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/M1XXXX
October 2009, Xian, China

Title	**Newspaper sequence - Results of 3DV/FTV Exploration Experiments with depths and view synthesis**
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 Experiment (EE4) performed on "Newspaper" sequence [2] and is in response to W10720 "Description of Exploration Experiments in 3D Video Coding" [1].

2 Experiments conditions

Experiments were performed basing on W10720 [1] guidelines:

2.1 EE1

- Select stereo pair from data set, i.e. an original left view OL and an original right view OR (OL=4, OR=6)
- Estimate depth corresponding to neighboring original views OL (left) and OR (right), from neighboring cameras with use of semi-automatic mode
- Synthesize views (synthesized left SL and synthesized right SR) at positions from OL+D and OR+D
- Compare OL-OR with SL-SR subjectively

2.2 EE2

Two view case:

- Original reference texture data for views 4 and 6 of "Newspaper" sequence were compressed using JMVM software version 5.0.6 with different QP values. GOP length was

set to 16 frames, to comply with the requirement of at least 0.5 second GOP length (Newspaper is a 30 fps sequence)

- Depth maps for views 4 and 6, provided for the purpose of this experiment, were compressed using JMVM software version 5.0.6 with different QP values. GOP length was set to 16 frames.
- Appropriate depth and texture data were selected to meet the 0.75, 1.5, 3, 6 Mbps stream requirements
- Reconstructed texture and depth data were fed to the view synthesis software VSRS version 3.0.1, together with camera system parameters and Znear, Zfar values to recreate view 5.
- Synthesized view 5 was compared in terms of PSNR and PSPNR with original view 5 as well as with view 5 synthesized using uncompressed data. For PSPNR calculation, default settings were used – borders of 30 pixels width were excluded from both sides of synthesized frames for purpose of quality calculation.

Three view case:

- Original reference texture data for views 2, 4 and 6 of Newspaper sequence were compressed using JMVM software version 5.0.6 with different QP values. GOP length was set to 16 frames, to comply with the requirement of at least 0.5 second GOP length (Newspaper is a 30 fps sequence)
- Depth maps for views 2, 4 and 6, provided for the purpose of this experiment, were compressed using JMVM software version 5.0.6 with different QP values as well. GOP length was set to 16 frames.
- Appropriate depth and texture data were selected to match with 1, 2, 4.8, 7.8 Mbps stream requirements
- Reconstructed texture and depth data were fed to the view synthesis software VSRS version 3.0.1, together with camera system parameters and Znear, Zfar values to create the following views: 2.5, 3, 3.5, 4.5, 5, 5.5.
- Synthesized views were compared in terms of PSNR and PSPNR with appropriate views synthesized using uncompressed data. Where applicable, synthesis results were compared with original views (views 3 and 5). For PSPNR calculation, default settings were used – borders of 30 pixels width were excluded from both sides of synthesized frames for purpose of quality calculation.

The tests were performed on ‘Newspaper’ [2] sequence with following views selected as O and NL-NR.

Table 1. The specification of views for EE experiment.

Data set	Two cameras case		Three cameras case	
	O	NL-NR	O	NL-NR
Newspaper	5	4 - 6	2.5, 3, 3.5,4.5, 5, 5.5	2 - 4 - 6

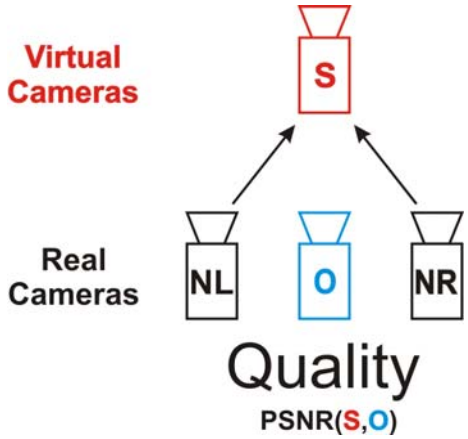


Figure 1. Two cameras case – test setup.

