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

ISO/IEC JTC1/SC29/WG11
MPEG2009/M16671
April 2009, London, England

Source: GIST (Gwangju Institute of Science and Technology)
Status: Report
Title: Results of EE4 on 'Newspaper' Sequence
Author: Cheon Lee, and Yo-Sung Ho

1. Introduction

This document reports experimental results of depth coding on 'Newspaper' sequence as a response to EE4 of 3D Video [1]. Using the results of depth video generated by DERS 4.0, we conducted coding experiments. At first, we tested existing coding standards for the depth coding. Using the reconstructed depth map, we synthesized the intermediate view images. In addition, we propose a new QP set for depth video coding. Finally, we compared two depth representations.

2. Results for EE4: Depth Video Coding

As discussed during the last Maui meeting, EE4 is included again into the EEs. We conducted the coding experiments following the description of EE in the document N10360. At first, we compared the performance of existing coding standards such as H.264/AVC and MVC (multiview video coding) for MVD representation. Then, we checked the performance of inter-view prediction using MVC tool. Next, we selected five different QPs to reduce the consuming bits for depth coding. To compare the two representations, we checked their consuming bits for whole data and synthesized images. As an additional experiment, we checked the bit rates for various depth estimation methods.

2.1. Comparison of Two Standards

As we mentioned above, we checked the performance of existing coding standards: H.264/AVC and MVC. General information of standards and coding conditions are described in Table 1. The quantization parameters are followed the description document of EE. Figure 1 describes the simulation cases for 3D video. At first, we encoded the texture video individually with no inter-view prediction. Then, we encoded the texture video with inter-view prediction using JMVC 5.0. Only the anchor frames are allowed for inter-view prediction. For depth coding, the same coding strategy is used.

Table. 1. Coding conditions

Coding standards	H.264/AVC	Multiview Video Coding
Reference software	JM 14.1	JMVC 5.0
Quantization parameters	22, 26, 30, 34, 38	
Coding structure	Hierarchical B	
GOP size	15	
Number of frames	31 (2 GOPs)	
Search range	16	
Target views	4, 6	

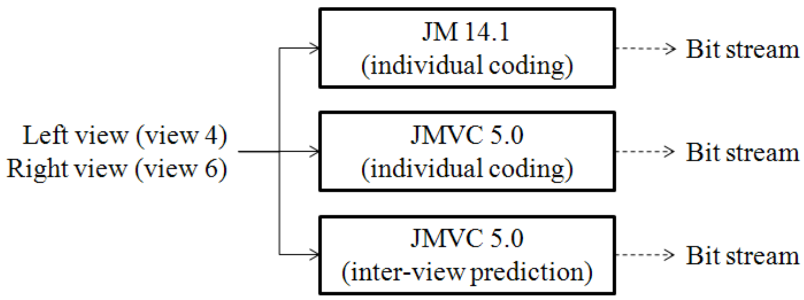


Fig. 1. Three coding strategies

Figure 2 shows the comparison of coding performances of texture video. The most efficient method is MVC with inter-view prediction. It gives 13.54% bit saving in BDBR comparing with H.264/AVC. On the other hand, in case of depth coding, H.264/AVC showed the best performance. Its saved 22.71%d bits comparing to the MVC. It is quite interesting result; it is needed more precise analysis.

