# A FOREGROUND BIASED DEPTH MAP REFINEMENT METHOD FOR DIBR VIEW SYNTHESIS

*Xuyuan Xu, Lai-Man Po, Kwok-Wai Cheung[#], Ka-Ho Ng, Ka-Man Wong, Chi-Wang Ting*

Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR, China
[#]Department of Computer Science, Chu Hai College of Higher Education, Hong Kong SAR, China

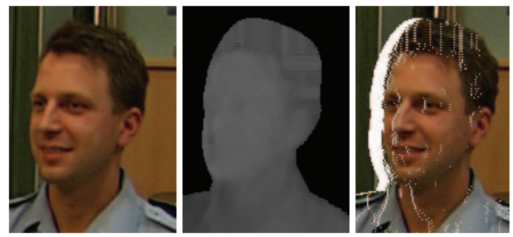
## ABSTRACT

The performance of view synthesis using depth image based rendering (DIBR) highly depends on the accuracy of depth map. Inaccurate boundary alignment between texture image and depth map especially for large depth discontinuities always cause annoying artifacts in disocclusion regions of the synthesized view. Pre-filtering approach and reliability-based approach have been proposed to tackle this problem. However, pre-filtering approach blurs the depth map with drawback of degradation of the depth map and may also cause distortion in non-hole region. Reliability-based approach uses reliable warping information from other views to fill up holes and is not suitable for the view synthesis with single texture video such as video-plus-depth based DIBR applications. This paper presents a simple and efficient depth map preprocessing method with use of texture edge information to refine depth pixels around the large depth discontinuities. The refined depth map can make the whole texture edge pixels assigned with foreground depth values. It can significantly improve the quality of the synthesized view by avoiding incorrect use of foreground texture information in hole filling. The experimental results show the proposed method achieves superior performance for view synthesis by DIBR especially for large baseline.

***Index Terms:*** View Synthesis, DIBR, Depth Map Refinement.

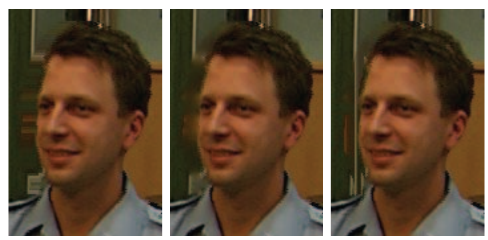
## 1. INTRODUCTION

Video-plus-depth format is a simple and efficient way to represent 3D video. It consists of 2D color texture video and depth map with per pixel depth information and is a very compact format especially suitable for mobile 3D video applications. It has high feasibility to render views with variable baseline by Depth Image Based Rendering (DIBR) [1]. Thus, stereo video and multiview video can be reconstructed on stereoscopic or auto-stereoscopic 3D display devices, respectively [2-4], based on DIBR. To synthesize new views using DIBR involves three major steps: Depth Map Preprocessing, 3D Image Warping, and Hole Filling. One of the major challenges to synthesize high quality virtual views is to reconstruct the large disoccluded areas after the 3D image warping process. The disoccluded regions after the warping process are called holes. They do not exist in the 2D texture image but are exposed in the synthesized view. The well-known 3D test image "Interview" with its texture and depth images are shown in Fig. 1(a) and 1(b), respectively. The holes in the 3D warped image are shown in Fig. 1(c) with white pixels. These disoccluded regions are usually the background regions. The most straightforward approach is to fill up these areas by linear interpolation [5] and depth-aid horizontal extrapolation [5] methods. Fig. 2(a) shows the synthesized view using depth-aid horizontal extrapolation hole filling method, in which artifacts can be easily observed in the large disocclused regions. Advanced hole filling approaches are multidirectional extrapolation [6] and image inpainting [7]. They analyze the surrounding texture information and use them to fill the holes in the synthesized views. Fig. 2(b) shows the view using advanced inpainting hole filling method, in which annoying artifacts still exist. The main reason is that the disocclusion regions normally involve large depth discontinuities. However, the hole filling techniques which only consider the planar image information cannot solve the problem.



