A Low Complexity Wavelet Based Depth Map Encoder for Low Bit Rate 3D Video Applications

Martínez-Rach, Miguel Onofre, De Silva, D. V. S. X., Fernando, W. A. C., Arachchi, Hemantha Kodikara, López Granado, Othoniel, and Malumbres, M.P.

Abstract —***Depth maps, which aid the virtual view generation process in Depth Image Based Rendering (DIBR) applications in 3D video, are generally piecewise smooth images with sharp depth discontinuities. Compression artifacts that occur along the depth discontinuities often lead to poor quality of generated virtual views. In line with the importance of depth discontinuities towards view generation, we propose a non-uniform quantization method for wavelet based encoding of depth maps. Average rendering quality improvements of up to 0.8dB could be achieved.***

I. Introduction

Depth Image Based Rendering (DIBR) utilizes a color texture image and a per-pixel depth map to generate virtual views representing stereoscopic views or intermediate viewpoints [1]. As the depth map can be encoded more efficiently than the independent encoding of two monoscopic video frames, it is bandwidth efficient to represent a 3D scene using color-plus depth representation.

Depth maps are not viewed by end users, but are used for virtual view generation. Therefore, it is important to compress depth maps in a way that it minimizes distortion in views rendered with them. In the recent past several compression methods of depth maps have been proposed. These methods can be broadly classified in to two categories. The first category proposes radical compression techniques such as platelet based coding [2], silhouette based coding [3] and 3D motion estimation based methods [4]. The other approach tries to optimize existing video codecs to encode depth maps, such as novel methods for mode selection [5] and reconstruction filter [6]. This paper presents a method that falls in to the second category, where the quantization step size is judicially selected. We experimentally identify the areas of the depth image that are less important towards virtual view generation process. Based on the experiments, a low complexity wavelet based encoding technique is proposed to adaptively compress the depth map based on prioritizing areas important for view generation process.

II. The Proposed Technique

DIBR algorithm mainly relies on vertical edges in the depth map or the depth discontinuities in the horizontal direction. Therefore, the quality of rendered virtual views is very sensitive to any errors along the vertical edges, and in contrast less sensitive to impairments on horizontal edges. As we propose the technique based on a wavelet encoder, in the following subsection we analyze the relative importance of individual sub bands towards the virtual view generation process. The wavelet transform performs a frequency and orientation decomposition of the image information which allows applying a non-uniform quantization based on the orientation of the respective sub bands.

A. Evaluation of relative importance of wavelet sub bands

For the following experiments a six level wavelet decomposition, as illustrated in Fig. 1, is considered. The vertical edges resides in the LH sub bands and horizontal edges reside in the HL sub bands.

Fig. 2 illustrates the Rate-Distortion (R/D) performance curves for the ballet sequence. For each scenario considered, a virtual view is generated with the uncompressed color image and its corresponding depth map that is compressed with one of the schemes. In the Uniform Quantization (UQ) scheme, given in a solid line in Fig. 2, all the sub bands are quantized with equal weight. The UQ scheme is compared with several Non Uniform Quantization (NUQ) schemes that set all the coefficients of selected sub-bands to be equal to zero. The NUQ-HL1 curve corresponds to a uniform quantization for all the wavelet sub bands except for the HL1 sub band whose coefficients have been set to zero (full quantized). In the same way the NUQ HH1 HL1 and the NUQ LH1 HH1 HL1 schemes set the relevant sub bands to zero and uniformly quantize all the remaining sub bands. Considering the R-D performance graphs, it is clear that vertical information is critical for the view generation process. Both horizontal and diagonal components can be discarded, for high compression, to reduce the required bit rate for the depth map.

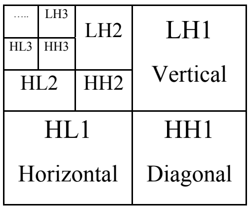

Fig. 1. Wavelet sub band decomposition schema

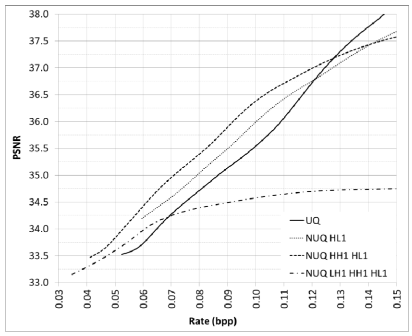

Fig. 2. R/D performance of the ballet sequence

B. Selection of Sub bands

Several experiments similar to the one described in section II.A are performed with different combinations of sub bands set to zero. Results indicate that the horizontal and diagonal sub bands of the first and second decomposition levels can be set to zero to effectively reduce the bit rate.

Therefore, we propose to set all the coefficients of HL1, HH1, HL2 and HH2 sub bands to zero, while the coefficients of all the remaining sub bands are given equal importance during quantization.

III. Experiments

The experiments are performed on a wavelet based image encoder [7] that operates to compute a dyadic six level wavelet decomposition of an image after which the wavelet coefficients are first quantized using a variable dead zone uniform quantizer. Finally, the zero-tree symbol map is encoded by means of an arithmetic coder. Simulations are performed on several test sequences and the results are presented for the 'Ballet', 'Breakdancers' (1024x768) from Microsoft Research and 'Interview' (720x576) sequences.

