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

ISO/IEC JTC1/SC29/WG11
MPEG/M23791
February 2012, San Jose, USA

Title	**Nonlinear depth representation – extended results**
Sub group	**Video**
Authors	**Olgierd Stankiewicz** (ostank@multimedia.edu.pl), **Krzysztof Wegner, Marek Domański** Poznań University of Technology, Chair of Multimedia Telecommunications and Microelectronics, Poznań, Poland

1 Introduction

This documents presents extended results attained by Poznan University of Technology related to Non-linear depth representation tool [1].
This tools has been integrated into 3D-HTM software (as a part of integration plan) and also has been implemented as tested in 3D-ATM as a part of core experiment CE3 [4] defined in [2].

2 Description of the tool

M22697 describes a normative tool named non-linear depth representation. The depth is internally represented in such a way that the closer objects are represented more accurately than distant ones. Internal depth sample values are defined by the following power-law expressions, similar as in the case of well known gamma correction:

$$depth\ value\ internal = \left(\frac{depth\ value\ external}{maximum\ value\ external}\right)^{exponent} \cdot maximum\ value\ internal \quad (1)$$

Exponent is automatically chosen by the encoder with use of base QP for the depth and sent to decoder in the encoded bitstream (one parameter is added to i.e. SPS):

$$\text{exponent} = \text{clip}\left(\left(QP_{depth} - 30\right) \cdot 0.0125 + 1.25\ ;\ 1.0\ ;\ 1.66\right) \quad (2)$$

Depth map samples are represented on increased number of bits with use of IBDI (Internal Bit Depth Increase) tool. Finally, in the decoder the original depth is reconstructed with inverse expression:

$$depth\ value\ external = \left(\frac{depth\ value\ internal}{maximum\ value\ internal}\right)^{1/exponent} \cdot maximum\ value\ external \quad (3)$$

This tool is designed to improve subjective quality of the synthesized views and thus assessment by PSNR is not adequate and should be done subjectively instead.

The parameter *exponent* should be transmitted to the decoder in order decode depth. A non-normative rule that describes an example of the choice of the *exponent* value.

3 Assessment of the tool

Various research on non-linear depth representation, performed by us during preparations to CfP, during core experiment AVC-CE3 [5], in which we were cross-tested by two companies, show that usage of non-linear depth representation has almost no impact on objective rate-distortion (PSNR) curves. This is **expected** result. The tool changes allocation of bits in the compressed depth representation - between the foreground and the background - to obtain higher quality of subjective results. Therefore, non-linear depth representation tool have to be evaluated **subjectively**.

The coding technology developed at Poznan University of Technology, including non-linear depth representation, has been subjectively assessed by 13 test laboratories during resolution of MPEG group Call for Proposals (CfP) on 3D Video Coding Technology as proponent P23, receiving very good results.
Unfortunately, CfP resolution did not bring any information about gains of the particular tools used in the proposals. Therefore, we have performed subjective evaluation in a way resembling MPEG tests that yielded in-depth perspective on performance of our tools coding [6]. We have limited our research to HD sequences in order to limit workload.
We have tested the following configurations:
- HEVC simulcast - MPEG anchors,
- Our proposed technology with all coding tools, as proposed to CfP (P23),
- single coding tools (including non-linear depth representation) turned on separately.

First thing that we were interested in, was whether out methodology was correct and if our results match those attained by MPEG. To find that out, we have compared common cases in our and MPEG evaluation - HEVC simulcast with our proposal P23. Figures 1-4 show this comparison. It can be noticed that most of the differences are not statistically important, because the confidence intervals overlap. In other cases, the differences are minor and thus it can be concluded that our evaluation matches those of MPEG.

