Joint Collaborative Team on 3D Video Coding Extensions of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11
5th Meeting: Vienna, Austria, July 27 – August 02, 2013

Document: JCT3V-E0269

Title:	**AHG14: Optimized QP/QD curve for 3D coding with half and full resolution depth maps**
Status:	Input Document
Purpose:	Information

Author(s) or Contact(s):	Olgierd Stankiewicz, Krzysztof Wegner, Marek Domański Polanka 3, 60-965 Poznań, Poland	Email:	ostank@multimedia.edu.pl kwegner@multimedia.edu.pl
Source:	Poznan University of Technology		

Abstract

In this document, an enhanced evaluation of impact of coded depth resolution on performance of the current AVC-based 3D video coding technology implemented in 3D-ATM [1] is presented.

Just like in case of previous works presented in JCT3V-D0080 [3], the evaluation has been done with respect to Common Test Conditions [2] in HP and EHP profiles. Two depth configurations has been tested - full depth resolution and half depth resolution. In this work, additionally, comparison have been done for operation point with maximized codec performance for each case. Maximization of coding performance was done by selection of optimal quantization parameter for both texture and depth views. Also, optimized quantization parameters curves were provided.

The results show that coding with half resolution of depth (currently used in CTC) provides better results in comparison to coding with full resolution of depth..

1 Introduction

In this document, we present an evaluation of depth resolution impact on performance of the current AVC-based 3D video coding technology implemented in 3D-ATM [1] with respect to Common Test Conditions [2] in HP and EHP profiles. We have tested two configurations regarding resolution of the associated depth maps:

- full resolution – where coded depth maps have the same resolution as the coded texture views,
- half resolution – where coded depth maps have been decimated by the factor of 2, resulting in two times less resolution compared to coded texture views (as in CTC).

During the Incheon JCT-3V meeting it has been noted, that comparisons of full resolution and half resolution depth maps coding presented in JCT3V-D0080 [3] require common reference in order to obtain meaningful results. At that time, the comparison has been performed under CTC conditions. Especially, quantization parameter for texture views (QP) and depth views (QD) were equal (QP=QD). Such condition is unfair when comparing coding with half and full-resolution of depth maps because it is not an optimal operation point for the codec. Therefore in this work, we have managed to find optimized quantization parameters for both texture views and depth view that maximize coding performance (QP-QD curves) for both of those cases separately.

2 Quantization parameter (QP-QD curve) optimization

One approach for finding quantization parameter that maximize coding performance is to test coding with all combinations of quantization parameters for texture and depth views (Fig. 1). Such an approach was successfully exploited in the past [4, 5]. Nevertheless such an approach is very time consuming and inefficient because only negligible number of quantization parameter pairs (from all possible) are the optimal pairs that maximize coding performance.

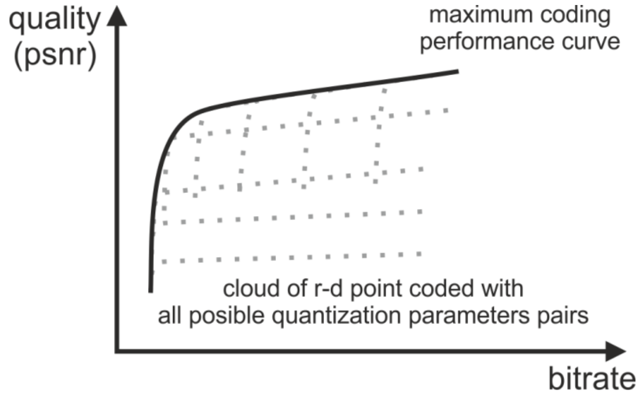


Figure 1. Results of coding performance with all combinations of quantization parameters for texture and depth views. Only negligible number of points are optimal pairs that maximize coding performance.

Therefore, we have proposed and tested a novel approach. Instead of testing all possible combinations of quantization parameters, we have used an iterative steepest-descent approach.

