Reliable depth-image-based rendering using parameter approximation in mobile devices

Cheolkon Jung^{a)} and Licheng Jiao

Key Lab of Intelligent Perception and Image Understanding of Ministry of Education of China, Institute of Intelligent Information Processing, Xidian University, Xi'an 710071, China

a) zhengzk@xidian.edu.cn

Abstract: This letter proposes a reliable depth-image-based rendering (DIBR) technique, which can provide users with reliable 3D services even if contents contain unknown or inappropriate values in some parameters. For the reliable DIBR, this letter presents how to set rendering parameters such as far/near clipping plane, convergence distance, and scale factor by taking mobile display size into consideration. Various experiments show that the proposed method can be effectively employed for 3D services based on DIBR over terrestrial-digital multimedia broadcasting (T-DMB).

Keywords: depth-image-based rendering (DIBR), reliable 3D services, parameter approximation, terrestrial-digital multimedia broadcasting (T-DMB)

Classification: Science and engineering for electronics

References

- [1] C. Fehn, "Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV," *Proc. SPIE Conf. Stereoscopic Displays and Virtual Reality Syst.*, pp. 93–104, 2004.
- [2] Y. K. Park, K. Jung, Y. Oh, S. Lee, J. K. Kim, H. Lee, K. Yun, N. Hur, and J. Kim, "Depth-image-based rendering for 3DTV service Over T-DMB," *Signal Processing: Image Communication*, vol. 24, pp. 122–136, 2009.
- [3] TTAS KO-07.0024: *Radio Broadcasting Systems, Specification of the Video Services for VHF Digital Multimedia Broadcasting (DMB) to Mobile*, Portable and Fixed Receivers, TTA, 2005.
- [4] L. Zhang and W. J. Tam, "Stereoscopic Image Generation Based on Depth Images for 3DTV," *IEEE Trans. Broadcast.*, vol. 51, pp. 191–199, 2005.
- [5] N. Hur, W. J. Tam, F. Speranza, C. Ahn, and S. I. Lee, "Depth-image-based Stereoscopic Image Rendering Considering IDCT and Anisotropic Diffusion," *Proc. IEEE Conf. Consumer Electron.*, pp. 381–382, 2005.
- [6] S. Choi, C. Jung, S. Lee, J. K. Kim, K. Jung, G. Lee, N. Hur, and J. Kim, "3D DMB Player and Its Realistic 3D Services over T-DMB," *Proc. IEEE Int. Symp. Multimedia*, pp. 440–441, 2009.

- [7] MSR 3D Video Download, Breakdancing and Ballet Sequences.
[Online] <http://research.microsoft.com/en-us/um/people/sbkang/3dvideodownload/>

1 Introduction

3D services over terrestrial-digital multimedia broadcasting (T-DMB) are very attractive because the single user environment of T-DMB is suitable for glassless 3D viewing with mobile displays. However, the T-DMB system has limitation of the bit budget for the transmission of additional video stream because of its limited bandwidth. Depth-image-based rendering (DIBR) can solve this problem because depth information can be compressed efficiently. The ATTEST project, which started in March 2002 as part of the European Information Society Technologies (IST), has proposed DIBR techniques because depth information as the additional data can be compressed efficiently below 64 Kbps [1, 2, 3].

In DIBR, left and right virtual views, which create auto-stereoscopic images, are rendered by reference image and its corresponding depth image in auto-stereoscopic displays [4, 5]. In the rendering procedure, exact parameters such as far/near clipping plane, convergence distance, and scale factor should be given for DIBR according to 3D contents and mobile displays. However, in a real environment, some parameters contain unknown values of actual contents when we create auto-stereoscopic images, because of the content provider's mistakes or transmission errors. Moreover, inappropriate values are given in some parameters, which make it extremely difficult to create auto-stereoscopic images in mobile displays. In this letter, we propose a rendering parameter approximation method which can create auto-stereoscopic images rapidly and provide users with reliable 3D services of various contents.

2 Proposed method

In DIBR, a depth distance, $Z(v)$, is calculated by a depth value of depth images, v , using z_{max} (far clipping plane) and z_{min} (near clipping plane) as follows [2].

$$Z(v) = \frac{1}{\frac{1}{z_{min}}(\frac{v}{255}) + \frac{1}{z_{max}}(1 - \frac{v}{255})}, \quad v \in [0, \dots, 255] \quad (1)$$

As shown in Eq. (1), two parameters, z_{max} and z_{min} , are requirements for getting the depth distance, $Z(v)$. However, most of actual contents do not provide these two parameter values. In this case, receivers should request transmitters to send them additionally, which can cause delays in rendering the contents. Moreover, if transmitters do not give the parameters, it is very difficult to generate auto-stereoscopic images in receivers. This problem prevents the popularization of 3D services based on DIBR over T-DMB.

Fig. 1. Geometry camera setup for generating two (left and right) virtual views in DIBR.