3 Semi-automatic data

In order to estimate the depth in semi-automatic DERS mode, an additional data has been created and provided.

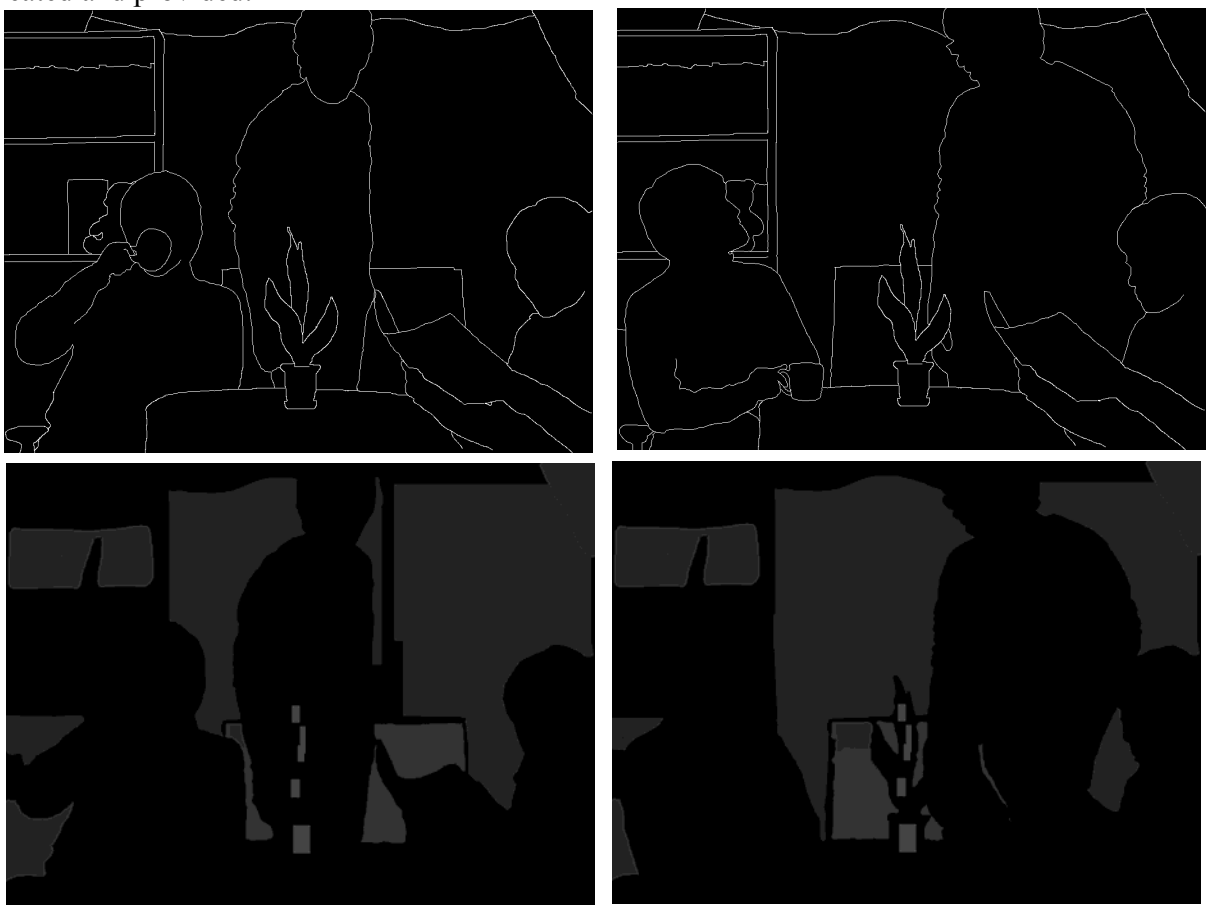


Figure 2. Example of new key frame (100 and 200) used to estimated depth with semi-automatic mode in DERS 5.0

