INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11
MPEG2011/N12036
Geneva, Switzerland
March 2011

Source: **Video and Requirement**
Title: **Call for Proposals on 3D Video Coding Technology**
Status: **Approved**

1 Introduction

This document is a Call for Proposals (CfP) on 3D Video Coding (3DVC) technology providing efficient compression and high quality view reconstruction of an arbitrary number of dense views. This CfP has been issued by ISO/IEC JTC1/SC29/WG11 (MPEG), and the evaluation of submissions will be carried out at the 98th MPEG meeting after formal subjective evaluation.

2 Purpose and Procedure

A new generation of 3D Video Coding (3DVC) technology that goes beyond the capabilities of existing standards to enable both advanced stereoscopic display processing and improved support for auto-stereoscopic multi-view displays is targeted. More background information, as well as information about applications and requirements, is given in [4][5][6].

The primary goal is to define a data format and associated compression technology to enable the high-quality reconstruction of synthesized views for 3D displays. It is recognized that technology for depth estimation and view synthesis, as well as the data format itself, has a significant impact on the reconstruction capability and quality of reconstructed views. Therefore, contributions on such technology are also of interest.

To evaluate the proposed technologies, formal subjective tests will be performed. Results of these tests will be made public, taking into account that no direct identification of any of the proponents will be made (unless it is specifically requested or authorized by a proponent to be explicitly identified). Prior to having evaluated the results of the tests, no commitment to any course of action regarding the proposed technology can be made.

Descriptions of proposals shall be registered as input documents to the 98th MPEG meeting (see timeline in section 3). Proponents need to attend this meeting to present their proposals. Further information about logistical steps to attend the meeting can be obtained from the listed contact persons (see section 11).

3 Timeline

The timeline of the Call for Proposals is as follows:

2011/01/28:	Draft Call for Proposals (public)
2011/04/15:	Final Call for Proposals
2011/04/15:	Formal registration period opens. Proponents shall indicate the number of submissions and test categories (i.e., AVC-compatible, HEVC-compatible, or unconstrained).
2011/07/20:	Formal registration period ends. An invoice for the testing fee (see section 9) will be sent by the test facility shortly after the 97th MPEG meeting.
2011/08/15:	Deadline to withdraw registration.
2011/09/01:	Coded test material shall be available at the test site. By this date, the payment of the testing fee is expected to be finalized.
2011/10/01:	Subjective assessment starts
2011/11/21:	Registration of documents describing the proposals[1]
2011/11/22:	Submission of documents
2011/11/26-12/02:	Evaluation of proposals at 98th MPEG meeting[2]

4 Test Material, Coding Classes and Anchors

Proponents are required to submit complete results for all test cases.

4.1 Test Material

The data sets from Annex A will be used for evaluation. All data sets have a linear camera arrangement and the data sets are rectified[3]. All video test material is progressively scanned and uses 4:2:0 color sampling with 8 bits per sample. The depth data and camera parameters for view synthesis and rendering are provided. The depth maps have the same pixel resolution and bit-depth as the video data. More information about file formats could be found in Annex B. Depth maps generated by other depth estimation algorithms are not permitted as additional input to the encoder.

[1] Contact persons will provide information about document submission process. Note that submitted documents will be made publicly available. Exceptions to public availability will be considered on a case by case basis upon request by the contributor.

[2] Proponents are requested to attend this standardization meeting. The starting date may be postponed by one day, participants will be notified.

[3] Rectified means that the images are properly registered by applying a homography matrix.

4.2 Coding Conditions of Submissions

All submission will be evaluated by a formal subjective testing that is fully described in section 7. The video data that is evaluated in the subjective viewing will originate from a dense set of synthesized views that are provided by the proponents and will be fed into 3D displays.

In order to evaluate the benefits of a proposed data format and corresponding compression technology, video data and associated depth maps are provided as input. The proposed data format shall be reconstructed from a bitstream as an output of a decoding process. The reconstructed data format is then used by a view synthesis algorithm to generate the synthesized views. Proponents may use the reference software that is provided [10] or a view synthesis algorithm that is optimized for the proposed data format.

The output of the view synthesis will be evaluated in the subjective tests in order to assess the combined merits of the proposed data format, compression technology and view synthesis algorithm. The objective quality of the reconstructed video of the decoder will also be measured as an indication of overall compression efficiency according to the PSNR of each individual view relative to the input video.

4.3 Test Scenarios and Requirement on Submissions

The following test classes, test scenarios and test categories are defined.

Test Classes

Class A:	1920x1088p 25fps: "Poznan_Street", "Poznan_Hall2", "Undo_Dancer", "GT_Fly"
Class C:	1024x768p 30fps: "Kendo", "Balloons", "Lovebird1", "Newspaper"

Test Scenarios

2-view:	refers to the 2-view input configuration
3-view:	refers to the 3-view input configuration

Test Categories

AVC-Compatible:	refers to submissions in which the compressed data format satisfy the requirement on forward compatibility with AVC [5]
HEVC-Compatible & Unconstrained:	refers to submissions in which the compressed data formats satisfy the requirement on forward compatibility with HEVC, or submissions without any compatibility constraints [5]

Submissions can be made for either or both of the test categories. For each test category, all test classes and test scenarios for all rate points associated with category as defined in section 5 must be submitted.

