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

ISO/IEC JTC1/SC29/WG11
MPEG2010/M17176
January 2010, Kyoto, Japan

Source **Poznan University of Technology,**
Chair of Multimedia Telecommunications and Microelectronics

Title **First version of depth maps for Poznan 3D/FTV test sequences.**

Authors Olgierd Stankiewicz **(ostank@multimedia.edu.pl),**
Krzysztof Wegner **(kwegner@multimedia.edu.pl),**
Marek Domanski.

1 Introduction

This documents presents preliminary depth maps produced for Poznan University of Technology test sequences: *Poznan_Hall1*, *Poznan_Hall2*, *Poznan_Street* and *Poznan_Carpark.* These sequences [4] have been provided to MPEG (and the scientific community in general) for research and standardization.

As posted on the MPEG-FTV mail reflector, the rectified and color corrected versions of our sequences are available on our FTP site:

ftp://multimedia.edu.pl/3DV/

Please note that the resolution of rectified sequences is 1920x1088 instead of the original 1920x1080. The original *Poznan_Hall* sequence has been split into two parts (*Poznan_Hall1* and *Poznan_Hall2*) in order to avoid problems with the pilar depth impressions.

The depth estimation has been performed on the geometrically rectified and color corrected versions of the sequences [7]. The results of depth map estimation, presented in this document, can be found in *depth* directory on our FTP site.

The **depth maps** and **test materials** are provided to MPEG and the scientific community in general for **research** and **standardization**. Of course, some words of acknowledgement are appreciated if the materials are to be used in publications. Any commercial use is prohibited unless an explicit permission is given by Poznan University of Technology, Chair of Multimedia Telecommunications and Microelectronics.

2 Poznan 3DV/FTV test sequences EE configuration

The depth map estimation has been performed basing on guidelines in N10926 [5]. for the sake of future exploration experiments. For each sequence, three depth maps for three selected views have been estimated, so that the results can be used for MVD (Multiple View Depths) representation in both 2-view and 3-view EE scenario. In these scenarios, original views (OL-OC-OR) with corresponding provided depth maps can be taken as an input for view synthesis of output Synthesized Views and assessed for quality.
Some of the parameters from [5] have been slightly changed to better suit selected frame ranges (modifications are grayed in tables Table 1 and Table 2). In particular, because *Poznan_Hall* sequence has been split into two sequences *Poznan_Hall1* and *Poznan_Hall2*, Original Views OL-OC-OR (Table 1, Table 2) have been altered accordingly from views 3-4-5 to 1-2-3 (*Poznan_Hall1*) and 5-6-7 (*Poznan_Hall2*). Thanks to that, selected views (in working frame range) are less concerned with the front-plane pilar, which has been identified as problematic.

Table 1: Input and output views for MVD representation format in 2-view configuration

Data set	**Original Pair OL-OR**	**Synthesized Pair SL-OR (OL-SR)**	**Frame Range for EE1**
Poznan_Hall1	1-2	1.5-2	0~199 (0~199 from original *Poznan_Hall*)
Poznan_Hall2	6-7	6.5-7	0~199 (350~549 from original *Poznan_Hall*)
Poznan_Street	3-4	3.5-4	150 ~349
Poznan_Carpark	3-4	3.5-4	200~399

Table 2: Input and output views for MVD representation format in 3-view configuration

Data set	**Original Views OL-OC-OR**	**Views to Synthesize for stereo viewing[1]**	**Views to Synthesize for 9-view display[2]**
Poznan_Hall1	1-2-3	1.875, 2.125	1.5, 1.625, 1.75, 1.875, 2.125, 2.25, 2.375, 2.5
Poznan_Hall2	5-6-7	5.875, 6.125	5.5, 5.625, 5.75, 5.875, 6.125, 6.25, 6.375, 6.5
Poznan_Street	3-4-5	3.875, 4.125	3.5, 3.625, 3.75, 3.875, 4.125, 4.25, 4.375, 4.5
Poznan_Carpark	3-4-5	3.875, 4.125	3.5, 3.625, 3.75, 3.875, 4.125, 4.25, 4.375, 4.5

[1] The selected baseline distances correspond to emulation of a 22-view display on a stereo display.
[2] The views to synthesize for a multiview display are subject to change according to specific display requirements that would be available at a given meeting.

