Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11
2nd Meeting: Shanghai, CN, 13–19 Oct. 2012

Document: JCT3V-B0124

Title:	**Analysis of View Synthesis Methods (VSRS 1D fast and VSRS3.5)**		
Status:	Input Document		
Purpose:	Report		
Author(s) or Contact(s):	Do-Young Kim, Woo-Seok Jang, Yo-Sung Ho 123 Cheomdangwagi-ro Buk-gu Gwangju, Korea 500-712	Tel: Email:	+82-62-715-2263 kimdo@gist.ac.kr jws@gist.ac.kr hoyo@gist.ac.kr
Source:	Gwangju Institute of Science and Technology (GIST)		

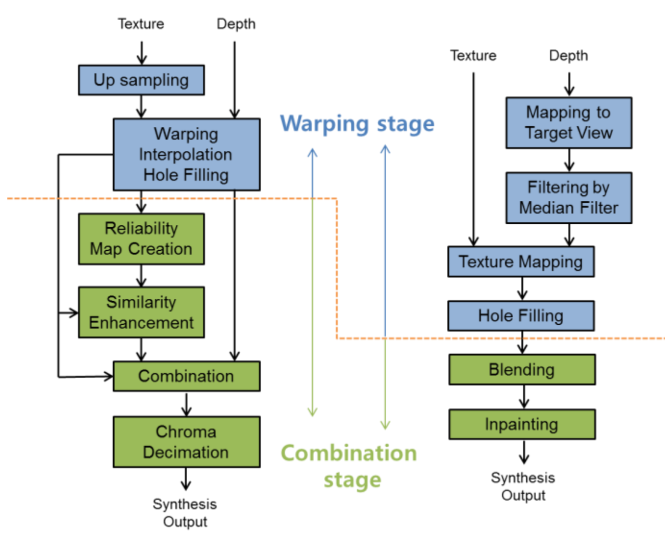
____

Abstract

In this document, to find better synthesis performance algorithm, we describe differences between VSRS-1D-fast in HTM and VSRS3.5. Processing steps of each algorithm is divided into two main parts. Then, by comparing results of each stage, we analyze the performance and effect of the stage. We test synthesis quality, performance of coding error compensation, and coding performance with MPEG 3DV test sequences. Also, we compare VSO performance in HTM. Experiments show VSRS-1D-fast is better algorithm than VSRS3.5.

1 View Synthesis Algorithms

Two main view synthesis algorithms can be utilized in HTM. One method is a fast 1-dimensional view synthesis algorithm called VSRS-1D-fast that is part of HEVC-based software. The other method is VSRS3.5 that was developed during the 3DV exploration experiments. In this document, we discuss "General mode" only among two modes of the software. Processing steps of the view synthesis algorithms are depicted in Fig. 1. The methods are divided into two main parts; warping stage consists of warping and hole filling, and combination stage is blending step.



(a) VSRS-1D-fast (b) VSRS3.5

Figure 1. Block diagram of view synthesis methods.

2 Experimental Environment

To analyze the performance and effect of each stage, we extract two kinds of synthesis image, and then compare the images of each algorithm. First, intermediate results from warping stage are obtained. In the VSRS-1D-fast, the reference view is reprojected to the target view using horizontal pixel shifting and the hole area is temporarily filled with background values. In VSRS3.5, the target view is generated by using pixel-by-pixel mapping based on 3D warping from the reference view. Originally, the hole filling method in the VSRS3.5 refer to the other reference view but we utilize inpainting algorithm for hole filling. Then, final synthesis results of combination stage are compared. We also test performance of view synthesis optimization (VSO) that is exploited only indirectly in the rendering process in VSRS-1D-fast. In short,

- Warping: disparity shifting/3D warping
- Hole filling: line wise background extension/inpainting algorithm
- Reliability map creation and similarity enhancement: on/off
- VSO in VSRS-1D-fast: on/off

3 Experimental Results

In the comparison experiments, to evaluate synthesis quality, quality of coding error compensation, and coding performance of each view synthesis algorithm, we have tested with the MPEG 3DV test sequences.

Table 1 shows the results for synthesis performance that is average value of differences in PSNR values between original view image and synthesis view image using original depth data. The comparative views of each sequence are as in the following.

- *Dancer*: 2 and 3 views
- *Kendo*: 2 and 4 views
- *Balloons*: 2 and 4 views
- *Newspaper*: 3 and 5 views

Most of the 1D-fast (VSRS-1D-fast) results of warping stage were lower than VSRS (VSRS3.5). But, after combination stage, the PSNR values were significantly improved. So 1D-fast gain that is a variance of each stage was higher than VSRS by maximum 2.4807 dB in *Newspaper* sequence as shown in Fig. 2. This means that 1D-fast algorithm generate intermediate views more similarly to the original view image.