The submissions will be evaluated on both stereoscopic as well as autostereoscopic displays. The 2-view test scenario will be evaluated on a stereo display, while the 3-view test scenario will be evaluated on both an autostereoscopic display as well as a stereo displays. For logistic reasons it may be necessary to reduce the number of cases that are subjectively assessed. This may be done by

showing only a subset of synthesis results on the autostereoscopic display, or omitting the viewing on autostereoscopic displays altogether.

It is not required that the bit stream for the 2-view test scenario is a subset of the bit stream for the 3-view test scenario.

Submissions shall include rendered views and bit streams for all sequences in all classes, and each decoding/rendering shall produce the full specified number of pictures and views for the sequence set (i.e., no missing pictures or views).

Submissions should follow constraints given in section 5.

4.4 Anchors

Anchors have been generated by encoding the above sequences using an MVC encoder (JMVC 8.3.1) and HEVC encoder with the high efficiency and random access encoder configuration (HM 2.0). The purpose of the anchors is to facilitate testing in accordance with BT.500 [9], providing useful reference points demonstrating the behavior of well-understood configurations of current technology. For submissions that satisfy forward compatibility, these anchors obey the same constraints as imposed on the proposals. The anchors will be among the encodings used in the testing process; however the purpose of the test is to compare the quality of video for proposals to each other rather than to the anchors.

For anchor coding in the AVC-Compatible test category, MVC is used for video data and separately for depth data. For anchor coding in the HEVC-Compatible test category, HEVC is applied independently to the video data and depth data.

Anchor bit streams, decoders, the view synthesis tool (VSRS), all necessary configuration files and utilities are provided at the ftp-site [10] for all test classes and scenarios.

5 Test Conditions & Parameters

5.1 Input Data

The input data for all submissions includes both video data and depth data as supplementary data for multiple views. The input views for both the 2-view and 3-view test scenarios are provided in Table 1.

Table 1: Input Views for Test Scenarios

Seq. ID	Test Sequence	2-view input	3-view input
S01	Poznan_Hall2	7-6	7-6-5
S02	Poznan_Street	4-3	5-4-3
S03	Undo_Dancer	2-5	1-5-9
S04	GT_Fly	5-2	9-5-1
S05	Kendo	3-5	1-3-5
S06	Balloons	3-5	1-3-5
S07	Lovebird1	6-8	4-6-8
S08	Newspaper	4-6	2-4-6

5.2 Coding Conditions

Submissions shall include encodings for all sequences in all classes and all test scenarios, and each decoding shall produce the full specified number of pictures for the sequence (no missing pictures).

Submissions shall also, for each of the test classes and scenarios defined above, submit results for the target rate points (which are not to be exceeded) according to the test category being submitted to. Submissions to the AVC-Compatible test category shall submit results for the rate points as listed in Table 2, while submissions to the HEVC-Compatible test category shall submit results for the rate points as listed in Table 3.

Table 2: Coding Conditions for 2-view and 3-view test scenarios for AVC-Compatible submissions

Seq. ID	**Test Sequence**	**2-view test scenario Bit rates** (kbps)				**3-view test scenario Bit rates** (kbps)			
		R1	R2	R3	R4	R1	R2	R3	R4
S01	Poznan_Hall2	500	700	1000	1500	750	900	1300	2300
S02	Poznan_Street	500	700	1000	1250	750	1100	1800	4000
S03	Undo_Dancer	1000	1300	1700	2200	1380	1750	2300	2900
S04	GT_Fly	1200	1700	2100	2900	2000	2380	2900	4000
S05	Kendo	400	500	800	1300	800	1000	1300	1900
S06	Balloons	320	430	600	940	500	600	800	1250
S07	Lovebird1	375	500	750	1250	500	800	1250	2000
S08	Newspaper	400	525	800	1300	500	700	1000	1350

Table 3: Coding Conditions for 2-view and 3-view test scenarios for HEVC-compatible and unconstrained submissions

Seq. ID	**Test Sequence**	**2-view test scenario Bit rates** (kbps)				**3-view test scenario Bit rates** (kbps)			
		R1	R2	R3	R4	R1	R2	R3	R4
S01	Poznan_Hall2	140	210	320	520	210	310	480	770
S02	Poznan_Street	280	480	800	1310	410	710	1180	1950
S03	Undo_Dancer	290	430	710	1000	430	780	1200	2010
S04	GT_Fly	230	400	730	1100	340	600	1080	1600
S05	Kendo	230	360	480	690	280	430	670	1040
S06	Balloons	250	350	520	800	300	480	770	1200
S07	Lovebird1	220	300	480	830	260	420	730	1270
S08	Newspaper	230	360	480	720	340	450	680	900

Submissions to the call shall obey the following additional coding constraints:

1. Pre-processing for all input video data (e.g. denoising, color correction, filtering) is highly discouraged. If it is necessary as part of a data format conversion, it is allowed and shall be documented in detail. An example pre-processing is down-sampling of a particular video such that the resolutions among input data can be different. Any non-automatic pre-processing is not permitted; this applies to both video and depth data.
2. Only use post-processing if it is part of the decoding and synthesis process or required for the data format conversion. Any such post-processing must be documented.
3. Quantization settings should be kept static. When change of quantization is used it shall be described.