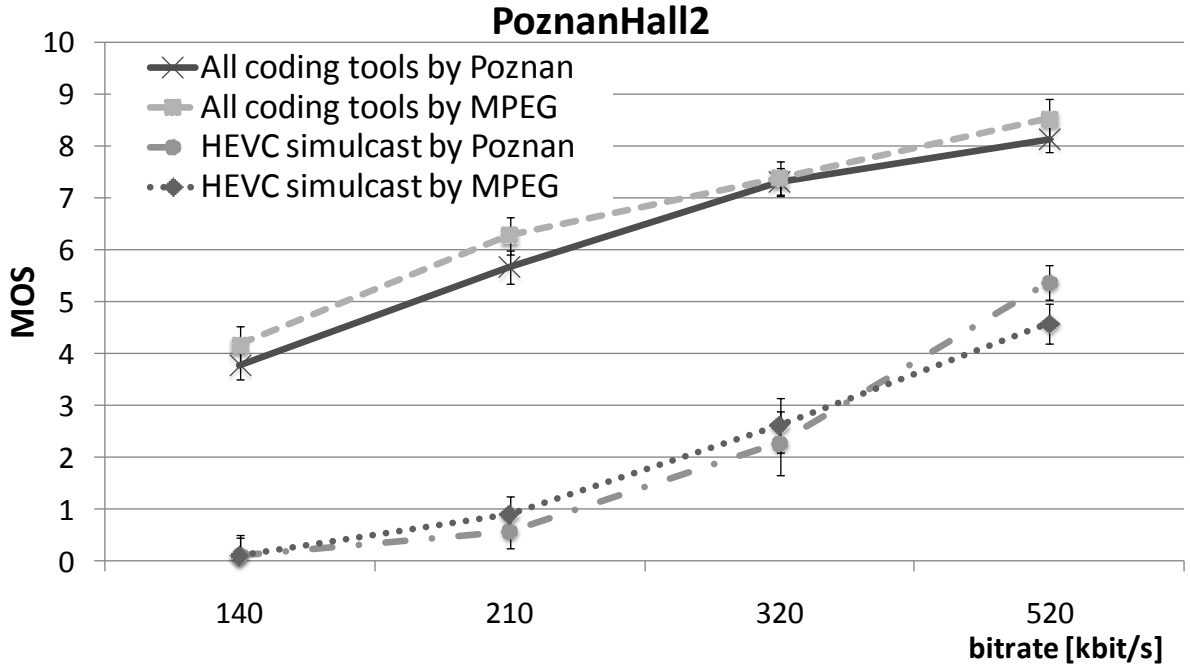

Fig. 1. Comparison of subjective test results performed by MPEG (CfP) and by Poznan University of Technology - Poznan Hall 2 sequence.

