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
2nd Meeting: Shanghai, CN, 13–19 Oct. 2012

Document: JCT3V-B1107

Title:	**Description of Core Experiment 7 (CE7) on Coded Depth Representation**
Status:	Output Document
Purpose:	Core Experiment Description

Author(s) or Contact(s):	Krzysztof Wegner	Tel: Email:	+48 61 6653893 kwegner@multimedia.edu.pl
Source:	CE coordinator		

Abstract

This document defines Core Experiment 7 on coded depth representation (CE7) to be performed for the 3rd JCT-3V meeting.

1 Introduction

The goal of this Core Experiment (CE) is to investigate the improved methods for coded depth representation (linear and non-linear) proposed to the 2nd JCT-3V meeting. This CE covers AVC-based proposals under experiment CE.A only.

The objectives of tools under test are both increasing the coding efficiency of depth and study coded depth representation influence on other coding tools already exists in Test Model.

2 Participants

Participant	Person	Email address	P	C
Samsung	I. Lim	ilsoon.lim@samsung.com	X	
Nokia	D. Rusanovskyy	Dmytro.Rusanovskyy@nokia.com	X	
Poznan	K. Wegner	kwegner@multimedia.edu.pl		X
MediaTek	Yu-Lin Chang Jian-Liang Lin	yulin.chang@mediatek.com jl.lin@mediatek.com		X

(P=proponent, C=crosss checker)

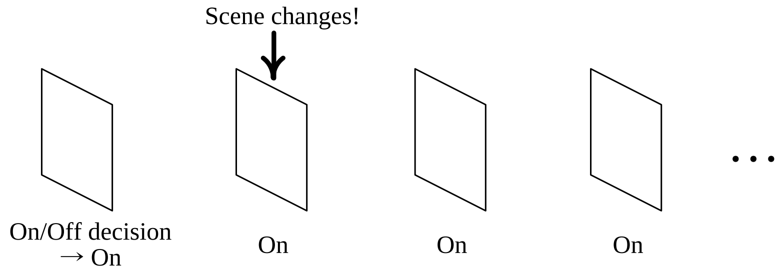
3 Tools under Test for experiment CE7.A

3.1 Improved Nonlinear Depth Representation [JCT3V- B0155]

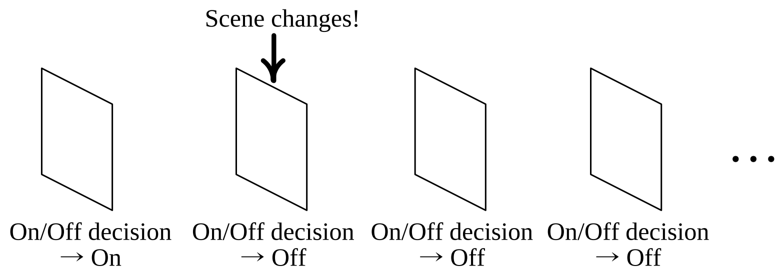
Although the present NDR (Nonlinear Depth Representation) that has already been adopted in 3DV ATM has good coding gain, there is still chance to improve error-resilience and coding gain. In this proposal, we address the improvement of the nonlinear depth representation. NDR is a kind of the non-uniform quantization so that distant objects are quantized more roughly than the closer one. Depth sample values were defined by the following power-law expressions, similar as in the case of well known gamma correction:

$$d_{out} = (d_{in})^{\alpha}$$

For the input depth map, NDR takes the forward transform (linear to nonlinear) and then transformed depth map is used in 3DV encoder and 3DV decoder. Finally, the coded depth map is inversely transformed in order to represent the nonlinear representation to initial linear representation. In current implementation to decide whether NDR enabling or not, encoder makes the analysis for the first frame of depth map and such decision is signaled in the bitstream for whole sequence in SPS. Thus one-time on/off decision suffer from a kind of scene change that often happens and is not efficient throughout a sequence.



(a) The present NDR



(b) The improved NDR

Fig. 1. The comparison of the present NDR and improved NDR

In the proposal, we change on/off signaling by which decision is made by not the first frame but every frame. Thus, we decide every slice whether NDR enabling or not and the on/off decision is explicitly signaled to the decoder on the slice basis. The on/off signaling was written inside the slice data but the LUT tables related to the nonlinear representation is written in SPS (EHP profile) or SEI (HP profile) as before. Compared to the present nonlinear depth representation, there are two differences as follows.

