

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/M16757
July 2009, London, UK

Title **An enhancement of Depth Estimation Reference Software with use of soft-segmentation**
Sub group **Video**
Authors **Olgierd Stankiewicz** (ostank@multimedia.edu.pl),
 Krzysztof Wegner (kwegner@multimedia.edu.pl) and
 Poznań University of Technology, Chair of Multimedia
 Telecommunications and Microelectronics, Poznań, Poland

1 Introduction

The current version of MPEG FTV 3DTV reference package (DERS), kindly provided by Nagoya University [1], incorporates block-matching and segmentation tools for better estimation of depth maps. In this document, we present an alternative approach: left and right images are matched with use of soft-segments, and thus block-matching and segmentation are merged in to a single tool. The proposed enhancement has been implemented and integrated with DERS software. The preliminary results (performed on most of the test sequences) show considerable improvement with respect to quality of synthesized image.

2 Soft-segmentation tool

A straight forward approach for matching of left and right image is to directly compare single pixels. Such similarity measure is very localized and vulnerable to noise. A more advanced technique that is currently implemented in DERS employs comparison with use of rectangular blocks. In this paper we propose an enhanced approach. Similarity between pixels is calculated in windows (like in DERS) with use of weighted comparison mask. This weighted comparison mask is computed from soft-segmentation masks generated around processed pixels, basing on following formula:

$$W(P, P') = e^{-\frac{|I(P)-I(P')|}{\gamma_c} - \frac{|P-P'|}{\gamma_d}}$$

where:

- P – center point in processed frame
- P' – processed point in processed frame
- $W(P,P')$ – soft-segmentation mask around center point P
- $I(P)$ – intensity of image at point P
- $|P-P'|$ – Euclidian distance between P and P'
- γ_c – color similarity parameter of the tool
- γ_d – distance similarity parameter of the tool

These soft-segmentation masks, computed for left and right image, are multiplied together to produce weighted comparison masks (Figure 1).