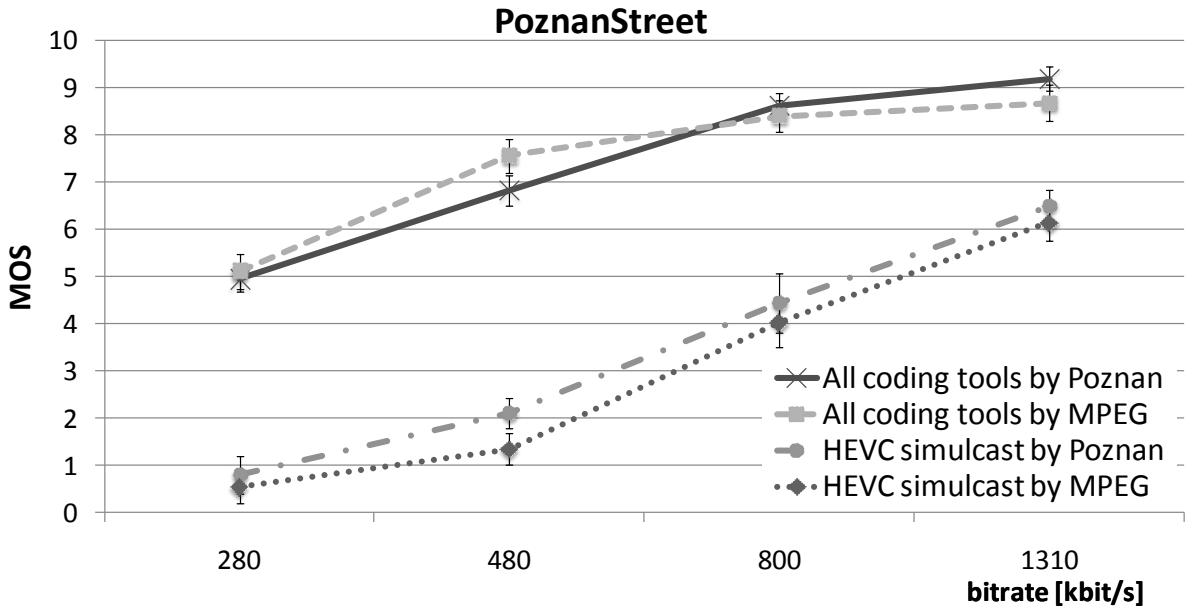

Fig. 2. Comparison of subjective test results performed by MPEG (CfP) and by Poznan University of Technology - Poznan Street sequence.

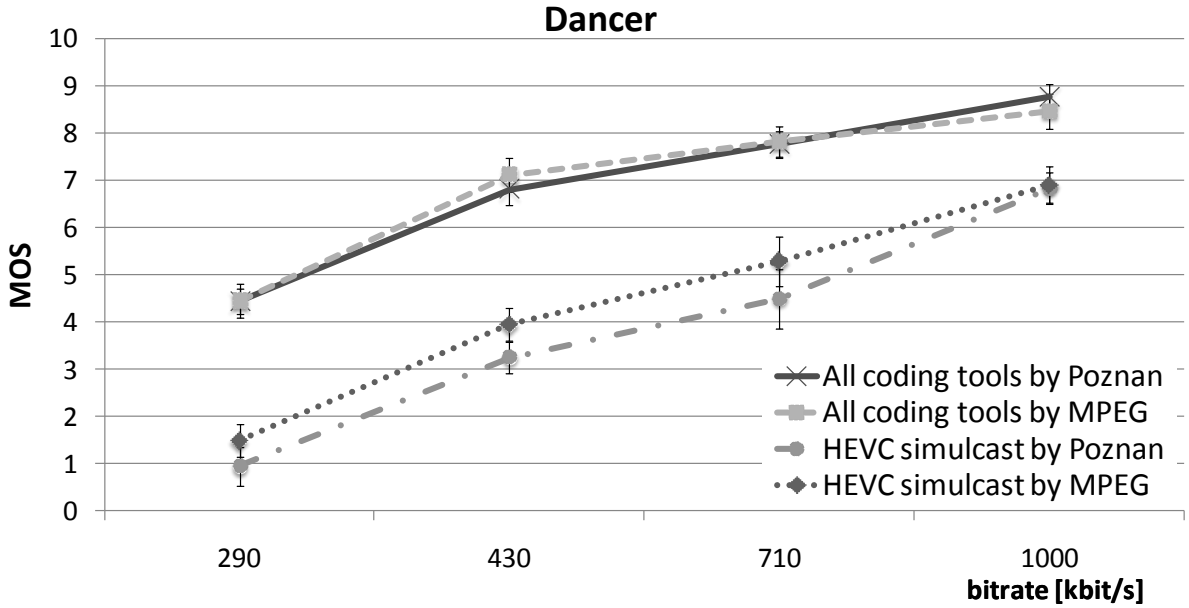

Fig. 3. Comparison of subjective test results performed by MPEG (CfP) and by Poznan University of Technology - Undo Dancer sequence.