The algorithm starts with largest possible value of the both quantization parameters (QP_0=50 and QD_0=50) as it relates to the lowest quality and the smallest bitrate (bottom-left corner at the RD-curve). Then, at each next iteration *i+1*, two possibilities of improving quality of the encoded 3D video are tested (at each iteration, required bitrate also increase):

a) increased quality of depth views (decreased quantization parameter for depth views) and unchanged quality of texture views:
($QP^a_{i+1}=QP_i$ and $QD^a_{i+1}=QD_i-1$),

b) increased quality of texture views (decreased quantization parameter for texture views) and unchanged quality of depth views:
($QP^b_{i+1}=QP_i-1$ and $QD^b_{i+1}=QD_i$),

Coding results of those two possibilities are evaluated and compared with respect to R-D performance (Fig. 2). We have chosen to evaluate the total bitrate ($bitrate_i$) with respect to image quality defined as average luminance PSNR ($psnr_i$) of six virtual views as defined in CTC.

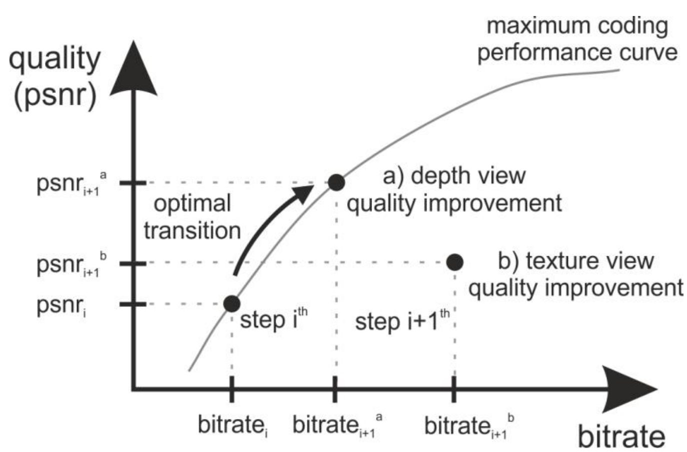


Figure 2. Steepest-descent optimization of quantization parameters for texture and depth view.

Basing on this evaluation, a single option ("a" or "b") that maximizes quality and minimizes bitrate is chosen.

$$\max_{a,b}\left[ArcTan\left(\frac{psnr_i^x - psnr_{i-1}}{bitrate_i^x - bitrate_{i-1}}\right)\right]$$

The option which is better is used for the next iteration. Process stops when either of two quantization parameters reach value 10.

In such an approach, maximally 2 x 41 (two options, QP,QD $\in$[10..50]) coder passes are sufficient to find quantization parameters pairs that maximize coding performance, instead of 41^2 coding passes (all possible QP/QD pairs).

3 Evaluation methodology

For each sequence, 3 videos along with 3 correspondent depth maps has been encoded with 3D-ATM in version 8.0r3 [1] according to the Common Test Condition [2] (only exception was quantization parameters used for texture and depth views). Based on the decoded videos and depth maps 6 views in the positions between the input views according to table 1 have been synthesized (CTC evaluation methodology). Synthesized views has been then compared via luminance PSNR with views synthesized at the same spatial positions with use of the original (uncompressed) data.

Table 1. Input view positions and synthesized views positions for 3 view case.

	Test Sequence	**Input views positions**	**Synthesized views positions**
S01	Poznan_Hall2	7-6-5	6.75 6.50 6.25 5.75 5.50 5.25
S02	Poznan_Street	5-4-3	4.75 4.50 4.25 3.75 3.50 3.25
S03	Undo_Dancer	1-5-9	2.00 3.00 4.00 6.00 7.00 8.00
S04	GT_Fly	9-5-1	8.00 7.00 6.00 4.00 3.00 2.00
S05	Kendo	1-3-5	1.50 2.00 2.50 3.50 4.00 4.50
S06	Balloons	1-3-5	1.50 2.00 2.50 3.50 4.00 4.50
S08	Newspaper1	2-4-6	2.50 3.00 3.50 4.50 5.00 5.50

4 Simulation

The simulations results were generated on a ~160 core cluster system. The cluster platform's processing units have the following specifications:

- Processor: Intel Xeon X5675
- Clock Speed: 3.06 GHz
- Memory: approx. 4 GB per Core
- OS: 64-bit Windows Server 2008
- Compiler: Microsoft Visual Studio 2008 (64 bit)

5 Simulation results

Figure 3 presents optimized quantization parameters pairs for texture (QP) and depth (QD) views. Dashed line present current common test condition – equal quantization parameters for both texture and depth views (QP=QD). On the basis of obtained results we have done linear regression with respect to least-square line fitting. This has yielded quantization curves presented in Table 2. Averaged quantization parameter values are presented in Table 3.