(a) (b) (c)

**Fig. 1: DIBR based view synthesis for "Interview" image in video-plus-depth format. (a) Original 2D texture image, (b) Original Depth Map, and (c) Synthesized Left-View Image after 3D Image Warping.**

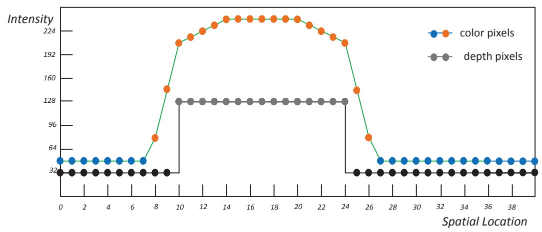


(a) (b) (c)

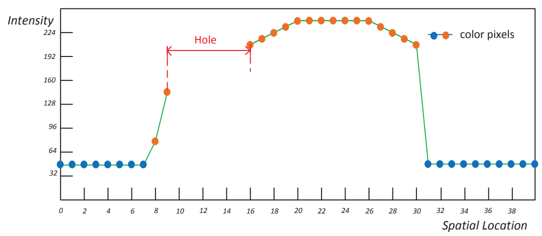
**Fig. 2: Synthesized Views using different hole filling methods. (a) Depth-aid horizontal extrapolation, (b) Image Inpainting, and (c) Multidirectional extrapolation.**

These annoying artifacts are mainly due to low depth map quality associated with incorrect depth values, especially for texture edge pixels that include foreground and background color pixels. In addition, object edges may be fuzzy and contain transitional edge pixels. As a result, unprocessed depth map usually cause artifacts after the hole filling process. These artifacts are mainly due to the fact that transitional edge pixels are mapped to background regions in the image warping process and these pixels' information are then used to fill up the holes. To tackle this problem, one approach is to use the smoothing filters such as average filtering, Gaussian filter [2], asymmetric filter [2] and adaptive filter [8] to blur the boundaries of depth map in order to eliminate holes or reduce the sizes of the large holes. The artifacts created in the hole filling process can be minimized, but the depth map may be highly degraded. The highly degraded depth map should correspond to poverty of 3D perception of the synthesized view. In this paper, a simple depth map preprocessing method with depth edge refinement is proposed for DIBR based view synthesis. The depth edges are refined with simple criteria to ensure that the whole texture edges are assigned as foreground depth values. Therefore, only background information is used in hole filling process. As a result, the depth map quality is not significantly degraded while the 3D perception quality of the synthesized view is significantly improved.

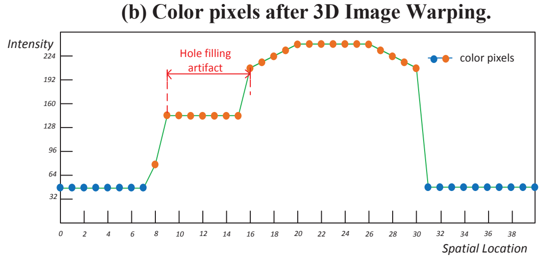
This paper is organized as follows. Section 2 describes how the artifacts are created in the DIBR view synthesis process. The details of the proposed depth map refinement method are presented in section 3. The experimental results with comparison with conventional preprocessing methods are provided in section 4. Finally, a conclusion is drawn in section 5.




**(a) Color pixel intensity values and depth values for a horizontal line in video-plus-depth image format.**




**(b) Color pixels after 3D Image Warping.**

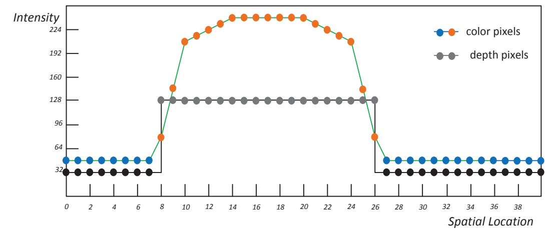



**(c) Effect of hole filling from neighbor pixels using depth-aid horizontal extrapolation hole fill method with annoying hole filling artifact.**

