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/M17049
October 2009, Xian, China

Source **Poznań University of Technology, Chair of Multimedia Telecommunications and Microelectronics**
Nagoya University
Status **Contribution**
Title **A soft – segmentation matching in Depth Estimation Reference Software (DERS) 5.0**
Author Olgierd Stankiewicz **(ostank@multimedia.edu.pl),**
Krzysztof Wegner **(kwegner@multimedia.edu.pl),**
Menno Wildeboer **(wildeboer@tanimoto.nuee.nagoya-u.ac.jp)**

1 Introduction

This document shortly describes a new soft-segmentation matching tool, added to the Depth Estimation Reference Software (DERS) 5.0 as released on September 24th 2009. The first version of the tool has been presented by Poznań University of Technology (m16757 [2]) on London meeting and has been improved by joint works of Nagoya University and Poznań University of Technology.

2 Soft-segmentation matching

One of the most crucial elements in depth estimation algorithm is matching of left and right images. First versions of DERS utilized a straightforward approach - direct compare of single pixels. Such similarity measure is very localized and vulnerable to noise. A more advanced technique, implemented in DERS 4.0 [1], employs comparison with use of rectangular blocks. Is that approach, small window around processed pixels is also taken for computation of similarity measure.

The described soft-segmentation matching enhances depth estimation by modification of image matching technique. The similarity between left and right image is computed basing on weighted comparison mask. This mask describes significance of pixels neighboring with processed pixels in comparison process. Devised weighted comparison mask is computed from two soft-segmentation masks (for left and right image separately) generated around processed pixels, basing on following formula:

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

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,
γ_d	– distance similarity parameter.

