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

ISO/IEC JTC1/SC29/WG11
MPEG/M16026
February 2009, Lausanne, Switzerland

Title	**Results of 3DV/FTV Exploration Experiments, described in w10173, for Alt Moabit sequence.**
Sub group	**Video**
Authors	**Olgierd Stankiewicz** (ostank@multimedia.edu.pl), **Krzysztof Wegner** (kwegner@multimedia.edu.pl) and **Krzysztof Klimaszewski** (kklima@et.put.poznan.pl) Poznań University of Technology, Chair of Multimedia Telecommunications and Microelectronics, Poznań, Poland

1 Introduction

This document presents results of Exploration Experiments (EE1,EE2 & EE4) performed on "Alt Moabit" sequence [2] and is in response to w10173 "Description of Exploration Experiments in 3D Video Coding" [1].

2 Experiments conditions

Experiments were performed basing on w10173 [1] guidelines (Figure 1):

- **Select stereo pair** from data set, i.e. an original left view OL and an original right view OR (OL=8, OR=9)
- **Estimate depth** corresponding to neighboring original views NL (left) and NR (right) (NL=7, NR=10), using any available camera
- **Synthesize views** (synthesized left SL and synthesized right SR) at positions of OL and OR from NL+D and NR+D
- Bring synthesized video to the meeting
- **Compare OL-OR with SL-SR** subjectively

The test were performed on 'Alt Moabit' [2] sequence with following views selected as OL-OR and NL-NR.

Table 1. The specification of view for EE experiment.

Data set	OL-OR	NL-NR
Alt-Moabit	8-9	7-10

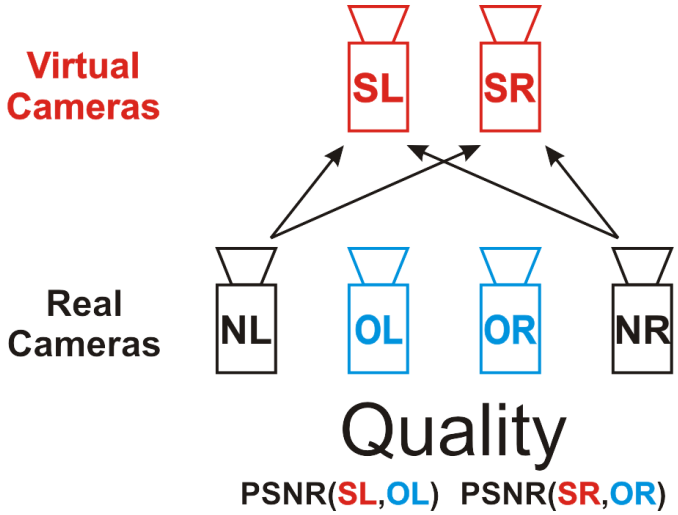
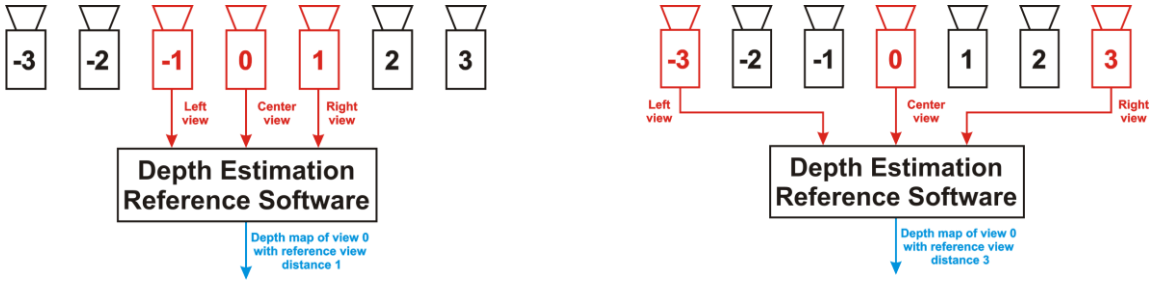


Figure 1. Setup of experiments for depth estimation and view synthesis software evaluation.

The depth estimation was performed with various Camera Distance (Figure 2) parameters of view synthesis software – from distance 1 to distance 5.



a) Depth map estimation with reference view distance 1

b) Depth map estimation with reference view distance 3

Figure 2. Setup of experiments for depth estimation and view synthesis software evaluation.

For EE4 experiment depth maps produced with Depth Estimation Reference software 2.0 and smoothing coefficient = 3.0 were used. Depth maps with various reference view distances were used. For coding and decoding of images and depth maps the last version of MVC [3] software was used.

3 Results – EE1 – Depth estimation precision improvement

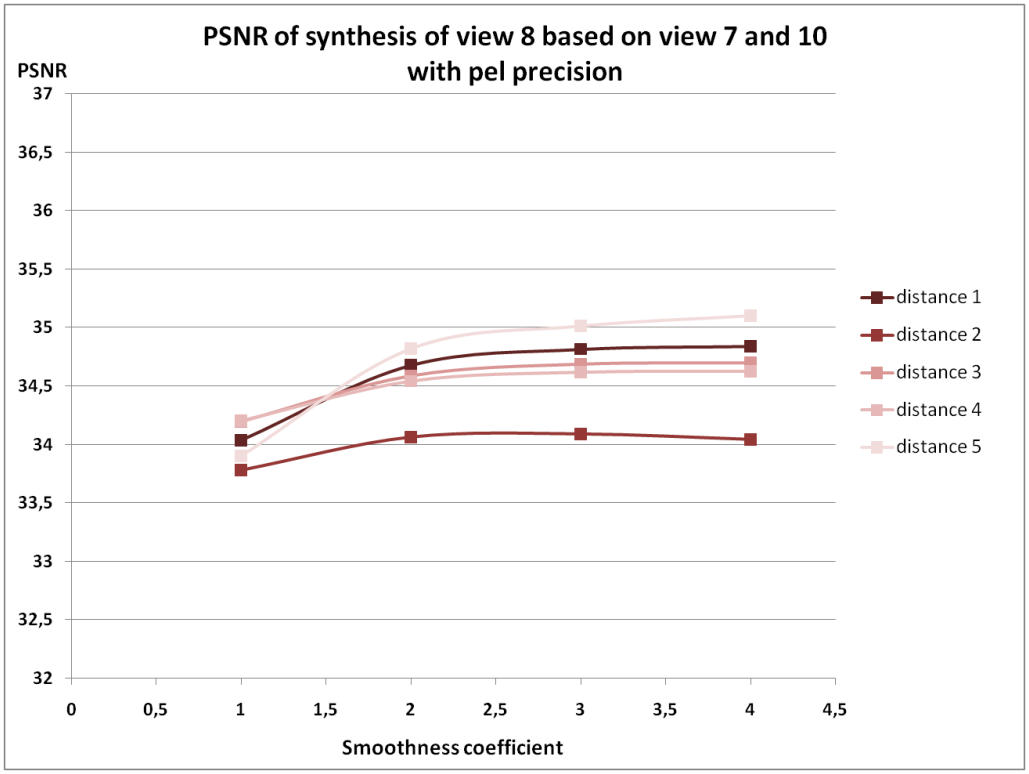


Figure 3. EE1 results, pixel precision, view 8.

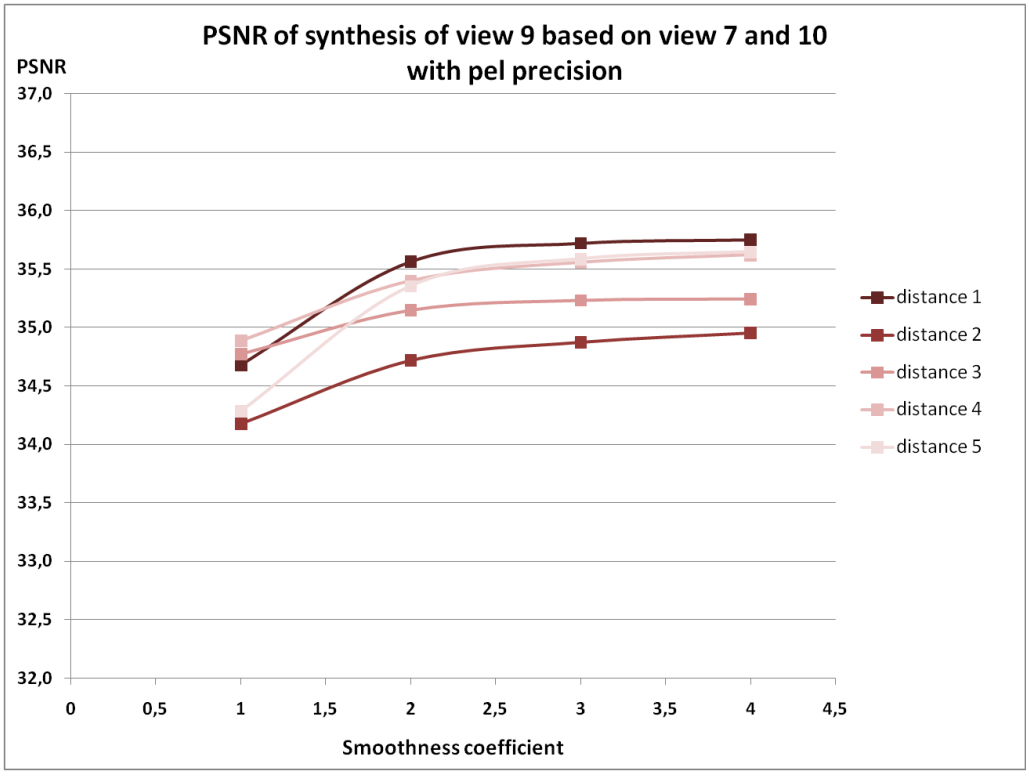


Figure 4. EE1 results, pixel precision, view 9.