Therefore, we try to find solutions of the problem mentioned above, in order to be able to provide reliable 3D services. If we assume that z_{max} is 255 and z_{min} is 1, then Eq. (1) is calculated as follows.

$$Z(v) = \frac{255}{\frac{254}{255}v + 1}, \quad v \in [0, \dots, 255] \quad (2)$$

Since $254/255$ can be approximated to 1, we can express Eq. (2) as the following form.

$$Z(v) = \frac{255}{v + 1}, \quad v \in [0, \dots, 255] \quad (3)$$

By using Eq. (3), depth values can be converted into depth distances on contents which have unknown values of the two parameters, z_{max} and z_{min} . Also, we can make virtual left and right views which generate auto-stereoscopic images implementing Eq. (3). Fig. 1 shows a virtual camera setup for generating virtual views. The parameters f and t_c denote the focal length and baseline distance between virtual cameras C_l and C_r , respectively. From the camera setup, the pixel positions (x_c, y) , (x_l, y) , and (x_r, y) of the reference view, and the two virtual views corresponding to each point P with the depth distance Z have the following relationships:

$$x_l = x_c + \frac{\alpha_u t_c}{2Z(x_c, y)} - \frac{\alpha_u t_c}{2Z_c} \quad (4)$$

$$x_r = x_c - \frac{\alpha_u t_c}{2Z(x_c, y)} + \frac{\alpha_u t_c}{2Z_c} \quad (5)$$

where Z_c and u are the convergence distance and scale factor between the two virtual cameras, respectively. Here, the scale factor α_u is obtained by dividing the focal length f into pixel size.

It has been reported that users may feel much eyestrain if differences between x_l and x_r are more than 3% of the image width. Accordingly, the maximum differences are adjusted to meet 3% of the image width. Based on the aforementioned assumptions, α_u can be calculated as follows.

$$x_l - x_r = \frac{\alpha_u t_c}{Z(x_c, y)} - \frac{\alpha_u t_c}{Z_c} < 0.03 \times w \quad (6)$$

where w is the image width. We assume that $t_c = 60$ mm because general distances between left and right eyes are about 60 mm. In order to maximize the differences between x_l and x_r , we set $Z(x_c, y) = 1$ and $Z_c = 255$. Then, α_u is computed by Eq. (7).

$$\alpha_u = 0.0005 \times w \quad (7)$$

Auto-stereoscopic images can be created reliably using Eqs. (4)-(7), even if contents contain unknown or inappropriate parameter values. Therefore, it is possible to provide users with reliable 3D services. In addition, this method can also reduce the computational cost of the rendering procedure in receivers because the depth distances are computed by Eq. (3). It is very effective for receivers with limited computational power and memory.

3 Experimental results

The experiments were performed by using the UMPC (SONY VAIO VGN-UX17LP, CPU: Intel CoreTM Solo Processor U1400) with VC++6.0 in [6]. To evaluate the effectiveness of the proposed method, five sequences, each consisting of the reference sequence and the corresponding depth sequence, were used for the tests. They are ‘Interview’ and ‘Orbi’ from Heinrich-Hertz-Institute (HHI) [1], ‘Ballet’ and ‘Breakdancer’ from Microsoft Research [7], and ‘ETRI.CG’ from Electronics and Telecommunications Research Institute (ETRI) [2], respectively. All sequences for the tests were adjusted to the size of 320×240 pixels. In the experiments, a depth distance $Z(v)$ corresponding to a depth value v is computed by Eqs. (1) and (3), respectively. The distribution of $Z(v)$ by Eq. (3) is nearly equivalent to that by Eq. (1). By Eq. (6), the maximum differences between x_l and x_r are about 10, and this means that users would feel much eyestrain if the differences between x_l and x_r are more than 10 pixels. Since the image width w is 320, the scale factor α_u is 0.16 by Eq. (7).

Fig. 2 shows two auto-stereoscopic images of the ‘Ballet’ sequence each created by Eqs. (1) and (3). It can be observed that the proposed method can create nearly identical auto-stereoscopic images such as the result obtained by Eq. (1). Therefore, reliable 3D services based on DIBR can be provided by the proposed method even if contents do not contain some of parameter values or contain inappropriate parameter values.

Moreover, the proposed method has the merit of reducing computational cost because of its approximated computation. To verify the effectiveness of the computational cost, we measured time it took to create the auto-stereoscopic images using separately Eqs. (1) and (3). As shown in Table I,

(a)

(b)

Fig. 2. Auto-stereoscopic images of the ‘Ballet’ sequence separately created by Eqs. (1) and (3). (a) Auto-stereoscopic image created by Eq. (1). (b) Auto-stereoscopic image created by Eq. (3).

Table I. Comparison of average processing time for creating auto-stereoscopic images by Eqs. (1) and (3).

Sequence	Eq. (1)	Eq. (3)
Interview	0.073	0.053
Orbi	0.071	0.053
Ballet	0.072	0.054
Breakdancer	0.072	0.053
ETRL_CG	0.074	0.053
Average	0.072	0.053

Experiments were performed by using the UMPC (SONY VAIO VGN-UX17LP, CPU: Intel CoreTM Solo Processor U1400) with Microsoft Visual C++ Version 6.0 using five sequences in [6]. The unit of this test is sec per image ($s/image$).

the average processing time of the proposed method is 0.053sec per image ($s/image$). Such results indicate that the proposed method reduces 26.4% (0.019 $s/image$) of the processing time as compared to Eq. (1). This means that our proposed method is very effective in terms of computational cost for creating auto-stereoscopic images.

4 Conclusion

We have proposed a rendering parameter approximation method to create auto-stereoscopic images reliably even if contents do not contain some parameter values, or if they contain inappropriate parameter values for DIBR. The proposed method reduces computational cost because of approximated computation as well. We believe that the proposed method can provide users with reliable and effective 3D services based on DIBR over T-DMB.

Acknowledgments

The partial work reported in this manuscript was done while the first author was with the Sungkyunkwan University in South Korea. The authors are grateful to Mr. Youngjin Oh and Prof. Joongkyu Kim for their helpful discussions and the anonymous reviewers for their useful comments. The work was completed during the study period of the Young Scientist Exchange Program between China and Korea, and we would also like to thank the China Postdoctoral Science Foundation and the National Research Foundation of Korea for their financial support.