formats and frame rates

Table 1: UHD SDR test sequence example pictures

 UHD1: Crosswalk1	 UHD2: FoodMarket3	 UHD3: Tango1	 UHD4: CatRobot1
 UHD5: DaylightRoad	 UHD6: BuildingHall1	 UHD7: ParkRunning2	 UHD8: CampfireParty

Table 2: HD SDR test sequence example pictures

 HD1: BQTerrace	 HD2: RitualDance	 HD3: Timelapse	 HD4: BasketBallDrive
 HD5: Cactus			

Table 3: SDR test sequences

Sequence ID	Sequence name	Resolution	Frame count	Frame rate	Chroma format	Bit depth
UHD1	Crosswalk1	4096×2160	470	60	4:2:0	10
UHD2	FoodMarket3	4096×2160	720	60	4:2:0	10
UHD3	Tango1	4096×2160	600	60	4:2:0	10
UHD4	CatRobot1	3840×2160	600	60	4:2:0	10
UHD5	DaylightRoad1	3840×2160	600	60	4:2:0	10
UHD6	BuildingHall1	3840×2160	500	50	4:2:0	10
UHD7	ParkRunning2	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	10

Sequence ID	Sequence name	Resolution	Frame count	Frame rate	Chroma format	Bit depth
HD3	Timelapse	1920×1080	600	60	4:2:0	10
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	Sequence name	MD5Sum
UHD1	Crosswalk1	978a5dea90fe9125f6bce42aade55b61
UHD2	FoodMarket3	a3cb399a7b92eb9c5ee0db340abc43e4
UHD3	Tango1	2ebe6dbf052d7decbd64dc398895a880
UHD4	CatRobot1	03a89792693fd9ecfd72ef2590025e97
UHD5	DaylightRoad1	165c70e3008d37b9ff476e997297fc5e
UHD6	BuildingHall1	836a5a0558b24e8dde6b9a256e7aa468
UHD7	ParkRunning2	9de83b1bc2bca1afedb5342a2df572ba
UHD8	CampfireParty	b676cf8de483c1b890379976323f92af
HD1	BQTerrace	efde9ce4197dd0b3e777ad32b24959cc
HD2	RitualDance	a3cb399a7b92eb9c5ee0db340abc43e4
HD3	Timelapse	3d0c4a356e092b401032a8a0a6b2b48e
HD4	BasketballDrive	d38951ad478b34cf988d55f9f1bf60ee
HD5	Cactus	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	180	280	480	800
HD4	800	1200	2000	3500
HD5	500	800	1200	2000

2.1.2 Coding conditions for HEVC SDR anchors

In this test case, a Random Access scenario is used for evaluation and follows 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 24fps, 25fps or 30fps, 48 for 50fps, 64 for 60fps, and 96 for 100fps.

HEVC anchors are generated using the HM16.15 software package. A static quantization parameter (QP) setting is applied for generation of the anchors, though 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 SDR submissions

Submissions to the Call for Evidence shall obey the following rules:

– Encoded to within +/−2% of the target bit rates defined for the test case

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

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

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

– No use of preprocessing

– No use of postfiltering, unless it is part of the decoding process

2.1.4 Coding conditions for JEM SDR anchors

The JVET maintains a Joint Exploration Test Model (JEM) software package containing coding tools that are developed or studied in a coordinated test model [5]. JEM anchor bitstreams will be generated using this software package and will obey the coding conditions in Section 2.1.3. It is planned that the JEM 6.0 software package will be used to generate the anchors, though a later version may be used if available.

2.2 HDR

2.2.1 Sequence formats and frame rates

Table 6: HDR test sequence example pictures

 HDR1: Market3	 HDR2: ShowGirl2	 HDR3: Hurdles	 HDR4: Starting
 HDR5: Cosmos1			

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	50	4:2:0	10
HDR4	Starting	1920×1080	500	50	4:2:0	10
HDR5	Cosmos1	1920×856	240	24	4:2:0	10

NOTE – The capture frame rate of the HDR3 (Hurdles) and HDR4 (Starting) sequences was 100fps. However, these sequences are treated as 50fps sequences for the evaluation processes defined in this document.