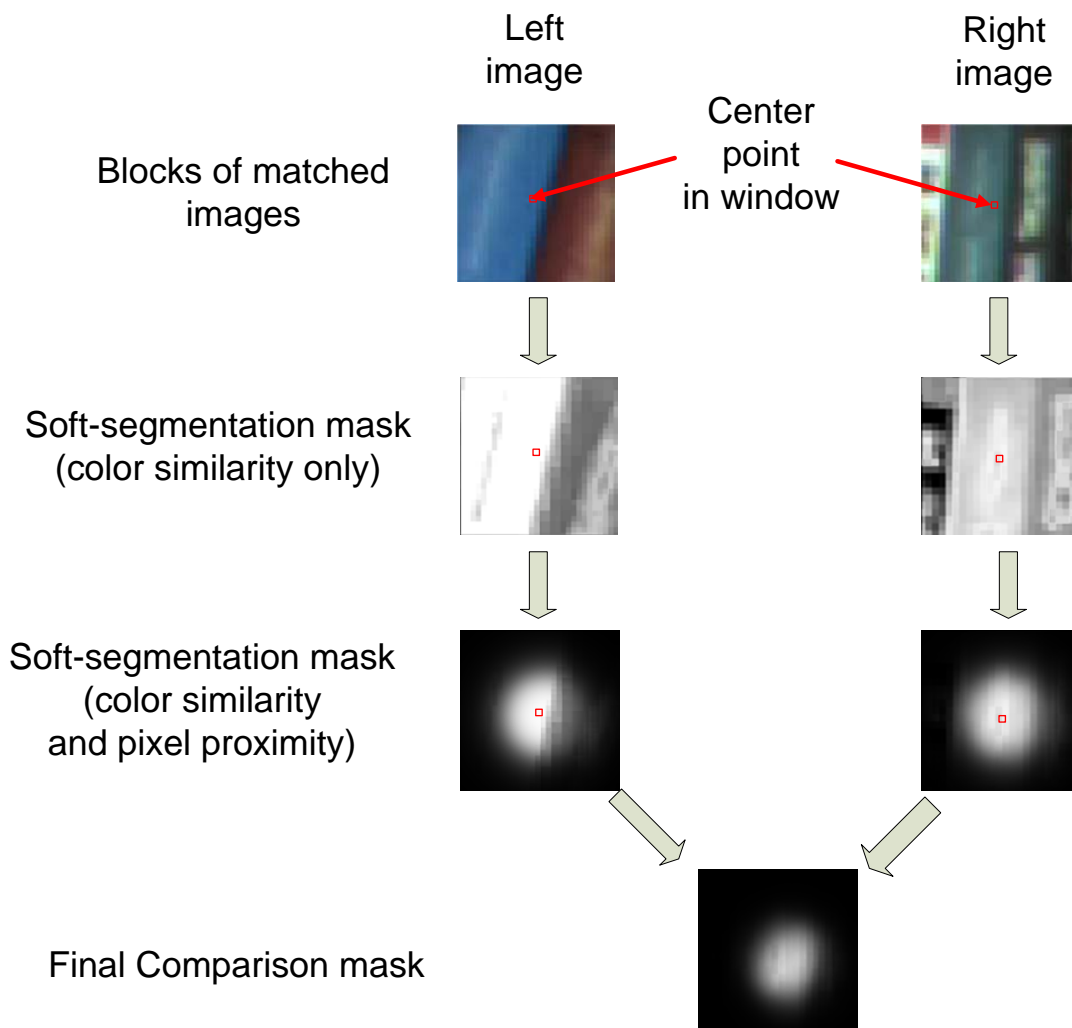


Figure 1. Generation of weighted comparison mask.

The desired similarity between pixels P_L and P_R is computed basing on modified SAD criterion. Similarity is formulated as sum of absolute differences between pixels of matched pixels, weighted by comparison mask, as follows:

$$Similarity(P_L, P_R) = \frac{\sum_{P_L', P_R' \in window} W_L(P_L, P_L') \cdot W_R(P_R, P_R') \cdot |I_R(P_R) - I_L(P_L)|}{\sum_{P_L', P_R' \in window} W_L(P_L, P_L') \cdot W_R(P_R, P_R')}$$

where:

- P_L, P_R – center point in processed frame (left/right image)
- P_L', P_R' – processed point in processed frame (left/right image)
- $W_L(P_L, P_L')$ – soft-segmentation mask around center point P (left image)
- $W_R(P_R, P_R')$ – soft-segmentation mask around center point P (right image)
- $I_L(P_L)$ – intensity of image at point P (left image)
- $I_R(P_R)$ – intensity of image at point P (right image)

In such approach, depth estimation process can be interpreted as soft-segmentation of images, followed by comparison of weighted windows, neighboring compared pixels.

3 Additional configuration parameters

The following parameters has been added to configuration files:

MatchingMethod	3	# 0...Conventional, 3...Soft segmentation matching
SoftDistanceCoeff	10.0	# SoftSegmentation Distance Coefficient
SoftColorCoeff	20.0	# SoftSegmentation Color Coefficient
SoftBlockWidth	11	# SoftSegmentation Block Width
SoftBlockHeight	11	# SoftSegmentation Block Height

4 Experiments conditions

Experiments were performed as follows (Figure 3):

- **Select stereo pair** from data set, i.e. an original left view OL and an original right view OR
- **Estimate depth** corresponding to neighboring original views OL (left) and OR (right), from neighboring cameras with various camera distances
- **Synthesize view SC**
- **Compare OC with SC** objectively and subjectively

The test were performed on ‘Newspaper’ [2] and ‘Book arrival’ [3] sequences with following views selected as SC and OL-OR.

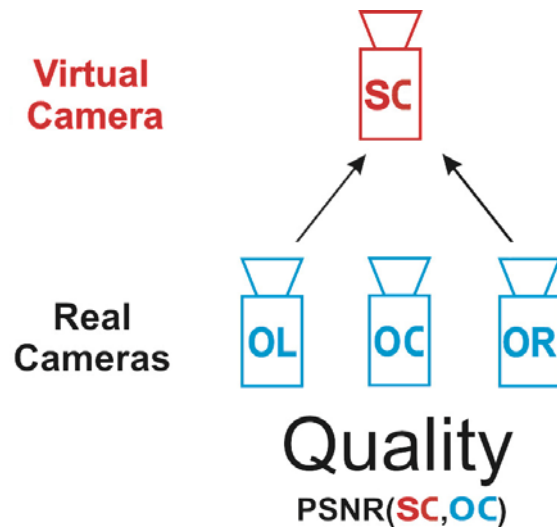


Figure 1. Setup of experiments for depth-estimation/view-synthesis software evaluation.