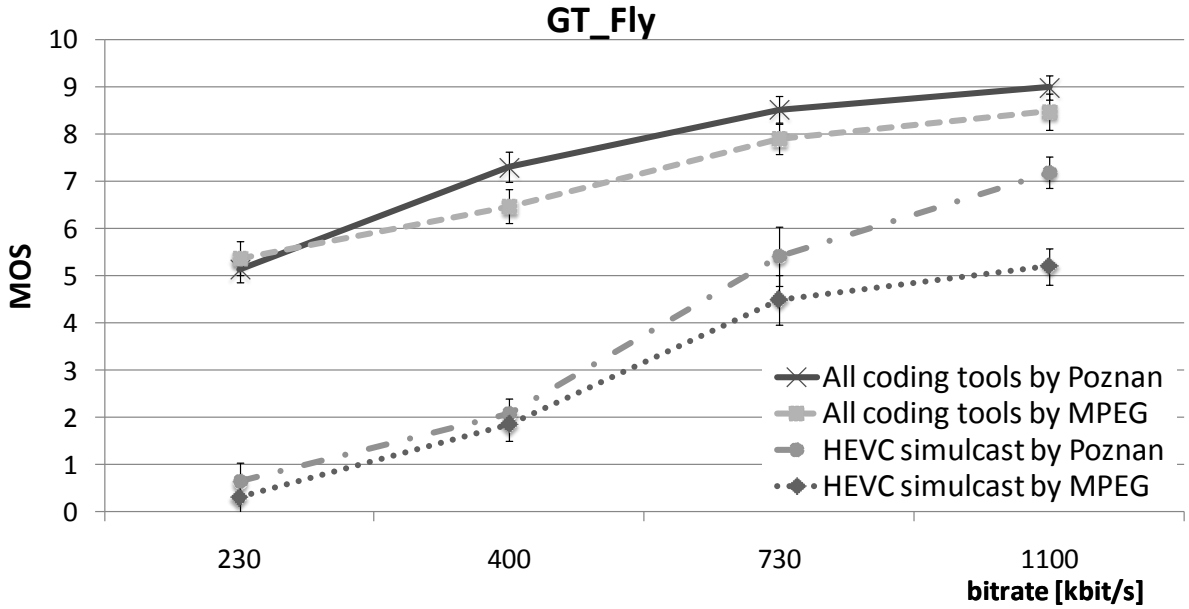

Fig. 4. Comparison of subjective test results performed by MPEG (CfP) and by Poznan University of Technology - GT_Fly sequence.

The further step was analysis of gains attained by distinct tools of our proposal. Here, we focus on non-linear depth representation. The results presented in figures 5-8 show how much gain is attained only by usage of non-linear depth representation, comparing to HEVC simulcast and to all coding tools (P23). Apart from single cases in which the confidence intervals overlap, in most of the cases non-linear depth representation provides from **10% to even 40%** of the whole gain attained by P23.
Such substantial gain is provided at almost no computational cost, because as marked during CE3 [5] implementation of non-linear depth representation consist of some low-cost pre- and post-processing of depth map and some modifications in the LUTs (Look-Up Tables) in the codec loop.

4 Conclusions and recommendations

The proposed non-linear depth representation tool provides substantial gains in subjective quality. Those gains cannot be measures by objective PSNR calculation.
We recommend to perform subjective assessment of performance of this tool - on the upcoming meeting or in a Subjectively-Oriented Core Experiment.
For such SO-CE it may be required to create Common Test Conditions in a similar way to descriptions provided in CfP.

