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

ISO/IEC JTC1/SC29/WG4/m69531
Oct. 2024, Kemer, TR

Source	**ETRI**
Status	**Input Contribution**
Title	**[INVR][EE3] Summary of MIV-DSDE anchor generation for 3D-INVR EEs**
Author	**Jun Young Jeong, Gun Bang (ETRI), Adrian Dziembowski (PUT), Jong-Beom Jeong (SKKU), Yiyi Liao (Zhejiang Univ.)**

1 Introduction

This document provides the summary of INVR EE3 (Anchor Generation for Compression of Radiance Fields), which aims to generate MIV-DSDE anchor for future INVR activities.

2 Test conditions and task allocations

TMIV 18 [N0409] and its' corresponding CTC [N0406] were selected as the baseline, with the below modifications for the 3D-INVR activity:

- Test sequences are different, and details can be found in [N0561]. Monocular contents capturing dynamic scenes are excluded, as they are incompatible with TMIV
- Selected test views are excluded from the atlases and synthesized on the decoder side using the training views and their estimated depths. For INVR, only the test views (without training views) are considered for evaluation
- The number of atlases is adjusted to include all training views for a fair comparison with radiance field techniques
- Objective metrics:
  - ✓ INVR uses PSNR, SSIM, LPIPS, IV-SSIM in the RGB domain. Therefore, source and rendered views are converted to loseless PNG sequences using FFmpeg
  - ✓ Tools used:
    - PSNR and IV-SSIM: QMIV [m68224]
    - LPIPS: Python-based tool by R. Zhang et.al, CVPR2018 (AlexNet) [LPIPS]
    - SSIM: Inria's 3D Gaussian Splatting code [SSIM]
- The MIV reporting template is simplified for 3D-INVR but will need further improvements, including adding metrics to be derived during the standardization phase

Additional considerations:

- YUV420p content was converted to YUV420p10le using FFmpeg before running TMIV
- For some sequences, test views and start frames were adjusted from those in [N0561]
  - ✓ For Choreo, test views were not listed at [N0561], so v7, v9, v10 and v12 were selected
  - ✓ For VRRoom1D, v07, v11, v20 and v24 were chosen due to the incorrect listings in [N0561]
  - ✓ For HauntedLamp, the initial frame was shifted to 265 due to lack of movement at frame 0. Also, test views were changed from to v01 and v11 to v01 and v10, as v01 and v11 are far-corner views where stable synthesis may be difficult

The second anchor generation trial was conducted with the following task allocation. Some participants provided incomplete results, such as missing IV-SSIM, RP0, cross-checking results,

Participant	**Sequences in charge**	**Cross-check**
Adrian Dziembowski	Choreo (M-NC2)	
Jong-Beom Jeong	VRroom1D(M-NC3)	Garage, Choreo
Gun Bang	Garage(M-CG2),	Mirror, KITTI-360
Jun Young Jeong	Mirror(M-CG1), Bartender(M-NC1) HauntedLamp(M-NC4)	VRoomID, HauntedLamp
Yiyi Liao	KITTI-360(S-NC2)	Bartender

3 Results

Below are the RD-curves for different sequences. Note that only Garage and VRroom1D had near-complete, and other results may contain errors.

PSNR RD-curves

- Y-axis ranges from 27 to 32dB
- For Mirror, Garage and Bartender, RP4 minus RP0 values are 1.87, 1.22 and 1.88dB. It seems RD-curves have a fair amount of curvature (Fig. 1 left) for calculating the BD-rate
- Except for Bartender, the other natural sequences show a flat curve (Fig. 1 right), with a maximum difference of 0.54dB between RP4 and RP0