Table 8: HDR test sequence md5sums

Sequence ID	Sequence name	MD5Sum
HDR1	Market3	c97abe47455fd12f6d6436cecfad7c7d
HDR2	ShowGirl2	44f1974d68f7799c71eea29fb72b245b
HDR3	Hurdles	bc3cba849d6f4ee74d39056600722aa5
HDR4	Starting	1cbc416696cb0dfcf4da9886eeb6a4a2
HDR5	Cosmos1	da4a2488c249720da0535f01c3693efa

Table 9: HDR target bit rates

	Target bit rates [kbit/s]			
Sequences	**Rate 1**	**Rate 2**	**Rate 3**	**Rate 4**
HDR1	750	1100	1500	2400
HDR2	380	550	900	1500
HDR3	450	700	1500	2000
HDR4	450	600	1000	1700
HDR5	500	900	1500	3000

2.2.2 Coding conditions for HEVC HDR anchors

In this test case, the Random Access scenario is used for evaluation. The description in Section 2.1.2 applies except that the generation of the anchor does not use a static quantization parameter (QP) setting. Instead, the configuration allows the QP value to vary spatially, where the variation is an explicit function of the average, local luma value. A one-time change of the quantization parameter from value QP to value QP+1 may also be applied in order to meet the defined target bit rates. The quantization parameter settings applied for the anchors will be reported.

NOTE – The configuration of the anchor also uses a static setting for the chroma quantization parameter settings that is different than the static configuration used for test sequences in Section 2.1.2.

2.2.3 Coding conditions for HDR submissions

Submissions to the Call for Evidence shall obey the constraints in Section 2.1.3 with the following exception:

– The quantization settings do not need to be kept static. Instead, the quantization settings may be adjusted within a frame as a function of the local, average luma value and/or the local, average chroma value. If local adjustment is used, a description of the adjustment scheme shall be provided in the descriptive document submission.

2.2.4 Coding conditions for JEM HDR anchors

The JVET maintains a Joint Exploration Test Model (JEM) software package containing coding tools that are developed or studied in a coordinated test model [5]. JEM anchor bitstreams will be generated using this software package and will obey the coding conditions in Section 2.2.3. It is planned that the JEM 6.0 software package will be used to generate the anchors, though a later version may be used if available.

2.3 360° Video

2.3.1 Sequence formats and frame rates

Table 10: 360º Test sequences

 VR1: SkateBoardInLot	 VR2: ChairliftRide	 VR3: KiteFlite
 VR4: Harbor	 VR5: Trolley	

Table 11: 360º video test sequences

Sequence ID	Sequence name	Input resolution	Anchor resolution	Coded luma sample count of anchors	Frame count	Frame rate	Chroma format	Bit depth
VR1	SkateBoardInLot	8192×4096	4096×2048	8388608	300	30	4:2:0	10
VR2	ChairliftRide	8192×4096	4096×2048	8388608	300	30	4:2:0	10
VR3	KiteFlite	8192×4096	4096×2048	8388608	300	30	4:2:0	8
VR4	Harbor	8192×4096	4096×2048	8388608	300	30	4:2:0	8
VR5	Trolley	8192×4096	4096×2048	8388608	300	30	4:2:0	8

NOTE – The sequences are omnidirectional 360º × 180º degree video and are stored in an equirectangular projection (ERP) format. The number of coded luma samples in the anchor is lower than the resolution of the input sequence.

Table 12: 360° video test sequence md5sums

Sequence ID	Sequence	MD5Sum
VR1	SkateboardInLot	e8eae04c43e959060f641fec4892fced
VR2	ChairliftRide	9126f753bb216a73ec7573ecc4a280c3
VR3	KiteFlite	18c0ea199b143a2952cf5433e8199248
VR4	Harbor	aa827fdd01a58d26904d1dbdbd91a105
VR5	Trolley	25c1082d1e572421da2b16530718156d

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	1200	2000	3300	6000
Chairlift	1500	2400	4000	7000
KiteFlite	1200	2400	4000	7000
Harbor	700	1200	2000	3300
Trolley	1500	2400	4000	7000

2.3.2 Coding conditions for HEVC 360º anchors

In this test case, the Random Access scenario is used for evaluation. The description in Section 2.1.2 applies.

2.3.3 Coding conditions for 360º submissions

Submissions to the Call for Evidence shall obey the constraints in Section 2.1.3 with the following exceptions:

– The quantization settings do not need to be kept static. Instead, the quantization settings may be adjusted within a frame as a function of the geometric position. If local adjustment is used, a description of the adjustment scheme shall be provided in the descriptive document submission.

– Pre-processing may be used to perform a projection mapping operation, and post-filtering may be used to perform an inverse projection mapping operation. The projection mapping algorithms may allow dynamic changes within a sequence if an automatic selection algorithm is used. The same projection mapping operation and inverse projection mapping operation shall be used for all test sequences in the test case. If projection mapping is used, a description of the projection mapping technique shall be provided in the descriptive document submission. Respondents are asked to provide information regarding at least: (i) the coded resolution of the projection map, (ii) the use of padding and blending, (iii) the use of global rotation, (iv) the use of multi-pass projection mapping, and (v) PSNR values comparing each test sequence to the result of applying the projection mapping algorithm and then converting this result back to the equirectangular projection format without compression.

