

A divide-and-conquer hole-filling method for handling disocclusion in single-view rendering

Jianjun Lei¹ · Cuicui Zhang¹ · Min Wu¹ · Lei You¹ ·
Kefeng Fan² · Chunping Hou¹

Received: 8 April 2015 / Revised: 31 January 2016 / Accepted: 29 February 2016 /

Published online: 14 March 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Large holes are unavoidably generated in depth image based rendering (DIBR) using a single color image and its associated depth map. Such holes are mainly caused by disocclusion, which occurs around the sharp depth discontinuities in the depth map. We propose a divide-and-conquer hole-filling method which refines the background depth pixels around the sharp depth discontinuities to address the disocclusion problem. Firstly, the disocclusion region is detected according to the degree of depth discontinuity, and the target area is marked as a binary mask. Then, the depth pixels located in the target area are modified by a linear interpolation process, whose pixel values decrease from the foreground depth value to the background depth value. Finally, in order to remove the isolated depth pixels, median filtering is adopted to refine the depth map. In these ways, disocclusion regions in the synthesized view are divided into several small holes after DIBR, and are easily filled by image inpainting. Experimental results demonstrate that the proposed method can effectively improve the quality of the synthesized view subjectively and objectively.

Keywords Single-view rendering · Disocclusion handling · Hole-filling · Depth map · 3D · Stereoscopic image processing

1 Introduction

Depth image based rendering (DIBR) is one of the most important technologies to synthesize different virtual views, using color images and their associated depth maps [3, 5, 15,

✉ Min Wu
wumin_tju@tju.edu.cn

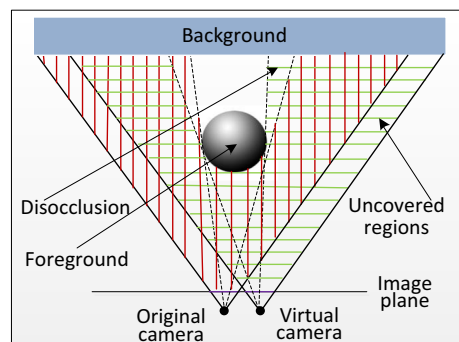
¹ School of Electronic Information Engineering, Tianjin University, Tianjin, China

² China Electronics Standardization Institute, Beijing, China

16, 31, 32]. However, some regions occluded by the foreground (FG) objects in the original views may become visible in the synthesized views, which are called disocclusion [2, 11, 12, 18, 28]. Figure 1 shows the occurrence of disocclusion in single-view rendering. Original camera represents the captured reference viewpoint, and virtual camera represents the virtual view synthesized by DIBR. The visible areas in the reference view are marked with vertical lines in red, while those visible areas in the virtual view are marked with horizontal lines in green. The disocclusion region, which is not visible in the reference viewpoint but visible in virtual image, is located at the boundary of the foreground and background as pointed out in Fig. 1. Disocclusion problem is particularly serious in single-view-plus-depth format. Therefore, the task of disocclusion-filling becomes more crucial in single-view rendering.

Several methods for solving this problem have been proposed, which can be categorized into two classes: (1) depth map preprocessing before rendering [10, 13, 14, 29, 30] and (2) image inpainting jointing depth and texture information after rendering [6, 17, 19, 21, 24]. In this paper, we focus on the former class, which mainly consists of depth filtering and depth boundary refinement. The main idea of depth filtering is to reduce the difference of depth values on the object boundary. Lee et al. [14] adopted adaptive smoothing filters which including an asymmetric smoothing filter and a horizontal smoothing filter for depth map preprocessing. The asymmetric smoothing filter was adopted to reduce the geometric distortions and the horizontal smoothing filter was adopted to reduce the time of computation and occurrence of holes. In this way, disocclusion region in the synthesized view becomes smaller. However, depth filtering usually introduces rubber-sheet artifacts and geometric distortions. Another method applied is the dilation of the FG objects of the depth map [13, 29, 30]. In [29], a dilation based method for view synthesis by DIBR was proposed. The depth pixels which located in the large depth discontinuities are refined with the neighboring foreground depth values, and thus improve the quality of the rendered virtual view. In [30], Xu et al. extended the method in [29] with a misalignment correction of the depth map for a better view synthesis result. Koppel et al. [13] proposed a depth preprocessing method which including the depth map dilation, adaptive cross-trilateral median filtering, and adaptive temporally consistent asymmetric smoothing filtering to reduce holes and distortions in the virtual view. Under the dilation based method, the FG objects in the depth map are slightly bigger than the corresponding objects in the textured image, but the disocclusion region is not reduced. When the big holes are filled through image inpainting, the diffusion region may be blurred.