Table 1: Input and output views for MVD representation format.

Data set	Original Pair OL-OR	Synthesized View SC
Book arrival	7-10	8

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

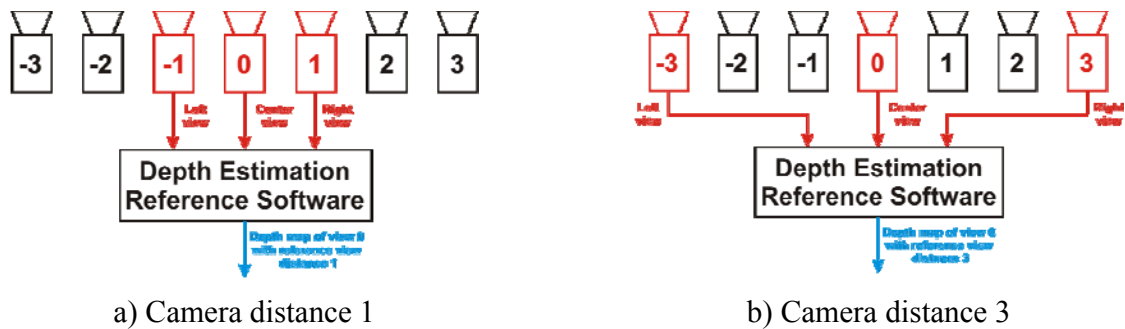


Figure 3. Setup of experiments for depth-estimation/view-synthesis software evaluation.

5 Results

5.1 Book Arrival sequence

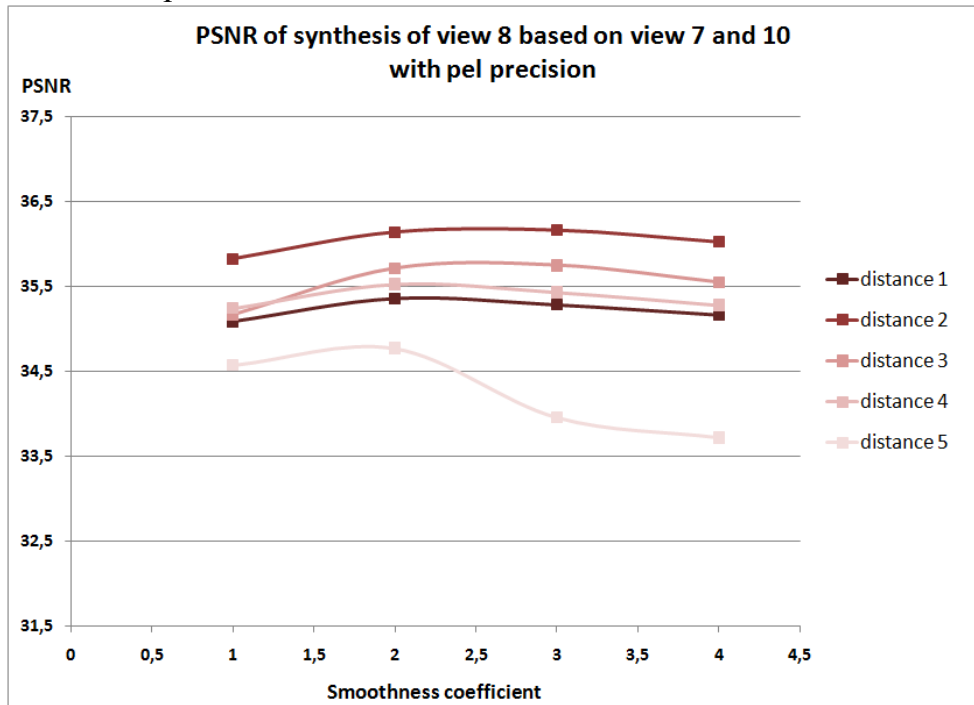


Figure 4. Results of view synthesis based on depth maps produced with **DERS** with **Pel** precision