4. Proponents are discouraged from optimizing encoding parameters and any processing steps using non-automatic means.
5. The video and depth data coding shall not be used as the training set for training large entropy coding tables, VQ codebooks, etc.
6. Usage of multi-pass encoding is limited to the picture level and must be documented.
7. The interval between two successive random access points must not be greater than 0.5 seconds.

5.3 Rendering Conditions (View Synthesis)

Submissions shall produce synthesized views using a view synthesis algorithm, which may be either the VSRS software [6] or their own method, for all sequences in all classes and all test scenarios, based on the decoded output. The views to be synthesized for both stereoscopic and autostereoscopic displays are specified in Table 4 (no missing or duplicate views).

The VSRS configuration files that are used for the anchors for each sequence, as well as the corresponding camera parameter files are provided on the ftp site [10]. The camera parameters are provided for each sequence along with the anchors with 5-digit floating point precision. The synthesized views shall represent the scene as if the cameras were positioned according to the views indicated in Table 4.

Table 4: Synthesized output views for stereoscopic and autostereoscopic displays

Seq. ID	Test Sequence	View to Synthesize from 2-view test scenario (and stereo pair)	Views to Synthesize from 3-view test scenario (and stereo pair)
S01	Poznan_Hall2	6.5 (6.5-6)	All 1/16 positions between views 7 and 5 (6.125-5.875)
S02	Poznan_Street	3.5 (3.5-3)	All 1/16 positions between views 5 and 3 (4.125-3.875)
S03	Undo_Dancer	3 (3-5)	All 1/4 positions between views 1 and 9 (4.5-5.5)
S04	GT_Fly	4 (4-2)	All 1/4 positions between views 9 and 1 (5.5-4.5)
S05	Kendo	4 (4-5)	All 1/8 positions between views 1 and 5 (2.75-3.25)
S06	Balloons	4 (4-5)	All 1/8 positions between views 1 and 5 (2.75-3.25)
S07	Lovebird1	7 (7-8)	All 1/12 positions between views 4 and 8 (5.75-6.25)
S08	Newspaper	5 (5-6)	All 1/12 positions between views 2 and 6 (3.75-4.25)

6 Submission Requirements

More information about file formats can be found in Annex B. Files of decoded sequences and bitstreams shall follow the naming conventions as specified in Annex C.

Proponents shall provide the following; incomplete proposals will not be considered:

A) Coded test material submission

The following material must be received by the test facility on hard disc by 2011 September 1 (17:00 CET). It must also be brought to the 98th MPEG meeting.

1. Bitstreams for all test classes and test scenarios that satisfy rate constraints specified in Section 5.2 for both the 2-view and 3-view test scenarios.

2. Decoded and synthesized sequences for all test classes, test scenarios and rate points, including decoded input views as well as synthesized views as specified in Table 4 for both stereoscopic and autostereoscopic viewing. Further details are specified below for each test scenario.

   For the 2-view test scenario, submissions must package the stereo pair indicated in Table 4 as a single AVI file for each sequence in all the test classes for all the rate points with the Stereo2AVI utility provided. Separate YUV files must also be provided.

   For the 3-view test scenario, all views shall be generated according to the positions indicated in Table 4[4]. Separate YUV files must be provided to facilitate both the stereo and autostereoscopic viewing described in section 7.

   For the 3-view test scenario and stereo viewing, one fixed stereo pair and one randomly selected stereo pair will be evaluated for each sequence. Submissions must package the center stereo pair indicated in Table 4 as a single AVI file for each sequence in all test classes and for all rate points with the Stereo2AVI utility that is provided [10]. For the randomly selected stereo pair, the random selection will be made by the test facility and the test facility will produce the single AVI file from the YUV files that are provided. The same random selection shall be used for all proposals. For the 3-view test scenario with autostereoscopic display, the test facility will package 28 views from the views that are provided. The same 28 views shall be used for all proposals.

3. Binary decoder executable. For submissions that use an alternative rendering algorithm, a rendering executable must also be provided.

4. Checksum files for 1.-3.

   B) Document to be submitted before the 98th MPEG meeting shall contain:

1. A technical description of the proposal sufficient for full conceptual understanding and generation of equivalent performance results by experts and for conveying the degree of optimization required to replicate the performance. This description should include all data processing paths and individual data processing components used to generate the bitstreams. It does not need to include complete bitstream format or implementation details, although as much detail as possible is desired.