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

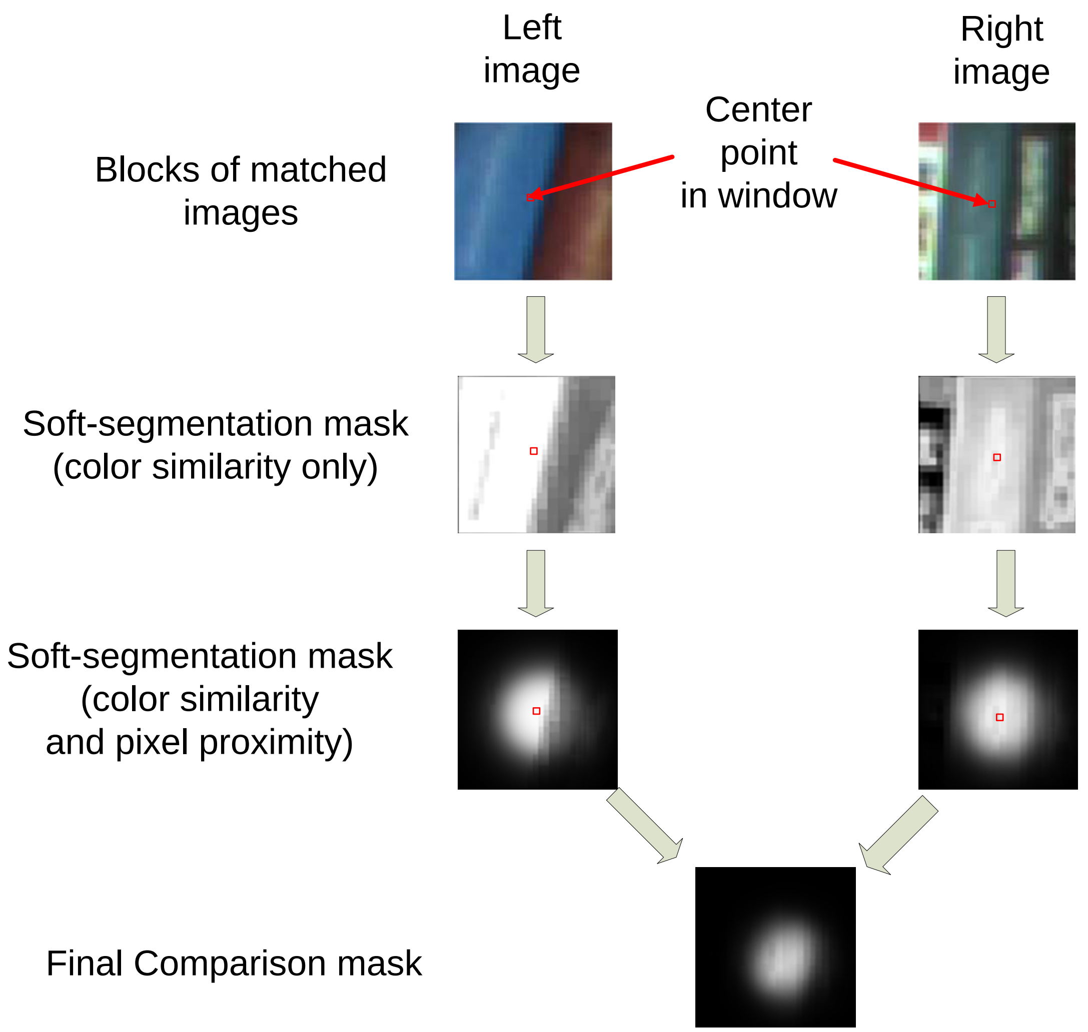


Figure 1. Generation of weighted comparison mask.

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:

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

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).

The resultant similarity value, computed basing on formula above, is feed as similarity metric instead of classical SAD computed in pixels (DERS 1x1) or blocks (DERS 3x3)

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

3 Configuration parameters

The integration of soft-segmentation matching into DERS imposed introduction of additional configuration parameters to configuration file. The following parameters (left column) 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

The second column presents typical (but arbitrarily chosen) parameter values, and the last column presents description of parameters meaning.

4 Preliminary assessment of technique

In order to evaluate quality of devised soft-segmentation matching, some experiments were performed. These experiments, which were following MPEG guidelines for exploration experiments, were basing on comparison of quality between synthesized view SC and original view OC. The whole experiment setup was as follows (Figure 2):

- **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**

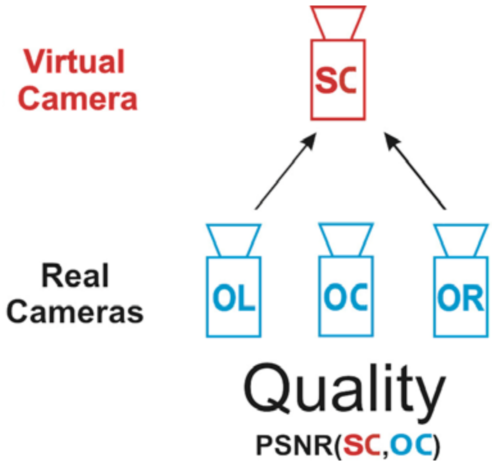
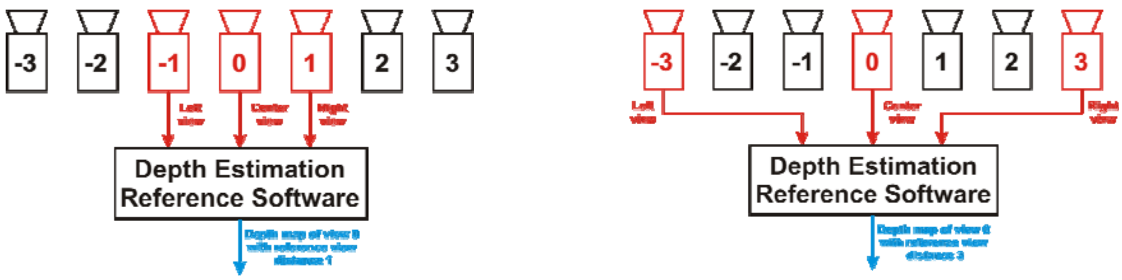


Figure 2. Setup of experiments for depth-estimation/view-synthesis software evaluation.

The depth estimation was performed with various Camera Distance (Figure 3) parameters– from distance 1 to distance 5.



a) Camera distance 1 b) Camera distance 3

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

For the sake of comparison between old block matching and new soft-segmentation matching, results (depth estimation and view synthesis) were produced for following cases:

- DERS in manual mode, 1x1 pixel matching,
- DERS in manual mode, 3x3 block matching,
- DERS in manual mode, soft-segmentation matching,
- DERS in semi-automatic mode, 1x1 pixel matching,
- DERS in semi-automatic mode, 3x3 block matching,
- DERS in semi-automatic, soft-segmentation matching.