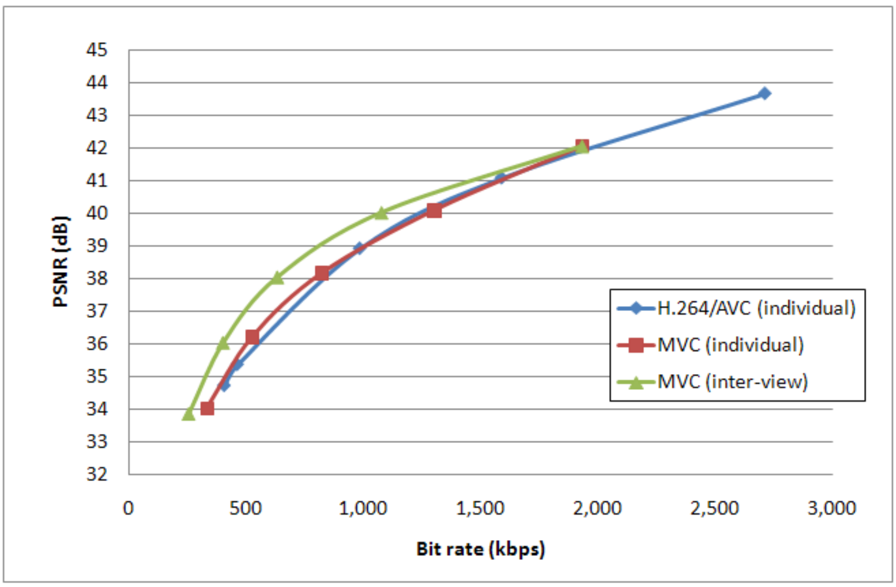


Fig. 2. Comparison of coding standards: Texture Coding

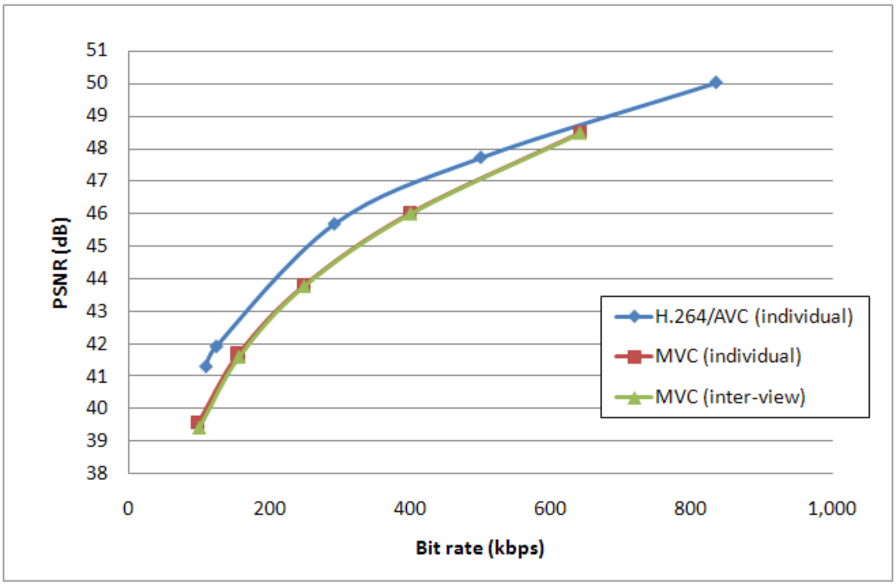


Fig. 3. Comparison of coding standards: Depth Coding

Using the reconstructed texture and depth data, we generated the intermediate view images at position 5 from position 4 and 6. Table 5 represents the PSNR values of the generated images using reconstructed texture and depth data coded by the same QP. Obviously, the quality of generated image was best when the QP was 22. The total bit rate is calculated by,

Total bit rate = Rate(L_texture) + Rate(R_texture) + Rate (L_depth) +Rate(R_depth)

The relationship between rate and distortion of synthesis image is shown in Fig. 6. The most effective method was MVC using inter-view prediction. Although MVC on depth coding was not efficient comparing to AVC, it outperformed AVC due to the high performance of texture coding.

Table. 5. Total bit rates and objective quality of synthesized images

QP	H.264/AVC (indiv.)		MVC (indiv.)		MVC (inter-view pred.)	
	PSNR of syn. (db)	Total bit rate (kbps)	PSNR of syn. (db)	Total bit rate (kbps)	PSNR of syn. (db)	Total bit rate (kbps)
22	32.31	7097.38	32.62	5149.87	32.62	5152.55
26	32.19	4178.66	32.54	3402.83	32.52	2957.25
30	31.95	2550.20	32.33	2143.21	32.26	1764.09
34	31.08	1168.43	31.95	1361.30	31.85	1115.27
38	30.84	1027.58	31.28	863.49	31.14	709.26