Fig. 1 Occurrence of disocclusion in single-view rendering



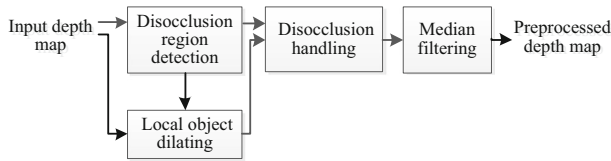


Fig. 2 Block scheme of the proposed method

In addition, the global dilating may causes distortions, as the boundary disparity information is changed in the common area. The disocclusion is mainly located between the FG and the background (BG). Furthermore, the size of the disocclusion region is quite large. Such challenging conditions prevent the use of traditional image inpainting methods and make such holes difficult to be filled.

In this paper, in order to resolve the above problems, we propose a divide-and-conquer hole-filling method to effectively solve the disocclusion problem. Due to the fact that reducing the sharp depth discontinuities can narrow the disocclusion region, we propose to modify the depth pixels around the depth discontinuities by dilating associated with a linear interpolation process. In particular, to avoid the distortion on other FG objects, the interpolation is limited to the neighboring BG. Finally, median filtering is adopted to remove isolated depth pixels. Experimental results demonstrate that our approach is effective in recovering the large disocclusion region.

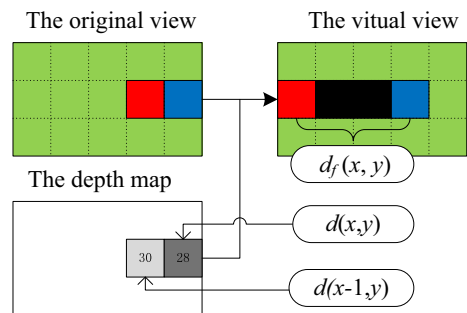
2 Proposed method

Figure 2 shows the block scheme of the proposed method, which includes disocclusion region detection, local object dilating, disocclusion handling, and median filtering. With the preparation of disocclusion region detection and local object dilating, the divide-and-conquer strategy is mainly reflected in the disocclusion handling section.

For the right synthesized view, disocclusion occurs on the depth map with high to low sharp depth transition areas, and vice versa for the left synthesized view as described in [14]. In this paper, we take the rendering of the right virtual view for example. The same process can be applied on the rendering of the left virtual view. The depth difference between two horizontal adjacent pixels can be expressed as:

$$d_f(x, y) = d(x - 1, y) - d(x, y) \tag{1}$$

Fig. 3 The relationship between the virtual view and the original view



where $d(x, y)$ represents the depth pixel value at position (x, y) , and $d_f(x, y)$ represents the horizontal depth difference between the depth pixels at positions (x, y) and $(x - 1, y)$. If $d_f(x, y)$ is larger than the specified threshold T_0 , there is a large depth discontinuity at position (x, y) .