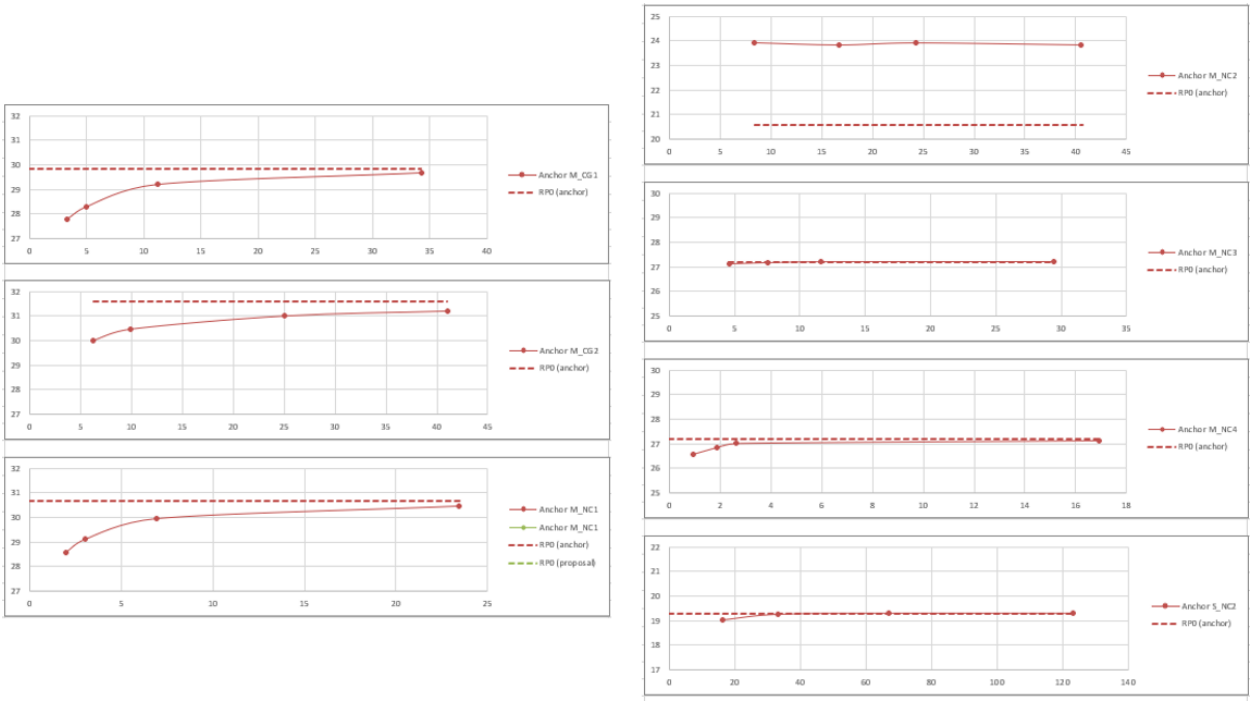

Figure 1. PSNR RD-curves of INVR sequences

SSIM and IV-SSIM

- Y-axis ranges from 0.7 to 1
- Like PSNR, only the RD-curves for Mirror, Garage and Bartender show some curvature (Fig. 2 & 3 left), but it's unclear if it's sufficient. RP4 minus RP0 values are 0.07, 0.04, 0.07 for SSIM and even lower for IV-SSIM
- The remaining sequences show a flat curve (Fig. 2 & 3 right), with a maximum difference of 0.02