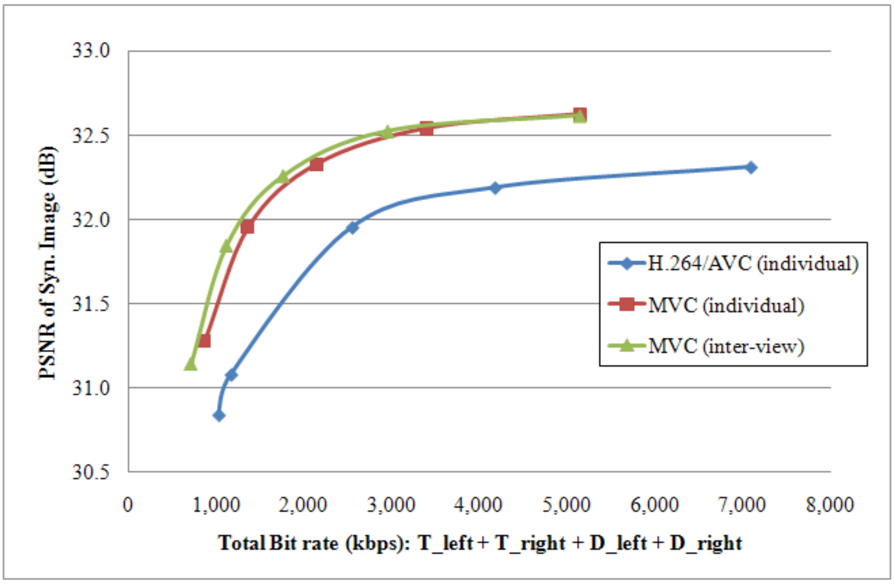


Fig. 6. Rate and synthesized image

2.2. New QP set for Depth Video Coding

To achieve the objectives of EE4, we selected another QP set for depth video coding. Considering the impact of distorted depth map by the lossy coding, we thought that the QP for depth coding would not be lower than that of the texture coding. Selecting proper QPs are very difficult problem because there is no evaluation method. Every expert may agree that the distortion of depth map affects the quality of synthesized image; however, it is hard to generalize the bits for depth map minimizing the distortion of the generated image. Leaving the problem in mind, we selected one QP set for depth coding as represented in the Table 6.

Table. 6. The proposed QP combination for 3D video coding

QPs for Texture	QPs for Depth
22	30
26	34
30	38
34	40
38	42

Using these QP combinations, we synthesized the intermediate images. Since we have newly reconstructed depth data and bits, we compared the quality of generated images and total bit rates as shown in Fig. 7. We can see that the proposed QP set shows better results except for the lowest bit point.

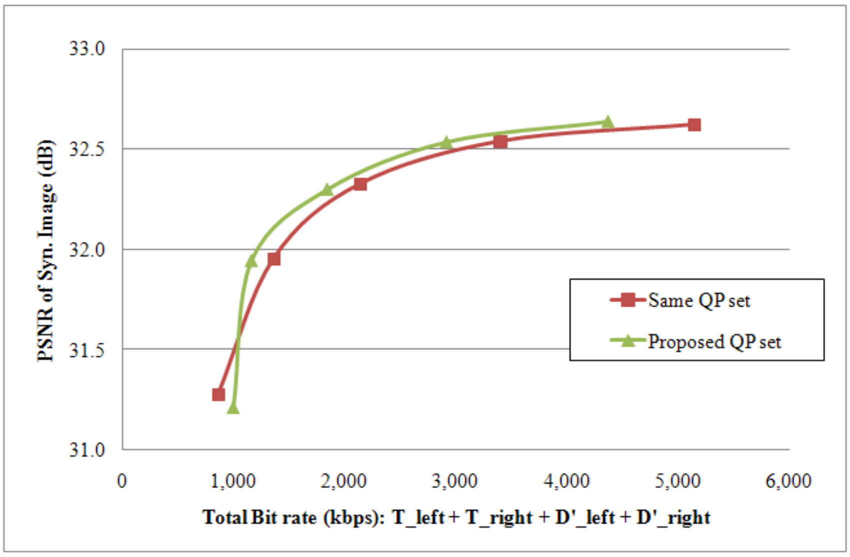


Fig. 7. The effect of the proposed QP set for depth coding

2.3. Results on LDV Data Format

In this subsection, we report the coding results of LDV representation. We chose different center view from the description of EE document N10552 because the synthesized view is position 5 in MVD format. We obtained BG and BD data from position 4, 6, and 7, and then synthesized the images of position 5 as illustrated in Fig. 8. Figure 9 shows the LDV data generated by the MVD2LDV 2.2.