Based on the above concept, the disocclusion regions are detected according to the degree of depth discontinuity and are labeled as $b(x, y) = 1$ as follows.

$$b(x, y) = \begin{cases} 1, & \text{if } d_f(x, y) > T_0 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

Then the target region for refinement is restricted to the neighboring BG and is marked as a binary mask as:

$$mask(x_0 + k, y) = \begin{cases} 1 & \text{if } b(x_0, y) = 1 \&\& \\ & d(x_0 + k, y) - d(x_0, y) < T_0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

$k \in \{0, 1, 2, \dots, 2 * d_f(x_0, y) - 1\}$

where x_0 represents the x-coordinate of a pixel where $b(x_0, y) = 1$. The condition $d(x_0 + k, y) - d(x_0, y) < T_0$ is used to avoid that other FG objects be marked by the binary mask.

Figure 3 shows the relationship between the virtual view and the original view. Assume that $d(x - 1, y)$ is 30 and $d(x, y)$ is 28, the depth difference between $d(x - 1, y)$ and $d(x, y)$ is equal to 2 (as shown in Fig. 3, the adjacent pixels are marked in red and blue). After rendering, the hole region is appeared in the virtual view (the black pixels between the red pixel and blue pixel). There are two hole pixels in the virtual view between the adjacent pixels in original view which are marked in red and blue, respectively. It is worthwhile to mention that the pixel value of depth map in Middlebury is equals to the value of disparity, and the parallel rendering is adopted in our paper.

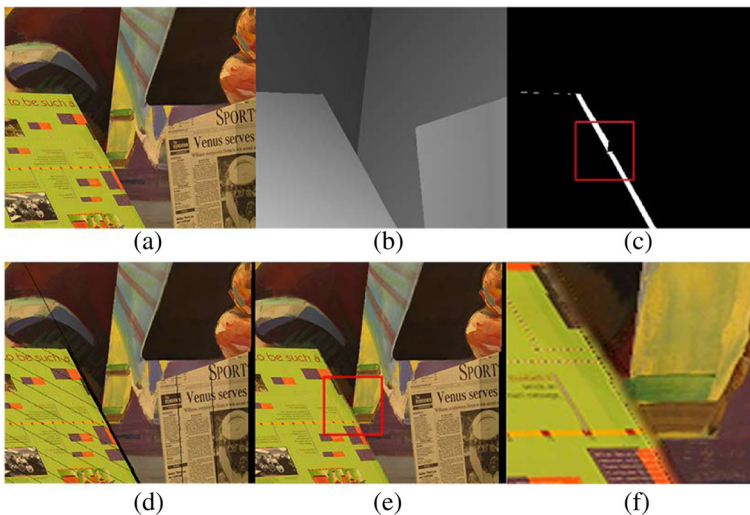


Fig. 4 Example of the disocclusion region detection. (a) The reference color image, (b) the corresponding depth map, (c) half of the binary mask, (d) the right virtual image generated by (a) and (b), (e) the hole-filling result without preprocessing, and (f) enlarging the red mark of (e)