2.3.4 Coding conditions for JEM 360º anchors

The JVET maintains a Joint Exploration Test Model (JEM) software package containing coding tools that are developed or studied in a coordinated test model [5]. JEM anchor bitstreams will be generated using this software package and will obey the coding conditions in Section 2.1.3, but will

use static quantization settings. It is planned that the JEM 6.0 software package will be used to generate the anchors, though a later version may be used if available.

3 Evaluation Methodology

Evaluation of the submissions in response to the Call for Evidence will be performed at the July 201s7 JVET meeting in Torino, IT.

Respondents are asked to make submissions including bitstreams, binary decoders, 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, as much as possible[1], documentation of the compression technology. Submissions must provide bitstreams for all sequences, and the binary decoder must be capable of decoding the bitstreams and storing the decoded data in the same format as the test sequence.

The evaluation methodologies to visually assess the quality of the received submissions are described below, detailing the assessment of SDR, HDR and 360 video content. Please note that some changes to the methodology may be employed in order to complete the evaluation by the end of the July 2017 meeting. For example, the evaluation may be modified if a large number of submissions are received.

3.1 SDR Video evaluation

The evaluation of the submissions to the Call for Evidence for SDR content (both UHD and HD formats) will be done by assessing a set of representative video clips that are selected from the submissions and determined by the JVET experts to properly represent a SDR content compression use case. The subjective assessment of the received submissions will be done by an expert panel either before or during the July meeting. The method used will tentatively be the EVP protocol as described in Recommendation ITU-R BT.2095; an alternative will be to use the DCR method as described in Recommendation ITU-T P.910. A panel of at least nine experts will be selected among the available volunteers to participate in the evaluations.

3.2 HDR Video evaluation

In addition to the evaluation methodology described in Section 3 and Section 3.1, respondents to the Call for Evidence for HDR content are further asked to make submissions including the following metrics: 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, as well as to provide the Bjøntegaard Delta-Rate and Delta-PSNR for each metric.Metric definitions are provided in the JVET common test conditions and evaluation procedures for HDR/WCG video [8].

3.3 360º Video evaluation

The evaluation of submissions to the Call for Evidence for 360º content will use the methodology described in Section 3 and Section 3.1 but with a slightly shorter viewing distance to better represent the field of view of head-mounted displays (HMDs). Two-dimensional rectilinear viewports will be extracted from the 360º × 180º omnidirectional video using bi-linear interpolation. The method will be similar to the default viewport extraction used in the 360Lib software [6].

Dynamic rectilinear viewports are expected to be used for the viewport extraction, 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 the bitstreams. If the projection and packing format used in a submission is not supported in the 360Lib software,

[1] Though it is not mandatory to describe the underlying technology in detail, the description should allow an assessment to understand its relevance for prospective standardization, e.g. in terms of processing complexity, memory usage, encoding/decoding delay, relation with existing video compression technology, licensing conditions, etc.

respondents are asked to provide a binary decoder that can either store the decoded data in an 8K, equirectangular projection (ERP) format or with the capability to generate a dynamic rectilinear viewport using the same metadata input file format as the 360Lib software.

Respondents to the Call for Evidence for 360º content are further required to make submissions including the following metrics: E2E WS-PSNR, E2E CPP-PSNR, E2E S-PSNR-I, E2E S-PSNR-NN, WS-PSNR, as described in [6]. Reporting of the CPP-PSNR, S-PSNR-I, and S-PSNR-NN metrics is further encouraged.

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, email 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, email ohm@ient.rwth-aachen.de

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

Expressions of interest to submit a response shall be made by contacting the people above on or before 16 June 2017. Interested parties are kindly invited to express their intent as early as possible.

Details on how to format and submit documents, bitstreams, and other required data will be communicated directly to those who express an interest of participation. Additionally, the JVET chairs will provide assistance to submitters from outside JVET in order to attend the JVET meeting.

Test sequences, anchors, and configuration files will also be made available by contacting one of the above individuals.

5 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)", Annex Q6.B of report of Q6/16 and TD 8R1/WP3, ITU-T SG 16, Geneva, 16-27 January 2017.

[5] "Algorithm Description of Joint Exploration Test Model 6 (JEM6)", Joint Video Exploration Team (JVET) of ITU-T VCEG (Q6/16) and ISO/IEC MPEG (JTC 1/SC 29/WG 11), 6th Meeting, Hobart, April 2017, Doc. JVET-F1001.

[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), 6th Meeting, Hobart, April 2017, Doc. JVET-F1003.

[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), 6th Meeting, Hobart, April 2017, Doc. JVET-F1020.

[9] "JVET common test conditions and evaluation procedures for 3600 video", Joint Video Exploration Team (JVET) of ITU-T VCEG (Q6/16) and ISO/IEC MPEG (JTC 1/SC 29/WG 11), 6th Meeting, Hobart, April 2017, Doc JVET-F1030.

[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