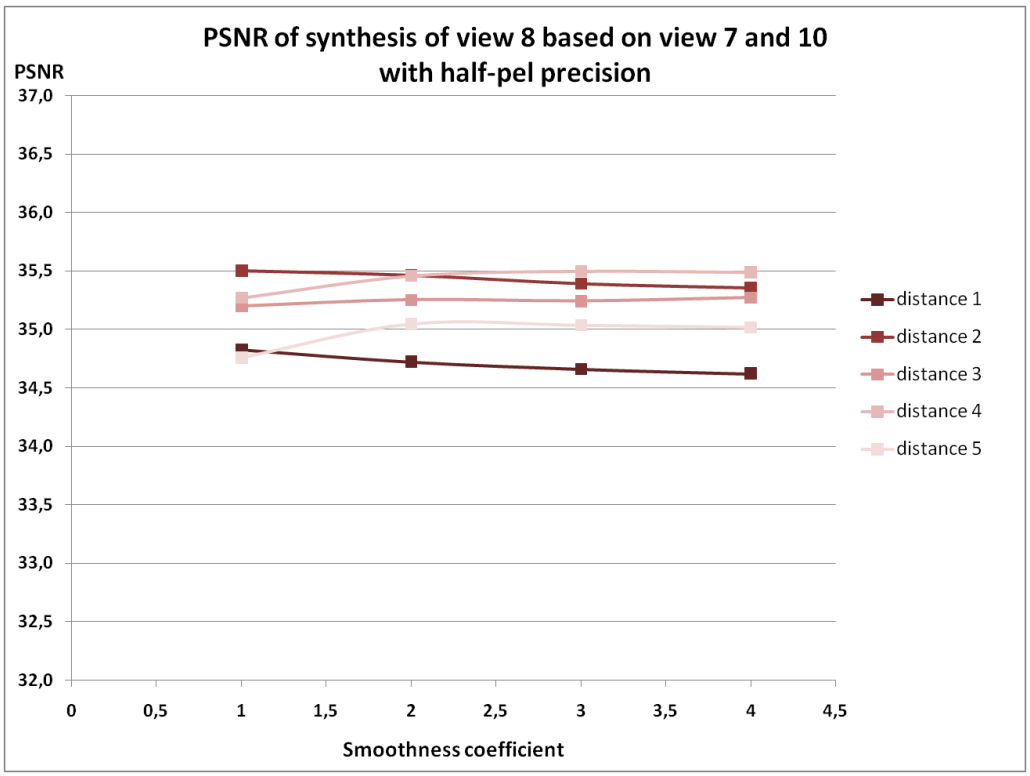


Figure 5. EE1 results, half-pixel precision, view 8.

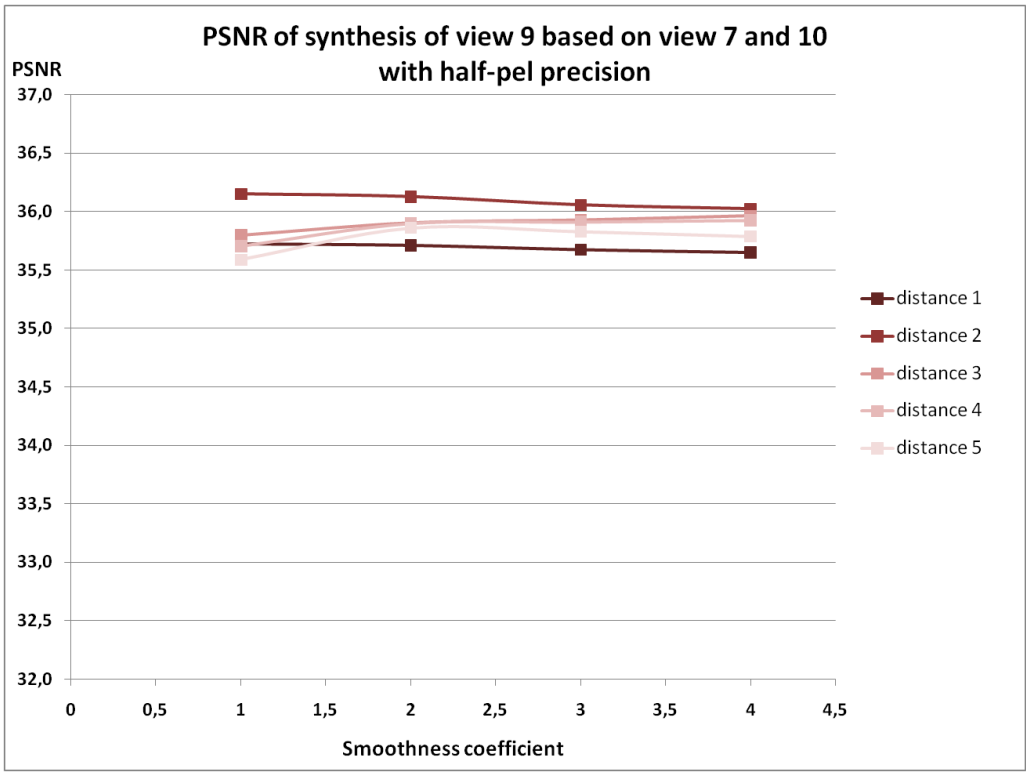


Figure 6. EE1 results, half-pixel precision, view 9.

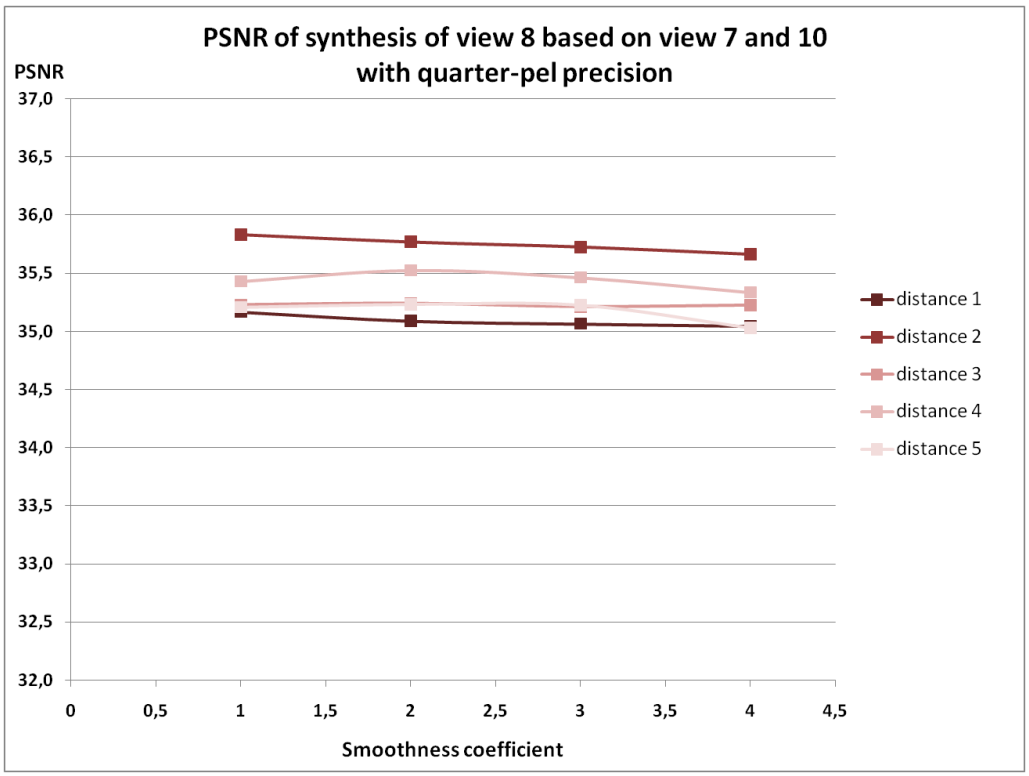


Figure 7. EE1 results, quarter-pixel precision, view 8.

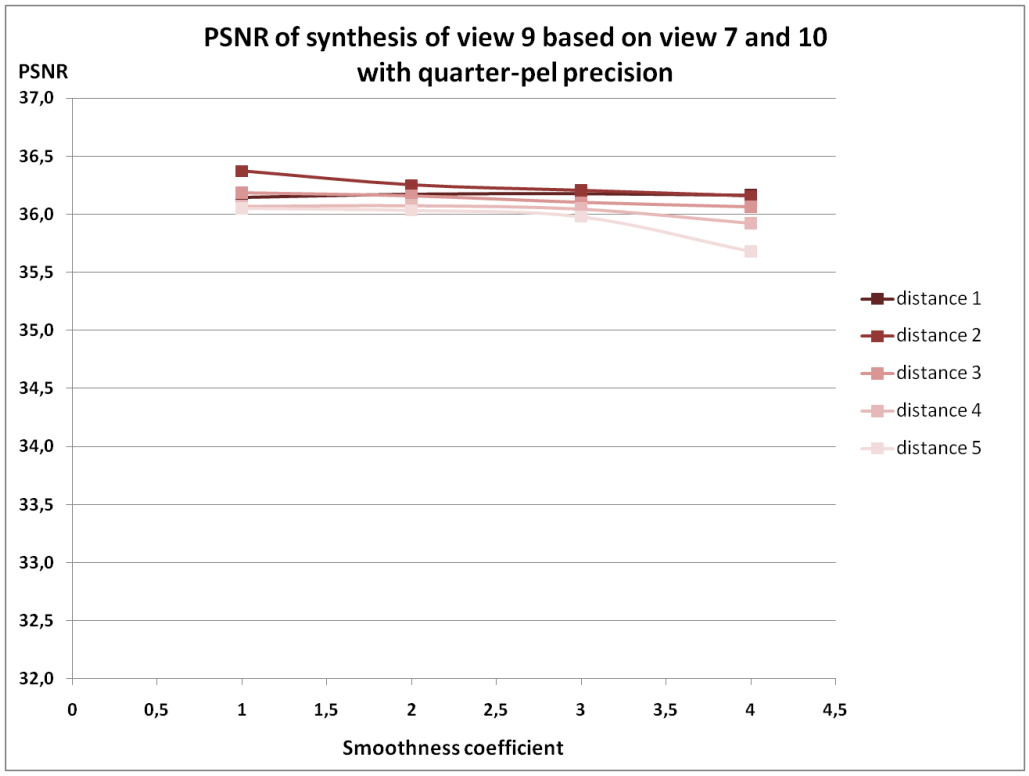


Figure 8. EE1 results, quarter-pixel precision, view 9.