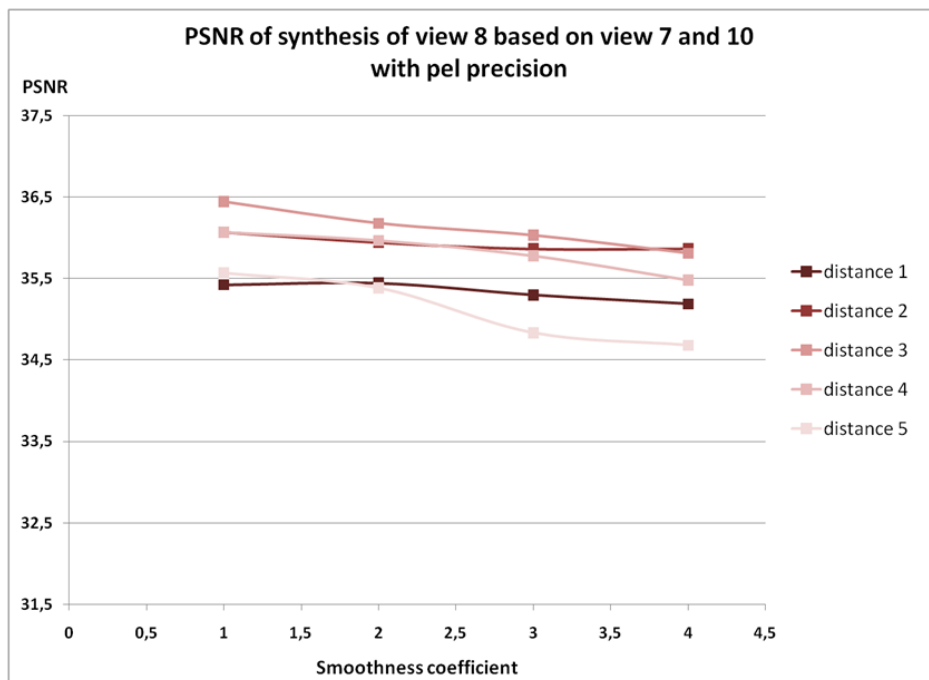


Figure 5 Results of view synthesis based on depth maps produced with use of **proposed soft-segmentation tool** with **Pel** precision.

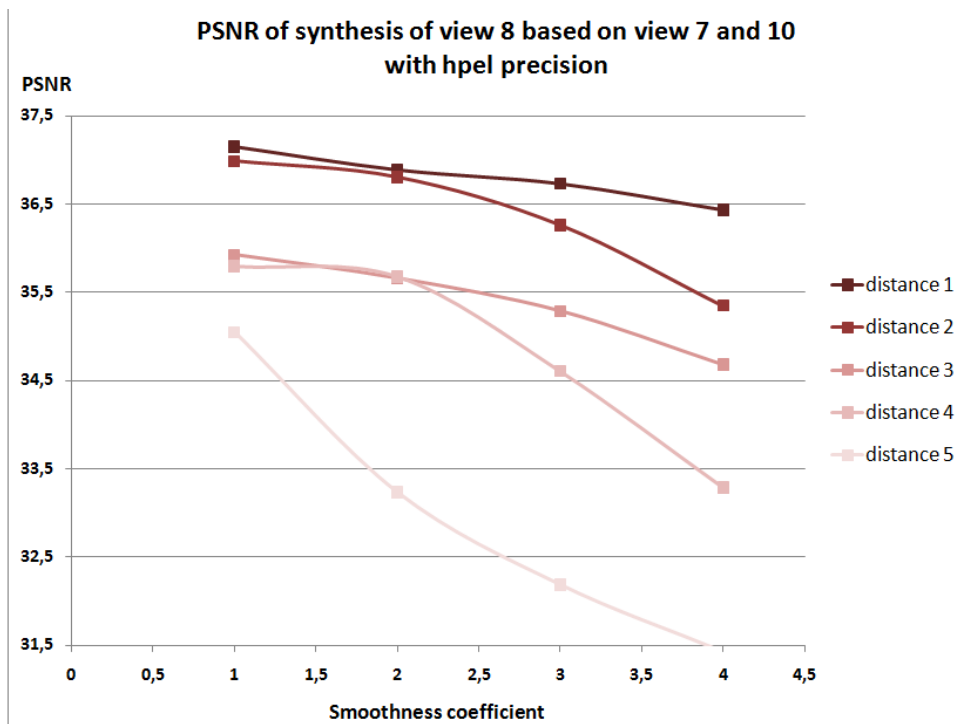


Figure 6. Results of view synthesis based on depth maps produced with **DERS** with **HPel** precision

