Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11
1st Meeting: Stockholm, SE, 16–20 July 2012

Document: JCT2-A0090
M26046

Title: **Impact of View Synthesis Optimization (VSO) on depth quality**

Status: Input Document

Purpose: Proposal

Author(s) or Contact(s): Olgierd Stankiewicz, Krzysztof Wegner, Marek Domański,

Tel: +48 61 665 3800
Email: ostank@multimedia.edu.pl

Source: Poznan University of Technology, Poland

---

1 Introduction

The 3D-HTM software includes a non-normative tool called View Synthesis Optimization (VSO). This document deals with the impact of this tools tool on the quality the depth reconstructed in the decoder.

2 The problem

This document relates to View Synthesis Optimization (VSO) in 3D-HTM reference software. The VSO is a tool that modifies classical rate-distortion-optimization mechanism for the depth, so that the distortion term is calculated basing on the quality of texture synthesized using coded depth, instead of basing on the quality of the coded depth itself [1].

VSO allows the encoder to discard depth information which appears to be irrelevant for the view synthesis process and provides a few-percent b-d rate gain over an encoder without VSO. On the other hand, the forthcoming 3D video coding standard should be generic and be applicable in the scenarios where high-fidelity depth maps representation is needed. Moreover, the future encoders on the market may not implement VSO or may have VSO implement it differently and thus current CTC may not resemble a real world solution.

The CTC should be aware about those observations in order to give the opportunity to obtain realistic results in various scenarios.

In the current version of Common Test Conditions (CTC) [2] used in core experiments, VSO is turned always on.

3 View Synthesis Optimization in the Common Test Conditions

For the current experiments with VSO, we need to be aware about the following observations:

a) **The quality of the synthesized texture is measured at arbitrarily selected positions,** which are not necessarily virtual camera positions in which view synthesis will be done.

c) **In the simulated (CTC) results significant depth information is discarded,** which might be unrealistic in case of some potential 3D standard applications.

d) **Evaluation of normative depth coding tools is inadequate**, not only because some of the depth information is discarded, but also because VSO influences operation of coding tools in unpredictable manner.

Figures 1-4 show some exemplary coded depth maps with VSO turned on and off for comparison. Case with VSO turned on has significant artifacts which may be considered undesirable. Of course such behavior can be observed in many more places.

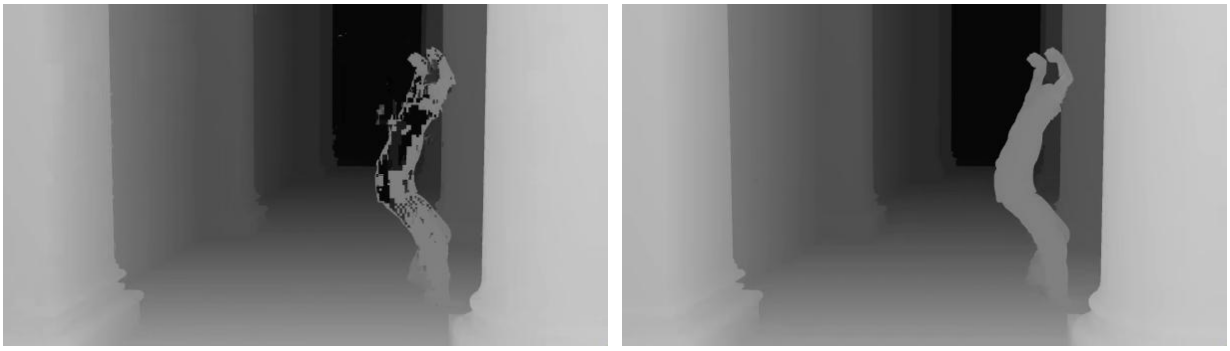

Figure 1: Coded depth of "Undo_Dancer" sequence with VSO turned on (left) and off (right).

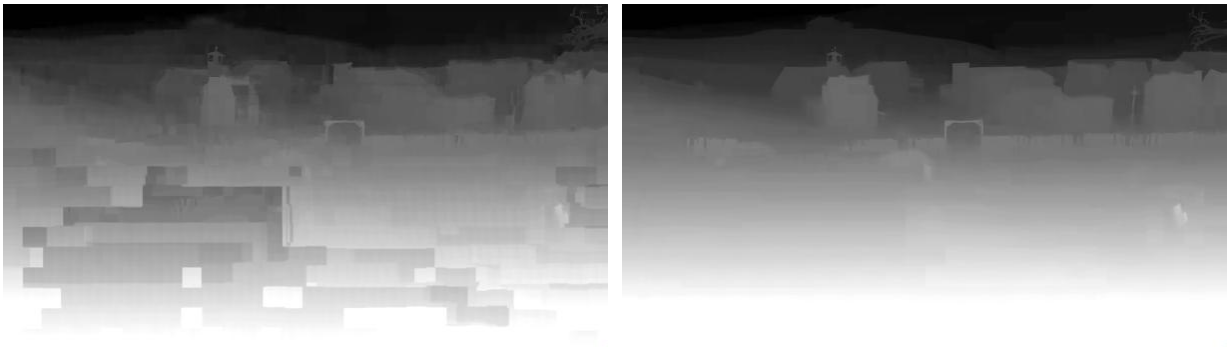

Figure 2: Coded depth of "GT_Fly" sequence with VSO turned on (left) and off (right).