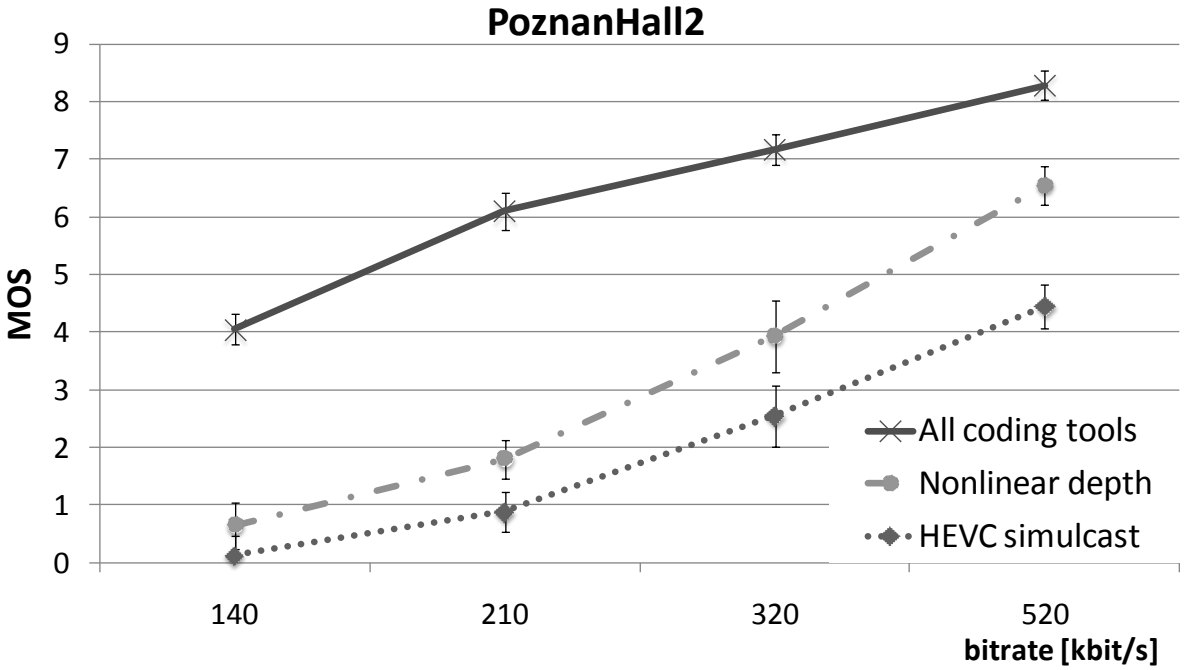

Fig. 5. Subjective test results for Poznan Hall 2 sequence.