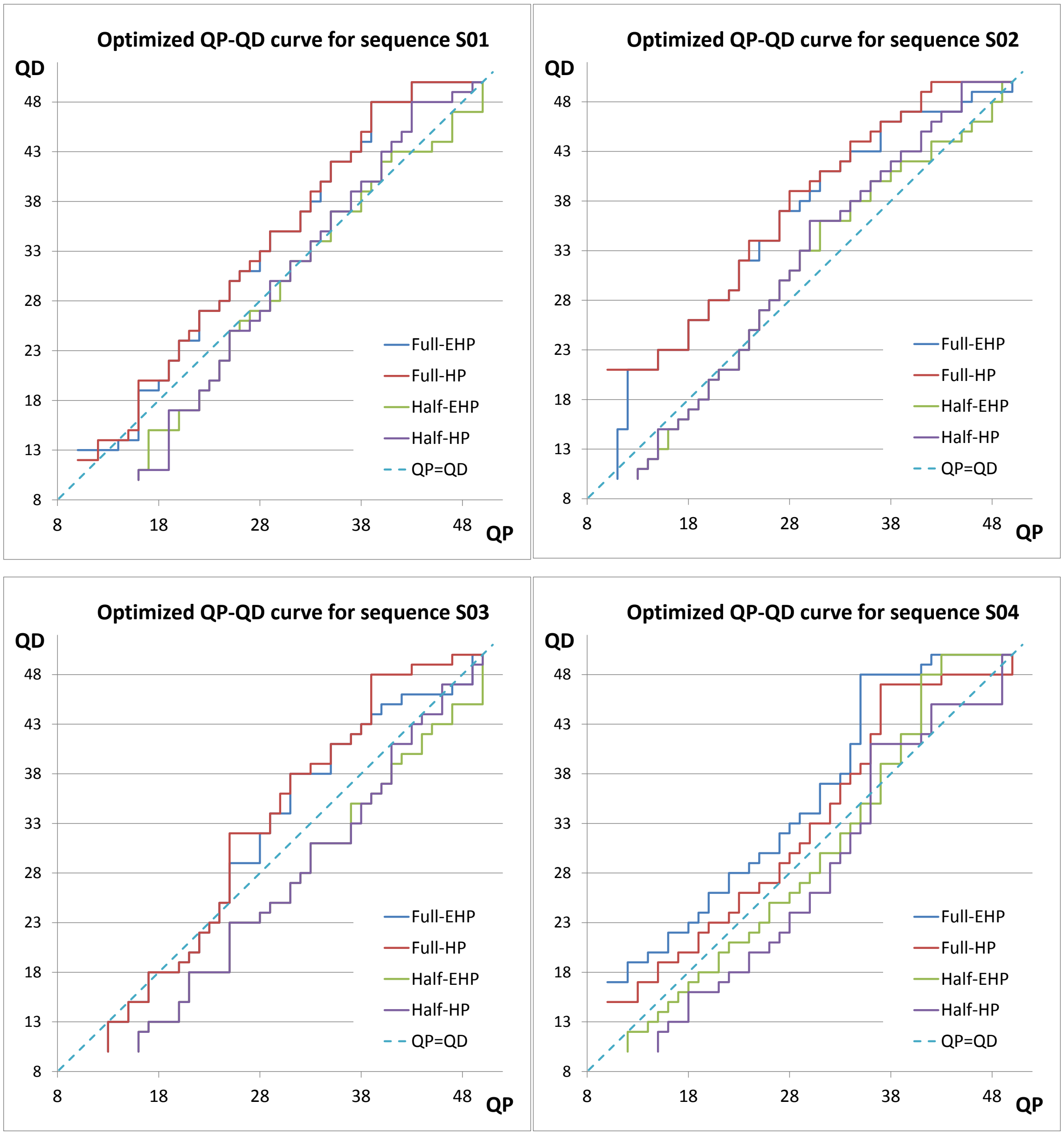
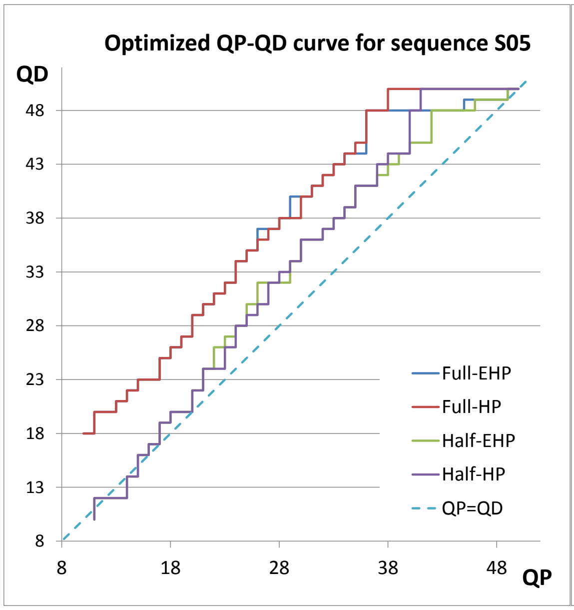
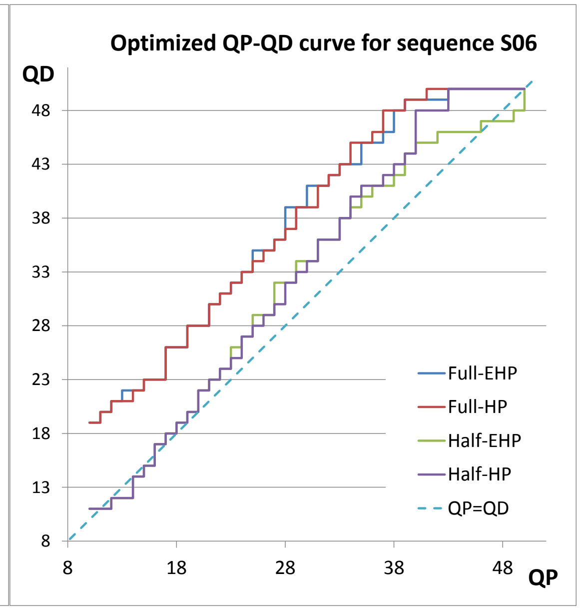