5 Results

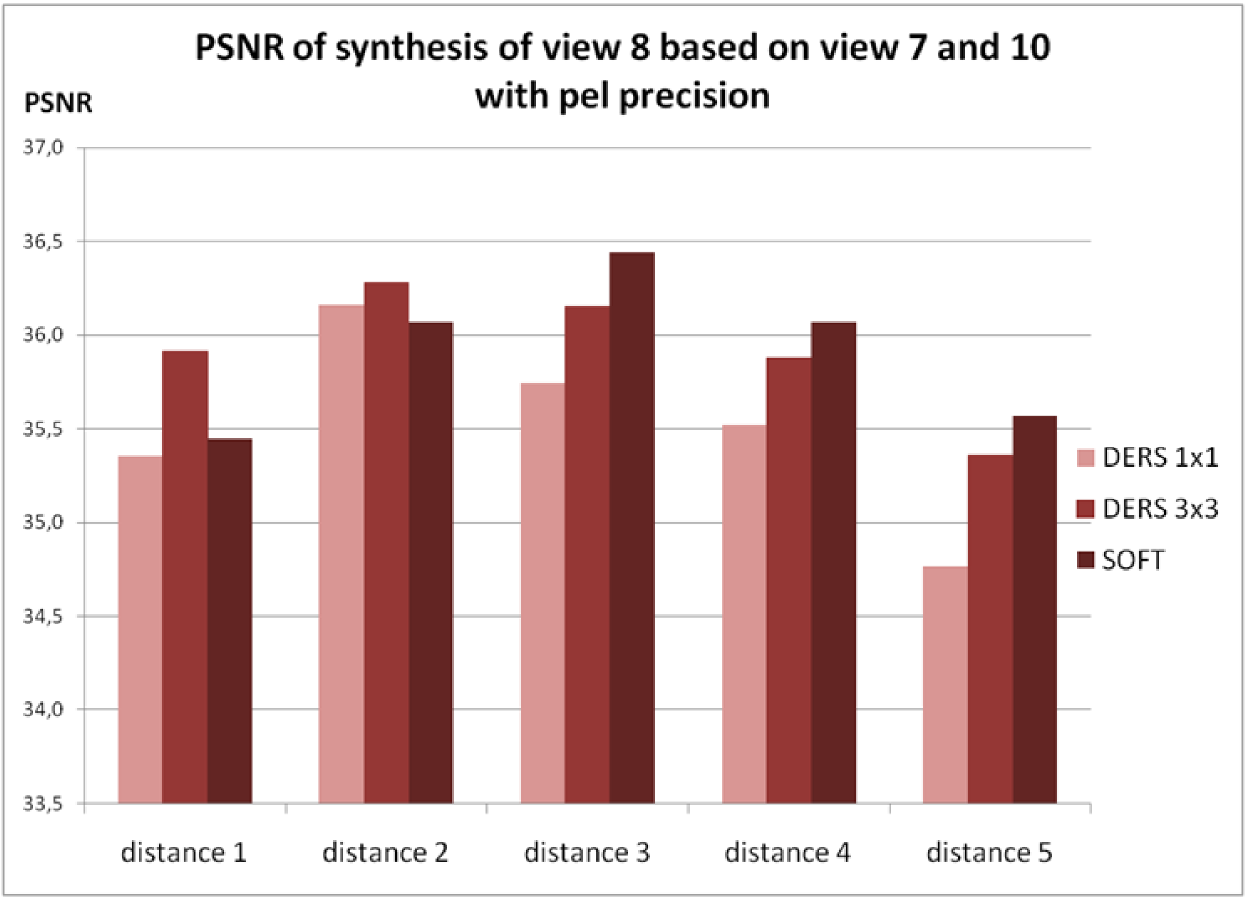


Figure 4. Comparison of synthesis results of 1x1 pixel matching, 3x3 block matching and new soft-segmentation matching. DERS was in manual mode/Pel precision. Only the best results, with respect to other parameters (smooting coefficient etc.), are presented.

As shown in Figure 4, soft-segmentation matching (SOFT) outperforms block matching modes (DERS 1x1 and DERS 3x3) for camera distances greater than 2, in case of pixel precision. The gain of new technique is of about 0.8÷1.5dB from DERS 1x1, and of about 0.3÷0.5dB from DERS 3x3. For remaining camera distances (1 and 2), the best-performing option is to use DERS in 3x3 block matching (it outperforms SOFT for about 0.8dB).