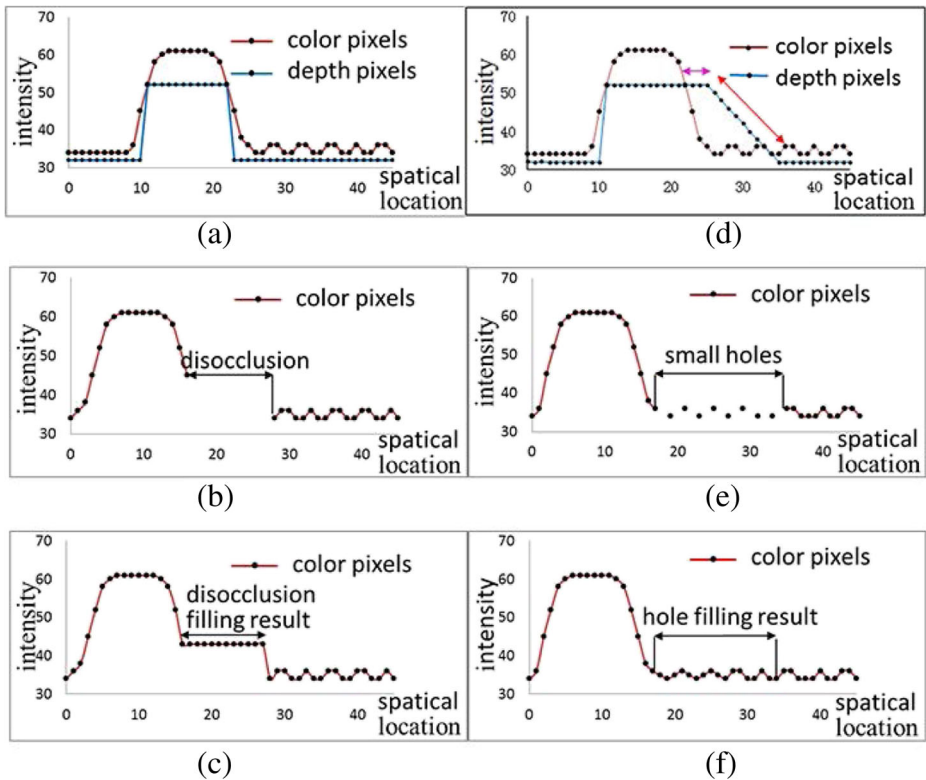


Fig. 5 Example for clarification of the divide-and-conquer disocclusion handling method. (a) Color pixel values and the corresponding depth values for a horizontal line, (b) color pixels after rendering by input depth map, (c) hole-filling result from (b), (d) our depth map preprocessing result, (e) virtual color pixels after rendering using (d), and (f) hole-filling result from (e)

A more generalize description is shown as follows. As the similar with [11], assume that the x -coordinates of two adjacent pixels in the original view are $x-1$ and x , the x -coordinates of the two pixels become $x-1-d(x-1, y)$ and $x-d(x, y)$ after rendering. Thus, the width of the disocclusion region can be formulated as follows.

$$(x-d(x, y)) - (x-1-d(x-1, y)) - 1 = d(x-1, y) - d(x, y) = d_f(x, y) \quad (4)$$

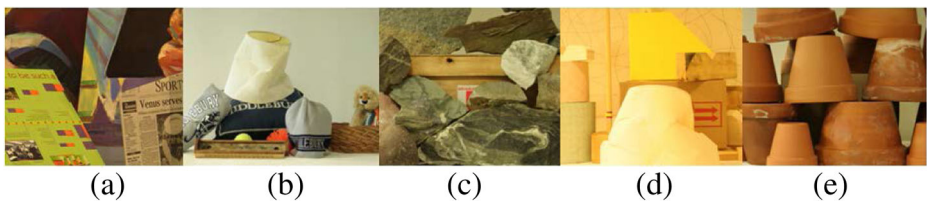


Fig. 6 The original images of Middlebury database. (a) Venus (434*383), (b) Midd11 (465*370), (c) Rocks1 (425*370), (d) Lampshade1 (433*370), and (e) Flowerpots (437*370)

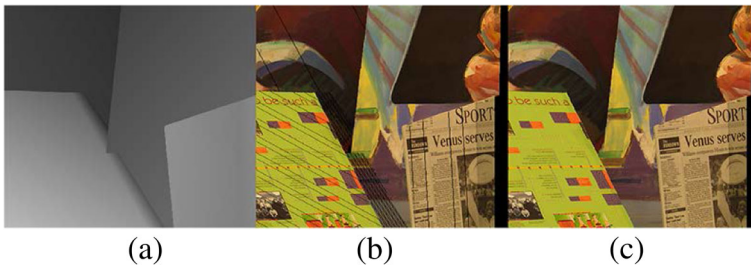


Fig. 7 Results of the proposed algorithm for Venus. (a) The preprocessed depth map, (b) the synthesized image using (a), and (c) the hole-filling result

For (x_0, y) , the width of the disocclusion region (which is the information loss area in the synthesized view) is equal to $d_f(x_0, y)$. In our proposed method, the neighboring background information with width of $d_f(x_0, y)$, is used to fill the disocclusion region. Thus, the maximum value of k is set as $2 * d_f(x_0, y) - 1$. In other words, the horizontal size of the refined region in the depth map is $2 * d_f(x_0, y)$.