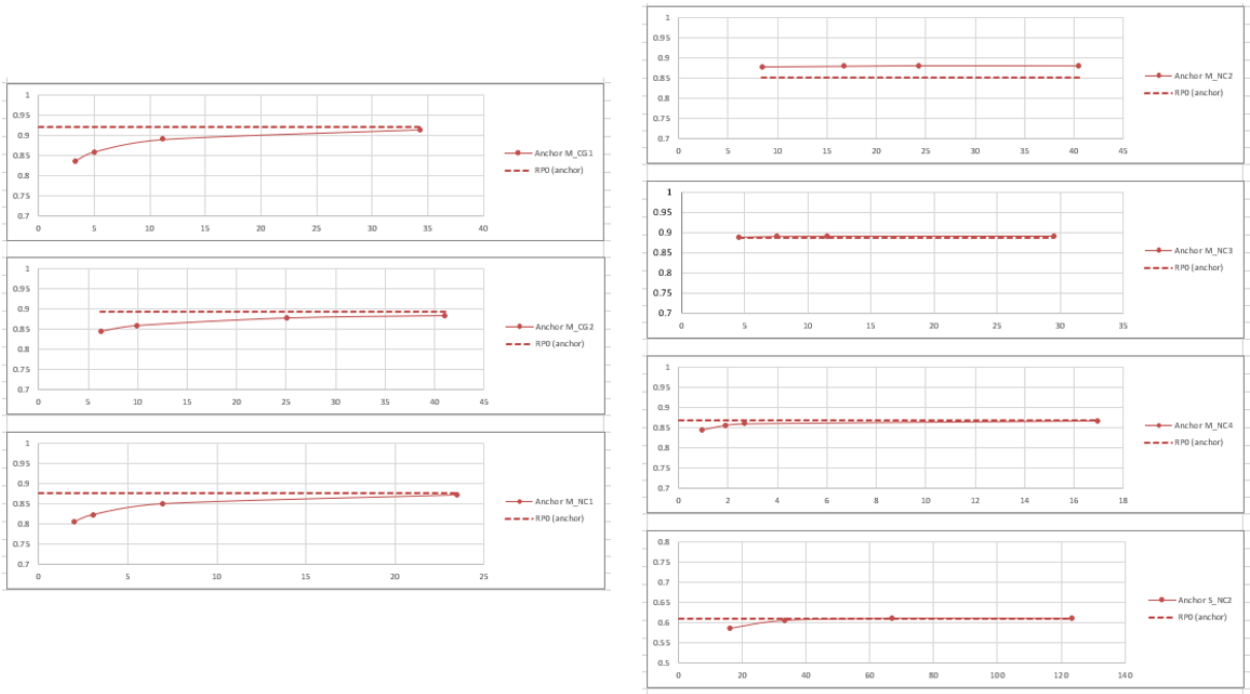

Figure 2. SSIM RD-curves of INVR sequences

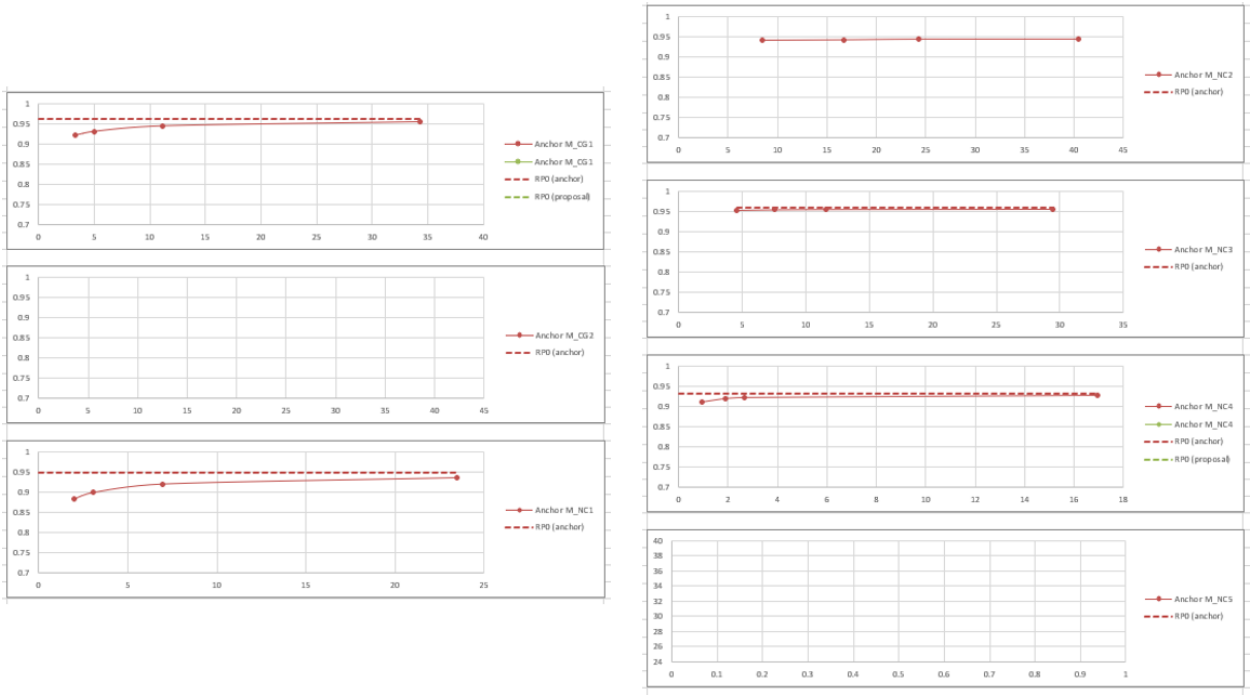

Figure 3. IV-SSIM RD-curves of INVR sequences

LPIPS

- Y-axis ranges from 0.5 to 1
- For Mirror, Garage and Bartender, RP4 minus RP0 values are 0.16, 0.12 and 0.22. It seems RD-curves have a fair amount of curvature (Fig. 4 left)
- Unlike PSNR and SSIM, the RD-curves for HauntedLamp and KITTI-360 show some