Figure 3. Optimized quantization parameters pairs for texture (QP) and depth (QD) views for each sequence.





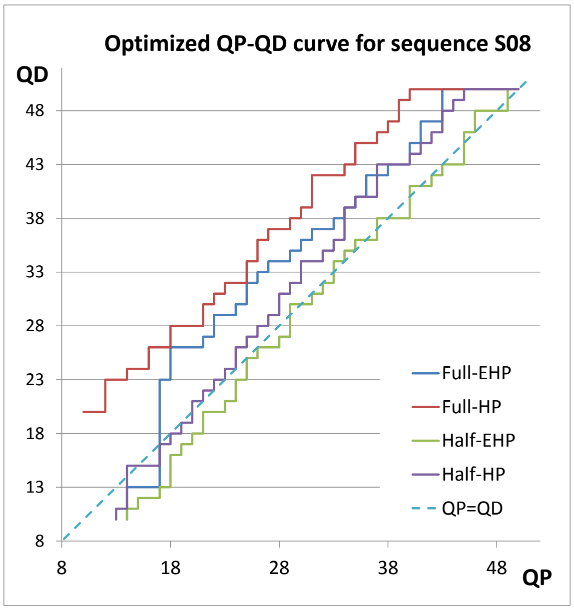


Figure 3. (continued) Optimized quantization parameters pairs for texture (QP) and depth (QD) views for each sequence.

Table 2. QD(QP) equations, derived based on linear regression with minimization of least square line fitting.