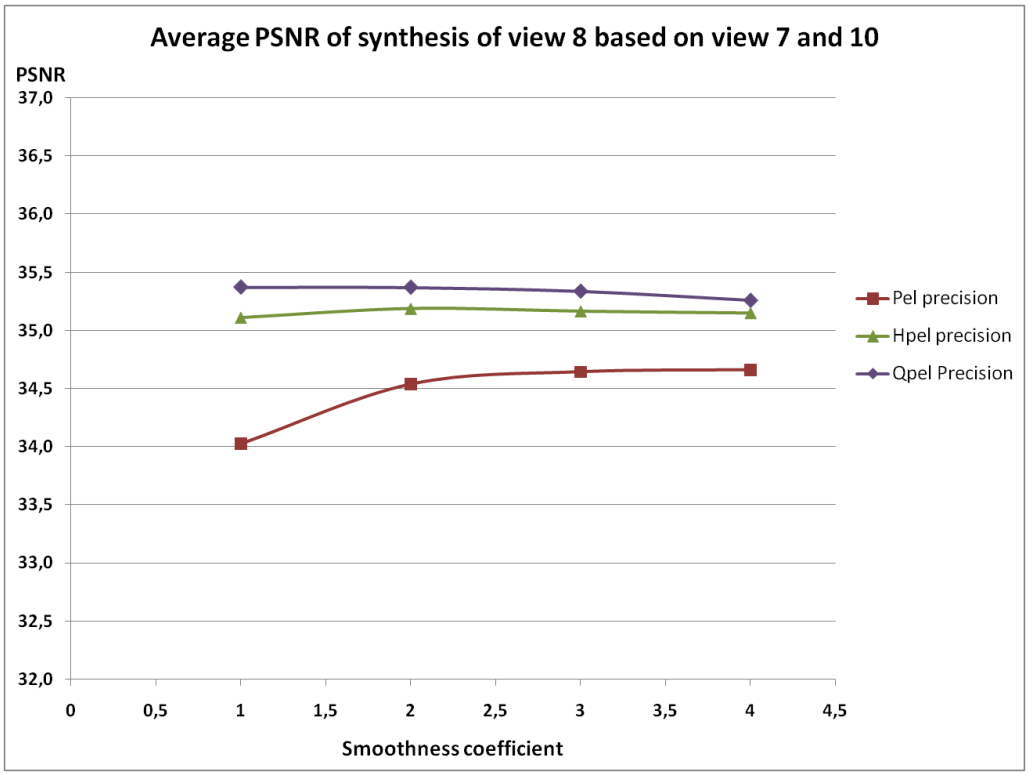


Figure 9. EE1 results, quality improvement over smoothing coefficient, view 8.

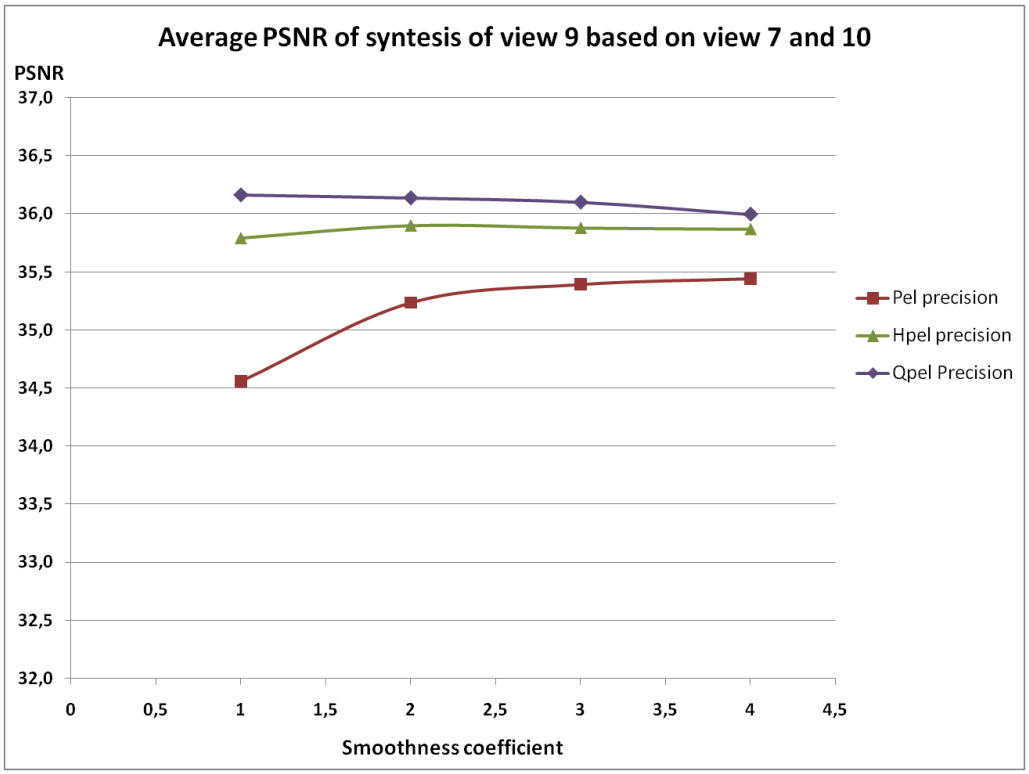


Figure 10. EE1 results, quality improvement over smoothing coefficient, view 9.

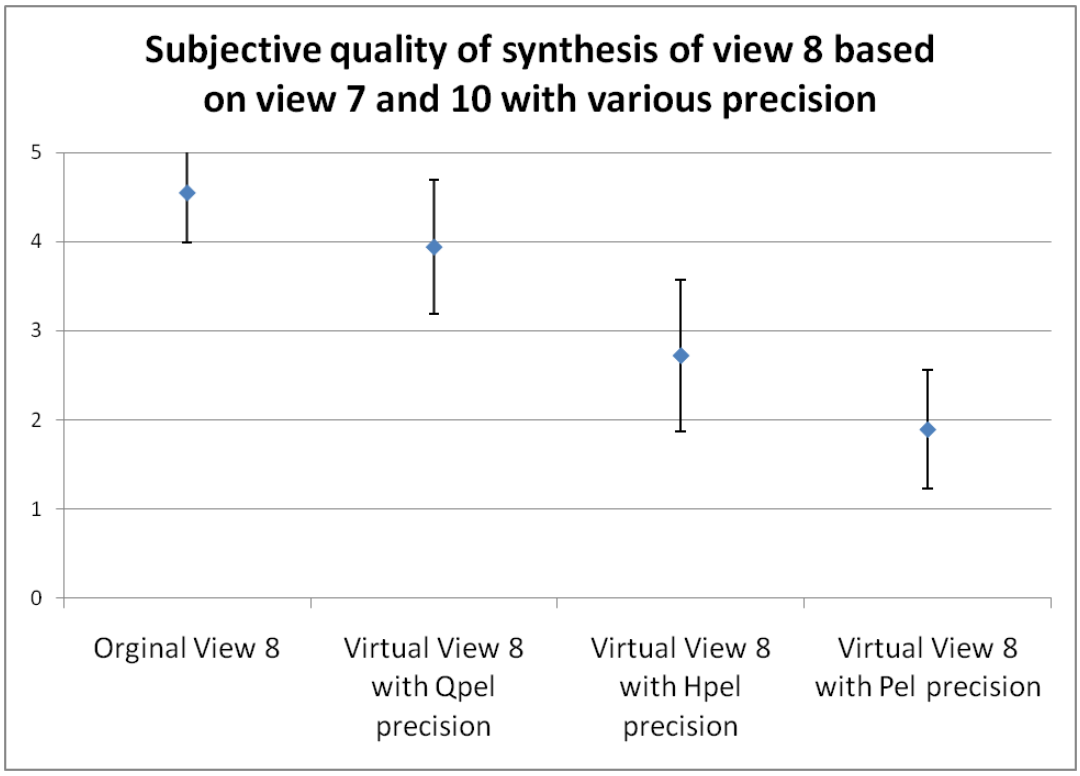


Figure 11. EE1 results, subjective quality evaluation for 60 subjects.

4 Results – EE1 – Depth estimation temporal consistency improvement

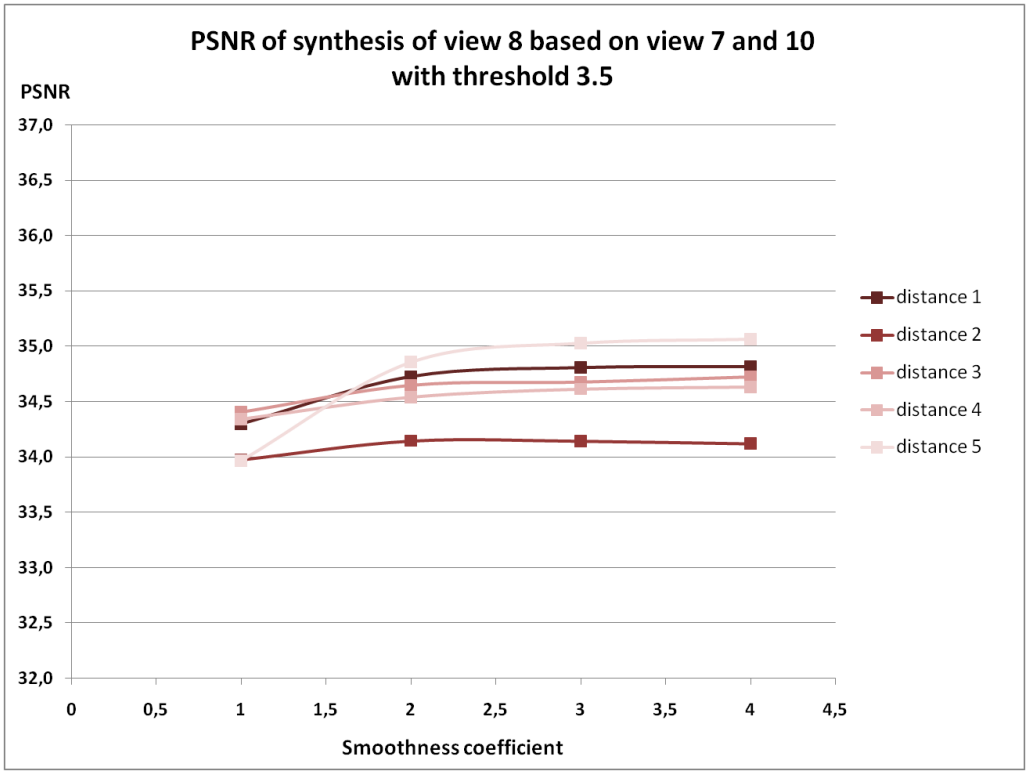


Figure 12. EE1 results, temporal enhancement, view 8, threshold 3.5.