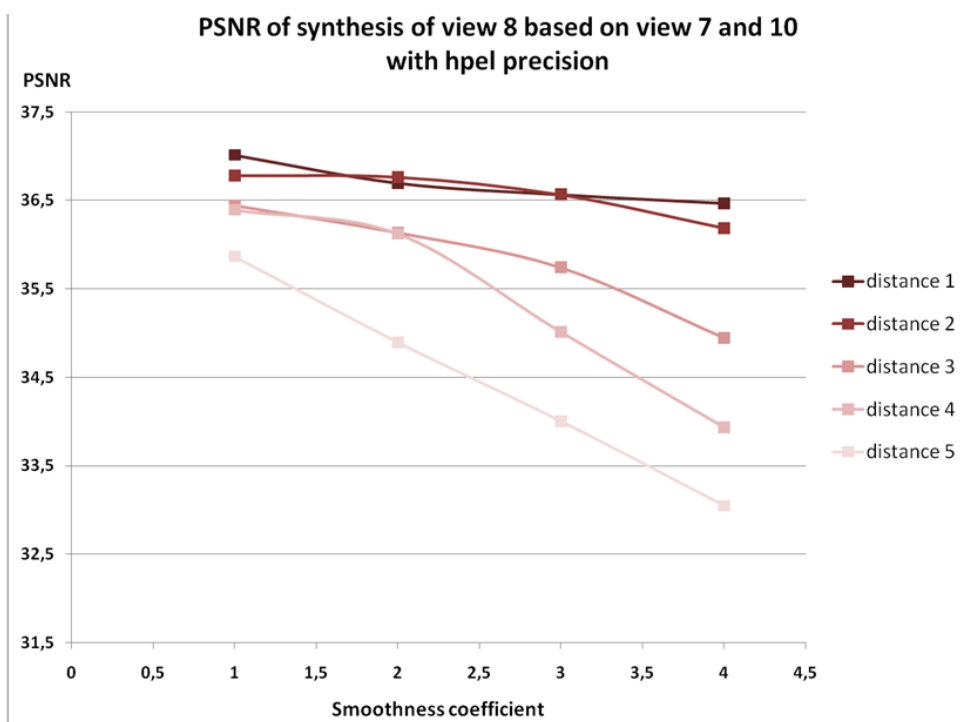


Figure 7. Results of view synthesis based on depth maps produced with use of **proposed soft-segmentation tool** with **HPel** precision

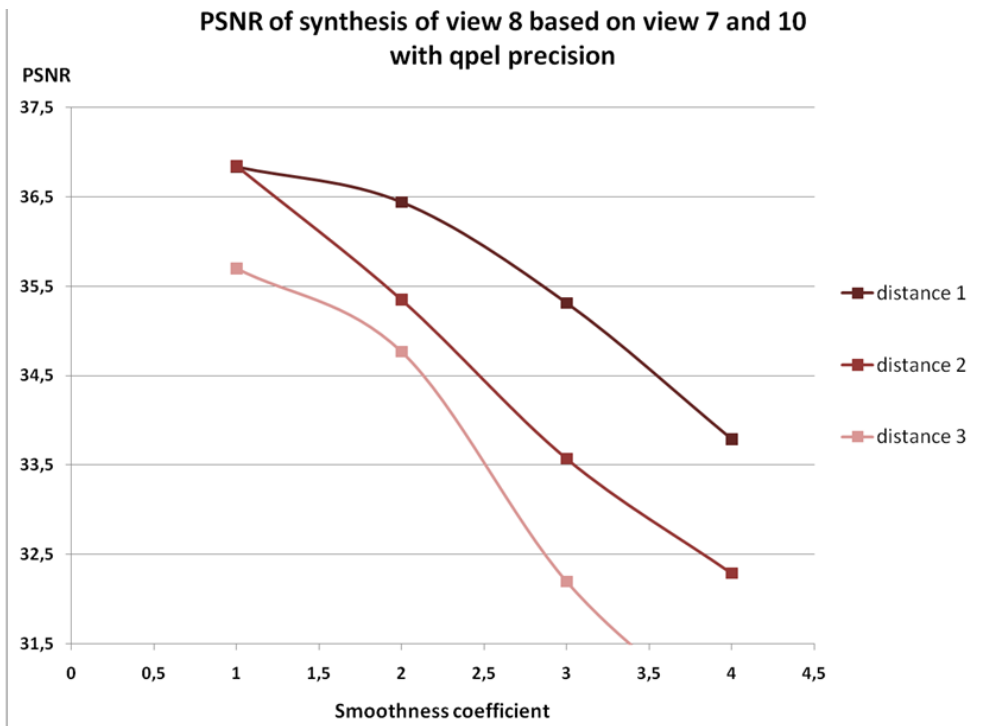


Figure 8. Results of view synthesis based on depth maps produced with **DERS** with **QPel** precision