Sequence	Full-Res EHP	Full-Res HP	Half-Res EHP	Half-Res HP
S01	QD=1,213*QP - 1,938	QD=1,202*QP - 1,436	QD=1,195*QP - 7,174	QD=1,317*QP - 10,504
S02	QD=1,161*QP + 3,268	QD=1,007*QP + 7,826	QD=1,203*QP - 4,641	QD=1,254*QP - 5,513
S03	QD=1,284*QP - 5,601	QD=1,326*QP - 6,092	QD=1,083*QP - 6,689	QD=1,124*QP - 7,724
S04	QD=1,117*QP + 2,893	QD=1,119*QP + 0,11	QD=1,082*QP - 3,422	QD=1,239*QP - 9,555
S05	QD=1,055*QP + 7,251	QD=1,114*QP + 5,98	QD=1,214*QP - 2,479	QD=1,231*QP - 3,017
S06	QD=1,029*QP + 7,846	QD=1,063*QP + 7,019	QD=1,201*QP - 2,788	QD=1,221*QP - 3,313
S08	QD=1,156*QP - 0,121	QD=0,983*QP + 9,305	QD=1,132*QP - 5,122	QD=1,228*QP - 4,476
Average	**QD=1,126*QP + 2,441**	**QD=1,108*QP + 3,424**	**QD=1,090*QP – 2,800**	**QD=1,145*QP - 3,973**

Table 3. Optimized quantization parameters pairs that maximize coding performance

Full resolution depth map coding EHP																																									
QP	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
QD	14	16	17	18	19	20	21	23	24	25	26	27	28	29	31	32	33	34	35	36	38	39	40	41	42	43	45	46	47	48	49	50	51	51	51	51	51	51	51	51	51
Half-resolution depth map coding EHP																																									
QP	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
QD	17	18	19	20	21	22	23	24	25	26	27	28	29	30	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	51	51	51	51	51	51	51
Full-resolution depth map coding HP																																									
QP	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
QD	7	8	9	10	12	13	14	15	17	18	19	20	21	23	24	25	26	27	29	30	31	32	33	35	36	37	38	39	41	42	43	44	45	47	48	49	50	51	51	51	51
Half-resolution depth map coding HP																																									
QP	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
QD	7	8	9	10	12	13	14	15	17	18	19	20	22	23	24	25	27	28	29	30	32	33	34	35	37	38	39	40	42	43	44	45	47	48	49	50	51	51	51	51	51

For the coding with optimized quantization parameters pairs we have plotted R-D curve for each sequence (Fig. 4). As can be seen in all cases half resolution depth map coding is superior compared to full resolution depth map coding. Bjontegaard bitrate delta with respect to average of psnr of the synthesized views have been provided in Table 3 and 4 for HP and EHP retrospectively. Half-resolution depth map coding is superior in comparison with full-resolution depth map coding by **-6.58% for HP and -7.93% for EHP** on average.

Moreover we have compared coding performance with optimized quantization parameters pairs with pairs from Common Test Conditions (equal QP and QD values). As can be seen in Tables 5-8 optimized quantization parameters gives better results that CTC by -1.57% HP half res, -1.25% EHP half-res, -8.55% HP full-res and -7.96% EHP full-res. Current quantization parameters for depth views in CTC are not the optimal ones, and there are still a lot of room for improvement by simply quantization parameter adjustment.