4 Results

3.1 EE 1:

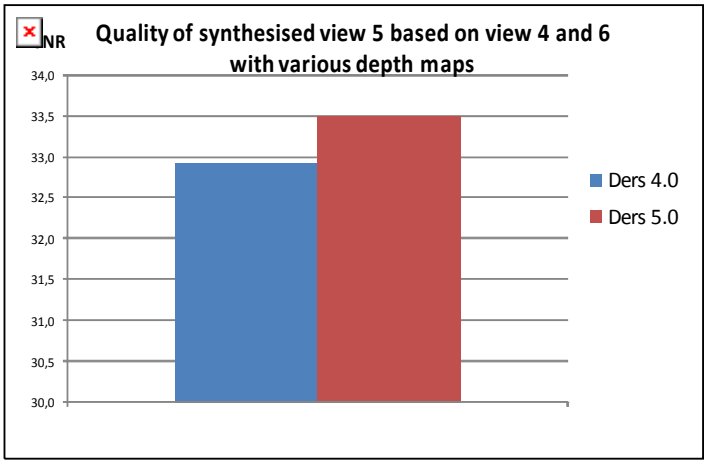


Figure 3. Comparison of quality of synthesized view 5 based on depth maps provided with sequence and new depth maps obtain by DERS 5.0

Figure 3 shows, that DERS in version 5.0 (which was also working with new semi-automatic data) outperforms previous version (4.0). The gain is of about 0.5dB.

3.2 EE 4 - Two view case:

To meet the requirements imposed on bitstream size, the following pairs of QP and QD (QP index for depth encoding) were selected:

Table 2. Selected QP-QD pairs for synthesis

Bitrate	**QP-QD pairs selected**
750 kbps	43-31, 37-36, 35-41, 38-34, 40-32, 39-33, 36-37, 35-40, 42-31, 44-30, 37-35
1.5 Mbps	30-32, 35-26, 41-24, 28-44, 29-35, 31-30, 28-43, 28-42, 37-25, 40-24, 28-41
3 Mbps	29-21, 25-26, 23-34, 33-19, 24-28, 23-33, 30-20
6 Mbps	28-13, 20-20, 19-23

With selected pairs, view 5 was synthesized and its quality was measured:

Table 3. Quality of synthesized views, 750 kbps case

Bitrate [kbps]	QP	QD	against uncompressed synthesis				against original view			
			PSNR [dB]	PSNR (pspnr) [dB]	PSPNR temporal [dB]	PSPNR spatial [dB]	PSNR [dB]	PSNR (pspnr) [dB]	PSPNR temporal [dB]	PSPNR spatial [dB]
730	43	31	31.33	31.49	45.31	33.82	29.62	29.74	44.8	31.38
733	37	36	34.29	34.49	48.43	38.64	31.39	31.56	48.09	34.02
737	35	41	34.59	34.87	49.23	39.4	31.65	31.87	49.12	34.54
747	38	34	33.83	34.01	47.91	37.85	31.17	31.31	47.47	33.65
752	40	32	32.94	33.1	46.92	36.35	30.59	30.72	46.43	32.8
753	39	33	33.46	33.63	47.49	37.19	30.93	31.07	46.99	33.3
753	36	37	34.61	34.84	48.92	39.23	31.6	31.79	48.6	34.36
754	35	40	**34.69**	**34.97**	**49.27**	**39.5**	**31.69**	**31.9**	**49.12**	**34.57**
755	42	31	31.86	32.01	45.89	34.6	29.93	30.04	45.36	31.81
762	44	30	**30.67**	**30.83**	**44.71**	**32.83**	**29.14**	**29.26**	**44.2**	**30.72**
764	37	35	34.34	34.54	48.45	38.7	31.41	31.58	48.08	34.03
	- the best result for given bitrate						- the worst result for given bitrate			