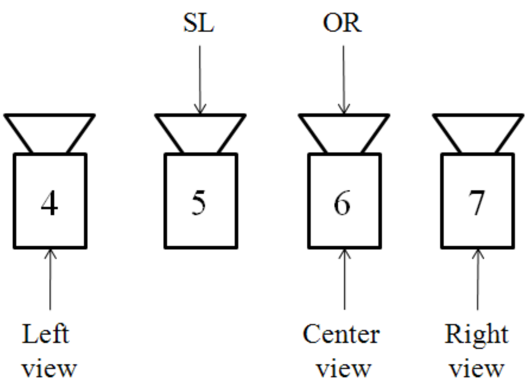
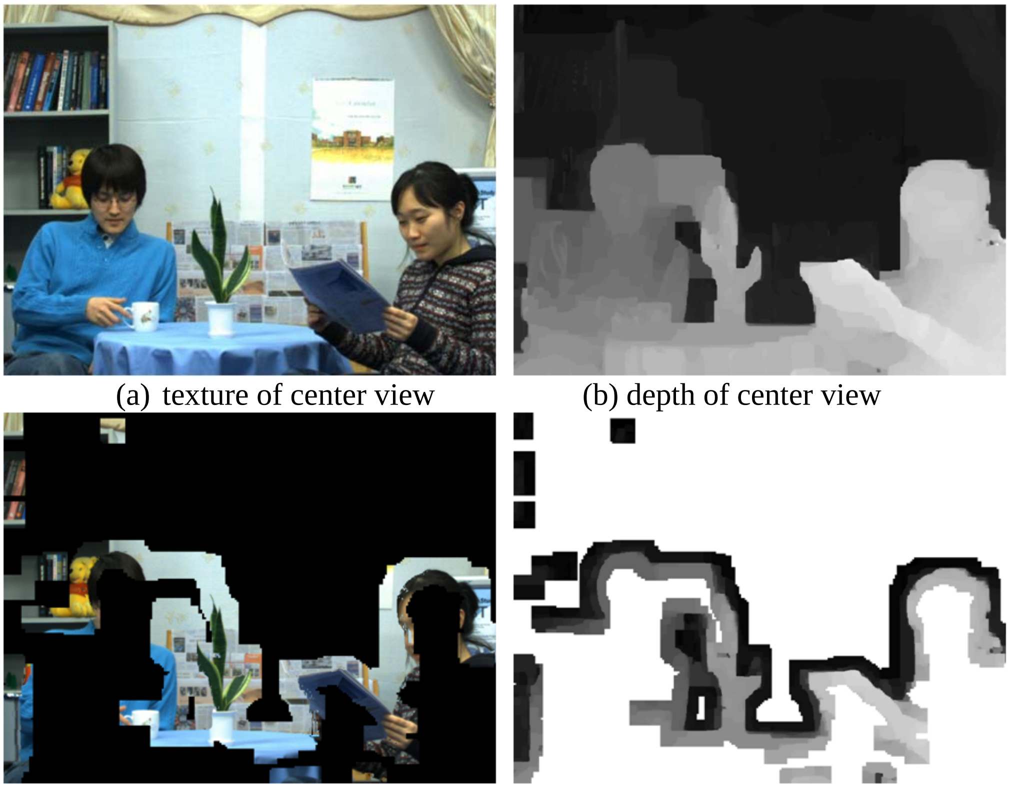


Fig. 8. Input and output views for LDV representation format

The coding condition is almost the same with the MVD coding except for the inter-view prediction. We coded 31 frames and synthesized the target image of position 5. To compare the bit rate of data, we summed the all bits such as;

Total bit rate = Rate(C_texture) + Rate(C_depth) + Rate (B_texture) +Rate(B_depth)



(a) texture of center view (b) depth of center view

(c) background texture (d) background depth

Fig. 9. LDV data of newspaper sequence

The objective quality is calculated by PSPNR software. Table 7 is the results of LDV coding. Figure 10 is still cut of the synthesized image at the position 6.