2. An Excel sheet with all fields for the respective test cases filled for unified reporting of a proposal's coding results (measuring bitrates, PSNR and execution time measures) is provided in [10]. BD measures [7][8] against the appropriate anchor will be automatically computed from the Excel sheets.

3. The technical description shall also contain a statement about the programming language in which the software is written, e.g. C/C++ and platforms on which the binaries were compiled.

[4] There will be a total of 33 views for S01-S06, and a total of 49 views for S07-S08.

4. The technical description shall also include a section about fulfilment of requirements from the requirements document [5]. Specifically, it must indicate the test category of the submission, i.e., compatible or unconstrained.

5. The technical description shall include a complexity assessment of the main modules used. In particular, similarities and differences between the core techniques and architecture to encode the video data and supplementary data should be provided in order to assist in the assessment of common technology used for different data components. The following information must also be provided:

- Encoding time[5] (for each submitted bitstream) of the software implementation of the proposal, and the anchor run on the same platform for corresponding conditions. Proponents shall provide a description of the platform and methodology used to determine the time. To help interpretation, a description of software and algorithm optimisations undertaken, if any, is welcome.
- Decoding time[6] for each bitstream running the software implementation of the proposal, and for the corresponding constraint case anchor bitstream(s) run on the same platform.
- Rendering time[7] (including any processing required for rendering of output views) for all synthesis results running the software implementation of the proposal, and for the corresponding reference view synthesis used for the anchors run on the same platform. Proponents shall provide a description of the platform and methodology used to determine the time. To help interpretation, a description of software optimisations undertaken, if any, is encouraged.
- Expected memory usage of encoder and decoder (including any algorithms utilized for rendering of output views).
- Complexity characteristics of encoder, decoder, view synthesis algorithms, and any kind of pre-processing or post- processing that is used in the submission.

6. The technical description shall indicate any format or architecture-specific dependencies to encode the supplemental data in order to assist in the assessment of potential compatibility with codecs different than the one being proposed for encoding the video data.

Furthermore, the technical description should point out any specific properties of the proposal (e.g., view scalability in the sense that the bit stream for the 2-view case is a subset of the bit stream for the 3-view case).

C) Optional information

Proponents are encouraged (but not required) to allow other committee participants to have access, on a temporary or permanent basis, to their encoded bitstreams and binary executables or source code.

---

[5] For example, using ntimer for Windows systems.

[6] The decoder source code to be used to process the anchor bitstreams will be provided to proponents and must be compiled as-is, without modification of source code, compiler flags, or settings

[7] The view rendering source code to be used to process the anchor view synthesis will be provided to proponents and must be compiled as-is, without modification of source code, compiler flags, or settings

7 Subjective Viewing Requirements

Submissions will be evaluated through formal subjective testing on both stereoscopic and autostereoscopic displays. A complete description of the testing environment and methodology is provided in Annex C.

Since submissions submitted to the different test categories have different constraints and the quality range is expected to be different, all submissions will be analyzed according to the test category.

Information on the specific viewing tests that will be conducted is described below.

7.1 Stereoscopic Viewing

For the stereoscopic viewing tests, stereo pairs generated from each data set are selected.

In the 2-view test scenario, the stereo pair is formed with one of the decoded views among the input views and the synthesized view as specified in Table 4. In this case, there is always one decoded view and one synthesized view.

In the 3-view test scenario, a stereo pair is formed from two distinct viewpoints within the range of the input views specified in Table 1. The two viewpoints shall be selected so they are comfortable to view. The baseline distance for the stereo pairs to be viewed will be fixed for each sequence. Two stereo pairs will be selected: one that is centered around the center input view specified in Table 1, and another that is randomly with the same baseline distance between left and right views. In contrast to the 2-view test scenario, both views in this test scenario may be synthesized.

It is expected that the results of the stereoscopic viewing produce the most meaningful results in terms of overall evaluation of proposals.

7.2 Auto-Stereoscopic Viewing

For autostereoscopic viewing tests, 28 views generated from each data set are selected. The 28 views are formed from the set of views including the decoded views and the synthesized views specified in Table 4. The 28 viewpoints are selected so that they are comfortable to view and provide a sufficient depth range, i.e., there is a sufficiently wide baseline distance between the 1st view and the 28th view.

8 Test Sites and Delivery of Test Material

The proposals submission material will be evaluated by means of a formal subjective assessment process. The tests will be conducted at several test labs to be determined by the test coordinator before the 97th MPEG meeting.

All proponents need to deliver, by the due date of 2011 September 1 (17:00 CET), a hard drive to the address of the Test Coordinator (see section 11). The disk shall contain all the material described in section 6.

The correct reception of the disk will be confirmed by the Test Coordinator. Any inconvenience caused by unexpected delivery delay or a failure of the disk will be under the complete

responsibility of the proponents, but solutions will be negotiated to ensure that the data can still be included in the test, which means that correct and complete data need to be available before the beginning of the test at latest.

Utilities for YUV to AVI file format conversion and check-sum program will be made available (see Annex B). Further technical details on the delivery of the coded material are also provided in Annex B.