An example of the disocclusion region detection is shown in Fig. 4. Figure 4a and b show the original left color image and its associated depth map. Figures 4c and d demonstrate that half of the binary mask (Fig. 4c) can accurately mark out the disocclusion region (the black areas in Fig. 4d). It is worthwhile to point out that the region marked by the red rectangle in Fig. 4c does not match the hole in Fig. 4d. This is because other FG object is present in this region. In this way, the FG depth pixels are preserved, and the distortion on the FG object after rendering is avoided. Figures 4e and f show the image inpainting result accompanied by annoying artifact distortion.

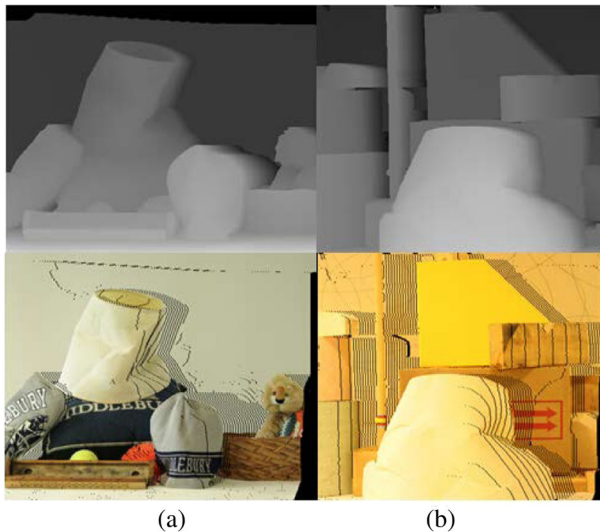


Fig. 8 Intermediate results of the proposed algorithm for Midd11 and Lampshade1. The top row shows the results of the preprocessed depth map, and the bottom row shows the synthesized image using the preprocessed depth map. (a) Midd11, (b) Lampshade1

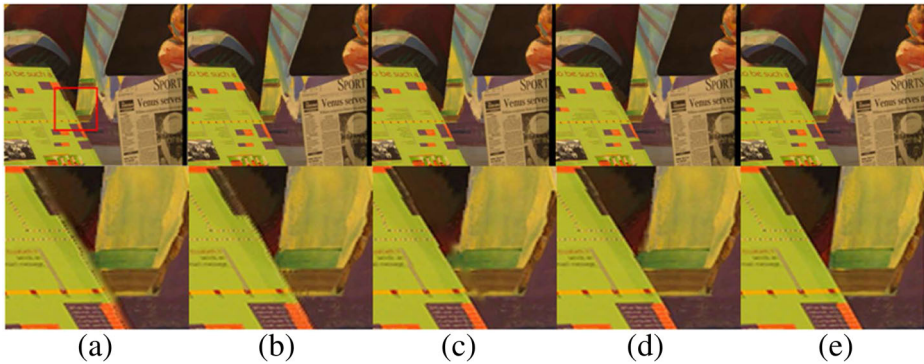


Fig. 9 Synthesized virtual images of Venus based on the produced depths. (a) Without preprocessing, (b) Lee's method, (c) Xu's method, (d) proposed divide-and-conquer hole-filling method, and (e) the groundtruth

The principle of the proposed divide-and-conquer disocclusion handling approach is shown in Fig. 5. Figure 5a shows the positional relationship in the horizontal direction between color and depth discontinuity. After rendering, the generated disocclusion around the depth discontinuity is shown in Fig. 5b. The size of the disocclusion region is determined by the value of $d_f(x, y)$, and the quantitative relationship between them can be found in [29]. In Fig. 5c, the disocclusion region is filled with the neighbor pixels. As the surrounding pixels of the disocclusion region contain both FG and BG pixels, annoying artifact will appear. To avoid this annoying artifact, the FG objects of the depth map, which neighbors with the sharp depth discontinuity, are firstly dilated as shown in Fig. 5d (marked by purple arrow). Thus, the depth boundary will be slightly wider than the corresponding color boundary. In such a way, the problem of annoying artifact is solved. A structure element $n \times n$ large is used for dilation in this paper.

