

**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/M21284  
July 2011, Torino, IT**

**Source: GIST (Gwangju Institute of Science and Technology)**  
**Status: Report**  
**Title: Comments on Evaluation of 3D Video Coding**  
**Author: Yunseok Song, Cheon Lee and Yo-Sung Ho**

## **1. Introduction**

This document contains our comments on Call for Proposals (CfP) on 3D Video Coding (3DVC) [1]. We address several aspects of evaluation described in the CfP. The main objective is to aid the overall evaluation and clarification of CfP.

## **2. Suggestions**

### **2.1 Depth information**

Regarding subjective evaluation, depth information as an extra characteristic can benefit evaluators in understanding how sequences are different. Numerically, depth information can be described by minimum and maximum disparities. According to Table 1, test sequences do not employ similar depth ranges. In particular, ‘Poznan\_hall2’ depth range is significantly greater than ‘Balloons’. We used three levels to indicate depth information: strong, intermediate and weak. ‘Strong’ implies steep depth range.

Table 1. Depth information for each test sequence (2-view)

<b>Test Sequence</b>	<b>Stereo Pair</b>	<b>Focal Length</b>	<b><math>Z_{near}</math></b>	<b><math>Z_{far}</math></b>	<b><math>d_{max}</math></b>	<b><math>d_{min}</math></b>	<b>Depth</b>
Newspaper	5-6	2929.494	2715.182	9050.605	50.0	15.0	Strong
Balloons	4-5	2241.256	448.251	11206.280	25.0	1.0	Weak
Kendo	4-5	2241.256	448.251	11206.280	25.0	1.0	Weak
Lovebird1	7-8	2017.807	2228.746	156012.207	35.0	0.5	Intermediate
Poznan_street	3.5-3	1732.876	34.506	2760.511	40.0	0.5	Intermediate
Poznan_hall2	6.5-6	1732.876	22.814	2760.513	60.5	0.5	Strong
Dancer	3-5	2302.853	2289.000	213500.000	40.2	0.4	Intermediate

## **2.2 Template sheet fields (Cfp\_Template\_3DV.xls)**

The template sheet includes fields for reporting coding results. Bitrates, PSNR and execution time measures are listed. We suggest two additional fields that would be beneficial.

### **A. PSNR between synthesized view and original view**

The current PSNR is defined to be measure of difference between reconstructed view and original view. In addition to that, the quality of resulting synthesized view is important as well. In order to take this into account, PSNR between synthesized view and original view would be suitable.

### **B. Bitrate ratio of texture and supplementary data (color and depth)**

Some minor computation with ‘Texture Rate’ and ‘Supl. Rate’, the ratio can be computed, providing useful information.

## **2.3 Unclear description**

There can be confusion from the description in ‘4.4 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.”

From the two sentences, the difference between ‘separately’ and ‘independently’ is not clear in the context. The explanation should be written in plain language.

## **3. Conclusion**

Concerning Cfp, we have pointed out three parts that could be revisited: providing depth information for subjective evaluation, adding two more fields in result template sheet, and clarifying description regarding anchor coding.

## **4. Acknowledgements**

This research was supported by a grant "Establishing the Foundation of International Standardization for Mobile Devices and Multimedia Convergence Technology" from Improvement of Standards Technology Program funded by Korean Agency for Technology and Standards, Ministry of Knowledge Economy of Korea.

## **5. References**

- [1] ISO/IEC JTC1/SC29/WG11 “Call for Proposals on 3D Video Coding Technology,” N12036, March 2011.