9 Testing Fee

Proponents will be charged a fee per submitted algorithm proposal for each test category. Such fee will be a flat charge for each proposal to cover the logistic cost (without any profit). The fee is non-refundable after the formal registration is made. The fee will be determined during the 97th MPEG meeting.

10 Source Code and IPR

Proponents are advised that, upon acceptance for further evaluation, it will be required that certain parts of any technology proposed be made available in source code format to participants in the core experiments process and for potential inclusion in the prospective standard as reference software. When a particular technology is a candidate for further evaluation, commitment to provide such software is a condition of participation. The software shall produce identical results to those submitted to the test. Additionally, submission of improvements (bug fixes, etc.) is certainly encouraged.

Furthermore, proponents are advised that this Call is being made subject to the patent policy of ISO/IEC (see ISO/IEC Directives Part 1, Appendix I) and the other established policies of the standardization organization.

11 Contacts

Contact person:

Prof. Dr. Jens-Rainer Ohm
RWTH Aachen University, Institute of Communications Engineering
Melatener Str. 23, 52074 Aachen, Germany
Tel. +49-241-8027671, Fax. +49-241-8022196, email ohm@ient.rwth-aachen.de

Test Coordinator:

Dr. Vittorio Baroncini
Senior Researcher, Audio Video Signal Processing Area
Fondazione Ugo Bordoni
Via B. Castiglione, 59 00142 – Rome - Italy
Tel. +39-06-54802134, Fax. +39-06-54804405, email vittorio@fub.it

12 References

[1] ISO/IEC JTC1/SC29/WG11, "Text of ISO/IEC 14496-10:200X/FDAM 1 Multiview Video Coding", Doc. N9978, Hannover, Germany, July 2008.

[2] ISO/IEC JTC1/SC29/WG11, “Updated Call for Proposal on Multi-view Video Coding”, Doc. N7567, Nice, France, October 2005.

[3] ISO/IEC JTC1/SC29/WG11, Call For Proposals on Scalable Video Coding Technology N6193, Waikoloa, HI, December 2003.

[4] ISO/IEC JTC1/SC29/WG11, “Vision on 3D Video”, Doc. N10357, Lausanne, Switzerland, February 2009.

[5] ISO/IEC JTC1/SC29/WG11, “Applications and Requirements on 3D Video Coding”, Doc. N12035, Geneva, Switzerland, March 2011.

[6] ISO/IEC JTC1/SC29/WG11, “Report on Experimental Framework for 3D Video Coding”, Doc. N11631, Guangzhou, China, October 2010.

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

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

[9] International Telecommunication Union Radio Communication Sector; Recommendation ITU-R BT.500-11

[10] ftp://ftp.merl.com/pub/avetro/3dv-cfp/

Annex A: Description of Test Sequences

Class A: 1920x1088, 25fps

Seq. ID	Seq. name	No. Frames	Camera Arrangement	Provider
S01	Poznan_Hall2	200 frames	9 cameras with 13.75 cm spacing, moving camera array	Poznan
S02	Poznan_Street	250 frames	9 cameras with 13.75 cm spacing	
S03	Undo_Dancer	250 frames	Computer generated imagery with ground truth depth data	Nokia
S04	GT_Fly	250 frames	Computer generated imagery with ground truth depth data	

Class C: 1024x768, 30fps

Seq. ID	Seq. Name	No. Frames	Camera Arrangement	Provider
S05	Kendo	300 frames	7 cameras with 5 cm spacing, moving camera array	Nagoya
S06	Balloons	300 frames	7 cameras with 5 cm spacing, moving camera array	
S07	Lovebird1	240 frames	12 cameras with 3.5 cm spacing	ETRI / MPEG Korea Forum
S08	Newspaper	300 frames	9 cameras with 5 cm spacing	GIST

The complete set of multiview video test sequences (include best available depth map data) that are being used for our experiments are listed below, along with ftp access information and copyright statements.

Poznan University of Technology
ftp://multimedia.edu.pl/3DV/
username: 3DV / password: ftvftv
directory: CFP

Nokia
ftp://mpeg3dv.research.nokia.com
username: mpegmember / password: S9"12#sHD)3

Copyright: The supplied data and content ("Supplied Data") is provided free of charge and made available for use by the Licensee, who shall be a member of the MPEG standardization committee or a respondent to a the Call for Proposals for the MPEG standard for 3D video coding that is to be issued March 2011, under the following conditions:

The Licensee agrees that the Supplied Data and all intellectual property rights therein remain the property of Nokia Corporation and its licensors (owners). The Supplied Data may only be used for the purpose of responding to the Call for Proposals that is to be issued March 2011, developing, testing and promulgating the MPEG technology standard for 3D video coding resulting from the Call for Proposals that is to be issued March 2011. The respective owners make no warranties with respect to the Supplied Data and expressly disclaim any warranties regarding its fitness for any purpose.