Table 4. Quality of synthesized views, 1500 kbps case

Bitrate [kbps]	QP	QD	against uncompressed synthesis				against original view			
			PSNR [dB]	PSNR (pspnr) [dB]	PSPNR temporal [dB]	PSPNR spatial [dB]	PSNR [dB]	PSNR (pspnr) [dB]	PSPNR temporal [dB]	PSPNR spatial [dB]
1 482	30	32	37.18	37.49	51.93	43.47	32.59	32.81	51.87	35.92
1 483	35	26	35.58	35.79	49.73	40.64	31.85	32.01	49.11	34.69
1 483	41	24	**32.52**	**32.67**	**46.5**	**35.6**	**30.26**	**30.38**	**45.85**	**32.33**
1 485	28	44	35.74	36.19	52.01	41.62	32.24	32.55	52.81	35.75
1 490	29	35	**37.21**	**37.57**	**52.26**	**43.59**	**32.66**	**32.88**	52.47	**36.08**
1 490	31	30	37.03	37.32	51.57	43.22	32.52	32.71	51.33	35.75
1 498	28	43	36.03	36.45	52.10	41.91	32.35	32.64	52.82	35.85
1 510	28	42	36.25	36.69	52.18	42.22	32.44	32.72	**52.87**	35.97
1 514	37	25	34.73	34.91	48.71	39.19	31.43	31.58	48.07	34.05
1 516	40	24	33.13	33.28	47.08	36.55	30.58	30.70	46.43	32.79
1 523	28	41	36.51	36.96	52.2	42.59	32.51	32.79	52.86	36.04
	- the best result for given bitrate						- the worst result for given bitrate			

Table 5. Quality of synthesized views, 3000 kbps case

Bitrate [kbps]	QP	QD	against uncompressed synthesis				against original view			
			PSNR [dB]	PSNR (pspnr) [dB]	PSPNR temporal [dB]	PSPNR spatial [dB]	PSNR [dB]	PSNR (pspnr) [dB]	PSPNR temporal [dB]	PSPNR spatial [dB]
2 950	29	21	38.40	38.7	53.14	45.62	32.66	32.86	52.41	36.05
2 959	25	26	**39.32**	**39.75**	54.59	**47.05**	32.97	33.20	54.62	36.8
2 975	23	34	38.64	39.11	54.45	45.40	33.00	33.25	**55.51**	36.94
3 014	33	19	**36.84**	**37.06**	**51.14**	**42.73**	**32.17**	**32.35**	**50.21**	**35.23**
3 016	24	28	39.30	39.75	**54.8**	46.81	**33.02**	**33.26**	55.08	36.90
3 018	23	33	38.80	39.24	54.57	45.62	33.00	33.25	55.51	**36.95**
3 040	30	20	38.06	38.34	52.64	45.00	32.55	32.75	51.82	35.84
	- the best result for given bitrate						- the worst result for given bitrate			

Table 6. Quality of synthesized views, 6000 kbps case

Bitrate [kbps]	QP	QD	against uncompressed synthesis				against original view			
			PSNR [dB]	PSNR (pspnr) [dB]	PSPNR temporal [dB]	PSPNR spatial [dB]	PSNR [dB]	PSNR (pspnr) [dB]	PSPNR temporal [dB]	PSPNR spatial [dB]
5 957	28	13	**39.04**	**39.47**	**54.24**	**47.06**	**32.46**	**32.75**	**52.85**	**35.96**
5 973	20	20	**41.1**	**41.65**	**56.84**	**49.47**	32.98	33.22	56.57	37.03
6 054	19	23	40.93	41.48	56.67	48.86	**33.01**	**33.26**	**56.93**	**37.13**