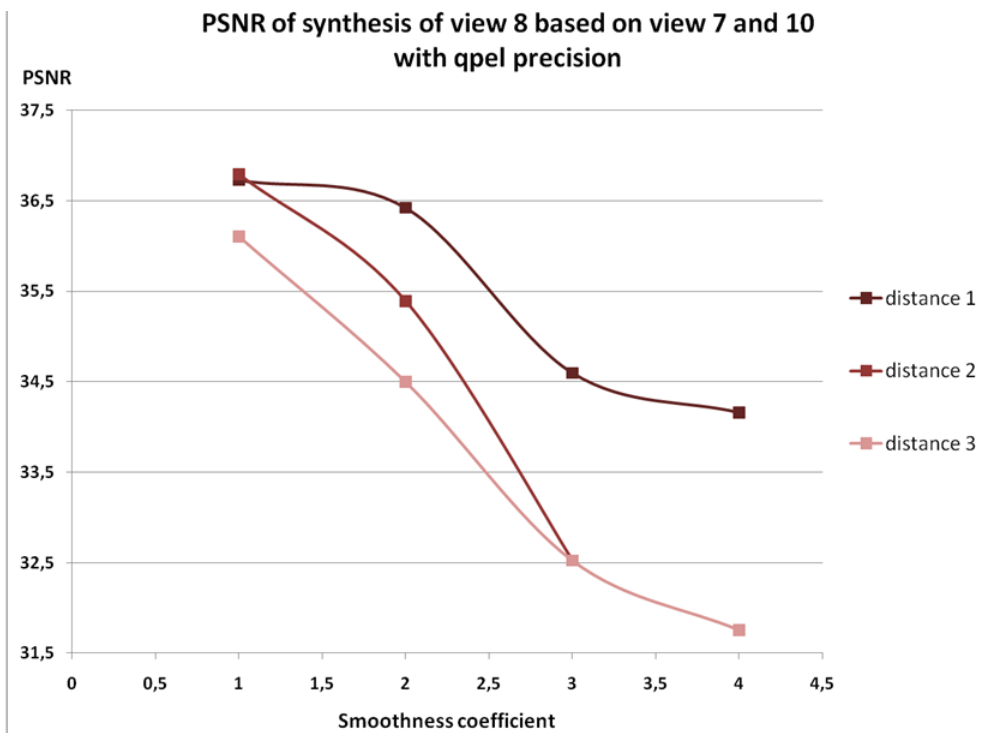


Figure 9. Results of view synthesis based on depth maps produced with use of **proposed soft-segmentation tool** with **QPel** precision