As a way of promulgating the MPEG technology standard for 3D video coding, the Licensee may present parts or modifications of the Supplied Data at academic conferences and publications.

The Licensee agrees not to provide the data to any third parties without permission from the owners and that the data shall not be sold, let for hire, or by way of trade, offered or exposed for sale or hire. This restriction shall apply to the original material or to any reproduction of it in whole or in part and to any modifications of the Supplied Data.

The texture views of the Supplied Data are copyright © UNDO. (See http://www.undo.fi)

Nagoya University
http://www.tanimoto.nuee.nagoya-u.ac.jp/~mpegftv/mpeg3dv/CfP/
username: mpegftv / password: fngOyfTv

Copyright: Individuals and organizations extracting sequences from this archive agree that the sequences and all intellectual property rights therein remain the property of the respective owners listed below. These materials may only be used for the purpose of developing, testing and promulgating technology standards. The respective owners make no warranties with respect to the materials and expressly disclaim any warranties regarding their fitness for any purpose.

Owners: Nagoya University, Japan.

GIST
ftp://203.253.128.142
username: 3DV / password: 3dvkr
directory: /GIST_Test_Sequence/Newspaper

Copyright: Individuals and organizations extracting sequences from this archive agree that the sequences and all intellectual property rights therein remain the property of the respective owners listed below. These materials may only be used for the purpose of developing, testing and

promulgating technology standards. The respective owners make no warranties with respect to the materials and expressly disclaim any warranties regarding their fitness for any purpose.

Owners: Gwangju Institute of Science and Technology (GIST), 1 Oryong-dong, Buk-gu, Gwangju, 500-712, Republic of Korea.

ETRI / MPEG Korea Forum
ftp://203.253.128.142
username: 3DV / password: 3dvkr
directory: /ETRI_MPEG_KOREA_Sequence/lovebird1

Owners: ETRI (Electronics and Telecommunications Research Institute), MPEG-Korea Forum, 161 Gajeong, Yuseong Gu, Daejeon, 305-700, Republic of Korea.

Annex B: Distribution formats for test sequences and decoded results, Delivery of Bitstreams and Binary Decoders, Utilities and Cross-check

Distribution of original video material files containing test sequences is done in YUV files with extension ".yuv". Colour depth is 8 bit per component. A description of the YUV file format is available at http://www.fourcc.org/ web site, designated as "planar iyuv". Depth map material files contain all the depth information in the luma component, the chroma components are set to 128 by default.

Anchor bitstreams are provided with extension ".bit". Bitstream formats of proposals can be proprietary, but must contain all information necessary to decode the sequences at a given data rate (e.g. no additional parameter files). The file size of the bitstream will be used as a proof that the bitrate limitations from Tables 2 and 3 have been observed. The file extension of a proposal bitstream shall be ".bit".

Decoded sequences shall be provided in the same ".yuv" format as originals, and additionally as AVI files (".avi" extension). A tool that converts YUV into AVI format is available from the test coordinator by request.

All files delivered (bitstreams, decoded sequences and binary decoders) must be accompanied by a checksum file to enable identification of corrupted files. MD5 checksum tools shall be used for that purpose. Such a tool is available typically as part of UNIX/LINUX operating systems where it should be run with option "-b" (binary). For Windows operating systems, a compatible tool can be obtained from http://www.pc-tools.net/win32/md5sums/. This tool should be run with additional option "-u" to generate the same output as under UNIX.

Hard disc should be shipped (for handling in customs) with a declaration "used harddisc for scientific purposes, to be returned to owner" and low value specification (e.g. 20€). The use of a harddisc with substantially larger size than needed is discouraged. The harddisc should be a 3½-inch SATA drive without any additional enclosure (no case, no power supply, no USB interface etc.), NTFS file format shall be used.

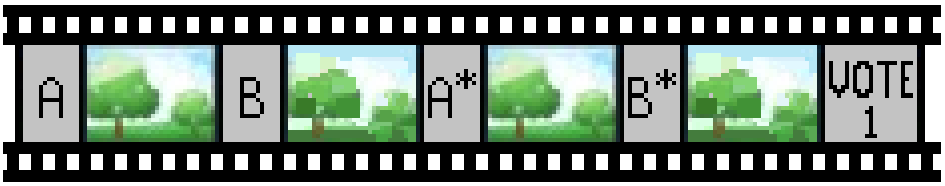
During the evaluation meeting, optional cross-checking of proposals may be done upon request. Proponents shall bring another hard-disc, which can be connected via USB 2.0 to a Windows PC, containing original and decoded sequences in YUV and AVI formats, bitstreams, binary decoder and view synthesis executables, and all related checksum files. Adequate computer systems will be made available at this meeting. Proponents shall specify the computing platform (hardware, OS version) on which the binary can be run. Should such a computing platform not be readily available, the proponent shall provide a computer adequate for decoder verification at this meeting. Further information will be exchanged with the proponents after the registration deadline.