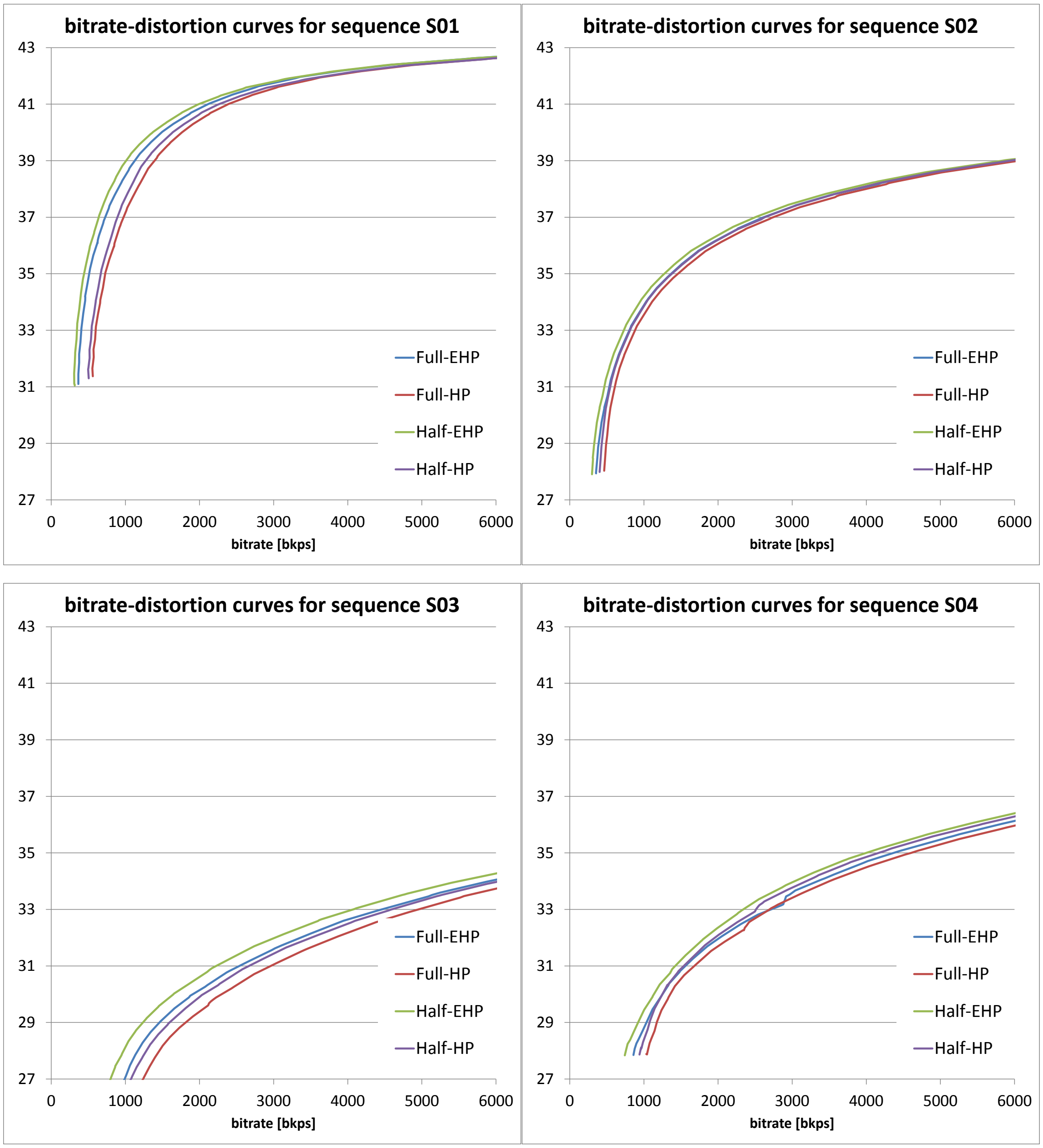
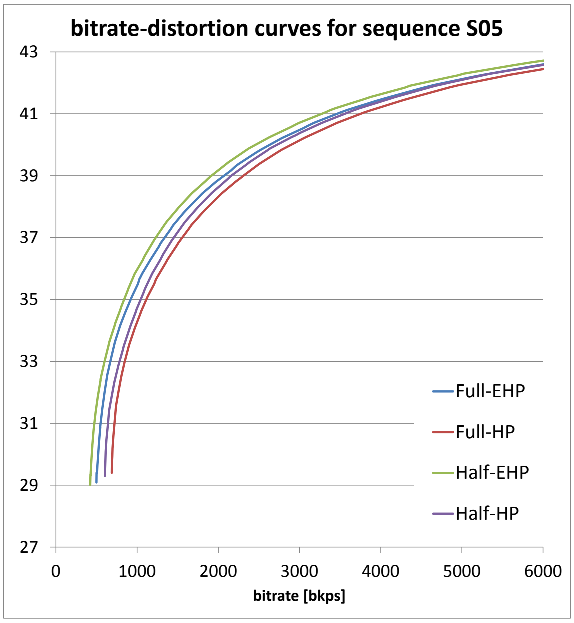
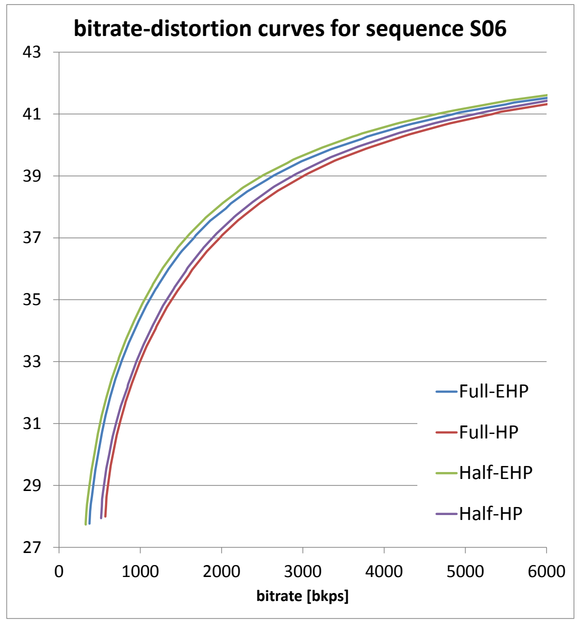