5.1 Comparison of the best synthesis results between DERS and DERS with use of proposed soft segmentation tool

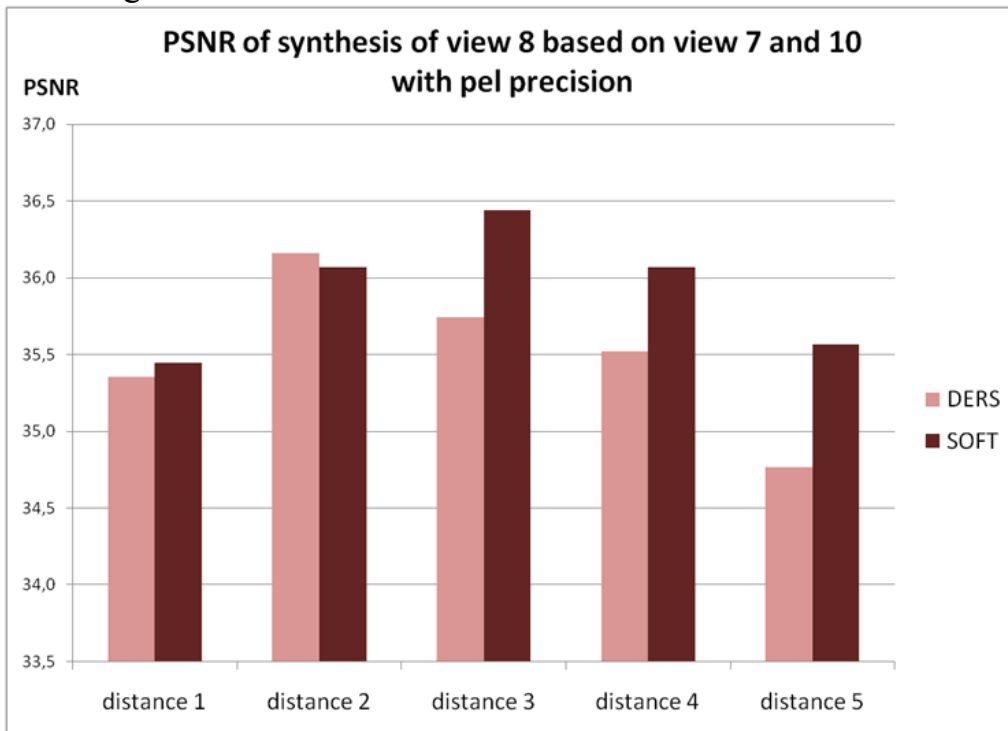


Figure 10. Comparison of the best synthesis results for given camera distance between DERS and DERS with use of proposed soft-segmentation tool (SOFT) – Pel precision

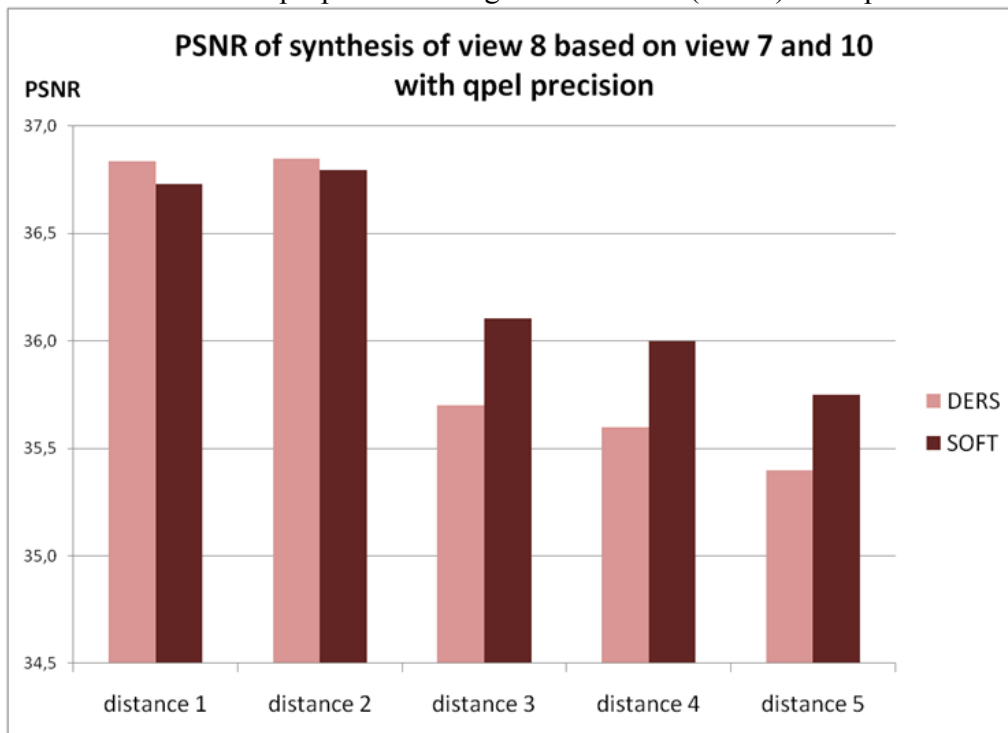


Figure 11. Comparison of the best synthesis results for given camera distance between DERS and DERS with use of proposed soft-segmentation tool (SOFT) – HPel precision