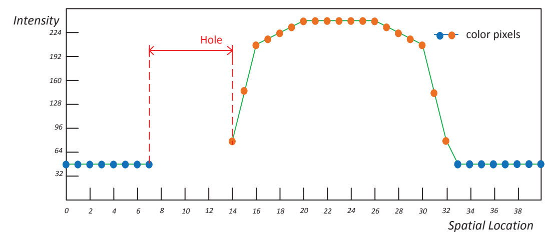
**Fig. 3: Depth map boundaries misalignment problem.**

## 2. DEPTH MAP BOUNDARIES MISALLGNMENT

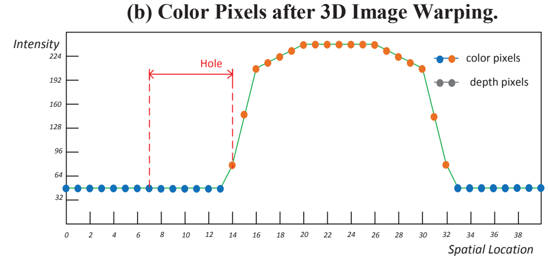
Annoying hole-filling artifacts usually appear in large disocclusion regions of the synthesized views when the holes surrounding pixel information is used for the hole filling. This is mainly due to the fact that object boundaries contain a combination of foreground and background color information. Incorrect depth values may assign to these edge pixels and some of the edge transitional pixels with similar foreground colors are treated as the background pixels. In the hole filling process, these background pixels are used to fill up the hole regions and it may result in the annoying artifacts. To illustrate this phenomenon, Fig. 3(a) shows color pixel intensive values and depth pixel values of a horizontal line in a texture image. There are two transitional edges of an object with sharp depth edges in the corresponding depth map. After the 3D image warping process, the pixel line of the synthesized left view is shown in Fig. 3(b). The background pixels and part of the transitional edge pixels are shifted to left and a large hole is created. In DIBR based view synthesis using one texture image, this hole is filled with the neighbor background pixels and Fig. 3(c) shows the effect of filling up this hole using depth-aid horizontal extrapolation method. Then, the hole is filled with a color of the transitional edge pixel, which creates the annoying hole filling artifact similar to Fig. 2(a).




**(a) Color pixel intensity values and refined depth values for a horizontal line in video-plus-depth image format.**




**(b) Color Pixels after 3D Image Warping.**




**(c) Effect of hole filling from neighbor pixels using depth-aid horizontal extrapolation method without annoying hole filling artifact.**

**Fig. 4: Foreground biased depth map refinement process.**

## 3. NEW DEPTH REFINMENT METHOD

Based on the observations in section 2, if the depth map can be refined in the preprocessing stage with the foreground region to cover the whole transitional region of texture edges, the annoying hole filling artifacts should be significantly minimized in the synthesized views. The main idea is to refine the depth values of these transitional edge pixels in order to make them become foreground pixels as shown in Fig. 4(a). After 3D warping, the whole texture edges are mapped to the foreground region as shown in Fig. 4(b). The artifacts is significantly minimized after hole filling as shown in Fig. 4(c).

To achieve this purpose, we only need to refine the depth values of transitional edge pixels with large depth discontinuity. Although the boundary artifacts appear around object boundaries, gradually changing depth values does not generate annoying artifacts since small depth discontinuities create only very small holes in the warped image. Artifacts usually exist in the large holes. In the proposed preprocessing method, we derive depth discontinuity thresholds based on the hole's size due to the horizontal depth values difference for triggering the refinement process. The relationship between hole's size and depth values difference between two horizontal adjacent pixels based on shift-sensor model for DIBR can be devised as

$$\Delta d = \frac{h}{t_c f} \cdot \frac{1}{\left(\frac{1}{255 z_n} - \frac{1}{255 z_f}\right)} \qquad (1)$$

where $\Delta d$ is the depth values difference between the horizontal adjacent depth pixels, $t_c$ and $f$ are the baseline distance and the focal length, respectively. $z_n$ and $z_f$ represent the nearest distance and the farthest distance in the scene. In our experiments, hole's sizes greater or equal to 3 ($h \geq 3$) are classified as large holes. Thus, the pre-defined depth discontinuity threshold $T_d$ is determined based on Eq. (1) with $h = 3$. For any absolute depth values difference larger than $T_d$, the hole's size in warped image will be larger than 3 pixels and the proposed foreground biased depth refinement will be performed around neighborhood's depth pixels.