Table. 7. Coding results of LDV data using JMVC 5.0

QP	Bit rate (kbps)					PSPNR (dB)		
	Center Texture	Center Depth	Back. Texture	Back. Depth	Total Bits	Av	Av_T	Av_S
22	2269.40	645.65	1535.47	646.75	5097.27	29.80	50.23	32.97
26	1367.98	409.64	1048.17	456.27	3282.06	29.77	50.63	32.98
30	864.80	259.06	731.75	320.05	2175.66	29.71	50.85	32.89
34	552.18	164.21	509.61	229.85	1455.85	29.45	51.09	32.58
38	350.19	104.97	343.74	171.42	970.33	29.09	51.51	32.23

In order to compare the coding performance between two representations, we collected the coding results and made a table as described in Table 8. To make a pair comparison, we calculate the bit rate of MVD case by multiplying 1.5 to the results of Table 5 because the total view angle of LDV is 1.5 times bigger than that of MDV. Looking through the table, we can know that the total bit rates of LDV are much less than MVD while the objective quality of LDV is obviously lower than that of MVD. However, this result does not mean that LDV format is better than MVD because MVD format renders much better synthesized images. Figure 10 shows the rate-distortion curves of the results. In conclusion, it is hard to tell which format is more efficient than the other.

Table. 8. Comparison of coding results for two data representations

QP	MVD		LDV	
	Av. PSNR (dB)	Total bit rate (kbps)	Av. PSNR (dB)	Total bit rate (kbps)
22	32.62	7728.83	29.80	5097.27
26	32.52	4435.88	29.77	3282.06
30	32.26	2646.14	29.71	2175.66
34	31.85	1672.91	29.45	1455.85
38	31.14	1063.89	29.09	970.33

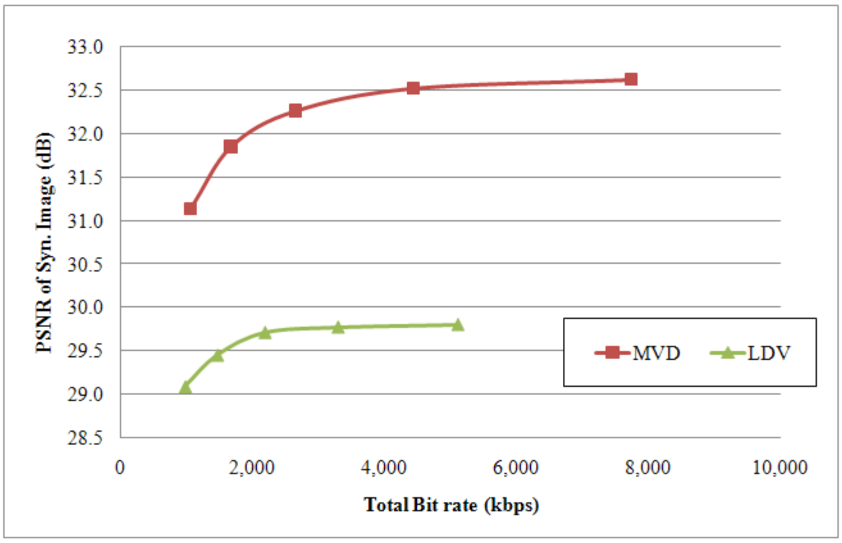


Fig. 10. R-D curves of the LDV format and MVD format

3. Conclusion

We conducted many experiments on depth coding. The multiview video coding tool outperformed H.264/AVC coder for texture coding. Also, we proposed a new QP set for depth coding; it reduced bits for depth video. The comparison of two depth representations has been dong in this experiment, but it is difficult to define a pair condition in comparison.

4. Acknowledgements

This work was supported in part by ITRC through RBRC at GIST (IITA-2009-C1090-0901-0017).

5. References

[1] ISO/IEC JTC1/SC29/WG11 "Description of Exploration Experiments in 3D Video Coding", N10360, February 2009.

[2] ISO/IEC JTC1/SC29/WG11 "Semi-automatic Depth Estimation for FTV", m16391, April 2009.

[3] ISO/IEC JTC1/SC29/WG11 "Depth Estimation Algorithm in SADERS 1.0", m16411, April 2009.