

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

ISO/IEC JTC1/SC29/WG11
MPEG2007/m14996
October, 2007 Shenzhen, China

Title: Depth Coding and Virtual View Synthesis for FTV
Source: GIST (Gwangju Institute of Science and Technology)
Author: Yo-Sung Ho, Sang-Tae Na, Kwan-Jung Oh, and Cheon Lee

1. Introduction

Increasing audio-visual (AV) applications are being designed to enable dynamic viewing of a captured event or scene[1][2]. Of the AV application scenarios being investigated they identified free viewpoint television (FTV) as being the most challenging. FTV enables unrestricted spatio-temporal navigation within a scene captured using multiple views. The multiple views can be virtual views as well as real views from multiple cameras. In order to obtain the virtual views we need corresponding depth views and camera parameters. In this document, we introduce the effects of down-sampled and median filtered depth. As a result, we found that down-sampling affects bad result in boundary region in synthesized view and median filtering can give a benefit to depth coding.

2. Depth Coding Virtual View Synthesis for FTV

Once we have multi-view video and its corresponding depth, we can generate the virtual view. When we map pixel positions in reference view to target view, the positions are determined by the depth values of the corresponding position in depth image. Therefore, in order to generate high quality of virtual views, it is necessary to preserve the depth values as far as possible. The followings are the experiments for the effects of down-sampled and median filtered depth.

2.1. The Effects of Down-sampled Depth in View Synthesis

We experiment the effect of using down-sampled depth in generating virtual views. In the experiment, we compare the quality of the synthesized image using the encoded original depth image and down-sampled/Up-sampled depth image with the similar bit-rate. For experiments, first ten frames of 'Breakdancers' view0 and view2 are used and we use JM 12.4 and IPPP structure. Table 1 shows the bit-rate comparison. Intentionally, we adjust the bit-rate for down-sampled depth with the bit-rate for encoded original depth. Fig. 1 and Fig. 2 illustrate the effect of down-sampled depth at high bit-rate and

low bit-rate. As shown the figures, the quality of virtual images is degraded as depth images are degrade. Therefore, in order to synthesize high quality virtual views, we should use high quality of depth. Furthermore, since down-sampling affects the edges which are important for view synthesis, down-sampling is not a good method for depth coding.

Table 1. Bit-rate Comparison (Breakdancers: view0 and view2) [bits, (QP)]

View0		View2	
Original	Down/Up-sampling	Original	Down/Up-sampling
1,089,564 (22)	1,011,084 (13)	969,444 (22)	925,392 (13)
590,460 (27)	580,164 (18)	530,928 (27)	521,832 (18)
306,132 (32)	291,576 (24)	280,872 (32)	260,952 (24)
161,640 (37)	158,940 (29)	144,720 (37)	142,560 (29)



(a) encoded original depth (b) original depth (c) down/up-sampled depth

Fig. 1 Synthesized Images at high bit-rate



(a) encoded original depth (b) original depth (c) down/up-sampled depth

Fig. 2 Synthesized Images at low bit-rate

2.2. The Effect of median filtered Depth in View Synthesis

We also experiments the effect of median filtered depth in view synthesis. To check the benefit of median filtering, we compare the quality of the synthesized image using original depth and median-filtered depth in Fig. 3. Fig. 4 shows the original depth and median-filtered depth. In addition, we compare the bit-rates in Table 2. The test conditions are the same as above. The quality of the synthesized image using median-filtered depth is quite similar to the original result, while we can save the bit-rate using median

Table 2. Bit-rate Comparison (Breakdancers: view 0)

QP	Original	Median filter (7 x 7)	Bit Saving (%)
22	969,444	776,592	19.89
27	530,928	447,012	15.81
32	280,872	246,888	12.10
37	144,720	136,044	5.60



Fig. 3 Synthesized Images (left: from original depth, right: from 7x7 median-filtered depth)



Fig. 4 Depth Comparison (left: original depth, right: 7x7 median-filtered depth)

3. Summary

In this document, we experiments the effects of down-sampled and median filtered on depth map. As a result, in order to generate high quality of virtual views we should use high quality of depth images. From the experiments we also demonstrated that down-sampling is a not good method for FTV and median filtering can give a benefit to depth coding.

4. Acknowledgements

This work was supported in part by the Information Technology Research Center (ITRC) through the Realistic Broadcasting Research Center (RBRC) at Gwangju Institute of Science and Technology (GIST), and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.

5. References

- [1] A. Smolic and P. Kauff, "Interactive 3d video representation and coding technologies," IEEE Special Issue on Advances in Video and Delivery, Dec 2004.
- [2] A. Smolic, K. Müller, P. Merkle, C. Fehn, P. Kauff, P. Eisert, and Thomas Wiegand, "3D Video and Free Viewpoint Video – Technologies, Applications and MPEG Standards", ICME 2006, IEEE International Conference on Multimedia and Expo, Toronto, Ontario, Canada, July 2006