3 Depth Maps

The presented first version of depth maps has been produced with use of Depth Estimation Reference Software 5.0 (DERS 5.0). The depths have been generated for original views (OL-OC-OR) specified in Table 1 and Table 2. DERS configured to work with Temporal Enhancement tool and in Semi-automatic Depth Estimation mode. Auxiliary data for semi-automatic estimation: edge maps (MEM), disparity constraints maps (MDM) and static maps (MSM) have been manually created for selected key-frames and fed to the software (examples: Figure 1 and Figure 2). Camera distance 1 has been used for estimation (depth estimated from the closest neighboring views).
Exemplary resultant depth maps are shown in Figure 3, Figure 4, Figure 5 and Figure 6.

All frame numbers in figures below are #0-based relatively to the start of a given sequence.

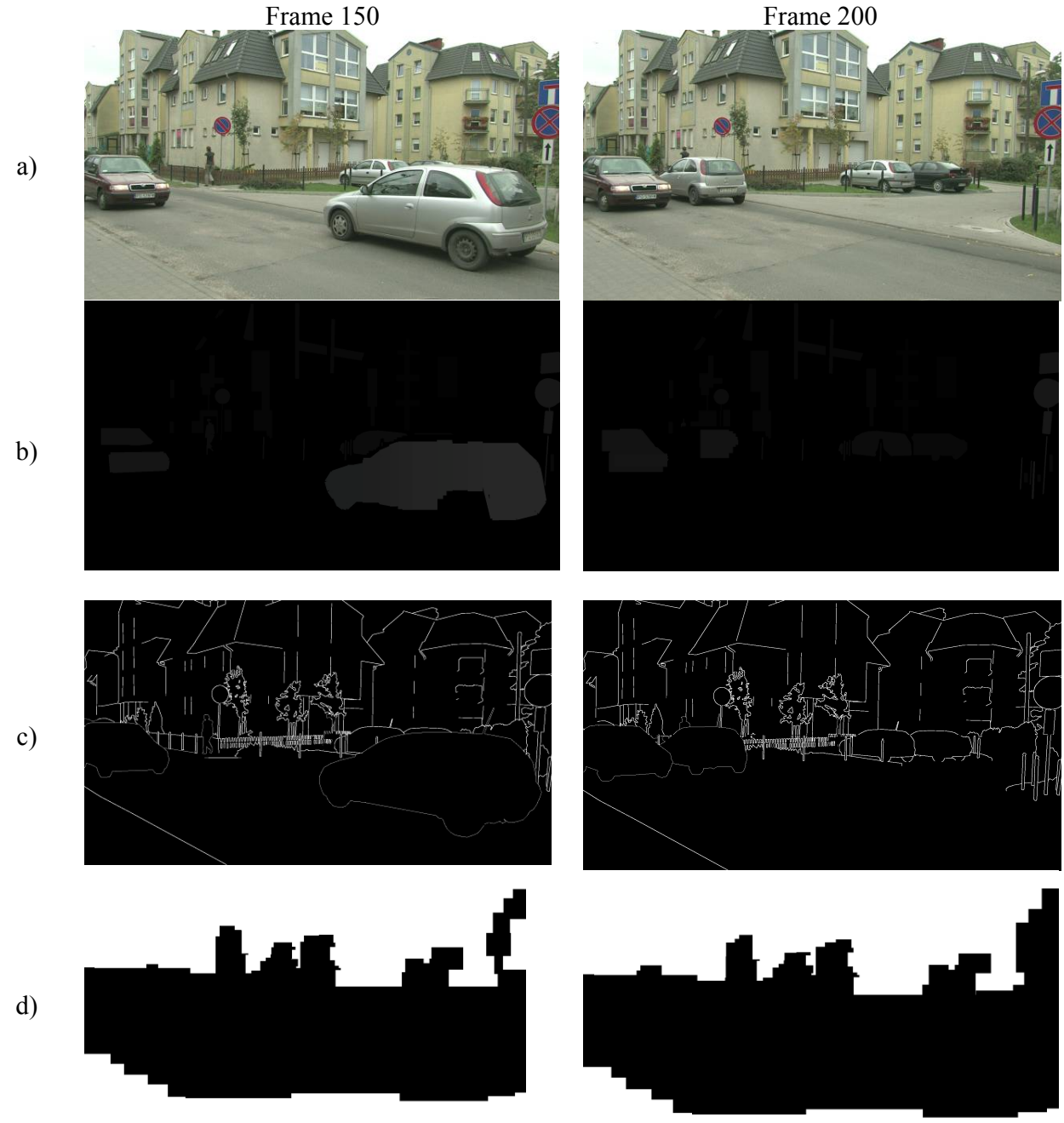


Figure 1. Examples of manually created auxiliary data for semi-automatic depth estimation for Poznan_Street Sequence (View 3, Frames: 150 and 200)
a) image b) disparity constrains map (MDM) c) edge map (MEM) d) static map (MSM)

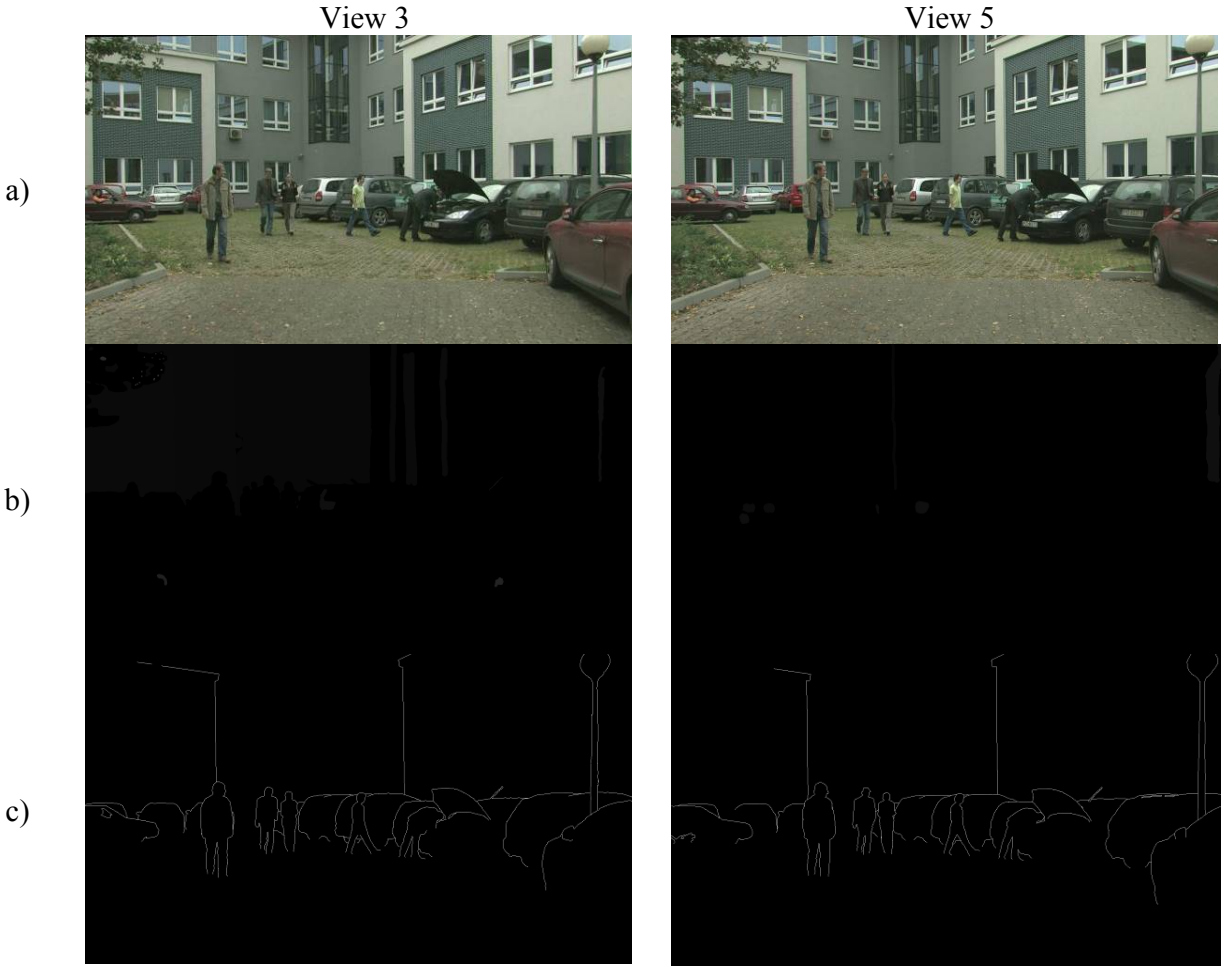
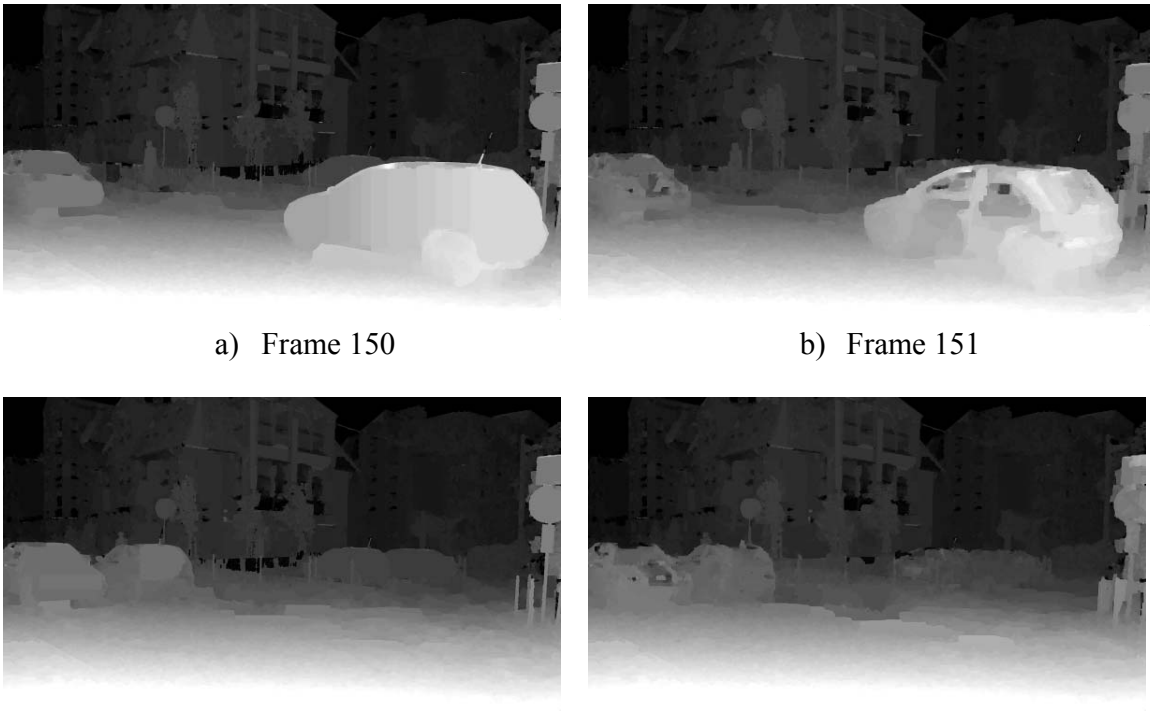