### 3.1 Foreground Biased Depth Value Refinement

The proposed refinement method is a line-by-line process aiming at extending the foreground depth values to cover the whole transitional region of texture edges. The idea is very simple. A refinement process will be applied when the horizontal depth values change from low to high and larger than the pre-defined depth threshold $T_d$ ($d_i - d_{i+1} < -T_d$) similar to the sharp depth edge on the left side of Fig. 3(a). The proposed refinement process will shift the foreground depth value to left side (setting $d_i = d_{i+1}$) if the horizontal gradient of the texture edge is greater than a pre-defined gradient threshold $G_h$. This step is repeated until the texture edge gradient is not greater than the threshold or the shifting is larger than a pre-defined window size $W$. In the proposed method, the horizontal gradient is defined based on the horizontal derivatives of Prewitt operator and the window size $W$ is set to 5. When the sharp depth edge is on the left side, the depth values will be shifted to the left by two pixels with the resulting depth values as shown in Fig. 4(a).

The refinement process will also be applied when the horizontal depth values change from high to low and larger than the pre-defined depth threshold $T_d$ ($d_i - d_{i+1} > T_d$) similar to the sharp depth edge on the right side of Fig. 3(a). The proposed refinement process will shift the foreground depth values to right size (setting $d_{i+1} = d_i$) if the horizontal gradient of the texture edge is greater than the gradient threshold $G_h$. When the sharp depth edge is on the right, the depth values will be shifted to the right by two pixels. Sample resulting depth values are shown in Fig. 4(a). The two transitional texture edges are assigned with the foreground depth values. Then, the artifacts are significantly reduced due to the use of background pixels for hole filling.

To further simplify the depth refinement process, only low to high or high to low refinement process is applied for synthesizing the virtual left or right view in the DIBR process. The proposed method can be extended to the general case of 3D image warping process with a small modification. We can replace the method of finding large hole of comparing depth difference with threshold by checking the depth values of neighbor pixels that create larger hole in the warping process. Thus, the proposed method can be easily integrated into the DIBR based 3D image/video systems.

## 4. EXPERIMENTAL RESULTS

The proposed method is compared with the pre-filtering approaches using adaptive filter and asymmetric Gaussian filter. The "Interview" sequence of video-plus-depth format is used as subjective evaluation with conventional DIBR rendering. The test sequences "Lovebird1" and "Breakdancers" of multi-view format are used for objective evaluations. VSRS is used to generate the synthesized view from one single input view and compare the PSNR of synthesized view and existing view.

From the view synthesis results of depth map without preprocessing for "Interview" as shown in Fig. 1, the artifact occurs after hole filling using depth aid horizontal extrapolation, multidirectional extrapolation and inpainting methods as shown in Fig. 2(a), Fig. 2(b) and Fig. 2(c), respectively. These artifacts are caused by the edge of color in Fig. 1(c) containing the foreground pixels and having the incorrect background depth values. The hole filling method uses these incorrect texture pixels and created these artifacts.