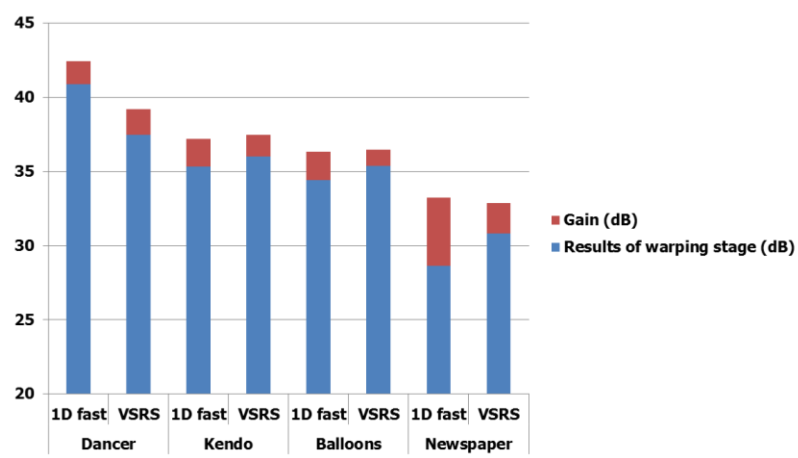
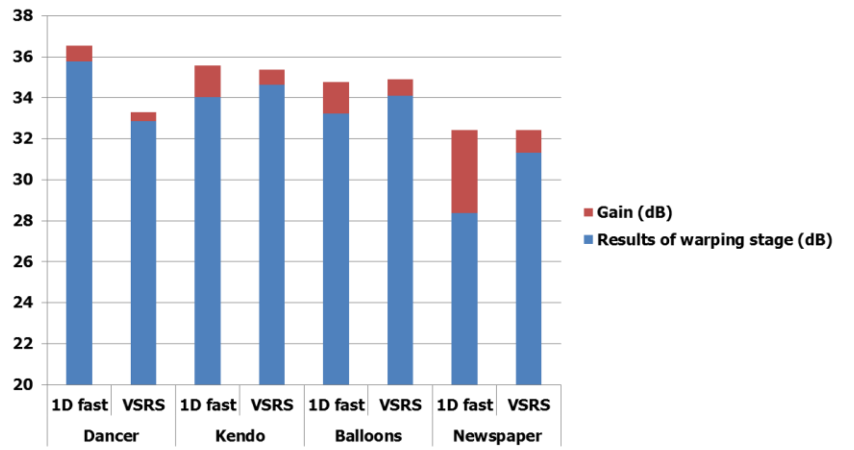


Figure 2. Comparison of synthesis performance

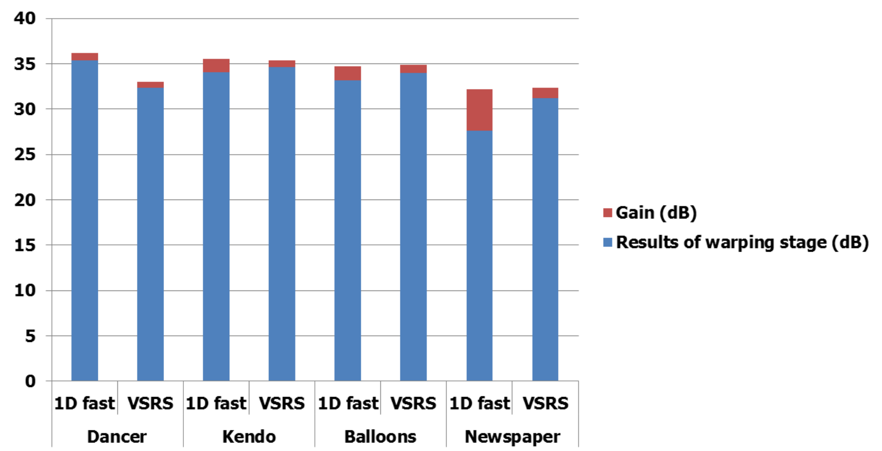
Table 1. Results for synthesis performance

	Results of warping stage (dB)			Final Results (dB)			Gain (dB)		
Method	1D fast	VSRS	Δ	1D fast	VSRS	Δ	1D fast	VSRS	Δ
Dancer	40.8872	37.4513	3.4359	42.4126	39.1757	3.2369	1.5254	1.7244	-0.1990
Kendo	35.3161	36.0073	-0.6912	37.1833	37.4609	-0.2776	1.8672	1.4536	0.4136
Balloons	34.3989	35.3905	-0.9916	36.3434	36.4584	-0.1150	1.9445	1.0679	0.8766
Newspaper	28.6552	30.8054	-2.1502	33.2104	32.8799	0.3305	4.5552	2.0745	2.4807

Table 2 shows the results for coding error compensation and performance of VSO. For this experiment, the original view images and synthesis view images using encoded depth data were compared by PSNR values. The sequences were encoded with four QPs: 25, 30, 35, and 40. We averaged the PSNR values of each view. The 1D-fast results of warping stage were lower than VSRS because of line-wise background extension for disocclusion area in 1D-fast while inpainting algorithm was utilized in VSRS. But, as shown in Fig. 3, combination stage in 1D-fast compensated the coding error effectively by using similarity enhancement and blending according to reliability map regardless of VSO on/off. But, in case of VSO on, the gain has been increased. The final results from 1D-fast were higher than VSRS by maximum 3.2659 dB in *Dancer* sequence with VSO on.