- the best result for given bitrate

- the worst result for given bitrate

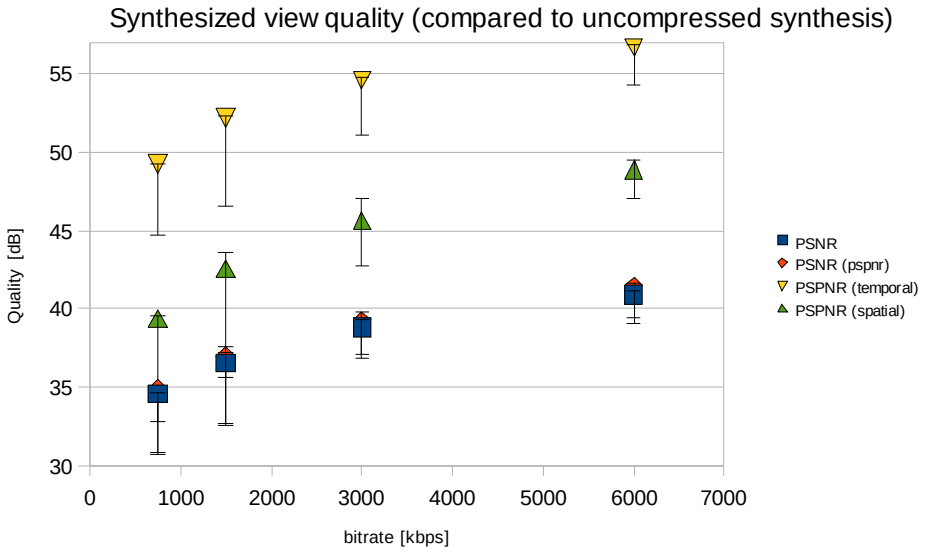


Figure 4. Quality of synthesized view as defined in [1] for 2 camera case. View synthesized with uncompressed data is used as reference. Points mark results for the pairs with minimal QP (and minimal QD, if more than one pair has the same QP). The bars show quality range from the best to the worst pair for each bitrate.

