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

Source	**Poznań University of Technology**
Status	**Input**
Title	**Summary report on Immersive Video CE-2: Pixel Pruning**
Author	Adrian Dziembowski (CE-2 Coordinator)

Abstract

This document reports the experiments performed during Core Experiment 2.

1   Introduction

The goal of the experiment was to study performance of pixel pruning. Given either re-projected or selected source views, pruning operation is performed in order to reduce the number of pixels that is to be finally packed into texture and depth atlases. It is also necessary that the resulting patches have temporal consistency to compress efficiently using a HEVC encoder.

Four participants registered for the experiments:

> 2.1. Philips,
> 2.2. Nokia,
> 2.4. ZJU (Zhejiang University),
> 2.5. ETRI (Electronics and Telecommunications Research Institute).

Two of the participants (Nokia and ZJU) have withdrawn their experiments for this meeting cycle. Philips and ETRI performed their experiments. Both participants' responses have been successfully crosschecked.

2   Summary of experiments

2.1   *m49962: Philips response to Immersive Video CE-2: Pruning*

This experiment was an update of the previous CE-2 response [M48271]. Expert viewing revealed a major problem: "Decapitation" of the guy with the black shirt. Philips has determined the cause to be as follows:

1.   There is a view A with the head
2.   Another view B is selected first
3.   In view B the head is outside of the viewport
4.   View B is warped to view A (denote B → A)
5.   In view B → A the floor behind the head is rendered
6.   Because those pixels are rendered, they are pruned

This update adds depth error as a 2nd pruning condition:

$$depthError = |renderedDepth / referenceDepth - 1| < maxDepthError = 10\%.$$

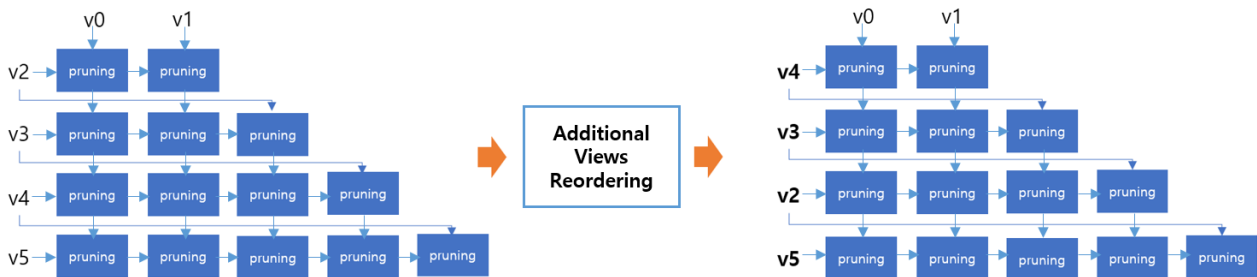Philips' response outperforms the MIV anchor for all test classes:
- on all quality metrics for low bitrate,
- on all metrics except for PSNR for high bitrate.

Philips recommends to:
- adopt their contribution into TMIV 3.0,
- keep CE-2 open for future responses.

2.5   m50030: ETRI response to Immersive Video CE-2: Optimized pruning order

The pruning order of additional views actually influences not only the general appearance and total pixel rate of atlas but also the quality of the rendered view. Based on this inference, the main aim of this experiment was to discover the optimized order of pruning that can minimize the total pixel rate of atlas while maximizing the quality of the rendered view.

The experiment was conducted with the updated algorithm of [M49150]. ETRI's solution clearly outperforms MIV anchor for all test classes on almost all quality metrics.

ETRI's response outperforms the MIV anchor for all test classes:
- on all quality metrics for all bitrate excepting VMAF at low bitrate of C1.

ETRI recommends to:
- adopt their contribution into TMIV 3.0,
- keep CE-2 open for future responses.

3   Summary of crosschecks

3.1   Crosscheck of m49962

Results of M49962 were partially crosschecked by Nokia. The crosschecker obtained **exactly** the same results as the results provided by Philips.

Nokia has checked sequences A, D, E and J for arbitrarily chosen QPs and source views.

3.2   Crosscheck of m50030

Results of M50030 were partially crosschecked by Philips. For each test sequence one QP was computed (A1: QP1, B1: QP3, C1: QP5, D1: QP2, E1: QP4, J1: QP1, L1: QP3), thus the coverage of the crosscheck is 20%.

The texture and depth coding results were **exact** except for the empty atlases (D1 and E1). The participant removed them manually from the reporting template. The crosschecker suggested to decrease number of atlases by decreasing 'MaxLumaSamplesPerFrame' parameter. According to this suggestion, ETRI performed additional test. The results for modified case were identical to initial ones except for slight decrease of metadata bitrate.

The synthesis results obtained by the crosschecker were not exact, but **very close**:
- The maximum per-view errors are:
  - Y-PSNR 0.140 dB (C1 v8),
  - U-PSNR 0.029 dB (L1 v01),
  - V-PSNR 0.016 dB (L1 v01).
- The maximum average error per sequence is:
  - Y-PSNR 0.0008 dB (C1),
  - U-PSNR 0.0028 dB (L1),
  - V-PSNR 0.0016 dB (L1).

The crosschecker also noticed, that there are unnecessary code changes (however they did not affect the crosscheck results):
- Added unused files,
- Added unused saveViewport overload,
- General improvements unrelated to pruning:
  - Keeping camera names for printing,
  - Parameter array as command-line argument.

4 Conclusions
- Two CE2 responses have been received.
- Both responses have been successfully crosschecked.
- Both responses outperform the MIV anchor on all quality for both test classes metrics (except for PSNR for high-bitrate anchor for m49962).
- This CE-2 should be continued.