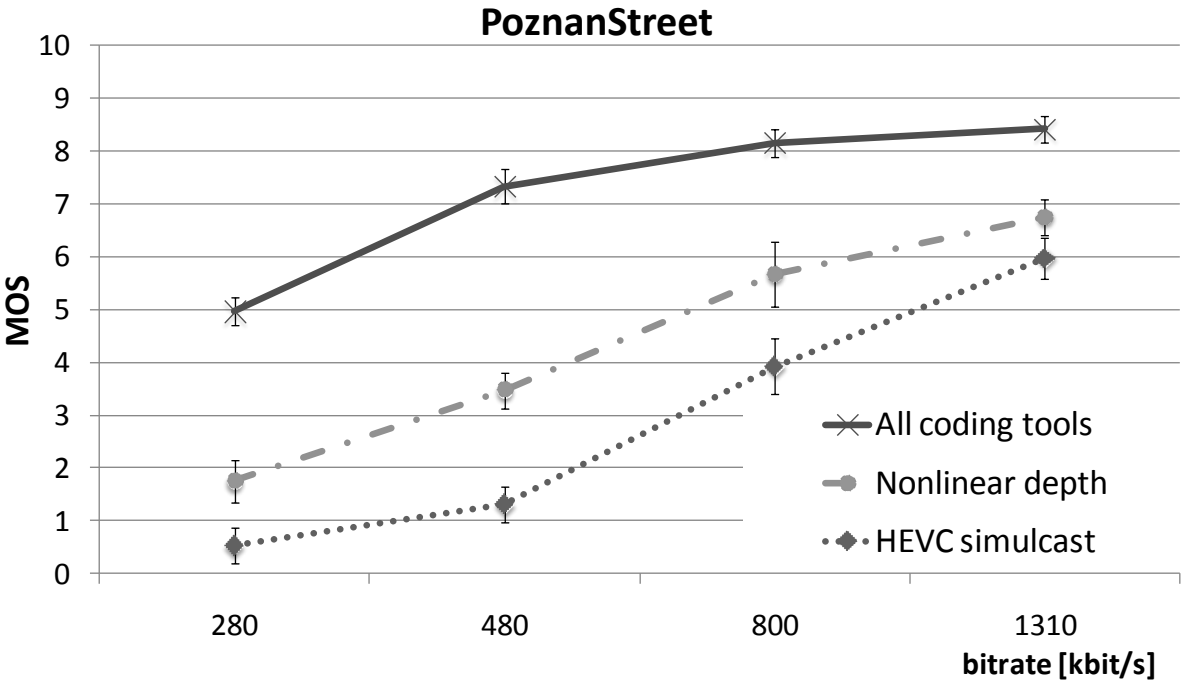

Fig. 6. Subjective test results for Poznan Street sequence.

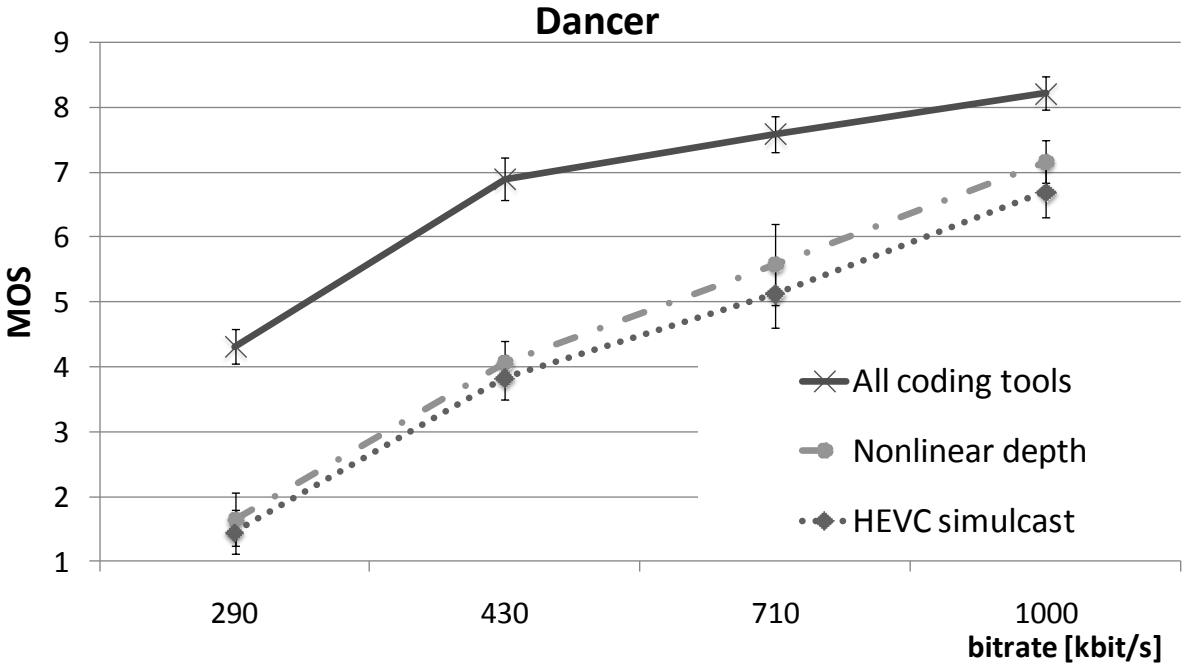

Fig. 7. Subjective test results for Undo Dancer sequence.

