Depth reconstruction filter for depth coding

K.-J. Oh, S. Yea, A. Vetro and Y.-S. Ho

Depth images represent the distances of scene elements from a camera in 3D space; their efficient coding is crucial for emerging applications such as free-viewpoint TV and 3D video. An in-loop reconstruction filter that improves the depth-coding performance as well as the rendering quality of virtual views based upon the coded depth is proposed.

Introduction: Most existing image-based rendering (IBR) methods [1] employ depth images in combination with stereo or multi-view videos to realise 3D and free-viewpoint TV (FTV) applications [2, 3]. The recently finalised multi-view video coding (MVC) standard supports inter-view prediction for improved coding efficiency for multi-view videos, but it does not include any particular coding tools for depth images. Recently, MPEG initiated a new standardisation activity envisaging applications such as FTV where efficient estimation and coding of depth are crucial in order to enable high-quality virtual view generation at the receiver end. In this Letter, an in-loop reconstruction filter to improve depth coding efficiency and the rendering quality of a virtual view based on the coded depth is proposed.

Depth reconstruction filter: Unlike common images, depth images are spatially monotonous except at object boundaries in general. Thus coding errors tend to be concentrated around object boundaries and failure to preserve them usually leads to the significantly compromised rendering qualities of virtual views. In this Letter, we propose a depth reconstruction filter that explicitly takes the characteristics of depth images into account to compensate for depth coding errors. It not only reduces the depth coding bit rate, but also improves the rendering quality of a virtual view based upon the coded depth.

The proposed depth reconstruction filter consists of the following two filters: 1) a frequent-close filter to recover object boundaries; 2) a bilateral filter to eliminate the remaining errors. The frequent-close filter is a kind of nonlinear filter and is designed as follows:

$$\begin{aligned}&\text{if } |FC_{first}(A) - I(x, y)| < |FC_{second}(A) - I(x, y)|\\&\quad I(x, y) = FC_{first}(A)\\&\text{else}\\&\quad I(x, y) = FC_{second}(A)\end{aligned} \tag{1}$$

where A represents a rectangular-shaped region of pixels and $I(x, y)$ is the intensity value of the pixel at (x, y) within A. FC_{first} and FC_{second} are the pixel intensity values of the highest and the second-highest frequencies of occurrence within A, respectively. In other words, $I(x, y)$ is replaced by the closer of the two representative values FC_{first} and FC_{second} in the proposed frequent-close filter.

The proposed frequent-close filter has the following advantages over other linear filters. First, it is more robust against outliers, a single pixel that does not have a similar intensity to those of neighbouring pixels will not affect the frequent-close value significantly. Secondly, since the frequent-close value must actually be the value of one of the pixels in the neighbourhood, the frequent-close does not create new unrealistic pixel values when the filter straddles an edge. However, some errors may still remain and look like noise. To eliminate the remaining errors, we apply a bilateral filter.

The bilateral filter removes the remaining errors around the object boundary while preserving the boundary itself by means of a nonlinear combination of nearby pixel values. The bilateral filter has two parameters, colour sigma (σ_1) and space sigma (σ_2), which determine the strengths of two filter kernels, each of which pertains to photometric and geometric distances of input pixels, respectively.

Experimental results: The depth reconstruction filter was implemented in the H.264/AVC reference software and was located after the deblocking filter process shown in Fig. 1. We tested the proposed algorithm on 'Ballet' and 'Breakdancers' sequences. Among the eight views available, views 3 and 5 were selected as reference views and view 4 was set as the virtual view to be synthesised. Depth views were independently encoded using JM 14.0 with QPs 22, 27, 32, and 37 with the IBP coding structure. In this Letter, the proposed depth reconstruction filter consisted of a 9×9 frequent-close filter and a 3×3 bilateral filter. The bilateral filter parameter σ_1 was set to 15 and 10 for 'Ballet' and 'Breakdancers', respectively, and σ_2 was set to 10 for both sequences.

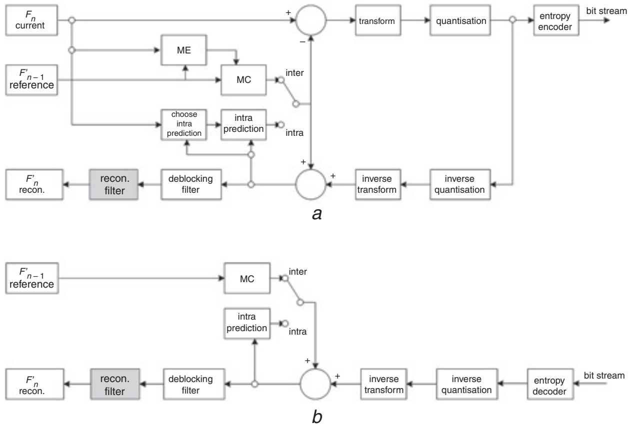

Fig. 1 *Implementation of depth reconstruction filter on H.264/AVC*
a H.264/AVC encoder
b H.264/AVC decoder

Figs. 2*c* and *d* show sample frames of the virtual views generated based upon the reconstructed depths shown in Fig. 2*a* (i.e. without the proposed filter) and Fig. 2*b* (i.e. with the proposed filter), respectively. It can be readily confirmed that the proposed depth reconstruction filter eliminates the depth coding errors effectively while preserving the object boundary, resulting in a better rendering result.