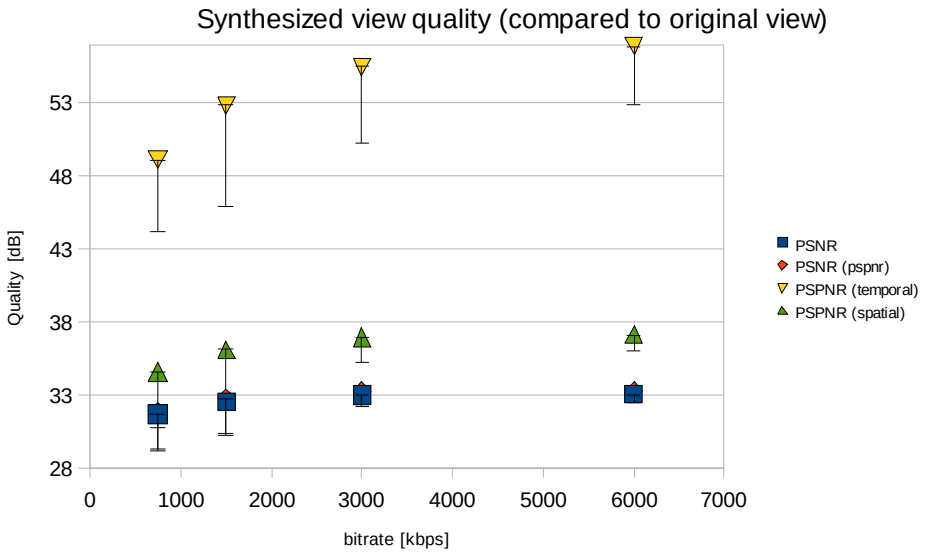
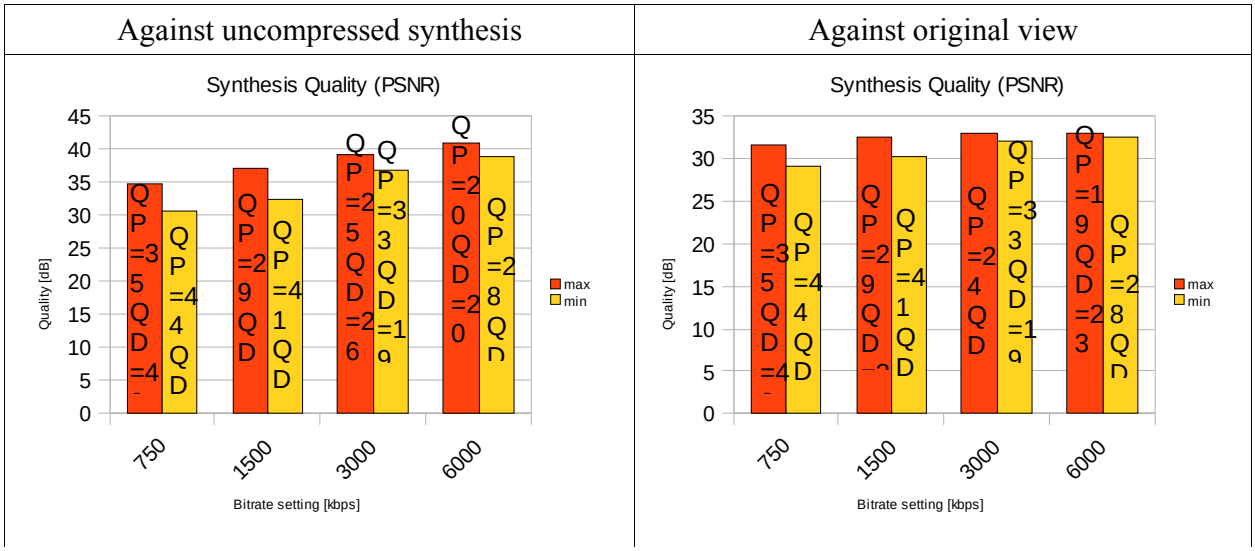


Figure 5. Quality of synthesized view as defined in EE4 for 2 camera case. Real view from camera 5 is used as a reference. Points mark results for the pairs with minimal QP (and minimal QD, if more than one pair has the same QP). The bars show quality range from the best to the worst pair for each bitrate.