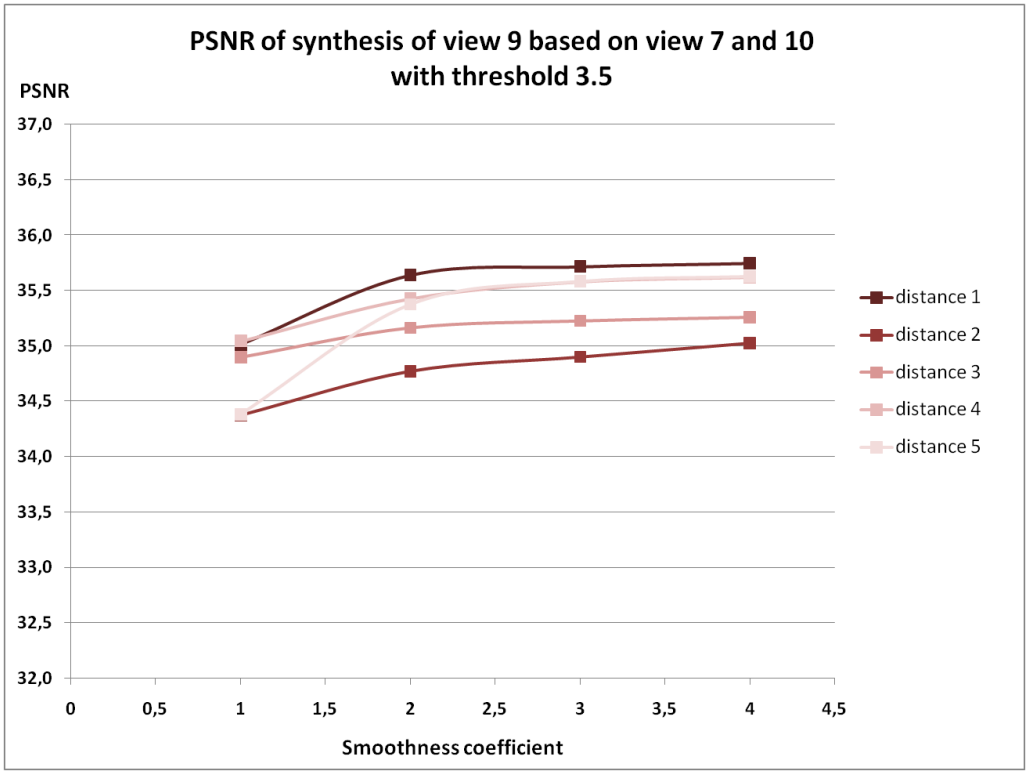


Figure 13. EE1 results, temporal enhancement, view 9, threshold 3.5.

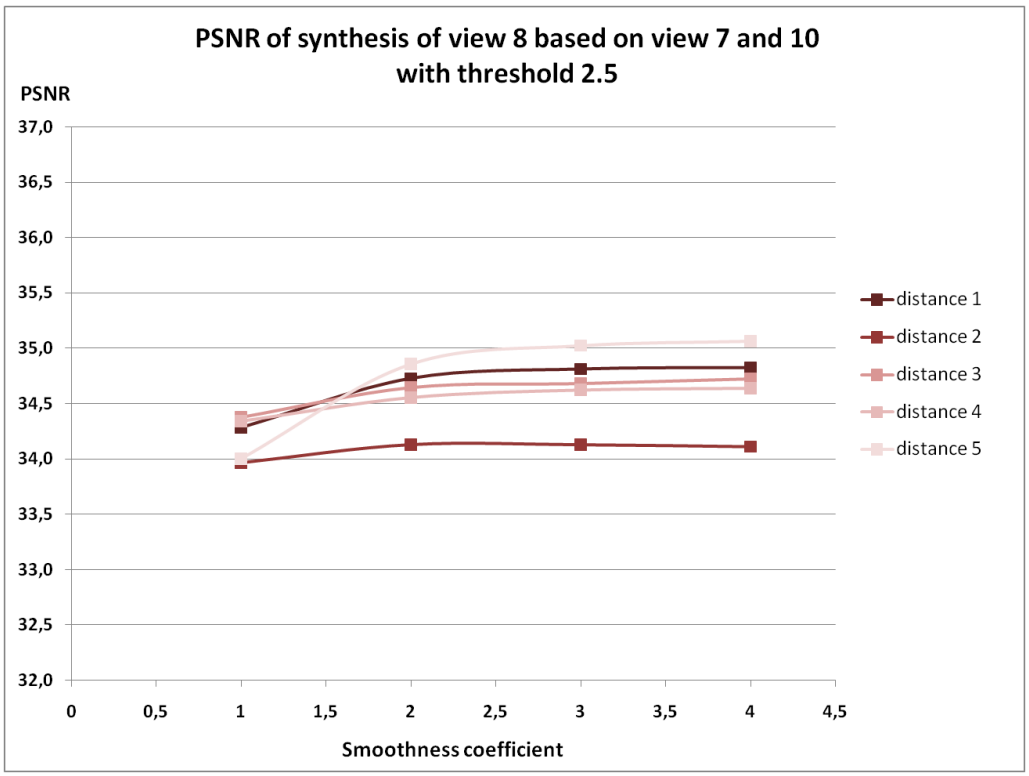


Figure 14. EE1 results, temporal enhancement, view 8, threshold 2.5.

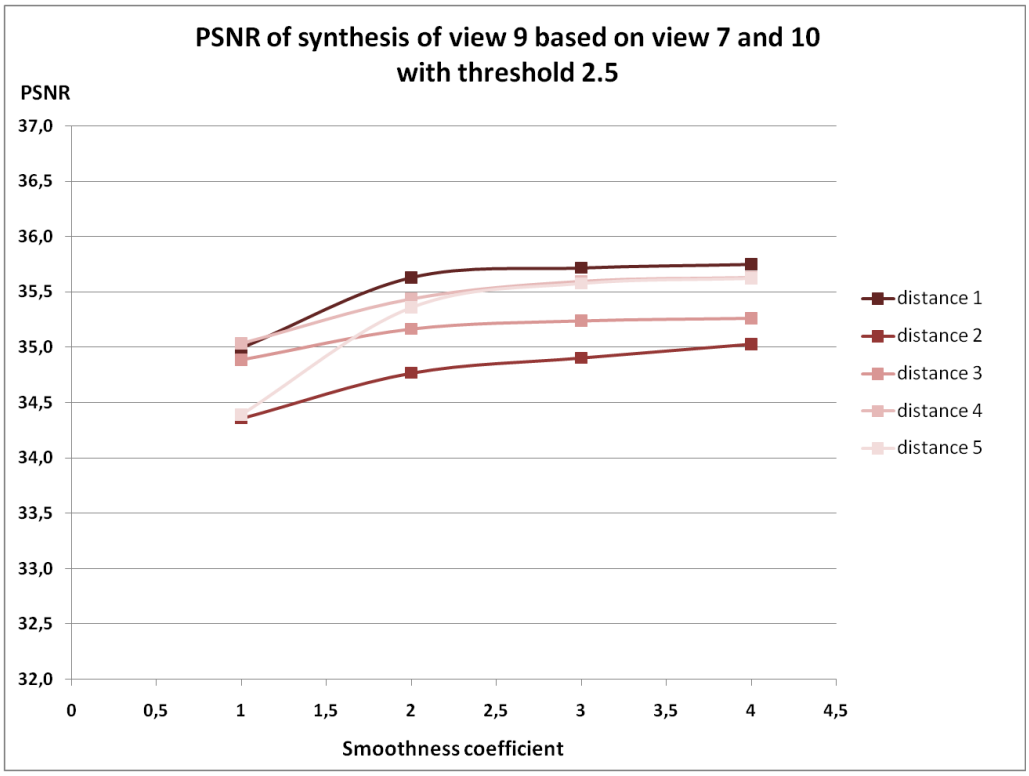


Figure 15. EE1 results, temporal enhancement, view 9, threshold 2.5.

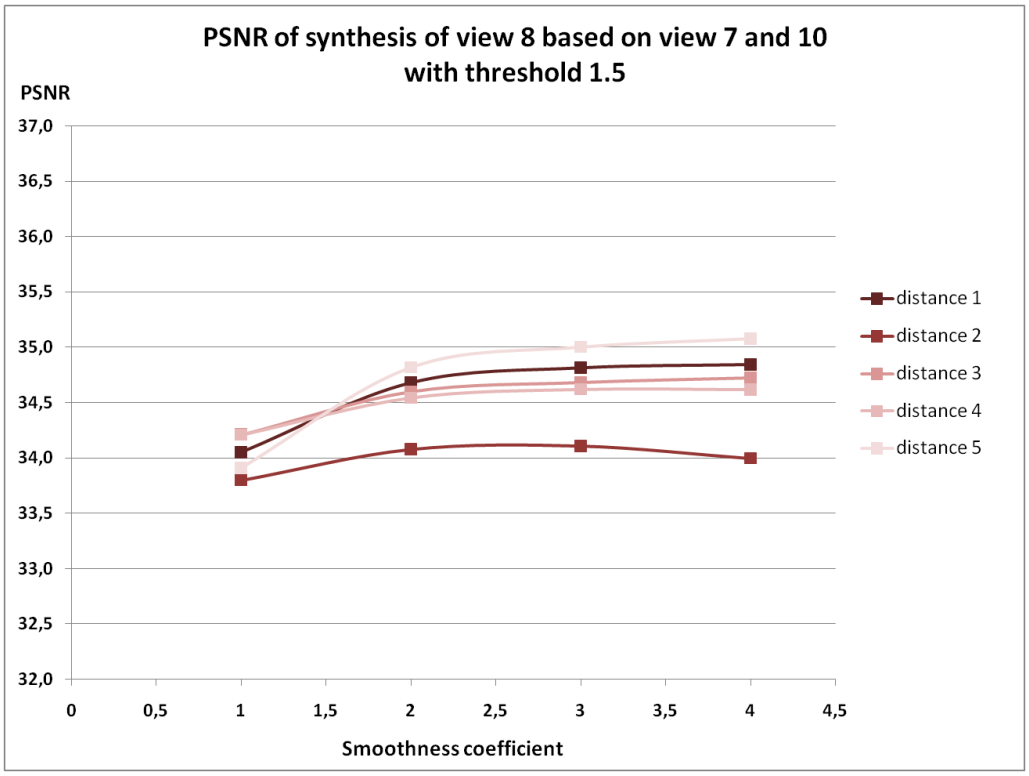


Figure 16. EE1 results, temporal enhancement, view 8, threshold 1.5.