Figure 2. Examples of manually created auxiliary data for semi-automatic depth estimation for *Poznan_CarPark* sequence (Views 3 and 5, Frame 200)
a) image b) disparity constrains map (MDM) c) edge map (MEM)



a) Frame 150 b) Frame 151

c) Frame 200 d) Frame 201

Figure 3. Examples of estimated depth maps for *Poznan_Street* sequence for key-frames (150 and 200) and for one frame after (View 3).

As can be seen in Figure 3, subjective quality of depths produced *for Poznan_Street* sequence is quite satisfying. The manual data, provided for key-frames 150 and 200, gives good results for cars and background as well. Unfortunately, this information seems to be faded in consecutive frames (151 and 201). Reflections on the cars' bodies and on the windows cause depth estimation to fail and produce artifacts.

Also, there are some problems with the pavement's depth in the center of view 3 estimated in key-frames (black region). This problems doesn't appear in next frames which is odd, because no such depth has been specified in key-frames.

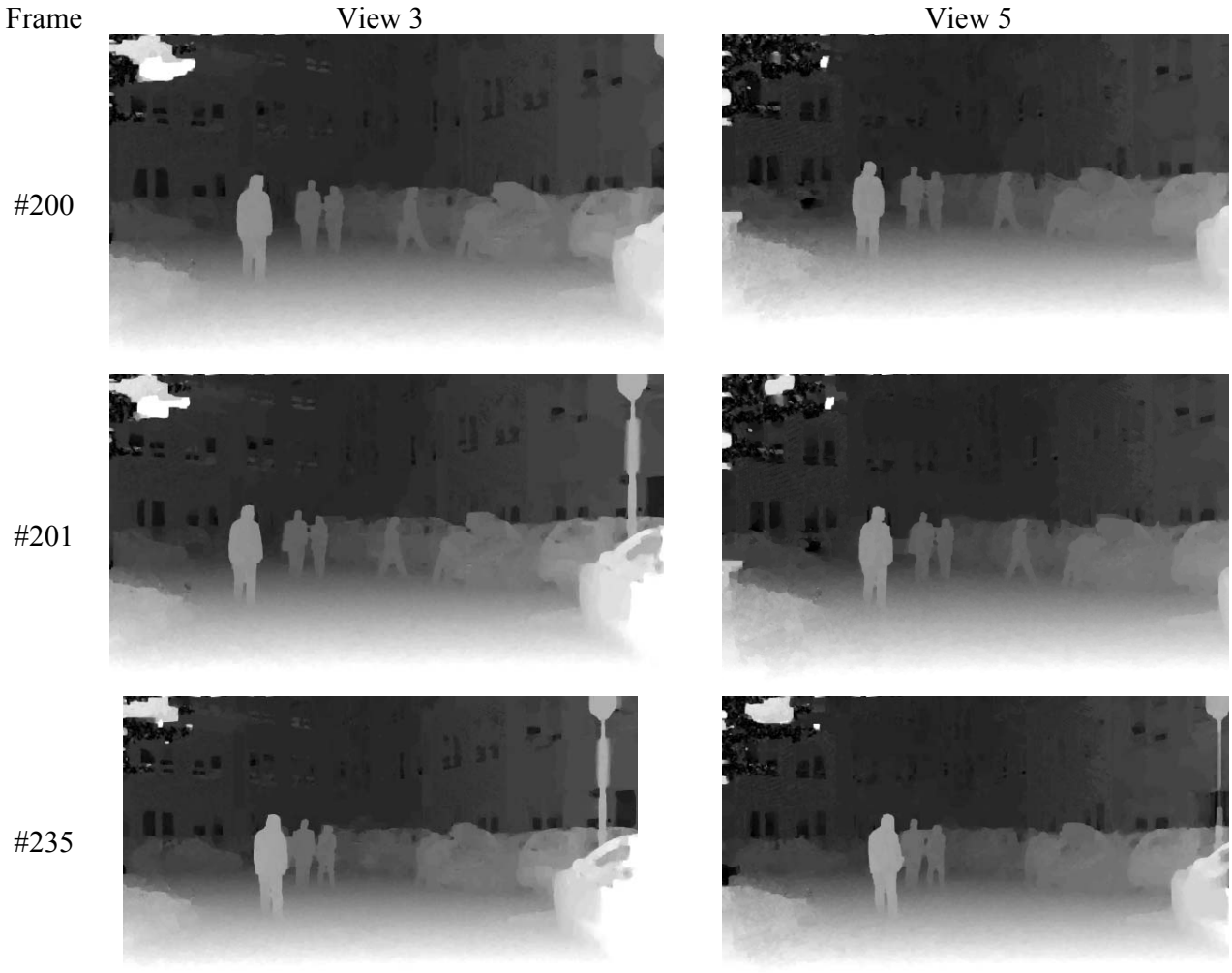
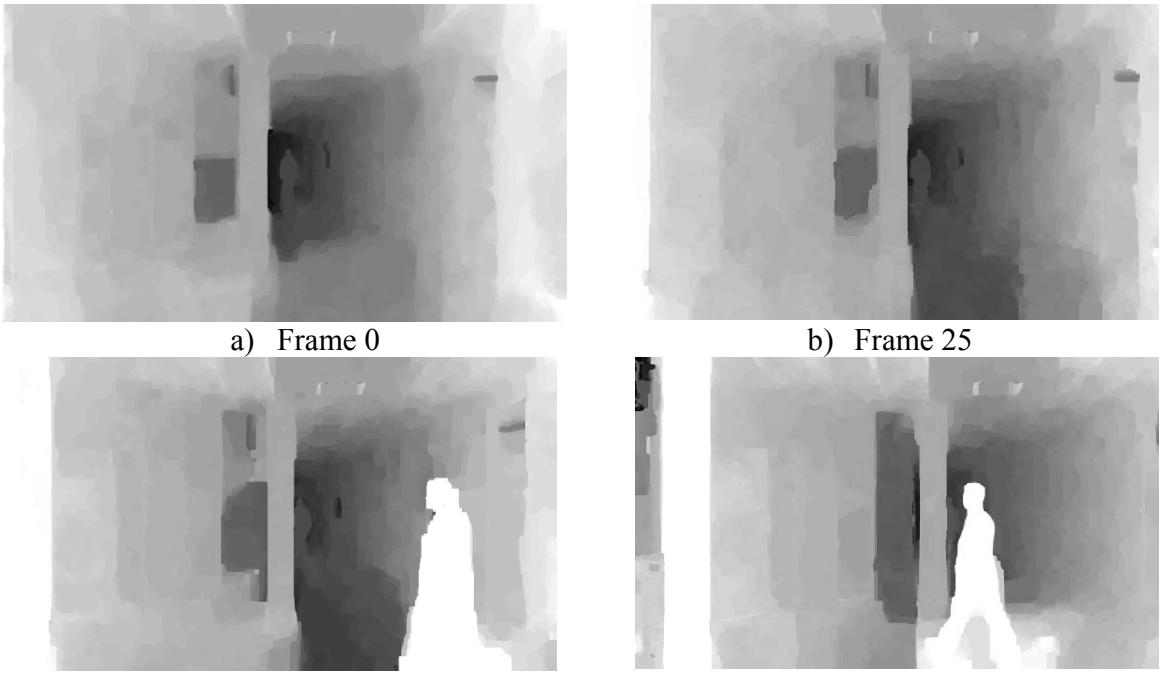