curvature, with RP4 minus RP0 values of 0.09 and 0.1 (Fig. 4 bottom right)
- The remaining sequences show a flat curve (Fig. 4 top right)

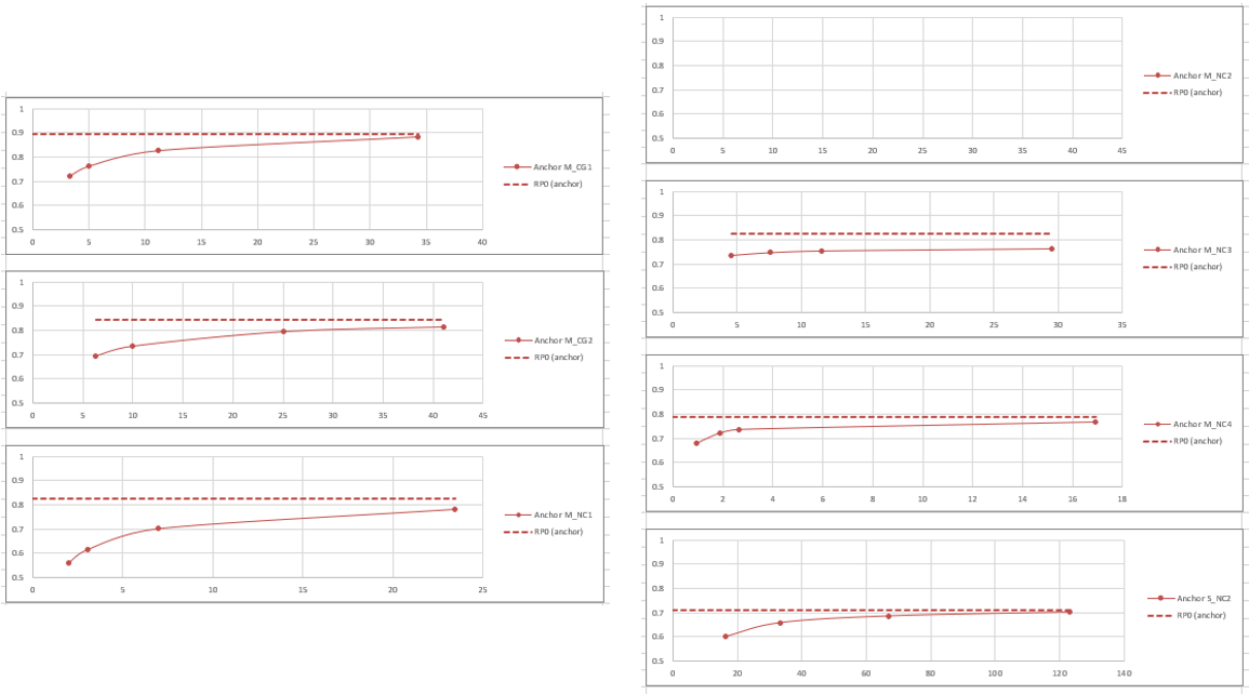

Figure 4. LPIPS RD-curves of INVR sequences

Cross-check

- A partial cross-check was performed for Garage, Choreo and VRroom1D. As shown in Fig. 5., Garage and VRroom1D have nearly identical (not exactly identical) RD-curves, while Choreo shows a significant gap