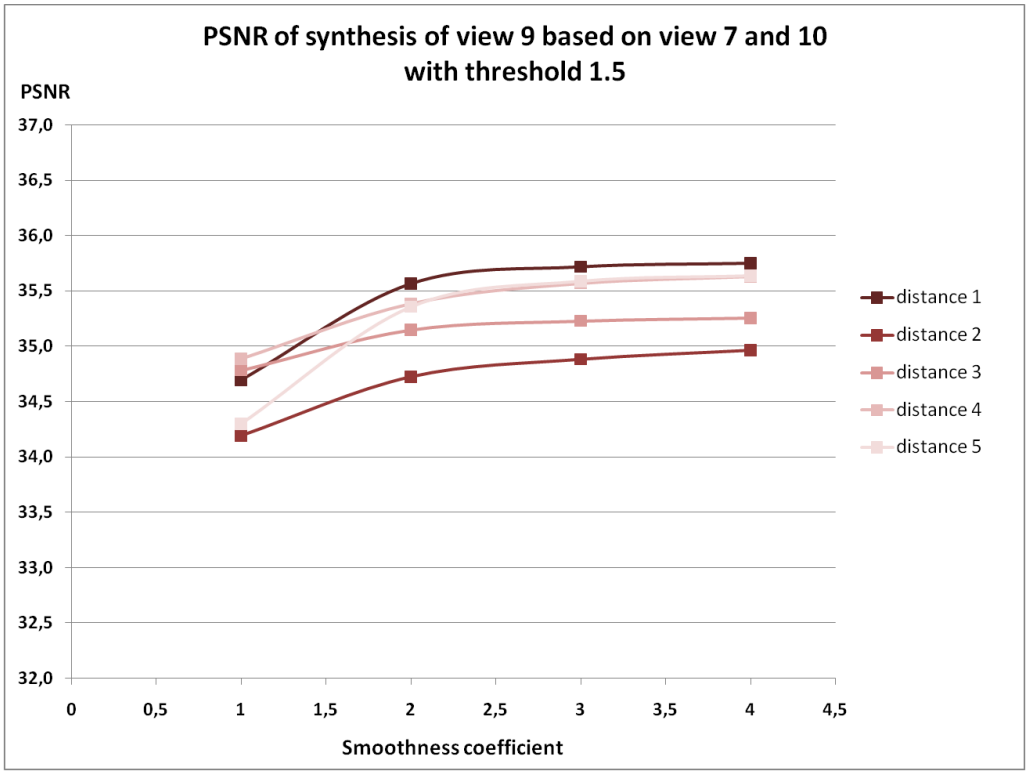


Figure 17. EE1 results, temporal enhancement, view 9, threshold 1.5.

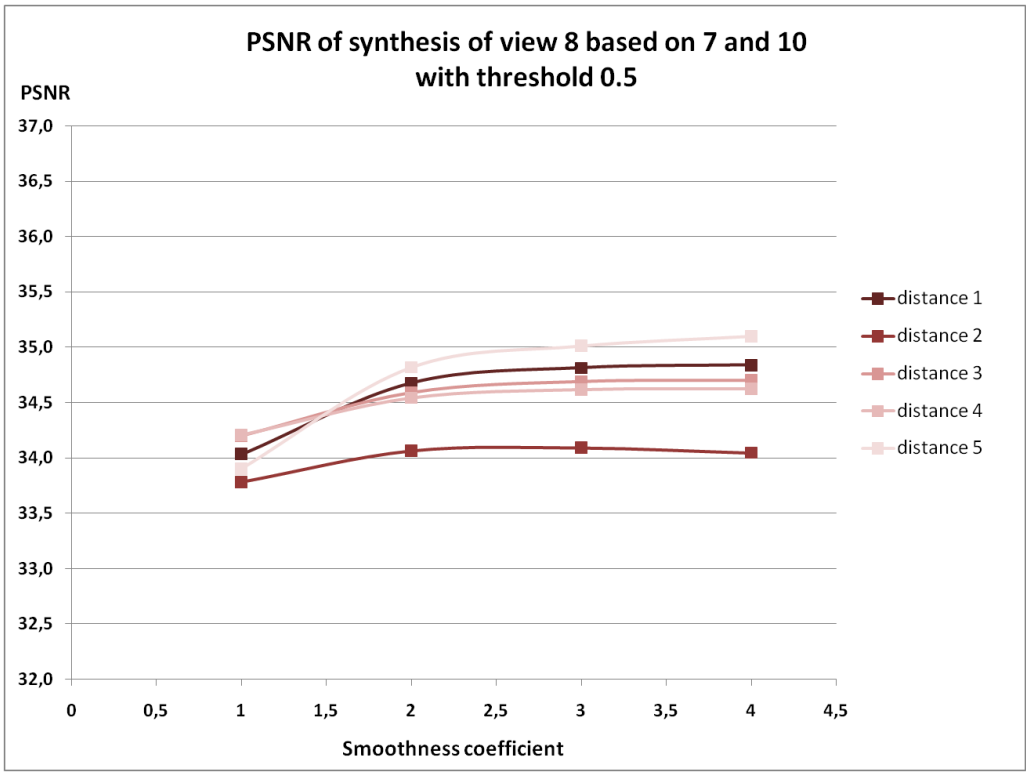


Figure 18. EE1 results, temporal enhancement, view 8, threshold 0.5.

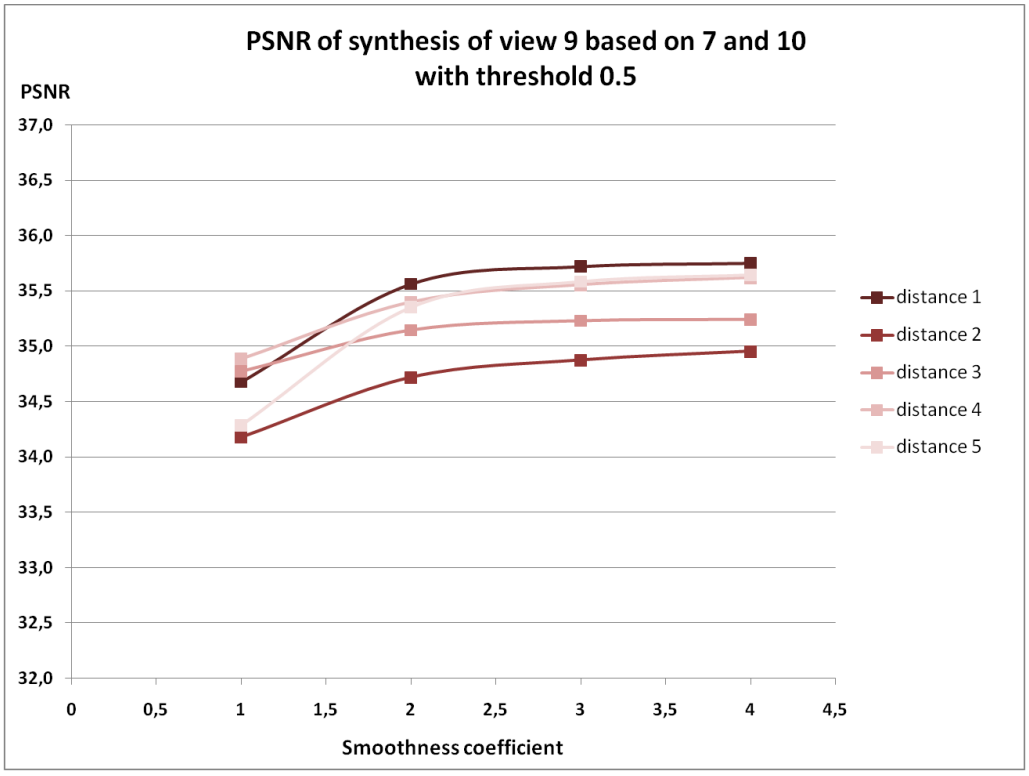


Figure 19. EE1 results, temporal enhancement, view 9, threshold 0.5.

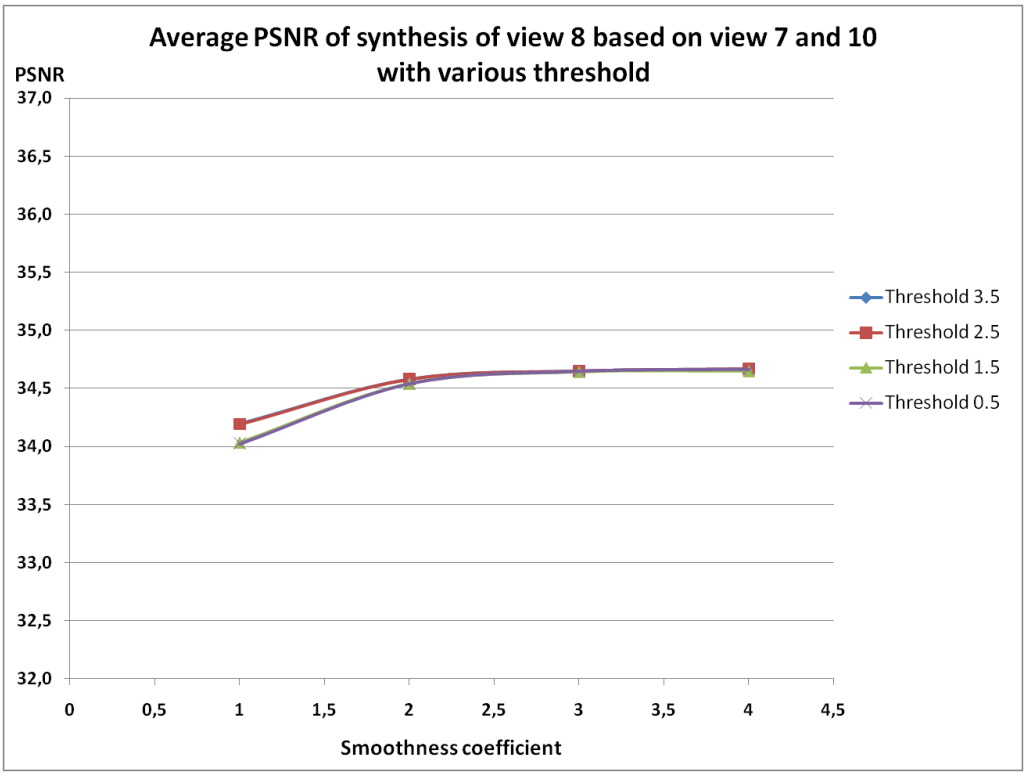


Figure 20. EE1 results, temporal enhancement averaged results, view 8.