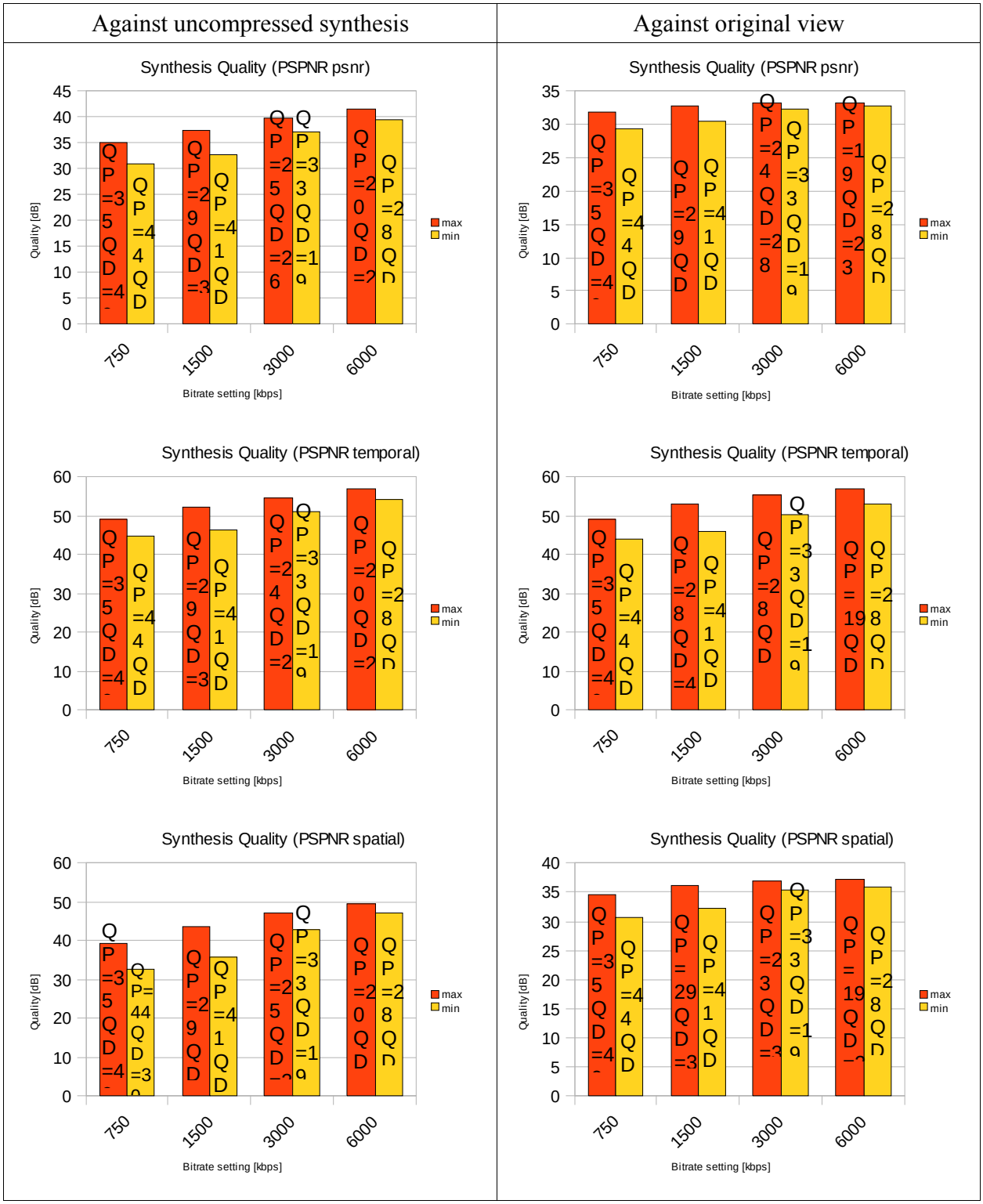


Figure 6. Best and worst results of synthesis for different measures and cases.

3.3 EE 4 - Three view case:

The three view case wasn't completed because depth map for view 2 wasn't delivered on time.

5 Conclusions

The conclusions are as follows:

4.1. EE 1

- Results attained with use of Semi-automatic DERS are about 2dB better than previous version of DERS.
- Subjective quality of synthesized views is also slightly better.

4.2. EE 4 - Two view case

- Quality of synthesized view depends more on quality of compressed/decompressed image (QP parameter) than on quality of compressed/decompressed depth (QD).
- The usual approach of choosing the minimal QP does not give the best results in some cases, but the differences of quality measures are negligible.
- When quality of synthesized view is calculated with original view as a reference, no quality improvement is observed when increasing bitrate from 3 Mbps to 6Mbps in terms of PSNR, PSNR (pspnr) and PSPNR (spatial).
- When quality of synthesized view is calculated with original view as a reference, the **PSPNR (temporal)** measure **does** increase when changing from 3 to 6 Mbps.
- Quality of synthesized views (PSNR as well as PSPNR) increases, when measured with reference synthesized from uncompressed data in the similar way as for 2D sequences – approximately constant increase of quality for increase of bitrate by a factor of two.

6 References

[1] "Description of Exploration Experiments in 3D Video Coding" MPEG 2009/W10720, London, England July 2009.

[2] Yo-Sung Ho, Eun-Kyung Lee, and Cheon Lee , „Multiview Video Test Sequence and Camera Parameters", MPEG 2008/M15419, Archamps, France, April 2008.