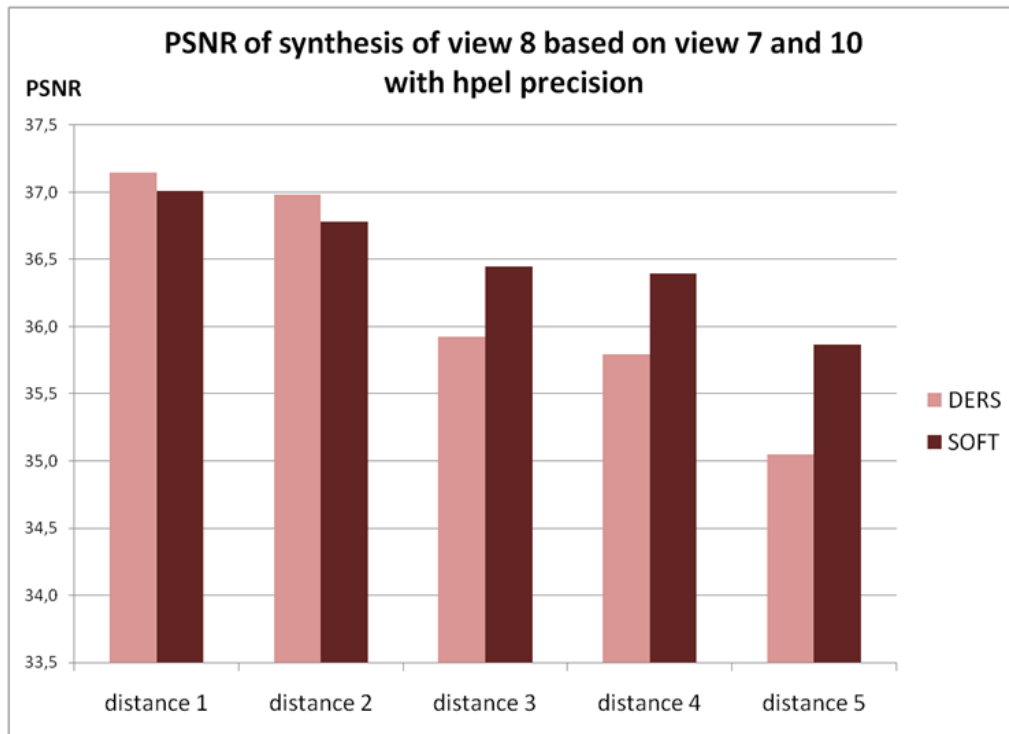


Figure 12. Comparison of the best synthesis results for given camera distance between DERS and DERS with use of proposed soft-segmentation tool (SOFT) – QPel precision

6 Conclusions

After taking the results into consideration we can conclude as follows:

- The quality of synthesis is better than the reference (gain in Pel up to 1,5dB, in HPel up to 1,7dB, in QPel up to 0,5dB).
- The gain of quality decreases with higher pixel precisions – probably due to lack of interpolation of center view in higher precision modes.
- Computation time is slightly longer than „block matching mode” in the DERS, but substantially shorter than with use of segmentation tool of DERS.
- The proposed tool is integrated in current DERS version 2.0 and can be easily added to the future versions.
- Use of soft-segmentation (instead of hard segmentation) prevents arbitrary decisions in estimation process – lesser flickering.
- The gains are higher in case of bigger camera distance – probably due to better usage of structural information.

7 Recommendations

We recommend the following:

- Integration of new tool in the future version of DERS.
- Fully fledged tests of the proposed tool on next pass of Exploration Experiments.

8 Acknowledgement

This work was supported by the public funds as a research project in years 2007 - 2009.

K. Wegner is the scholarship holder from Foundation for Polish Science.

9 References

- [1] "Description of Exploration Experiments in 3D Video Coding" MPEG 2008/W10173, Busan, Korea, October 2008.
- [2] Yo-Sung Ho, Eun-Kyung Lee, and Cheon Lee "Multiview Video Test Sequence and Camera Parameters", ISO/IEC JTC1/SC29/WG11, MPEG 2008/M15419, Archamps, France, April 2008.
- [3] I. Feldmann, M. Müller, F. Zilly, R. Tanger, A. Smolic, P. Kauff, T. Wiegand, „HHI Test Material for 3D Video”, ISO/IEC JTC1/SC29/WG11, MPEG 2008/M15413, Archamps, France, April 2008.