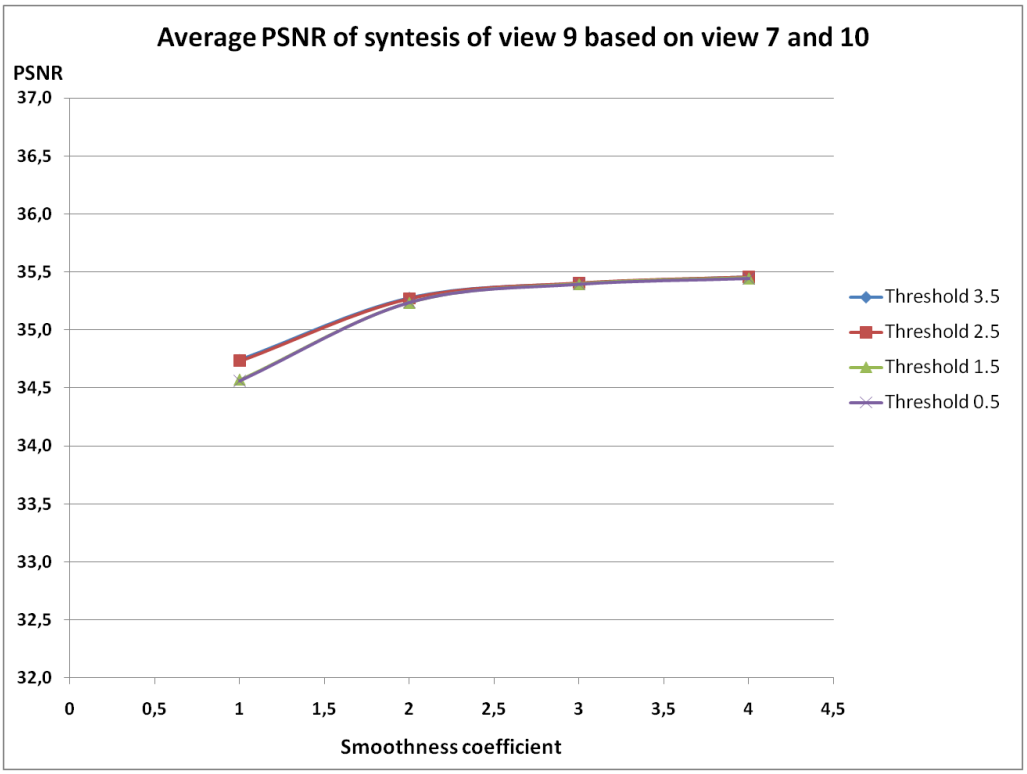


Figure 21. EE1 results, temporal enhancement averaged results, view 9.

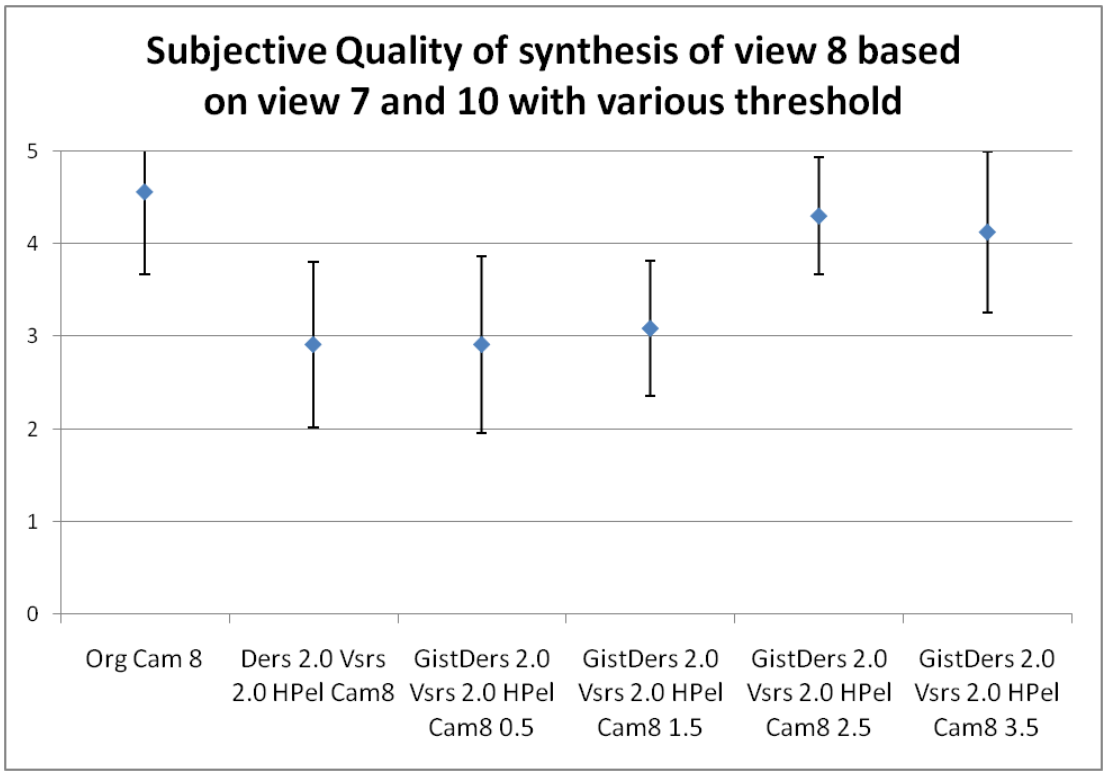


Figure 22. EE1 results, subjective evaluation of temporal enhancement improvement.

5 Results – EE2 – ViSBD

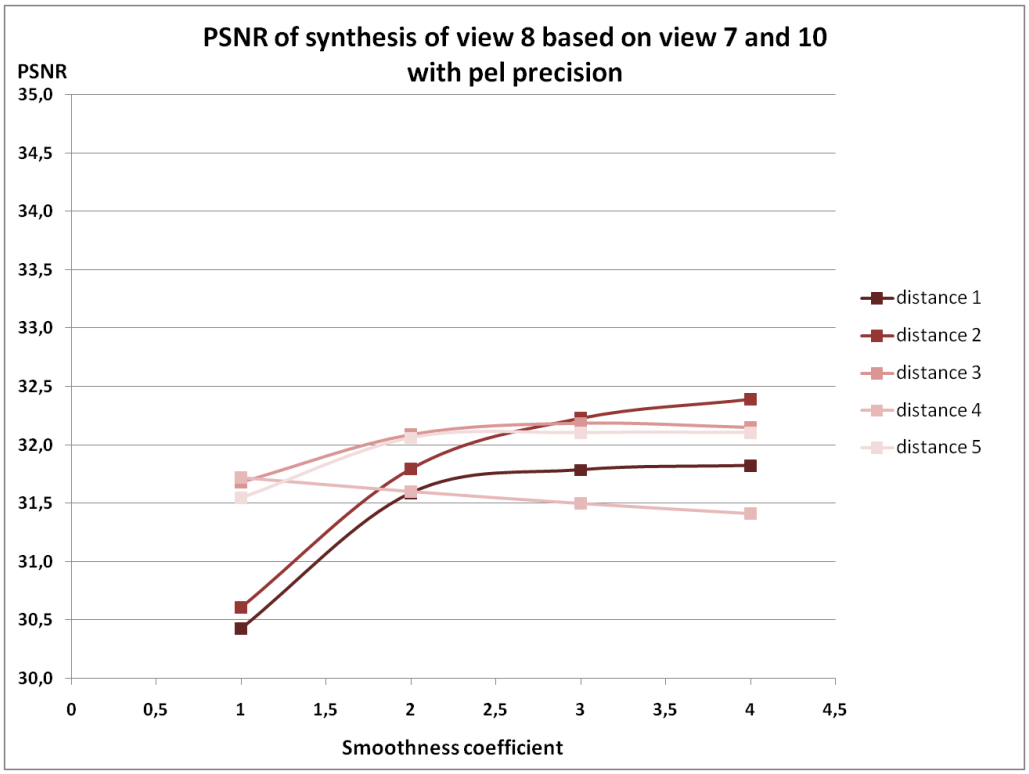


Figure 23. EE2 results, ViSBD, pixel precision, synthesis of view 8.

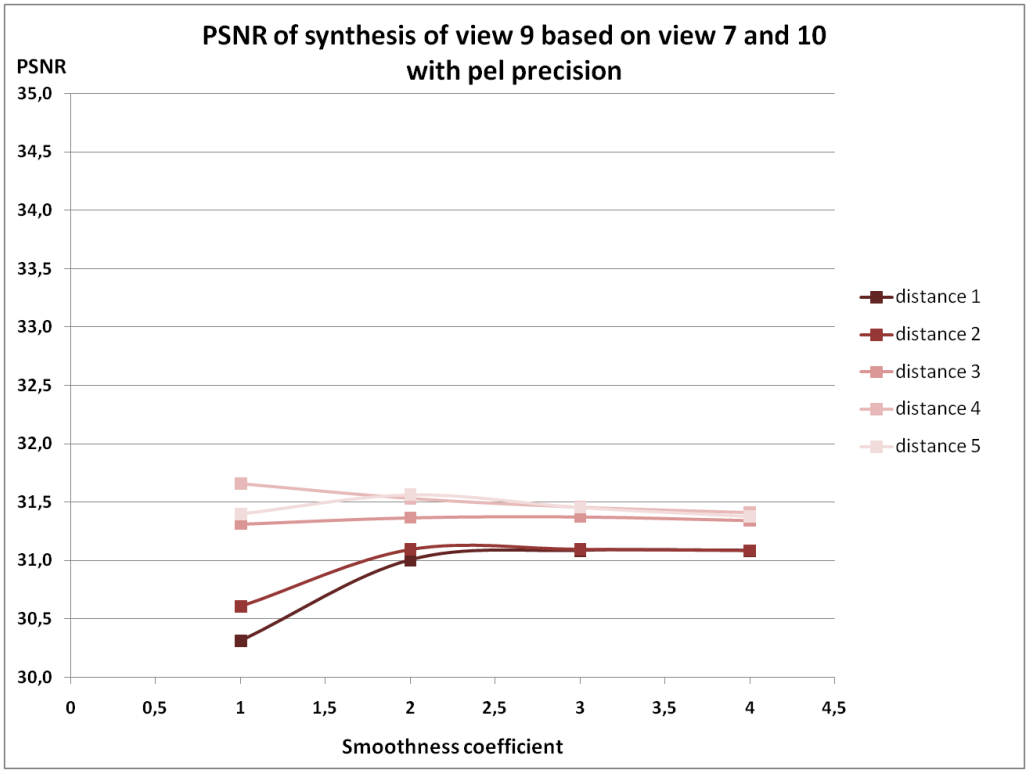


Figure 24. EE2 results, ViSBD, pixel precision, synthesis of view 9.