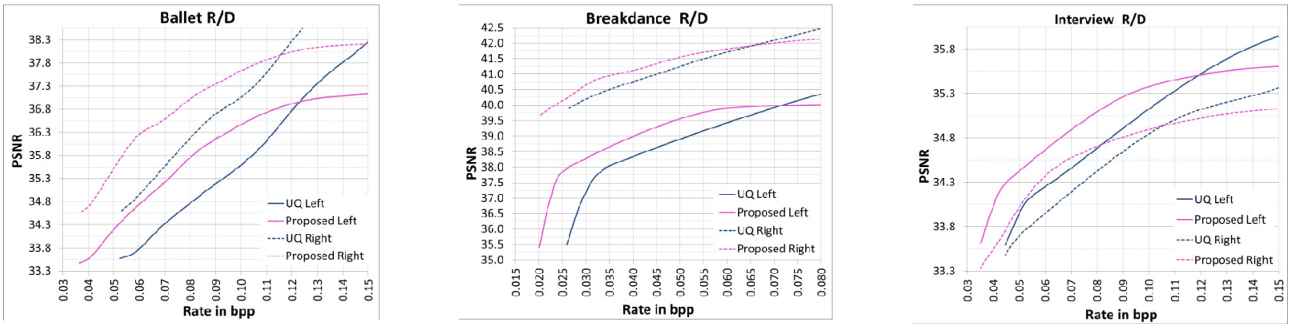

Fig. 3. Rate-Distortion performance graphs for three test sequences

Fig. 4. Ballet sequence: left view rendered with Uniformly Quantized depth map at 0.075 bpp, and right view with Proposed method.

Fig. 5. Interview sequence: left view rendered with Uniformly Quantized depth map at 0.070 bpp, and right view with Proposed method.

Virtual stereoscopic views (Left eye view and Right eye view) are generated according to MPEG informative recommendations [8] and the virtual stereoscopic camera positions are selected in a way that the epipolar lines are horizontal. The quality of the rendered views is measured relative to the views rendered with the original (uncompressed) color image and the corresponding depth map.

IV. RESULTS AND DISCUSSION

Fig. 3 illustrates the R/D performance curves for three of the sequences considered. The proposed non-uniform quantization technique is compared against uniform quantization technique that does not change the quantization step size based on the spatial variations in the depth map. Furthermore, we justify the use of the proposed technique by visual examples provided in figures 4 and 5.

It should be noted that for all the test sequences considered the proposed technique outperforms the reference technique by an average of 0.5dB gain in the quality of the rendered views. For the 'Ballet' sequence the gains are slightly higher than for both 'Breakdancers' and 'Interview' sequence. This difference in gains can be explained by the luminance/color variation in the sequences. In the 'Ballet' sequence the luminance variation of the corresponding color image is significantly higher than in both the other sequences. Therefore, even slight errors in the depth map will result in large PSNR losses in the rendered views. In contrast, in the 'Breakdancers' sequence, almost all the objects in the color image are of a dark luminance value. Therefore, even if a depth error occurs in the depth map, the PSNR is affected less.

The proposed method performed well for two other high definition sequences with disparity maps rather than depth maps. The discontinuities in these disparity maps are not as sharp as the depth maps. Nevertheless, for sequences 'Beergarden' and 'Badminton' average gains of 0.4dB and 0.1dB were obtained respectively.

Finally, we measured the coding delay of the overall process for both, the reference method and for the proposed method. In general, the proposed method is slightly faster than the uniform quantization method due to the fact of saving quantization and coding time of coefficients belonging to the subbands that are discarded.

V. CONCLUSION

A non-uniform quantization schema is proposed in this paper to improve the R/D performance of depth map coding at low bit-rates. The proposed technique based on a zerotree wavelet encoder discards coefficients in the horizontal and diagonal sub bands in the two high frequency decomposition levels, which are less relevant for the DIBR process. The remaining sub bands are given equal weighting during encoding. An average gain of 0.5 dB in the quality of rendered views is obtained for most of the sequences that are tested.

ACKNOWLEDGEMENTS

This work was funded by Spanish Ministry of Education and Science under grant DPI2007-66796-C03-03.

REFERENCES

[1] Christoph Fehn, "Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV," *Proc.of the SPIE*, vol. 93 & 5291, 2004.

[2] Y. Morvan, D. Farin, and P. H. N. de With, "Depth-Image Compression Based on an R-D Optimized Quadtree Decomposition for the Transmission of Multiview Images," in *IEEE International Conference on Image Processing*, 2007, vol. 5, p. V - 105-V - 108 *ICIP 2007*.

[3] S. Milani and G. Calvagno, "A Depth Image Coder Based on Progressive Silhouettes," *Signal Proc. Letters, IEEE*, vol. 17, no. 8, pp. 711-714, 2010.

[4] D. V. S. X. De Silva, W. A. C. Fernando, and S. L. P. Yasakethu, "Object based coding of the depth maps for 3D video coding," *Consumer Electronics, IEEE Transactions on*, vol. 55, no. 3, pp. 1699-1706, 2009.

[5] D. V. S. X. De Silva and W. A. C. Fernando, "Intra mode selection for depth map coding to minimize rendering distortions in 3D video," *Consumer Electronics, IEEE Transactions on*, vol. 55, no. 4, pp. 2385-2393, 2009.

[6] "Improved depth map filtering for 3D-TV systems," , in *IEEE International Conference on Consumer Electronics,* pp. 645-646, Jan. 2011.

[7] J. Oliver, M. P. Malumbres, "Fast and efficient spatial scalable image compression using wavelet lower trees," in Proc. IEEE Data Compression Conference, Snowbird, UT, March 2003

[8] ISO/IEC JTC 1/SC 29/WG 11, Doc. N8038. "Committee Draft of ISO/IEC 23002-3 Auxiliary Video Data Representations." Montreux, Switzerland, Apr. 2006.