(a) (b) (c) (d)

**Fig. 5: Synthesized view using adaptive filter in depth map preprocessing: (a) Preprocessed Depth map by adaptive filter, (b) Warpped Image, and (c) Synthesized Image using Inpainting hole filling, (d) Synthesized Image using multidirectional extrapolation.**

(a) (b) (c) (d)

**Fig. 6: Synthesized view using asymmetric Gaussian filter in depth map preprocessing: (a) Preprocessed Depth map by adaptive filter, (b) Warpped Image, (c) Synthesized Image using Inpainting hole filling, and (d) Synthesized Image using multidirectional extrapolation.**

(a) (b) (c) (d)

**Fig. 7: Synthesized view using proposed method: (a) Preprocessed Depth map by the proposed method, (b) Warpped Image, (c) Synthesized Image using Inpainting hole filling, and (d) Synthesized Image using multidirectional extrapolation.**

In Fig. 5(a) and Fig. 6(a), the depth maps are preprocessed with the adaptive filter and asymmetric filter. Both filters blur the transitional edge of depth map to match the smooth transitional edge in the texture image. In the smoothing process, hole sizes are reduced while artifacts surrounding the big hole are minimized.

However, some annoying artifacts still exist in Fig. 5(c), Fig. 5(d), Fig. 6(c) and Fig. 6(d) at the left side of the man's face. The blurring depth map degrades the depth map quality and sometimes destroys the texture in non-hole regions. In addition, the adaptive filter has less smoothing effect on the depth transitional region and preserves the sharp transition of depth edge. The depth map quality after adaptive filtering is better than the results of using asymmetric Gaussian filter and a better synthesis view obtained. . The results of using the proposed foreground biased refinement method are shown in Fig. 7, in which the annoying artifacts are significantly minimized as shown in Fig. 7(c) and Fig. 7(d).

(a) (b) (c) (d)

**Fig. 8: Synthesized view under difference situations: (a) Un-preprocess depth map (b) Adaptive Filter (c) Asymmetric Gaussian (c) Proposed Method.**

| Hole filling \ pre-filters | Without preprocess | Adaptive filter | Asymmetric Gaussian filter | Proposed method |
|---|---|---|---|---|
| Inpainting | 30.0766 | 30.0940 | 29.8620 | **30.2151** |
| Common-Hole (GIST) | 30.0764 | 30.0686 | 29.8389 | **30.1706** |
| Extra Inpainting | 30.0647 | 30.0585 | 29.8044 | **30.1576** |
| Spiral exemplar inpaing | 30.0591 | 30.0865 | 29.8556 | **30.2047** |
| Spiral weight average for extra | 30.0771 | 30.0945 | 29.8624 | **30.2138** |

**(a) Results of lovebird1 sequence**

| Hole filling \ pre-filters | Without preprocess | Adaptive filter | Asymmetric Gaussian filter | Proposed method |
|---|---|---|---|---|
| Inpainting | 32.4938 | 32.6564 | 31.1862 | **32.7082** |
| Common-Hole (GIST) | 32.9204 | 32.9517 | 31.1183 | **32.9915** |
| Extra Inpainting | 31.9722 | 31.8666 | 30.4965 | **31.9729** |
| Spiral exemplar inpaing | 33.5970 | 33.7293 | 31.8255 | **33.7649** |
| Spiral weight average for extra | 33.4658 | 33.7123 | 31.8879 | **33.8273** |

**(b) Results of breakdancers sequence**

**Table 1: PSNR measurement of synthesized view and existing view under four situations: (a) Without preprocess depth map (b) Adaptive Filter (c) Asymmetric Gaussian (c) Proposed Method**