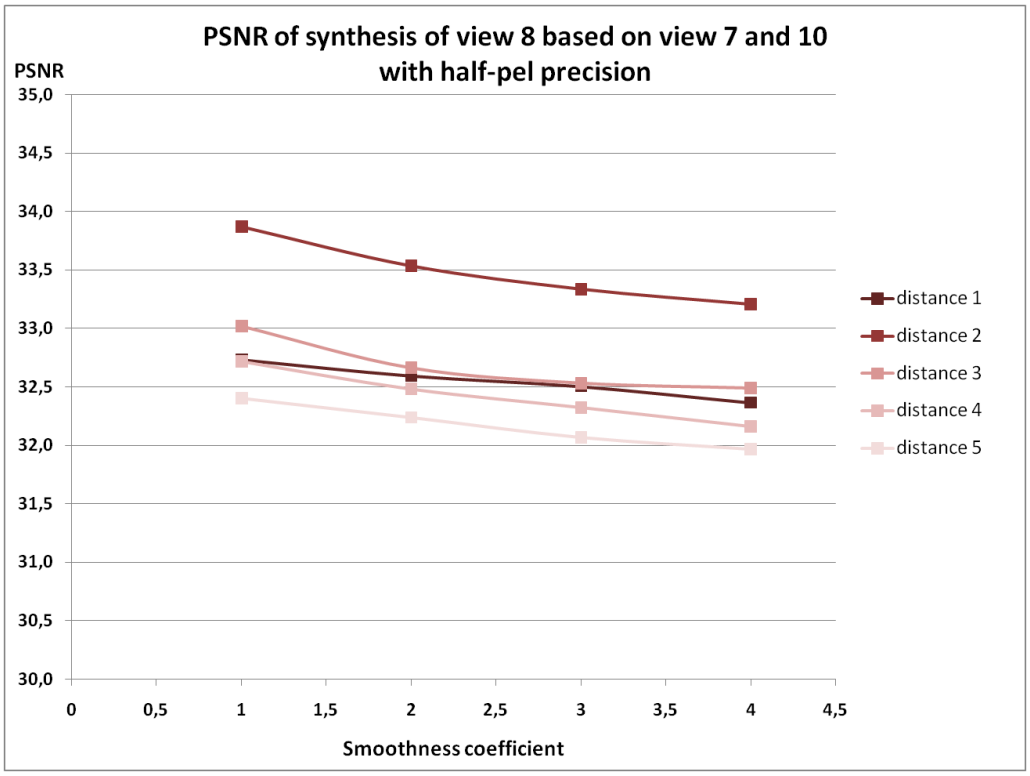


Figure 25. EE2 results, ViSBD, half-pixel precision, synthesis of view 8.

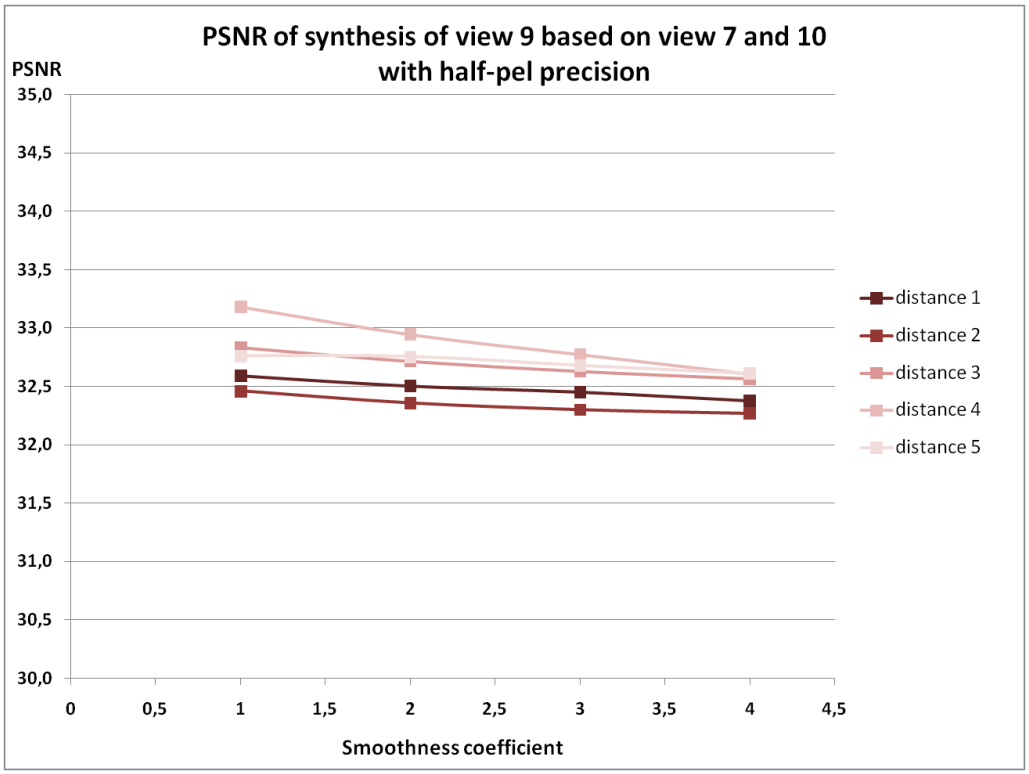


Figure 26. EE2 results, ViSBD, half-pixel precision, synthesis of view 9.

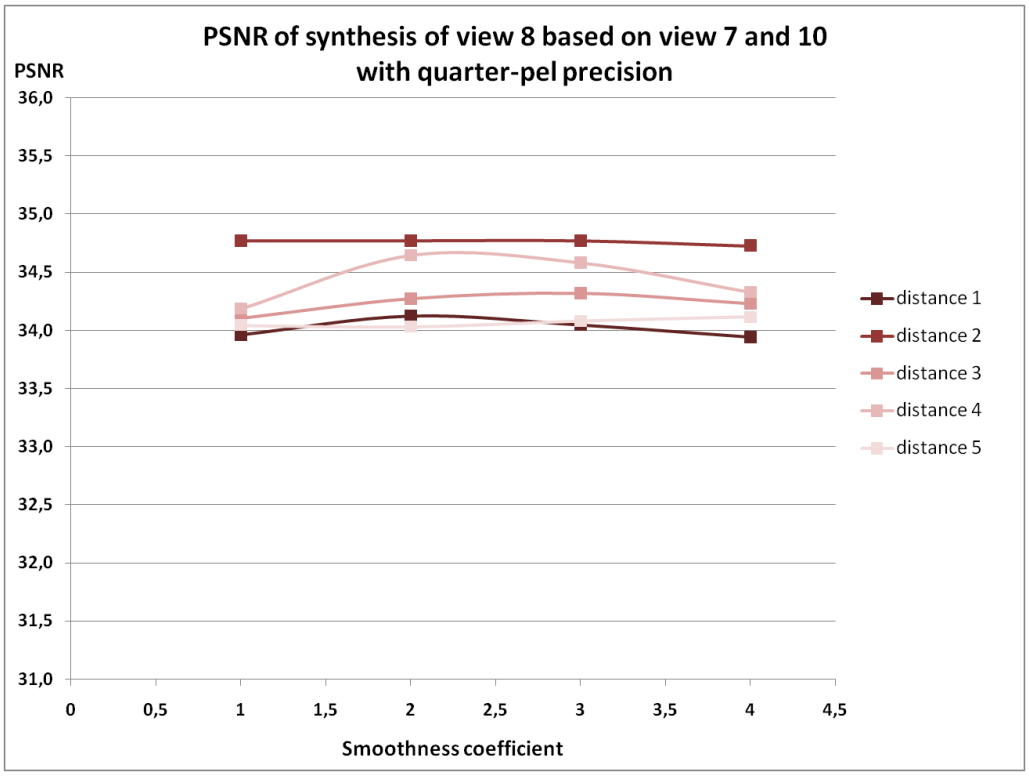


Figure 27. EE2 results, ViSBD, quarter-pixel precision, synthesis of view 8.

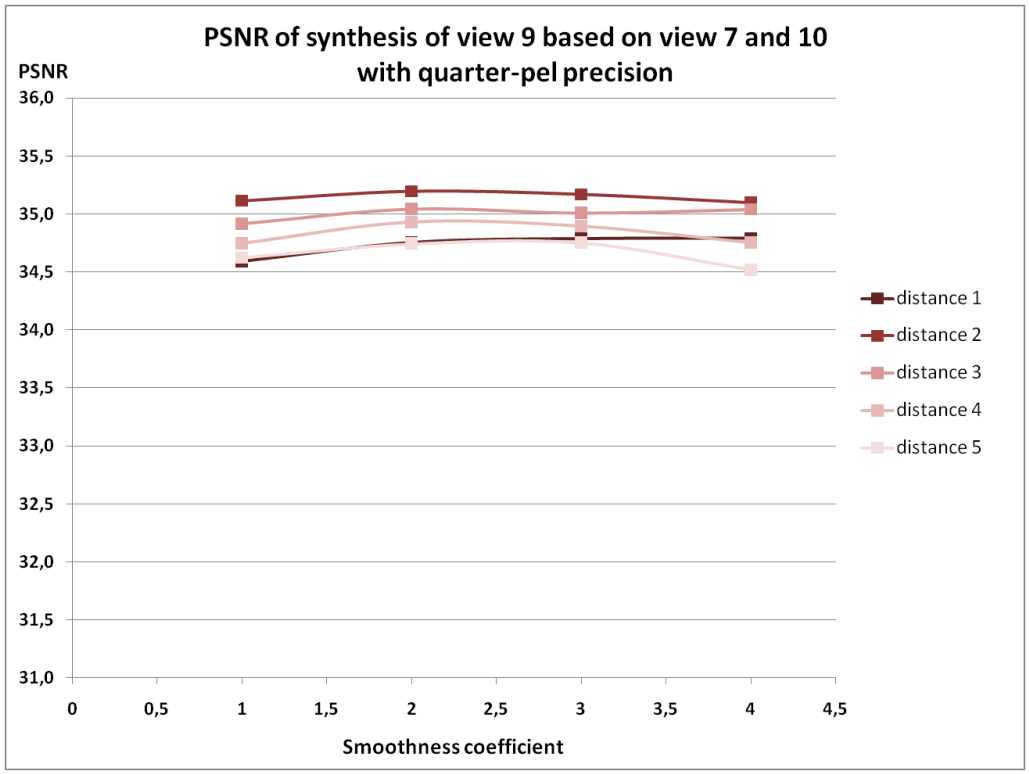


Figure 28. EE2 results, ViSBD, quarter-pixel precision, synthesis of view 9.