(a) VSO on



(b) VSO off

Figure 3. Comparison of performance of coding error compensation

Table 2. Results for coding error compensation

VSO ON									
	Results of warping stage (dB)			Final Results (dB)			Gain (dB)		
Method	1D fast	VSRS	Δ	1D fast	VSRS	Δ	1D fast	VSRS	Δ
Dancer	35.7805	32.8516	2.9289	36.5520	33.2861	3.2659	0.7715	0.4345	0.3370
Kendo	34.0252	34.6378	-0.6126	35.5679	35.3742	0.1937	1.5427	0.7364	0.8063
Balloons	33.2376	34.0776	-0.8400	34.7637	34.8866	-0.1229	1.5261	0.8090	0.7171
Newspaper	28.3772	31.3274	-2.9502	32.4221	32.4098	0.0123	4.0449	1.0824	2.9625
VSO OFF									
	Results of warping stage (dB)			Final Results (dB)			Gain (dB)		
Method	1D fast	VSRS	Δ	1D fast	VSRS	Δ	1D fast	VSRS	Δ
Dancer	35.3704	32.3884	2.9820	36.1940	33.0445	3.1495	0.8236	0.6561	0.1675
Kendo	34.0483	34.6384	-0.5901	35.5322	35.4012	0.1310	1.4839	0.7628	0.7211
Balloons	33.1821	33.9807	-0.7986	34.7082	34.8429	-0.1347	1.5261	0.8622	0.6639
Newspaper	27.6487	31.2357	-3.5870	32.1963	32.3898	-0.1935	4.5476	1.1541	3.3935

Table 3 shows the results for coding performance. The figures are differences in PSNR values between synthesis view images using original depth data and synthesis view images using encoded depth data. The sequences were synthesized with 6-view images as in the following and encoded with four QPs: 25, 30, 35, and 40.

- *Poznan_street*: 3.25, 3.5, 3.75, 4.25, 4.5, and 4.75 views
- *Dancer*: 2, 3, 4, 6, 7, and 8views
- *Kendo*: 1.5, 2, 2.5, 3.5, 4, and 4.5 views
- *Balloons*: 1.5, 2, 2.5, 3.5, 4, and 4.5 views
- *Newspaper*: 2.5, 3, 3.5, 4.5, 5, and 5.5 views

We averaged the PSNR values of each sequence. The results from 1D-fast had PSNR gain by average 3.69% in VSO on and 2.95% in VSO off. VSO reduces coding error of the encoded depth data more effectively. Therefore, when we use the 1D-fast algorithm for view synthesis, we can decrease residual data because the differences in PSNR values were lower than VSRS.

Table 3. Results for coding performance

	VSO ON (dB)			VSO OFF (dB)		
Method	1D fast	VSRS	Δ	1D fast	VSRS	Δ
Poznan_street	37.0828	36.8761	0.2067 (0.56%)	36.8421	36.6479	0.1942 (0.53%)
Dancer	37.4853	34.0236	3.4617 (10.1%)	36.8878	33.8773	3.0105 (8.89%)
Kendo	39.5698	39.3596	0.2102 (0.53%)	39.3302	39.2675	0.0627 (0.16%)
Balloons	39.1229	38.7785	0.3444 (0.89%)	38.8649	38.6223	0.2426 (0.63%)
Newspaper	36.7973	34.6244	2.1729 (6.28%)	36.1561	34.5867	1.5694 (4.54%)
Average (%)	3.69%			2.95%		

4 Conclusion

We compared 1D-fast and VSRS view synthesis algorithm to find a better synthesis performance algorithm. Overall, the results of warping stage were lower than VSRS. Similarity enhancement and blending according to reliability map in combination stage of 1D-fast compensated the coding error of depth data and improved the quality of synthesized view images. Consequently, 1D-fast is better than VSRS. Also, in terms of coding performance, the residual data can be decreased because the differences in PSNR values with results from 1D-fast were smaller than VSRS.

5 Acknowledgment

This research was supported a grant “Establishing the Foundation of International Standardization for Mobile Devices and Multimedia Convergence Technology” from Improvement of Standards Technology Program funded by Korean Agency for Technology and Standards, Ministry of Knowledge Economy of Korea.

6 Patent rights declaration(s)

Gwangju Institute of Science and Technology (GIST) does not have any current or pending patent rights relating to the technology described in this contribution.