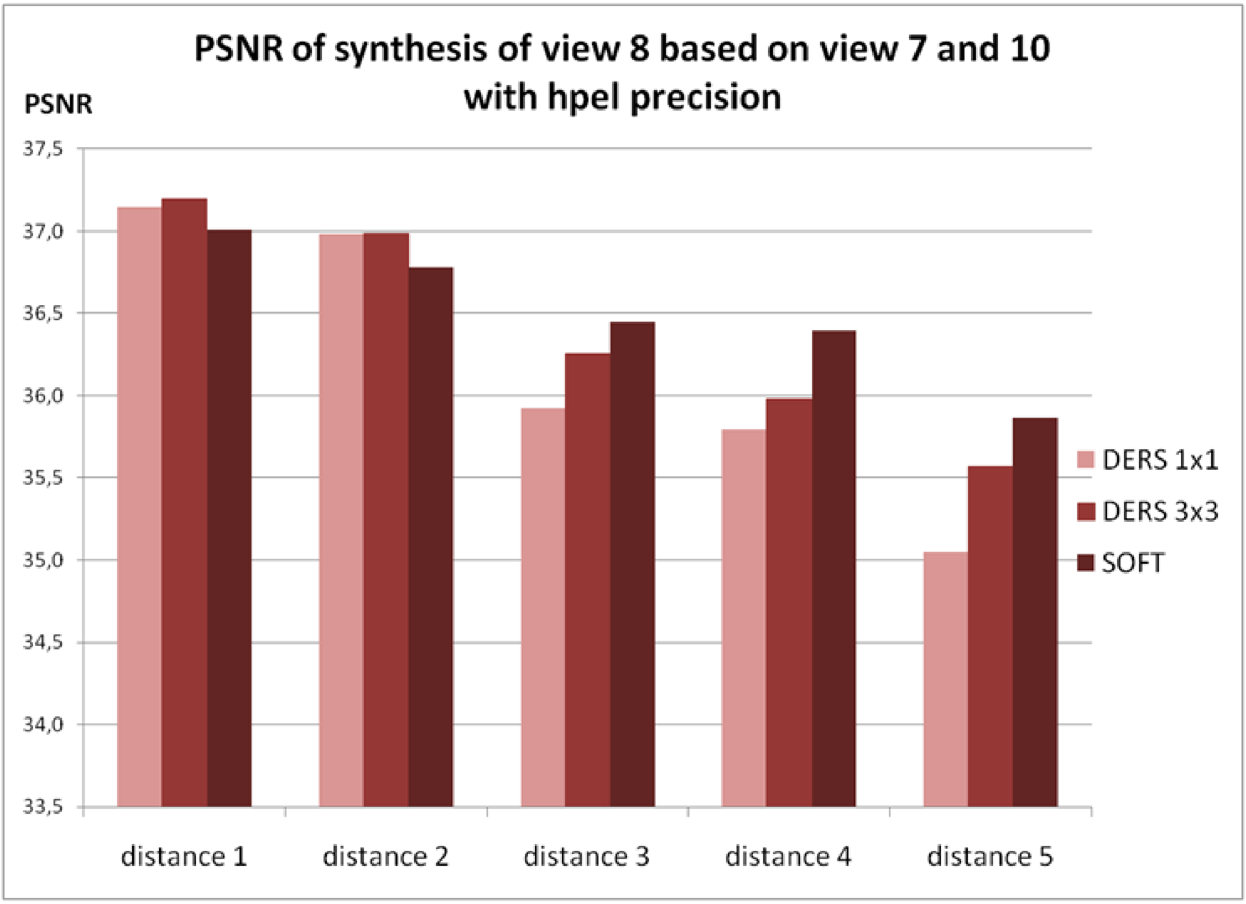


Figure 5. Comparison of synthesis results of 1x1 pixel matching, 3x3 block matching and new soft-segmentation matching. DERS was in manual mode/HPel precision. Only the best results, with respect to other parameters (smooting coefficient etc.), are presented.

Also in case of quarter-pixel precision mode (Figure 5) soft-segmentation matching performs slightly better than block matching modes for camera distances greater than 2. The gain is also comparable - about 0.6dB÷1.8dB from DERS 1x1 and about 0.3dB÷0.8dB from DERS 3x3. Just like in case of pixel precision, DERS using 3x3 block matching gives the best results for smaller camera distances, but here the difference is smaller (DERS 3x3 is better that SOFT for about 0.4dB).