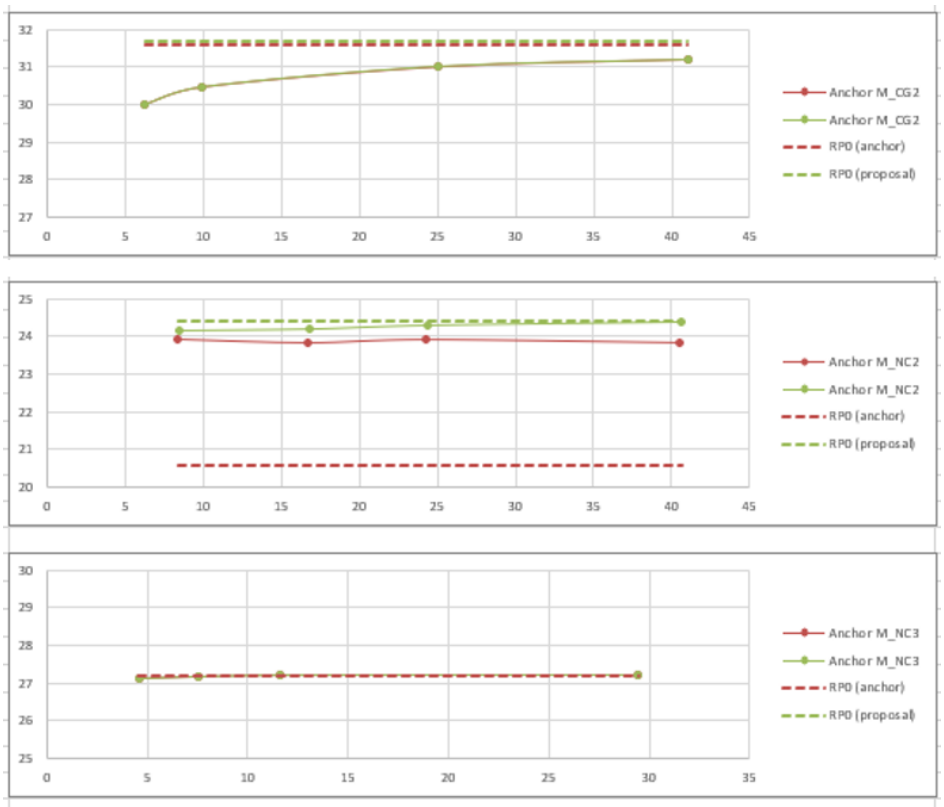

Figure 5. PSNR RD-curves with crosschecker's RD-curves

4 ETRI's proposals for anchor generation process

- Perform one more cycle of anchor generation to see if a lower bitrate can produce good, non-flat RD-curves for all metrics. If not, let's stick to the results with the maximum or no curvature
- Reduce the number of experimenters to 2 or 3, as aligning schedules and conditions among multiple people is not easy. ETRI will volunteer, but one or two more volunteers are needed until the next meeting
- Let's remove the unaligned dataset information from the INVR EE document [N0560] and only keep the CTC [N0561] one to avoid confusion
- Within CTC document, update the test views and start frame of some sequences (Choreo, VRroom1D, HauntedLamp) according to the conditions used in this EE
- Let's remove the datasets that doesn't provide two or three posetraces before the Kemer meeting

5 References

N0409 Test model 18 for MPEG immersive video, ISO/IEC JTC 1/SC 29/WG 04 N0409, Hannover, Oct. 2023.

N0406 Common test conditions for MPEG immersive video, ISO/IEC JTC 1/SC 29/WG 04 N0406, Hannover, Oct. 2023.

N0561 Common test conditions on radiance field representation and compression, ISO/IEC JTC 1/SC 29/WG 04 N0561, Sapporo, July. 2024.

M68224 [MIV] Quality Metrics for Immersive Video – QMIV framework, ISO/IEC JTC 1/SC 29/WG 04 M68224, Sapporo, July. 2024.

LPIPS https://github.com/richzhang/PerceptualSimilarity

SSIM https://github.com/graphdeco-inria/gaussian-splatting/blob/main/utils/loss_utils.py#L33