Results on using "Lovebird1" sequence for generating synthesized view images are shown in Fig. 8. By performing the foreground biased depth map refinement, the transitional edge pixels containing the foreground samples are classified as foreground. The background pixels are used for the prediction in the hole filling process. Boundary artifacts are significantly reduced. Since the proposed refinement does not involve with smooth filter, the sharp edges in depth map are maintained. It avoids texture distortion in non-hole region and retains high 3D perception of the synthesized views. For objective evaluation, "Lovebird1" sequence with 200 frames of view 6 and "Breakdancers" sequence with 100 frames of view 0 are used to generate the synthesized view 7 and synthesized view 1, respectively. The proposed method is compared with the pre-filtering approaches. Results are shown in Table 1 based on four different cases: (1) Input view to generate synthesized view; (2) Input view using adaptive depth map pre-filter to generate synthesized view; (3) Input view using asymmetric Gaussian filter to generate synthesized view; and (4) Input view using the proposed depth map refinement method to generate synthesized view. Five different hole filling methods of VSRS3.5 alpha are used for evaluations: (1) Inpainting [9]; (2) Common-Hole (GIST) [10]; (3) Extra Inpainting [10]; (4) Spiral exemplar inpainting [11]; and (5) Spiral weight average for extra [11]. The generated virtual view is compared with the original view based on the PSNR. From the result of Table 1, the average PSNR improvement is around 0.1 dB compared with case (1) and case (2) in both Lovebird1 and Breakdancer sequence. While compared with case (3), the average PSNR improvements are 0.31dB and 1.75dB in Lovebird1 and Breakdancer, respectively. After the proposed refinement method, image quality is improved for all hole filling methods. The refinement of depth map can reduce the annoying boundary artifacts and increases the prediction accuracy of hole filling methods to fill the disocclusion by using the background information.

## 5. CONCLUSION

The video-plus-depth representation of 3D format usually has the problem of boundary artifact in the rendering process caused by the inaccurate boundary matching between texture image and depth map. Preprocessing approach uses the blur filter to smooth the boundary region. It downgrades the 3D quality of depth map and sometimes destroys the texture mapping of background. The proposed method refines the depth value in the edge transitional region. Based on the texture edge information, the whole region of transitional edge is classified as foreground region by refining their depth values. Experimental results show that the proposed foreground biased depth map refinement method can significantly improve the rendered image quality. Most annoying artifacts are removed and the texture distortion can also be avoided without using any smooth technique compared with preprocessing approach.

## 6. REFERENCES


[1] C. Fehn, 2004. "Depth-Image-Based Rendering (DIBR), Compression, and Transmission for a New Approach on 3D-TV," SPIE, 5291, pp. 93-104, 2004.

[2] L. Zhang and W.J. Tam, "Stereoscopic Image Generation Based on Depth Images for 3D TV," IEEE Trans. On Broadcasting, vol. 51, No. 2, pp. 1191-199, June 2005.

[3] K. Müller, P. Merkle, T. Wiegand, "3-D Video Representation Using Depth Maps," Proceedings of the IEEE , vol.99, no.4, pp.643-656, April 2011

[4] K. Müller, A. Smolic, K. Dix, P. Merkle, P. Kauff, and T. Wiegand, "View Synthesis for Advanced 3D Video Systems," EURASIP Journal on Image and Video Processing, vol. 2008, Article ID 438148, 11 pages, 2008.

[5] C. Vázquez, W. J. Tam and F. Speranza, "Stereoscopic Imageing: Filling Disoccluded Areas in Depth Image-Based Rendering", Proc. Of SPIE, Vol. 6392, 2006.

[6] L.M. Po, S. Zhang, X. Xu, Y. Zhu, "A new multidirectional extrapolation hole-filling method for depth-image-based rendering" Processing of ICIP2011, pp. 2637-2640, Sep. 2011.

[7] M. M. Oliveira, B. Bowen, R. McKenna, and Y.S. Chang , "Fast Digital Image Inpainting", Proc. Inter. Conf. Visualization, Image and Image Processing, 2001, pp. 261-266

[8] Y. K. Park, K. Jung and ect, "Depth-Image Base Rendering for 3DTV Service over T-DMB", Signal Process. Image Commun, voi 24(1-2), pp. 122-136, Jan 2009.

[9] A. Telea, "An image inpainting technique based on the fast marching method", J. Graphics Tools, vol.9, no.1, pp.25–36, 2004.

[10] ISO/IEC JTC1/SC29/WG11 "Implementation of Hole Filling Methods for VSRS 3.5.alpha," m20005, Mar 2011.

[11] ISO/IEC JTC1/SC29/WG11 "Boundary noise removal and hole filling for VSRS3.5 alpha," M19992, Mar 2011