Fig. 2 *Example of depth reconstruction filter for 'Ballet' QP 37*
a Decoded depth
b Synthesised image from *a*
c Reconstructed depth
d Synthesised image from *c*

Table 1: Experimental results for 'Ballet'

QP	Depth PSNR (dB)		Depth bit rate (kbit/s)		rendering PSNR (dB)	
	original	proposed	original	proposed	original	proposed
QP 22	50.00	50.30	1032.82	1022.72	32.15	32.20
QP 27	46.95	47.57	610.00	599.13	32.03	32.10
QP 32	43.22	43.93	357.28	353.87	31.72	31.94
QP 37	39.24	39.76	185.36	188.95	31.37	31.73

Next, the objective performance of the proposed depth reconstruction filter was evaluated by the rate against PSNR curve of depth coding as well as the PSNR of the synthesised view with respect to the original view 4. Table 1 and Table 2 provide the relevant numbers and Fig. 3 and Fig. 4 show the corresponding rate-distortion (RD) curves for 'Ballet' and 'Breakdancers', respectively. It can be confirmed that the proposed depth reconstruction filter not only improves the rate-distortion performance of the depth coding as shown in Fig. 3a and Fig. 4a, but also significantly reduces the required depth bit rate for a given rendering quality level as shown in Fig. 3b and Fig. 4b.

Table 2: Experimental results for 'Breakdancers'

QP	Depth PSNR (dB)		Depth bit rate (kbit/s)		Rendering PSNR (dB)	
	original	proposed	original	proposed	original	proposed
QP22	50.59	51.33	1327.51	1315.28	32.03	32.11
QP27	46.82	47.54	748.97	747.45	31.92	32.06
QP32	42.83	43.23	378.96	382.34	31.74	31.92
QP37	39.40	39.48	181.87	185.11	31.50	31.68

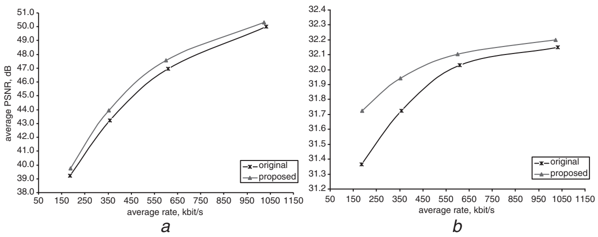

Fig. 3 *Rate-distortion curves for 'Ballet'*

a Depth rate and depth quality
b Depth rate and rendering quality

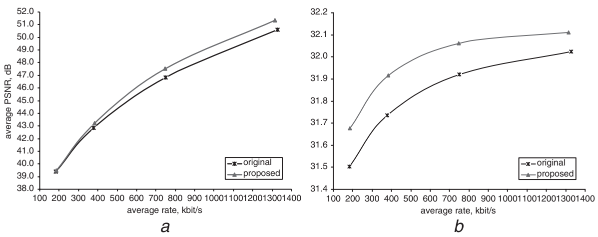

Fig. 4 *Rate-distortion curves for 'Breakdancers'*

a Depth rate and depth quality
b Depth rate and rendering quality

Conclusion: In this Letter we propose a depth reconstruction filter. The depth reconstruction filter consists of the frequent-close filter and the bilateral filter. The frequent-close filter is a nonlinear filter, designed with full consideration of the characteristics of depth images to reduce the coding errors, while the bilateral filter is adopted to eliminate the remaining outliers while preserving the object boundary. The experimental results demonstrate that the proposed scheme significantly improves the depth coding performance as well as the quality of the rendered views compared with the conventional scheme.

Acknowledgments: This work was supported in part by ITRC through RBRC at GIST and in part by MOE through the BK21 project.

5 November 2008
doi: 10.1049/el.2009.3182

K.-J. Oh and Y.-S. Ho (*School of Information and Mechatronics, Gwangju Institute of Science and Technology (GIST), 1 Oryong-dong Buk-gu, Gwangju, 500-712, Korea*)

E-mail: kjoh81@gist.ac.kr

S. Yea and A. Vetro (*Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139, USA*)

References

1 Shum, H.Y., Kang, S.B., and Chan, S.C.: 'Survey of Image-based representations and compression technique', *IEEE Trans. Circuits Syst. Video Technol.*, 2003, **13**, (11), pp. 1020–1037

2 Tanimoto, M.: 'Overview of free viewpoint television', *Signal Process. Image Commun.*, 2006, **21**, (6), pp. 454–461

3 Smolic, A., Kimata, H., and Vetro, A.: 'Development of MPEG standards for 3D and free viewpoint video', Optics East 2005: Communications, Multimedia & Display Technologies, 2005, **6014**, pp. 262–273