|  |  |
| --- | --- |
| **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-A1101 |

|  |  |  |
| --- | --- | --- |
| *Title:* | **Description of Core Experiment 1 (CE1) on View Synthesis Prediction** | |
| *Status:* | Output Document | |
| *Purpose:* | Core Experiment Description | |
| *Author(s)):* | Fabian Jäger Shinya Shimizu | [jaeger@ient.rwth-aachen.de](mailto:jaeger@ient.rwth-aachen.de) [shimizu.shinya@lab.ntt.co.jp](mailto:shimizu.shinya@lab.ntt.co.jp) | |
| *Source:* | JCT-3V CE Coordinators | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This document describes Core Experiment 1 (CE1) of the development of 3D extensions for HEVC and AVC. This Core Experiment investigates coding methods based on view synthesis prediction, which uses texture and depth information of the independent view to construct a synthesized prediction signal for coding the dependent views.

# Participants

|  |  |  |  |
| --- | --- | --- | --- |
| ***Participant*** | ***Contact*** | ***Email*** | ***Proponent*** |
| RWTH Aachen University | Fabian Jäger | [jaeger@ient.rwth-aachen.de](mailto:jaeger@ient.rwth-aachen.de) |  |
| NTT | Shinya Shimizu | [shimizu.shinya@lab.ntt.co.jp](mailto:shimizu.shinya@lab.ntt.co.jp) | P/C |
| Mitsubishi Electric | Dong Tian | [tian@merl.com](mailto:tian@merl.com) | P/C |
| MediaTek | Yu-Lin Chang | [yulin.chang@mediatek.com](mailto:yulin.chang@mediatek.com) | P/C |
| Nokia | Dmytro Rusanovskyy | [Dmytro.Rusanovskyy@nokia.com](mailto:Dmytro.Rusanovskyy@nokia.com) | P/C |
| Qualcomm | Xin Zhao | [xinzhao@qualcomm.com](mailto:xinzhao@qualcomm.com) | P/C |
| ETRI | Gun Bang | [gbang@etri.re.kr](mailto:gbang@etri.re.kr) | C |
| KHU | Y.K.Kim | [kimky@khu.ac.kr](mailto:kimky@khu.ac.kr) | C |
| Samsung | Jin Young Lee | [jinyoung79.lee@samsung.com](mailto:jinyoung79.lee@samsung.com) |  |
| Zhejiang University | Yin Zhao | [zhaoyin@zju.edu.cn](mailto:zhaoyin@zju.edu.cn) | C |
| Orange Labs | Joel Jung | [joelb.jung@orange.com](mailto:joelb.jung@orange.com) |  |

(P = proponent, C = cross checker)

# Tools under Investigation

The tools to be investigated in this Core Experiment use the decoded texture and depth information from the independent base view to compute a prediction signal for the texture and depth component of the dependent views. This process is done by means of a warping algorithm, which is similar to the non-normative algorithms used in the final view synthesis stage at the receiver.

For the case of view synthesis prediction (VSP), as tested in this CE, the view synthesis algorithm is part of the coding loop and therefore a normative tool.

## CE1.a: View Synthesis Prediction based on 3D-ATM

The tools to be investigated can be categorized into the following three categories. All separable tools should be evaluated separately according to the following categorization. Reports on combined tools are also encouraged.

1. Block-based view synthesis using a backward warping (JCT2-A0050, JCT2-A0107)
2. Skipping motion vector information coding (JCT2-A0107)
3. Removal of VSP picture from the reference picture list (JCT2-A0103, JCT2-A0050)
4. Sub-MB VSP skip/direct (JCT2-A0050, JCT2-A0103)

### JCT2-A0050: Interview Skip mode with sub-partition scheme

The inter-view skip mode is proposed to eliminate full frame view synthesis for reducing the computational complexity, memory bandwidth, and buffer. In contrast to the current VSP design, the inter-view skip mode is to utilize the depth from the current view and pick up sample predictors from a reference view by using a backward warping. No sample level vector is used, but a maximum depth value in the 4x4 is used for the whole 4x4 block.