Annex C – Description of Testing Method and laboratory set-up

Test Method

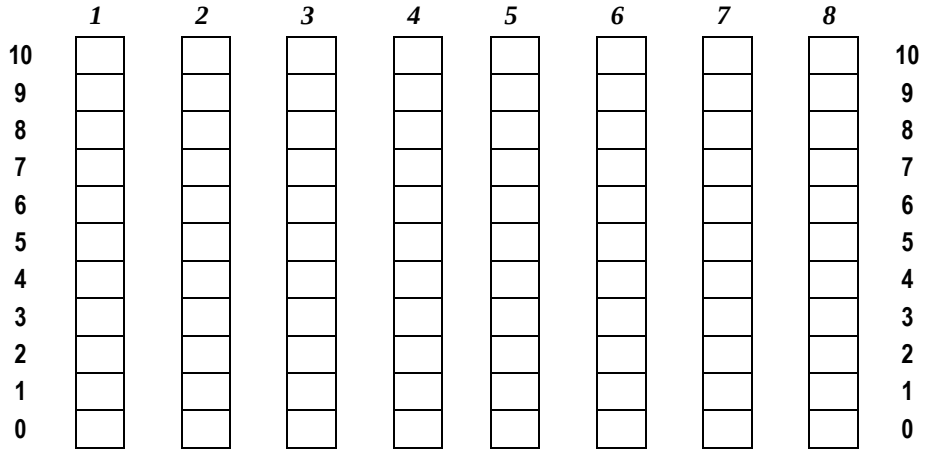
The Double Stimulus Impairment Scale (DSIS) test method will be used with 11 quality levels, where 10 indicates the highest quality and 0 indicates the lowest quality. The tests will be carried out with naïve viewers.

The Variant II of the DSIS method will be adopted; timing of the Basic Test Cell (BTC) is shown in figure here below.



Example of DSIS Variant II Basic Test Cell

The messages showing the letters A and A* will indicate to the viewers that the uncompressed original[8] is going to be shown, while the messages showing the letters B and B* will indicate to the viewers that the video to be evaluated is going to be shown. All the messages showing the letters will stand on the screen for 1 second. The message showing the sentence "Vote N" will stand for 6 seconds. The test subjects will be allowed to express their vote during that time period only. The quality votes will be expressed by the test subjects using paper scoring sheets (see example below).



Example of DSIS with 11 quality levels Scoring Sheet

[8] In the context of this 3D testing, an "uncompressed original" refers to video clips in which the synthesis result is generated without compression.

The scoring sheets will foresee a section (made of a column with 11 boxes) dedicated to the evaluation of each BTC; each column is numbered to allow an easy synchronization with the message present on the screen and the column to be used.

The whole test will be organized in test sessions. The design of the test sessions will be done in a way the first part of each session is made of a "stabilization" phase; the stabilization phase will be made of 4 BTC selected from those in that session, taking care of choosing one with high quality one with low quality and two of mid quality. This will allow subjects to create their quality impression in a more clear way before the actual test begins. The existence of the stabilization phase is not known to the test subjects.

Considering the particular care that a 3D visual test requires, the test sessions will be of a length never exceeding 20 minutes. The length and the design of the test sessions will be set by the Test Coordinator, taking into account the number of submission received by the Proponents and the number of testing laboratories available to run the tests. In this regard it is foreseen to create (for this testing effort only) an Independent Laboratory Group that will made of highly reliable test laboratories, managed by Test Managers whose past experience in visual testing (and in particular in 3D visual testing) provides a solid confidence of reliable results. The selection of the laboratories will be done taking to obtain a good balance of sites in the 3 world geographical areas (America, Europe and Far East). This is also done to perform 3D tests that will take into account also the different morphological characteristics of people of different continents.

The tests will be conducted for stereoscopic viewing and auto-stereoscopic viewing based on the viewing scenarios described in Section 7.

Raw Data Processing

Mean Opinion Scores (MOS) Standard Deviation (SD) and Confidence Interval (CI) with 95% precision will be computed on the raw data. Overlapping of CI values, centered on the MOS values, will indicate that, in a rather acceptable way, no Statistical Significant Difference exists between two test points.

Tables with the MOS and CI values for each test points will be provided together with graphs.

In order to respect of the privacy of the Proponents, the test points will be represented with codes that will indicate the coding condition and the Proponent, without any reference that allow to identify the Proponent; for this reason, after the formal registration will be confirmed, each Proponent will be assigned with a Test Code (e.g. P10, P11, … , PNN); the code P00 and P01 are reserved to the Anchor in the AVC-Compatible and the HEVC-Compatible & Unconstrained test category, respectively.

Test Subjects selection, screening and training

The tests will be conducted using naïve subjects only; the subjects will be selected among students aging form 18 to 30 years; all test subjects will be pre-screened for visual acuity, color blindness and stereo vision, according to what foreseen by the current literature.

The screening of the test subjects will consist of two steps: a first visual screening (as described earlier) and a post screening. The post screening will be done after the training phase is completed.