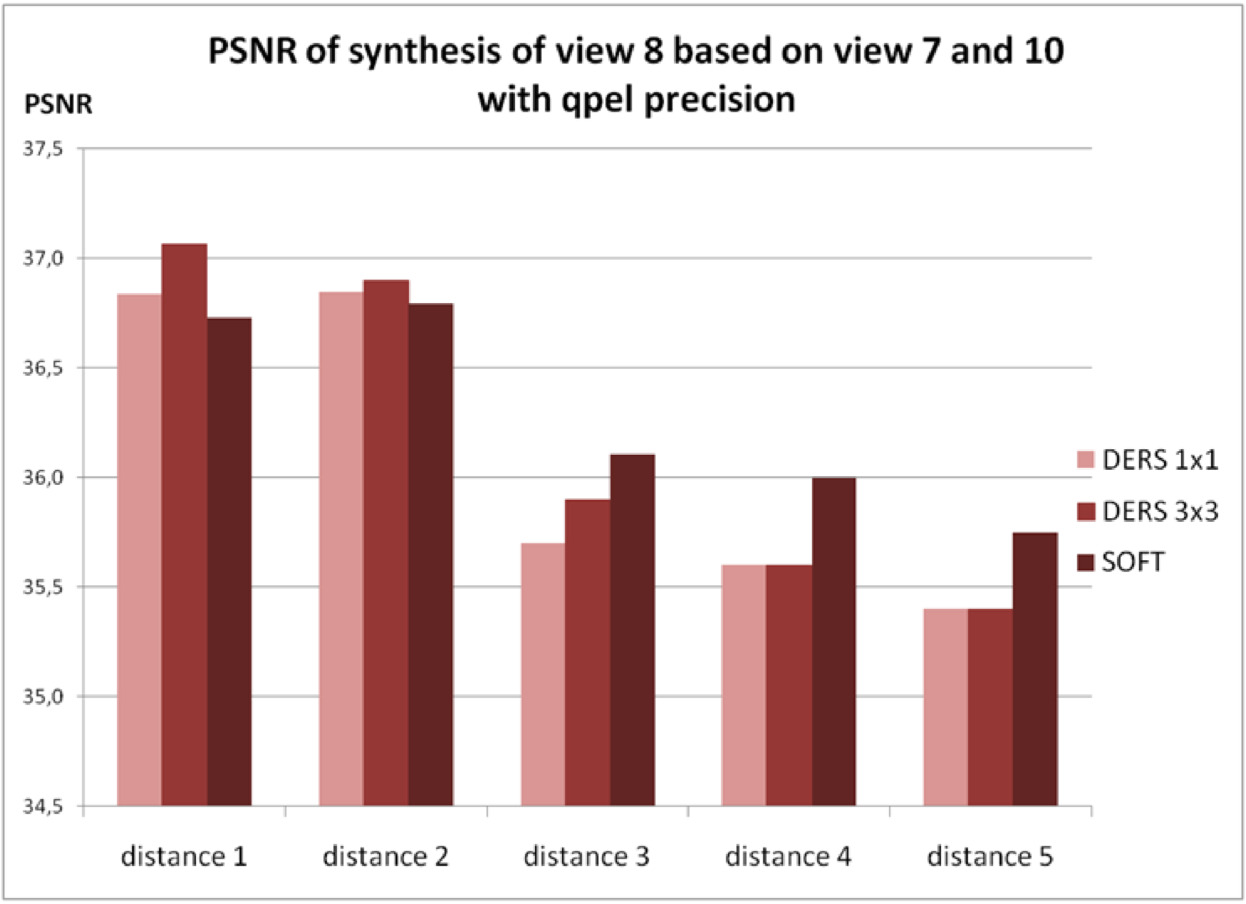


Figure 6. Comparison of synthesis results of 1x1 pixel matching, 3x3 block matching and new soft-segmentation matching. DERS was in manual mode/QPel precision. Only the best results, with respect to other parameters (smooting coefficient etc.), are presented.

Similar to pixel and quarter-pixel precision modes, results of quarter pixel mode (Figure 5) show that soft-segmentation matching can be a better option if camera distances are greater than 2. The gain, of about 0.3dB÷0.6dB (from both DERS 1x1 and DERS 3x3), is occupied by increased computational complexity. However, DERS working in 3x3 block matching mode still gives the best results for optimal, smaller camera distances, it is very small advantage of about 0.4dB.