**🡪This contribution should be evaluated in terms of the tools in categories (a), (c), and (d).**

### JCT2-A0103: Generalized view synthesis prediction (GVSP) mode

GVSP is a way to enable sub-MB skip/direct using VSP reference pictures. In addition to a vsp\_flag at MB level, a smaller partition, up to 8x8 blocks, can be signaled whether it is predicted from a VSP reference picture. It is also proposed to remove VSP picture from the reference picture list.

**🡪This contribution should be evaluated in terms of the tools in categories (c) and (d).**

### JCT2-A0107: Block-based View Synthesis Prediction for 3DV-ATM

A block-based VSP implementation (B-VSP) is proposed, which is based on a backward warping process. With the B-VSP, sample predictors are derived directly from the texture pixels of the reference view, not the synthesized VSP-frame. The displacement vectors required for this process are produced for each pixel from the depth map data of the current view. No motion information is coded when the reference picture index, which corresponds to the VSP, is coded. Zero motion vectors are assumed for such blocks.

**🡪This contribution should be evaluated in terms of the tools in categories (a) and (b).**

## CE1.h: View Synthesis Prediction based on 3D-HTM

### JCT2-A0018: View Synthesis Prediction for 3D-HTM

#### Basic VSP scheme

Within a basic implementation of VSP, a synthetic picture is first generated using a decoded texture image and a decoded depth image. The rendering module in HTM is re-used to generate the synthetic picture in our current implementation.  
In a second step, the synthetic picture is appended to the reference picture lists (LIST\_0 and LIST\_1). In principle, the synthetic picture can be inserted at the end of the reference picture lists or configured at any position in the lists.

**🡪This VSP scheme should be further investigated in this CE to find out to what extent the proposed method can reduce the bitrate of the depending views.**

#### VSP skip mode

On top of the basic VSP scheme, VSP skip/direct [2] is proposed to enable the use of a synthesized picture for skip/direct mode even when the synthesized picture is not placed at the beginning of the reference picture list. Furthermore, an adaptive signaling method was proposed by re-defining the semantics of the skip flag.

**🡪In this CE the improvements of the proposed VSP skip mode should be investigated and evaluated how it performs compared to the basic approach (2.2.1.1) of simply putting the synthesized picture into the reference picture lists.**

#### Simplification of in-loop view synthesis process

It is noticed that the decoding complexity is significantly increased by the introduction of VSP prediction. This is because of the rendering process for the synthesized picture. The proposed simplification is applied to the warping process: No interpolation is applied on the disparity filed in contrast to the bi-linear interpolation as applied in VSRS-1D-Fast.

**🡪The proposed simplification of the synthesis algorithm is to be investigated in this CE. Results concerning complexity reductions should be reported in terms of runtimes, but also in terms of number of operations and memory requirements.**

# Core Experiment Conditions

## Software

The proposed methods for view synthesis prediction (VSP) will be implemented into the 3D-HTM 4.0 or 3D-ATM 5.0 software, respectively. Proponents are requested to provide software that can be compiled under Windows and Linux platforms.

## Test Sequences

Test sequences will be used according to common test conditions [[1](#DRu12)].

## Coding Conditions

The experiments are conducted under the common test conditions [[1](#DRu12)].

## Evaluation Criteria

Evaluation will be done according to the common test conditions [[1](#DRu12)].

**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. The 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.

A corresponding Excel sheet for reporting the simulation results will be provided by the CE coordinators.

Proponents are encouraged to report details about decoder complexity in terms of number of operations and memory consumption.

## Timeline

2012/10/02 Make source code and simulation results available for cross check  
2012/10/09 Register documents for the 2nd JCT-3V meeting  
2012/10/09 Upload simulation and cross check results to the JCT-3V document repository

# References

x

|  |  |
| --- | --- |
| [1] | D. Rusanovksyy, A. Vetro, and K. Müller, "Common Test Conditions for 3D Video Extensions Development," Joint Collaborative Team on 3D Video Coding Extension Development (JCT-3V) of ITU-T VCEG and ISO/IEC MPEG, Stockholm, Sweden, JCT2-A1100 2012. |

x