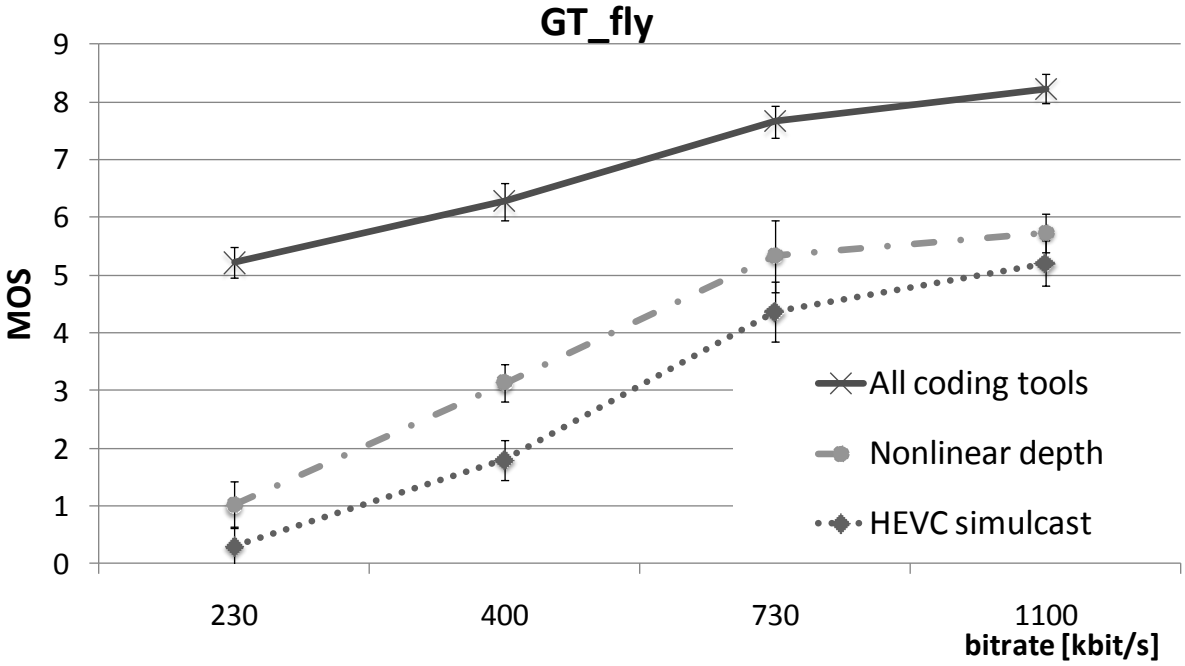

Fig. 8. Subjective test results for GT_Fly sequence.

5 References

[1] Marek Domański, Tomek Grajek, Damian Karwowski, Krzysztof Klimaszewski, Jacek Konieczny, Maciej Kurc, Adam Łuczak, Robert Ratajczak, Jakub Siast, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof zWegner, "Technical description of Poznan University of Technology proposal for Call on 3D Video Coding Technology", ISO/IEC JTC1/SC29/WG11 document M22697, Geneva, Switzerland, November 2011.

[2] "Description of Core Experiments for 3D-AVC ", ISO/IEC JTC1/SC29/WG11 N12353, December 2011, Geneva, Switzerland.

[3] Heiko Schwarz, Dmytro Rusanovskyy, "Common Test Conditions for HEVC- and AVC-based 3DV", ISO/IEC JTC1/SC29/WG11 MPEG2011/N12352, December 2011, Geneva, Switzerland.

[4] Olgierd Stankiewicz "3D-AVC-CE3 summary report: Depth Representation & Coding", ISO/IEC JTC1/SC29/WG11, MPEG/M23785, February 2012, San Jose, USA

[5] Olgierd Stankiewicz, "3D-AVC-CE3 summary report: Depth Representation & Coding", ISO/IEC JTC1/SC29/WG11 MPEG/M23785, February 2012, San Jose, USA.

[6] Filip Lewandowski, Mateusz Paluszkiewicz, „Quality subjective assessment for stereoscopic video", BSc Thesis, Poznan University of Technology, Poland, January 2012, in polish.