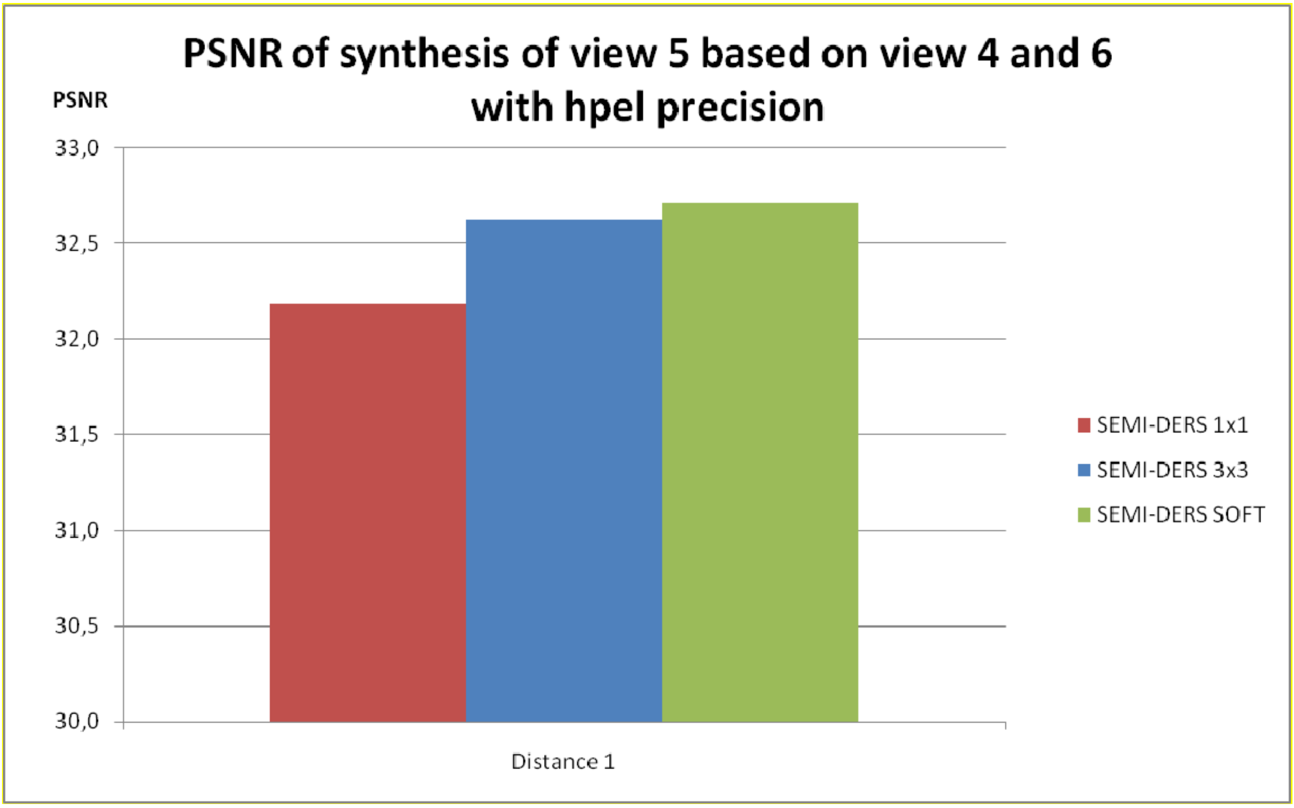


Figure 7. Comparison of synthesis results of 1x1 pixel matching, 3x3 block matching and new soft-segmentation matching in various pixel-precision modes. DERS was working in semi-automatic mode.

As shown in Figure 7, new soft-segmentation matching increases quality of produced depth maps and synthesized views by about 0.7dB in semi-automatic mode. This is important, because semi-automatic mode is considered to be the most promising technique for current works of MPEG-FTV group.

6 Conclusions

A new soft-segmentation matching has been added to DERS. It can be used both in manual and semi-automatic mode. The quality of produced depth maps is improved (for about 0.5dB) for camera distances greater than 2. For smaller camera distances it is still recommended to use DERS in 3x3 block matching mode.

7 References

[1] M. Tanimoto, T. Fujii, M. P. Tehrani, M. Wildeboer „Depth Estimation Reference Software (DERS) 4.0", ISO/IEC JTC1/SC29/WG11, MPEG 2008/M16605, London, UK, June 2009.

[2] O. Stankiewicz, K. Wegner „An enhancement of Depth Estimation Reference Software with use of soft-segmentation", ISO/IEC JTC1/SC29/WG11, MPEG 2008/M16757, London, UK, June 2009.