Figure 4. R-D curve for coding with optimized quantization parameters pairs, for each sequence.





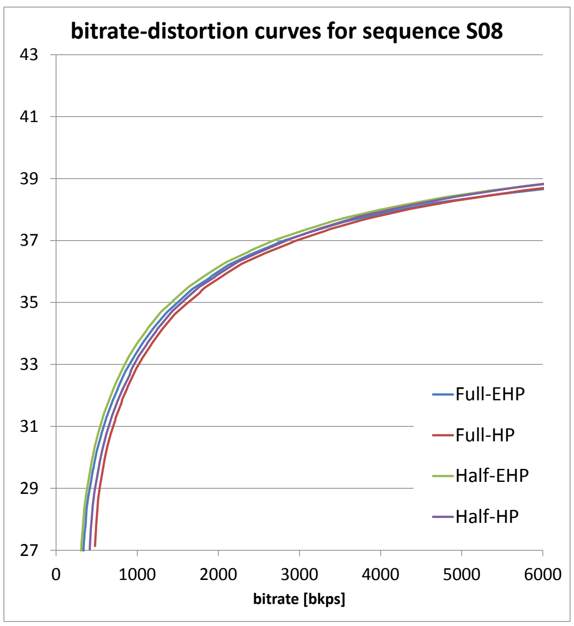


Figure 4.(continued) R-D curve for coding with optimized quantization parameters pairs, for each sequence.

Table 3. Comparison of half- vs full-resolution depth coding in HP
(both with use of proposed maximized performance QP/QD)
(half resolution depth coding is better)

	Synthesized Views	
	BD-rate (piecewise cubic)	BD-rate (cubic)
S01	-7,72%	-7,63%
S02	-5,50%	-5,50%
S03	-7,81%	-7,81%
S04	-8,56%	-8,55%
S05	-6,56%	-6,53%
S06	-4,59%	-4,54%
S08	-5,34%	-5,34%
Average	**-6,58%**	**-6,56%**

Table 4. Comparison of half- vs full-resolution depth coding in EHP.
(both with use of proposed maximized performance QP/QD)
(half resolution depth coding is better)

	Synthesized Views	
	BD-rate (piecewise cubic)	BD-rate (cubic)
S01	-10,18%	-10,19%
S02	-6,45%	-6,45%
S03	-9,45%	-9,44%
S04	-9,75%	-9,73%
S05	-8,11%	-8,10%
S06	-5,52%	-5,48%
S08	-6,02%	-6,07%
Average	**-7,93%**	**-7,92%**

Table 5. Comparison of proposed maximized performance QP/QD vs CTC
(half- resolution depth coding in HP)

	Synthesized Views	
	BD-rate (piecewise cubic)	BD-rate (cubic)
S01	-0,30%	-0,30%
S02	-0,99%	-1,00%
S03	-1,57%	-1,56%
S04	-1,90%	-1,88%
S05	-3,05%	-3,05%
S06	-1,65%	-1,67%
S08	-1,49%	-1,51%
Average	**-1,57%**	**-1,57%**

Table 7. Comparison of proposed maximized performance QP/QD vs CTC
(half- resolution depth coding in EHP)

	Synthesized Views	
	BD-rate (piecewise cubic)	BD-rate (cubic)
S01	-0,29%	-0,33%
S02	-0,70%	-0,70%
S03	-2,22%	-2,21%
S04	-0,49%	-0,48%
S05	-3,23%	-3,24%
S06	-1,78%	-1,78%
S08	-0,06%	-0,06%
Average	**-1,25%**	**-1,26%**