Figure 4. Examples of estimated depth maps for *Poznan_Carpark* sequence for key-frame (200), one frame after (201) and frame 235 (View 3 and View 5).

The depth for *Poznan_Carpark* sequence is of good quality, which can be seen in Figure 4. Mos of the scene is estimated properly with exception of tree leaves in the upper-left corner.
Also, there are artifacts over windows' surfaces, but these are less troubling than in case of *Poznan_Street* sequence.



a) Frame 0 b) Frame 25

c) Frame 50 d) Frame 75

Figure 5. Examples of estimated depth maps for *Poznan_Hall1* sequence for key-frames (150 and 200) and one frame after (View 3).

The problem of reflections seems not to appear in case of *Poznan_Hall* sequence. Both in case of *Poznan_Hall1* (Figure 5) and *Poznan_Hall2* (Figure 6) the depth seem to be acceptable.
The are some issues though:

- The floor in the *Poznan_Hall1* sequence is estimated to be (Figure 5, b,c) farer that it is in real,
- The moving people depth are very blurry (Figure 5cd ,6 the lowest row)
- The rails in the *Poznan_Hall2* sequence (Figure 6) are blurred and are modeled as a flat surface, instead of separate rods.

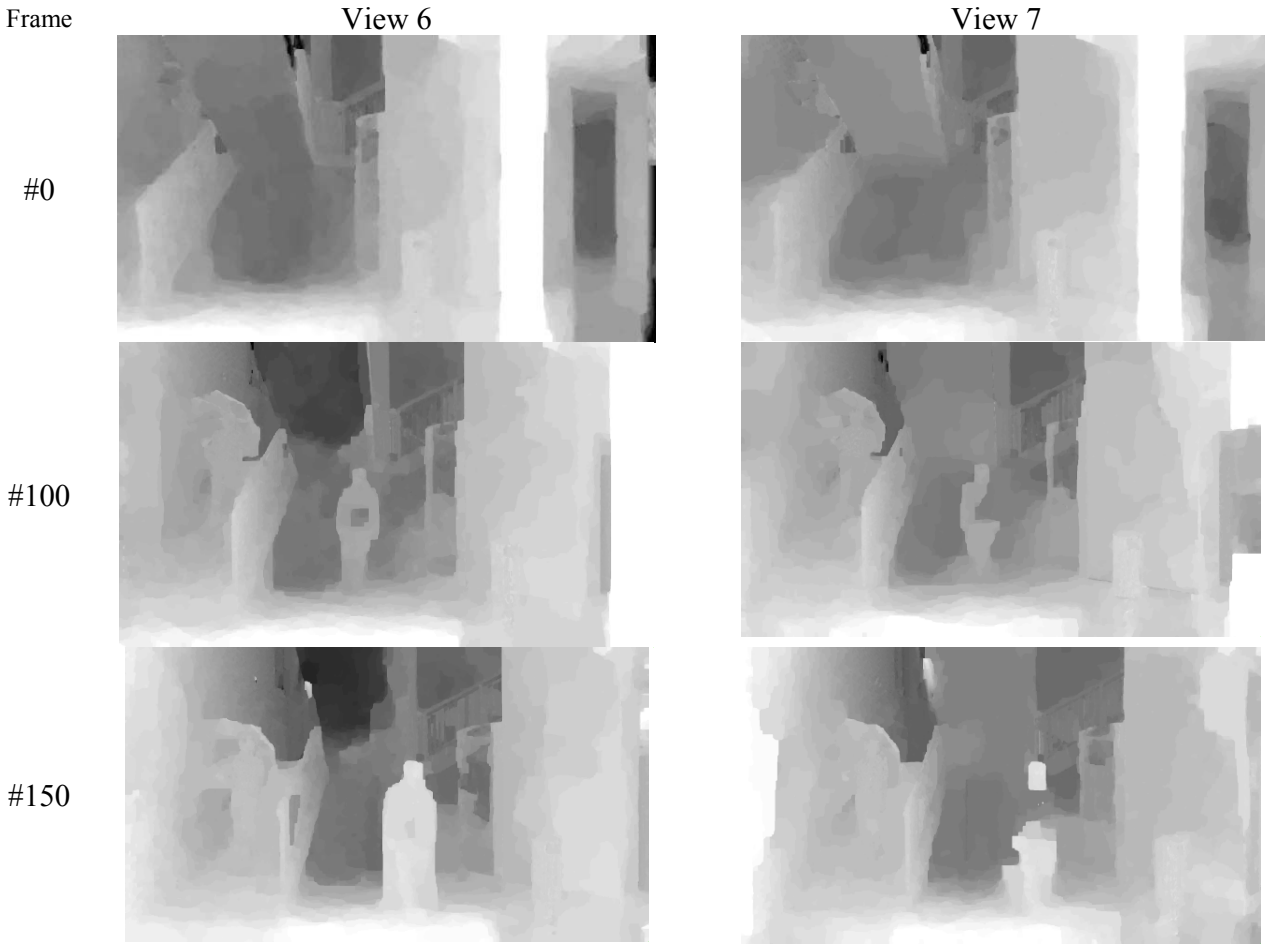


Figure 6. Examples of estimated depth maps for Poznan_Hall2 sequence for frames: 0, 100 and 150, views: 6 and 7.

4 View synthesis

Estimated depth maps have been used for synthesis of virtual view in place of original view for objective quality assessment (Table 3). The PSNR between synthesized virtual view (SC) and original view (OC).

Table 3: View confuration for synthesis-based objective quality assessment.

Data set	Original Views OL-OR	Synthesized view SC
Poznan_Hall1	1-3	2
Poznan_Hall2	5-7	6
Poznan_Street	3-5	4
Poznan_Carpark	3-5	4

As can be seen in Figure 7, PSNR levels of *Poznan_Street* and *Poznan_Hall2* sequences are above 35 dB, which is satisfying. *Poznan_Hall1* quality is lower – it is about 32,3dB. The lowest quality is achieved by *Poznan_Carpark* sequence. We believe that it results from huge amount of reflective windows in that scene.
Subjectively, most of the synthesis results are quite promising, especially in case of key-frames, but still there are many artifacts. This means that quality of depth maps should be improved.