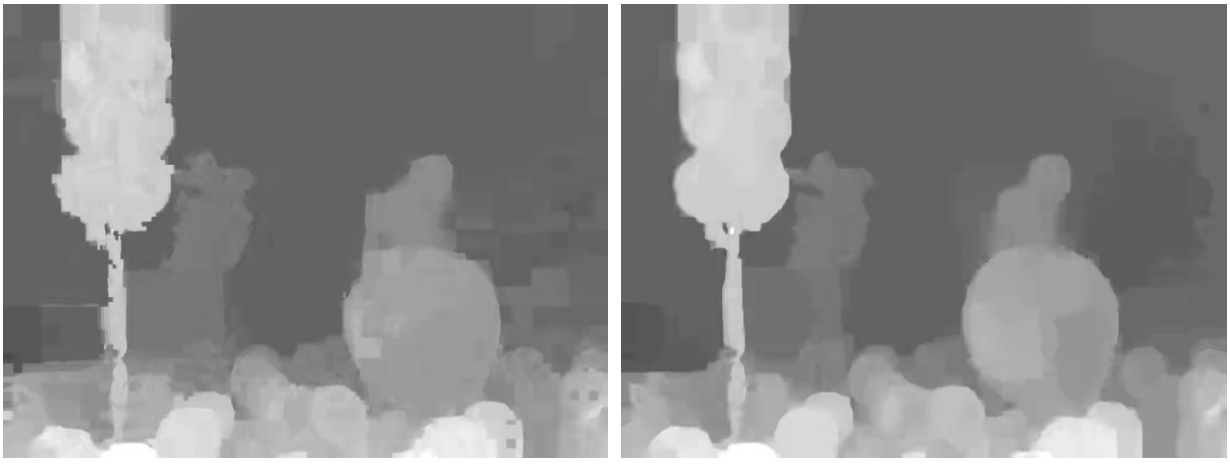

Figure 3: Coded depth of "Balloons" sequence with VSO turned on (left) and off (right).

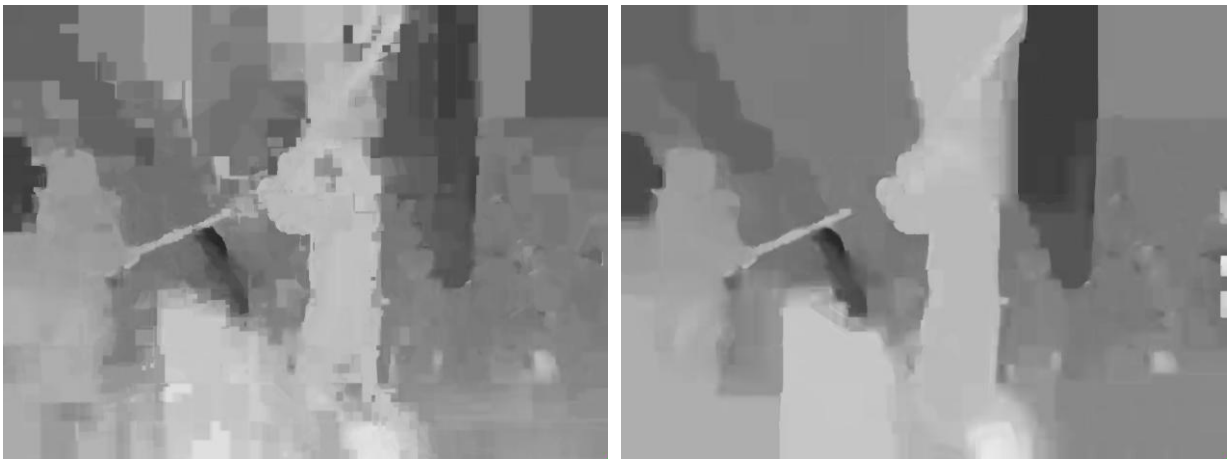

Figure 4: Coded depth of "Kendo" sequence with VSO turned on (left) and off (right).

4 Conclusions and recommendations

With the aim to study the behavior of individual coding tools, in particular those designed for depth coding, we propose to modify Common Test Conditions. It might be done in one of the following ways:

a) Disable VSO in some CEs related to depth coding only.

b) Disable VSO in CTC to allow good evaluation of depth coding tools.

c) **Use two variants - VSO enabled and disabled - to allow wider perspective like in case of RDQ in HEVC.**

We believe that careful study on the conditions of the use of VSO in CTC would be beneficial for further standard development.

References

[1] Krzysztof Wegner Heiko Schwarz, "Test Model under Consideration for HEVC based 3D video coding v3.0," ISO/IEC JTC1/SC29/WG11 MPEG, N12744 2012.

[2] Dmytro Rusanovskyy Heiko Schwarz, "Common Test Conditions for 3DV experimentation," ISO/IEC JTC1/SC29/WG11 MPEG, N12745 2012.