Table 6. Comparison of proposed maximized performance QP/QD vs CTC
(full-resolution depth coding in HP)

	Synthesized Views	
	BD-rate (piecewise cubic)	BD-rate (cubic)
S01	-3,43%	-3,58%
S02	-9,38%	-9,39%
S03	-2,08%	-2,08%
S04	-1,88%	-1,89%
S05	-15,78%	-15,81%
S06	-12,43%	-12,51%
S08	-14,57%	-14,59%
Average	**-8,51%**	**-8,55%**

Table 8. Comparison of proposed maximized performance QP/QD vs CTC
(full- resolution depth coding in EHP)

	Synthesized Views	
	BD-rate (piecewise cubic)	BD-rate (cubic)
S01	-3,87%	-3,96%
S02	-8,87%	-8,87%
S03	-2,03%	-2,03%
S04	-2,54%	-2,57%
S05	-17,55%	-17,58%
S06	-14,70%	-14,76%
S08	-6,00%	-5,98%
Average	**-7,94%**	**-7,96%**

Conclusions

Observations on optimized quantization parameter pairs that maximize coding performance
(QP-QD curve):

- The current condition, that quantization parameters for texture and depth views are equal (QP=QD) used in 3D-ATM CTC is not the optimal one – both for half and full-resolution depth maps coding.
- Generally, for coding with full resolution of depth maps – quantization parameter for depth views (QD) should be increased.
- Generally, for coding with half resolution of depth maps – quantization parameter for depth views (QD) should be decreased.
- **Optimized quantization parameters gives better results than CTC by -1.57% HP half resolution depth, -1.25% EHP half-resolution depth, -8.55% HP full resolution depth and -7.96% EHP full-resolution depth.**

Observations on half vs. full-resolution depth maps coding:

- Conclusions and observation made in JCT3V-D0080 [3] have been confirmed up to the level of relation (the values are different)
- **Bjontegaard bitrate delta (BD-Rate) for synthesized views of half-resolution depth map coding compared to full-resolution depth map coding is -6.58% for HP and -7.93% for EHP.**
- In current 3D-ATM it is better to encode depth with use of half-resolution compared to texture views both in HP and EHP.
- Thought, the difference in coding performance is smaller than estimated before
  (from D0014 BD-rate -17,9 % for HP and -12,6% for EHP)

Work will be continued for 3D-HTM

6 Recommendations

- Revise Common Test Condition according to optimized quantization parameters in order to evaluate new tools at maximum coding performance operation point of the current 3D-ATM
- Revise condition of subjective tests of 3D-AVC

7 Patent rights declaration(s)

Poznan University of Technology may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).

8 Acknowledgments

Research project was supported by National Science Centre, Poland, according to the decision DEC-2012/05/B/ST7/01279.

Olgierd Stankiewicz and Krzysztof Wegner are scholarship holders of Scholarship support for Ph.D. students specializing in majors strategic for Wielkopolska's development, Sub-measure 8.2.2 Human Capital Operational Programme, co-financed by European Union under the European Social Fund.

References

[1] D. Rusanovskyy, J.Y. Lee, J. Lin, D. Tian, O.Stankiewicz, ”AHG4: 3D-AVC Software Integration (AHG4)”, JCT3V-E0004, m30147.

[2] Dmytro Rusanovskyy, Karsten Müller, Anthony Vetro, “Common Test Conditions of 3DV Core Experiments”, JCT3V-D1100, m29521.

[3] Krzysztof Wegner, Olgierd Stankiewicz, Marek Domanski, “AHG14 Comparison of half resolution depth map coding versus full resolution depth map coding in 3D-ATM (Poznan)”, JCT3V-D0080, m28906.

[4] Krzysztof Klimaszewski, Krzysztof Wegner, Marek Domański, “Distortions of synthesized views caused by compression of views and depth maps” 3DTV-Conference 2009 The True Vision Capture, Transmission and Display of 3D Video, Potsdam, Germany, 4-6 May 2009

[5] Krzysztof Klimaszewski, Krzysztof Wegner, Marek Domański, “Influence of views and depth compression onto quality of synthesized views”, ISO/IEC JTC1/SC29/WG11 MPEG 2009/M16758, London, UK, July 2009