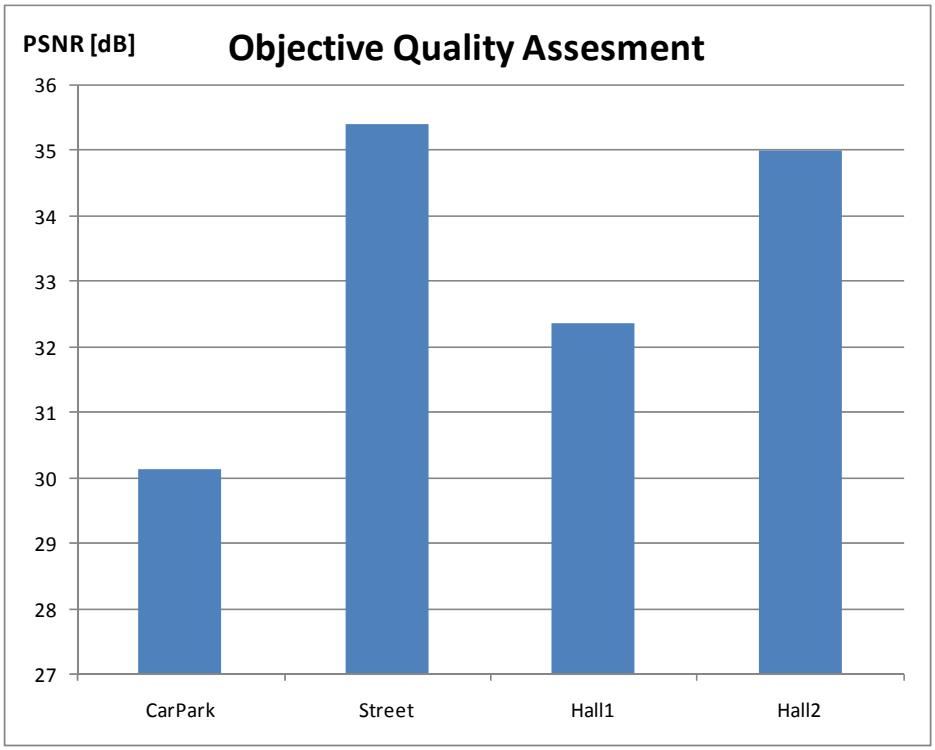


Figure 7. Synthesis-based objective quality assessment results.

5 Conclusions

- Preliminary versions of depth maps for Poznan sequences are available on FTP site.
- Quality of the depth maps is good, but there is still some work to be done.
- The Depth Estimation Reference Software should be improved to allow better usage of key-frames in sequence time-line, because currently key-frame influence only single frame and it is awkward to produce key-frames for each frame of sequences.

6 References

[1] "Call for Contributions on 3D Video Test Material", ISO/IEC JTC1/SC29/WG11 N9468, October 2007.

[2] "Call for Contributions on 3D Video Test Material", ISO/IEC JTC1/SC29/WG11 N9595, January 2008.

[3] "Call for 3D Test Material: Depth Maps & Supplementary Information", ISO/IEC JTC1/SC29/WG11 N10359, February 2009.

[4] M. Domanski, T. Grajek, K. Klimaszewski, M. Kurc, O. Stankiewicz, J. Stankowski, K. Wegner, "Poznan Multiview Video Test Sequences and Camera Parameters", ISO/IEC JTC1/SC29/WG11 m17050, Xian, China, October 2009.

[5] "Description of Exploration Experiments in 3D Video Coding" ISO/IEC JTC1/SC29/WG11 N10926, Xian, China, October 2009.

[6] M. Domański, T. Grajek, K. Klimaszewski, M. Kurc, O. Stankiewicz, J. Stankowski, K. Wegner, "Poznan Multiview Video Test Sequences and Camera Parameters", ISO/IEC JTC1/SC29/WG11 m17050, Xian, China, October 2009.

[7] J. Stankowski, K. Klimaszewski, O. Stankiewicz, K. Wegner, M, Domanski, "Preprocessing methods used for Poznan 3D/FTV test sequences", ISO/IEC JTC1/SC29/WG11 m17174, Kyoto, Japan, January 2010.