Training phase will be done of a short test session using test material and coding conditions similar to those of the actual test and will be made of just 5 or 6 BTC; The goal of the "training" session is to allow subjects to learn about the task they have to do and the timing of the voting. During the training session the test administrators are present and help subjects in their behavior; as training sessions is done, questions and doubts from subjects will be answered.

The "screening" 3D test session will be conducted, made of test material different form that of the actual test, but same length and similar content and impairments of an actual test session.

After the subjects have run the "screening" test sessions, the post screening will begin with the following two steps:

- analysis of the results of the "screening" test session;
- questions to the subjects about their physical and mental status (e.g. Any equilibrium disturbance? Any nausea? Which level of fatigue? Loss of concentration?); the questionnaire will be designed by the Test Coordinator and distributed to all the test sites.

After the above steps are completed the approved subjects will run the actual test sessions during the same day.

Laboratory set-up

The test areas in the test laboratories will be done of wide rooms with no reflecting ceiling walls and floor, and insulated by any external audible and visual pollution.

The dimension of the test area must allow for the following:

- deep enough to seat subjects at 4H for a 52' monitor (around four meters)
- have additional depth to put the monitor at least at 1 meter form the wall behind the monitor
- wide enough to seat 5 subjects in comfortable way

The wall behind the monitor has to be illuminated with a uniform light source, not directly hitting the viewers; the light level must not exceed the 5% of the monitor peak luminance (when functioning as stereo monitor); the light source has to be at 6500 K° and driven with an high frequency switching device (in any case not at power frequency; e.g. 50 or 60 HZ). The ambient illumination has to be provided only by the light illuminating the wall behind the monitor.

The subjects will be seated in front of the monitor in two different set-up according to the kind of monitor used (stereo or auto-stereo monitor) as follows.

Stereo monitor viewing set-up
Monitor: Hyundai 46" stereo display (interlaced stereo IF) with passive glasses. (Model: S465D)
Subjects per monitor: 2
Viewing distance: 4H
Angle of vision: < 60°

Class C test sequences will be displayed at the sequence resolution (1024x768) with a mid-grey frame. Viewers will be positioned according to the height of the active part of the screen.

Auto-stereo monitor viewing set-up
Monitor: Dimenco 52" auto-stereoscopic display with no glasses (Model: BDL5231V3D)
Subjects per monitor: 5 (central seat perpendicular to monitor)
Viewing distance: 4H
Angle of vision: < 120°

All test sequences shall be displayed at the full native resolution of the auto-stereoscopic display. The position of the viewers will therefore not change according to the class of the test sequence.

A set of common requirements complete the laboratory set-up, as follows:

3D server minimum requirements
PC based on I7 Intel CPU (1388 socket)
6G RAM (1333)
System disk (with OS and low speed storage) 2T
500 G SSD raid set (minimum 3 SSD disks)
Video Board (N-Vidia GTX 560 ti 1G RAM)
19" service monitor (full HD)

Proponents identification and file names

Each Proponent submitting to the CfP will be identified with a two digit code preceded by the letter "P" (e.g. P10 P11 … Pnn). The P00 code is reserved to the Anchor.

Each coded video file provided for a submission will be identified by a name formed by the below listed combination of letters and numbers:

For YUV Files [decoded output or synthesis result]

PnnCzSxxRyTmV%6.4f.YUV
where:

- Pnn identifies the Proponent;
- Cz identifies the z-view case (z=2 for 2-view case, z=3 for the 3-view case);
- Sxx identifies the original video clip used to produce the coded video, as identified in the tables of Annex A;
- Ry identifies the rate index y, as identified in **Błąd! Nie można odnaleźć źródła odwołania.** (i.e., R1, R2, R3, R4) – R0 is used to indicate an uncoded reference;
- Tm identifies the type of content, where m=t refers to texture, m=d refers to depth, m=s refers to a synthesized view; and
- V%6.4f identifies the view index, where %6.4f is a formatted floating point number of fixed length, e.g., V3.8750, V3.500, etc.

For AVI Files [bundled data for stereo or autostereoscopic viewing]

PnnCzSxxRyDi.AVI
where:

- Pnn identifies the Proponent;
- Cz identifies the z-view case (z=2 for 2-view case, z=3 for the 3-view case);

- Sxx identifies the original video clip used to produce the coded video, as identified in the tables of Annex A;
- Ry identifies the rate index y, as identified in **Błąd! Nie można odnaleźć źródła odwołania.** (i.e., R1, R2, R3, R4) R0 is used to indicate an uncoded reference; and
- Di identifies whether the file is for stereo (i=s) or autostereoscopic (i=a) viewing.

For BIT Files

PnnCzSxxRy.BIT
where:

- Pnn identifies the Proponent;
- Cz identifies the z-view case (z=2 for 2-view case, z=3 for the 3-view case);
- Sxx identifies the original video clip used to produce the coded video, as identified in the tables of Annex A; and
- Ry identifies the rate index y, as identified in **Błąd! Nie można odnaleźć źródła odwołania.** (i.e., R1, R2, R3, R4).

Any additional information that is required to uniquely identify the file should be added.