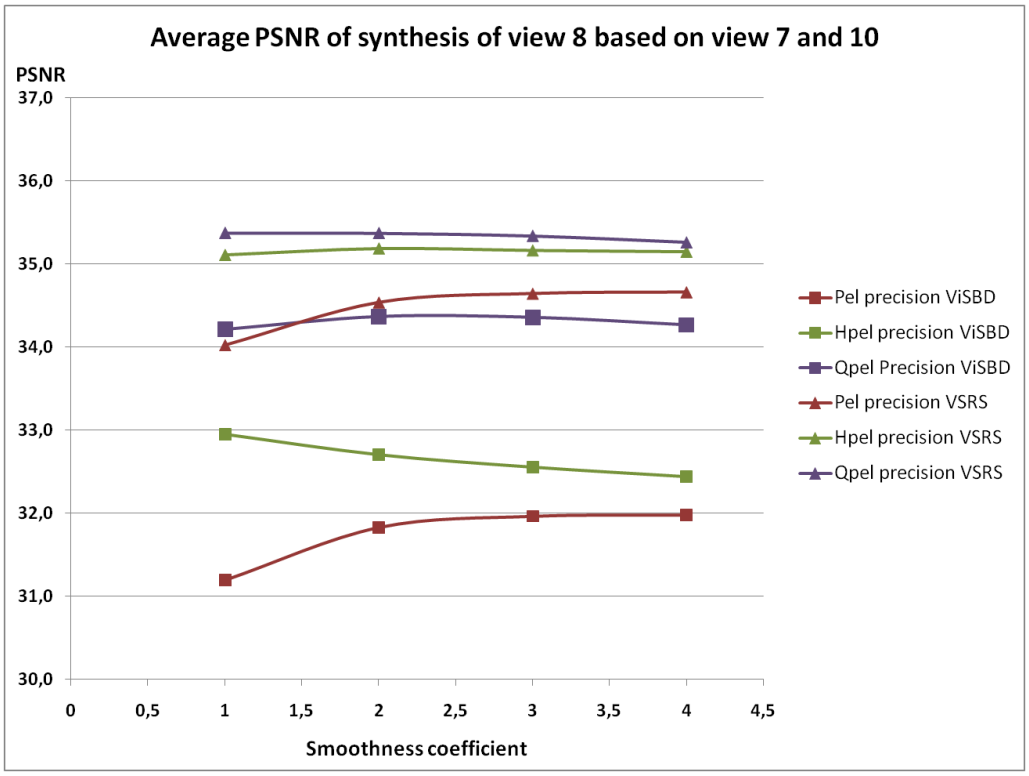


Figure 29. EE2 results, ViSBD, averaged results for view 8.

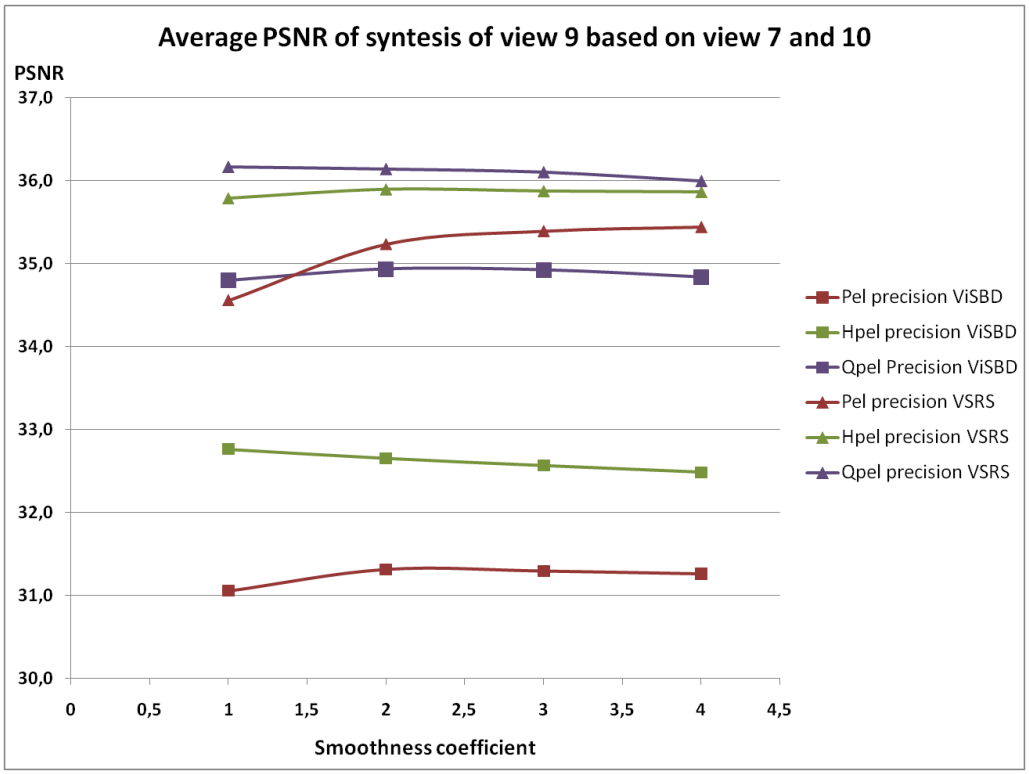


Figure 30. EE2 results, ViSBD, averaged results for view 9.

6 Results – EE4 – Coding Experiments

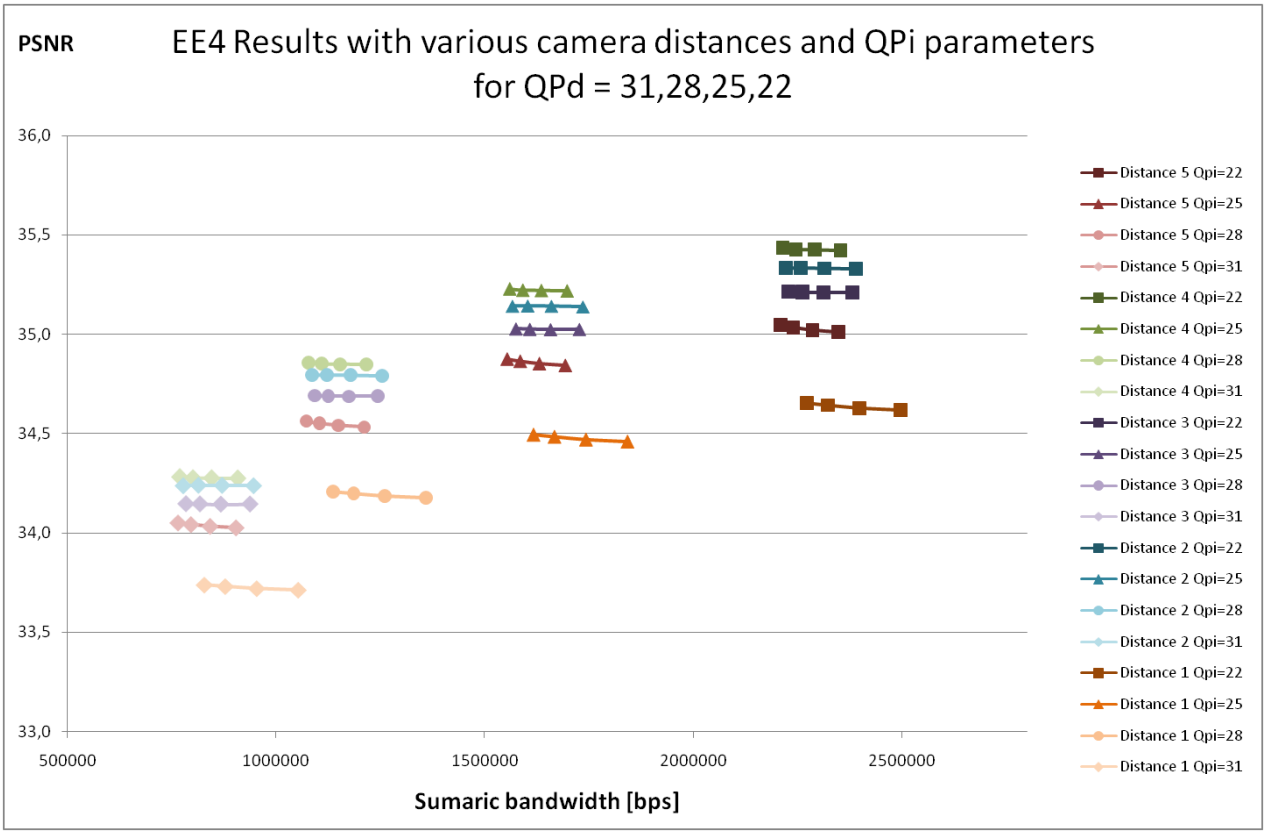


Figure 31. EE4 results, PSNR of virtual view synthesized from compressed/decompressed image and depth map streams of given bandwidth, for various depth quantization parameters (points on curves, starting from left: QPd=31,28,25,22),various image quantization parameters QPi and various camera distances (1..5).

7 Conclusions

EE1 – pixel precision:

- Half-pixel precise depth estimation and synthesis gives a gain of about **1.5dB** over pixel-accuracy ,
- Quarter-pixel precise depth estimation and synthesis gives another gain of about **0.5dB** over half-pixel-accuracy (current reference),
- Quarter-pixel has definitely the best results (except the original view) during both **objective and subjective** tests,
- New version of DERS outperforms the old one,
- The conclusions are the same for both views (8 and 9).

EE1 – temporal consistency:

- There is no PSNR improvement,
- The **subjective** quality of synthesized view is better with temporal consistency improvement than simple DERS for about **1** point (in 0...5 scale),
- The best threshold parameter for temporal consistency improvement is **2.5** (as GIST proposed).

EE2 – ViSBD:

- There is no significant improvement between ViSBD 2.0 and ViSBD 2.1,
- ViSBD performs worse than VSRS for about 1 dB (quarter-pixel precision) to 3 dB (pixel precision) of PSNR,
- The conclusions are the same for all camera distances and for both views (8 and 9),
- It is suggested that configuration files for ViSBD tool should be compatible with VSRS.

EE4 – coding experiment:

- Quality of output synthesized view depends more on quality of compressed/decompressed image (QPi parameter) than on quality of compressed/decompressed depth (QPd),
- QPd has almost no impact on output quality,
- Output quality DECREASES with greater fidelity of coded depth map – quality is the best for QPd=31 and the worst for QPd=22 – probably coding effects in more smooth depth map,
- The conclusions are the same for all camera distances 1..5 and for both views (8 and 9).

8 References

[1] "Description of Exploration Experiments in 3D Video Coding" MPEG 2008/W10173, Busan, Korea, October 2008.

[2] Feldmann, M. Mueller, F. Zilly, R. Tanger, K. Mueller, A. Smolic, P. Kauff, T. Wiegand „HHI Test Material for 3D Video", MPEG 2008/M15413, Archamps, France, April 2008.

[3] „Joint Draft 5.0 on Multiview Video Coding", ITU-T and ISO/IEC JTC1, JVT-Y209, October 2007.