- NDR on/off flag is signaled to the decoder every slice.
- Decision on NDR on/off is based on not only disparity center but also disparity concentration. (encoder issue only)

Both aspect will be study separately and in combination to asses coding gain source.

From this paragraph, we will describe the disparity center and disparity concentration in detail. First, disparity center has already been presented in [1] by Poznan University of Technology and is defined as follows:

$$\text{center\_disparity} = \frac{\sum_{i=0..255} \text{disparity\_histogram}(i) \cdot i}{\sum_{i=0..255} \text{disparity\_histogram}(i)}$$

where disparity_histogram(i) is a histogram of the first frame of depth map in center view. Second, disparity concentration is newly defined as follow:

$$\text{concentration\_disparity} = \frac{\text{The number of the effective bins in base histogram}}{\text{Total bin inside base histogram}}$$

where base histogram is defined as the histogram where the trivial tail and head are eliminated and the effective bin is defined as the bin with the more than some quantity. The disparity concentration is introduced to measure the disparity density in the histogram.

In the present NDR, if the disparity center is more than 100, then NDR is turned on. In the improved NDR, if disparity center is more than 50 and the disparity concentration is less than 60% , then NDR is turned on.

slice_header( ) {	**C**	**Descriptor**
…		
if(DepthFlag) {		
Nonlinear_depth_representation_flag	**0**	**u(1)**
}		
…		
}		

Nonlinear_depth_representation_flag specifies whether the current slice is represented in nonlinear form or not. Nonlinear depth representation flag equal to 1 specifies that the current slice is transformed in nonlinear representation.

3.2 Disparity map coding [JCT3V-B0152]

The document, propose a concept of mixed depth map/disparity coding for 3DV-ATM.

This proposal is based on the assumption, that depth map to disparity map conversion is performed as specified in JCT3V-B0150. Conversion is implemented through a linear model of correspondence between depth map value d and disparity D as it shown in (1):

$$D = (w * d + o) >> n, \qquad (1)$$

where d is a depth sample value, w is a scale factor, o is an offset value , and n is a shift parameter that depends on the required accuracy of the disparity vectors. Parameters of conversion w and o are transmitted within a bitstream with conventional variable length coding. For every pair of views (source view and target view) utilized in joint coding an independent set of parameters w and o are transmitted. In addition to this, parameters of w and o can be re-send for every coded frame if parameters of depth map are varying, e.g. GT_Fly sequence.

With this representation, a depth map becomes a special case of the disparity (1) with w=1 and o = n = 0.

Thus, depth map and disparity map coding support in 3DV-ATM is straightforward under the assumption that disparity is represented with 255 level of quantization.

However, study of CTC test material, show that parameters of disparity are temporally varying and for some frames it cannot be represented with 8 bits.

An optimal representation of the ranging data is selected at the encoder side for every frame, e.g. depth map or disparity map of a particular accuracy. Following the selection, an image of input depth map data may be converted to a disparity image and coded as a disparity map. The type of utilized depth map representation is signaled to the decoder. Decoder up-dates the parameters w and o utilized for in-loop disparity conversion and for output, based on the type of coded data.

4 Core Experiment Conditions

4.1 Software

Experiment CE7.A will use the ATM-6.0 software that is recommended in JCT3V-B1100. Proponents are requested to provide software that can be compiled under Windows and Linux platforms.

4.2 Test Sequences, Bit Rates and Coding Conditions

The CE will use the test sequences, configuration and conditions that are recommended in JCT3V-B1100. Moreover, proponents and cross checkers are required to provide simulation results for the Random access configuration as specified in JCT3V-B1100.

4.3 Evaluation of CE Results

The performance measurements are evaluated by switching on and off individual tools to identify their relative performance. The following measurements are considered to be used in this core experiment. A corresponding Excel sheet for reporting the simulation results will be provided by the CE coordinator.

Coding Performance Measurements: PSNR values shall be computed for the decoded texture views relative to original texture views and for the synthesized views relative to the synthesized views based on uncompressed texture and depth. 4-point BD-rate according to common test conditions is used to report the overall simulation results.

Complexity measurements: For the complexity measurement, the reference software and the reference software with the proposed method implemented will be executed on the same machine with the same configuration and the computational time will be measured. A time ratio will then be calculated between the reference software and the reference software with the proposed method implemented.

4.4 Timeline

2012/11/12	Reference software ATM-6.0 available
2013/01/04	Make source code and simulation results available for cross check
2013/01/10	Official document registration and upload deadline (register documents and upload simulation and cross check results for CE and related proposals)