The marked target region except for the dilation area in the depth map, is refined to reduce the sharp depth discontinuities through a linear interpolation process, whose pixel values decrease from the FG depth value to the BG depth value as shown in Fig. 5d (marked by red arrow). The linear interpolation process can be described as:

$$d^*(x_0 + k, y) = d(x_0 - 1, y) - \frac{d_f(x_0, y)}{2 * d_f(x_0, y) - n} * k \quad (5)$$

$$k \in \{n, n + 1, \dots, 2 * d_f(x_0, y) - 1\}$$

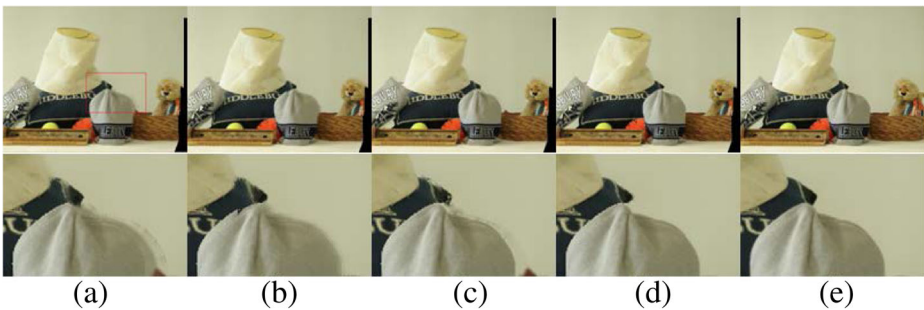


Fig. 10 Synthesized virtual images of Midd1 based on the produced depths. (a) Without preprocessing, (b) Lee's method, (c) Xu's method, (d) proposed divide-and-conquer hole-filling method, and (e) the groundtruth

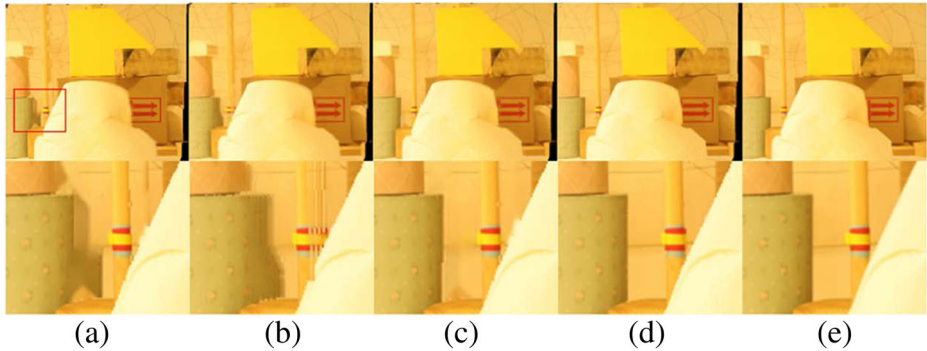


Fig. 11 Synthesized virtual images of Lampshade1 based on the produced depths. (a) Without preprocessing, (b) Lee's method, (c) Xu's method, (d) proposed divide-and-conquer hole-filling method, and (e) the groundtruth

where d^* represents the refined pixels in the depth map. Specifically, as we limited the refined areas, the minimum value of the linear interpolation will not reach the BG depth value when the horizontal distance between the boundaries of two FG objects is less than $2 * d_f(x_0, y)$.

Finally, in order to get a better result of DIBR, a two-dimensional median filtering algorithm [6] is used to remove isolated depth pixels. In this way, the error pixels on the synthesized viewpoints can be reduced.

3 Experimental results

In order to evaluate the effectiveness of the proposed method, the color images and corresponding depth maps of Venus, Middl1, Rocks, Lampshade1, and Flowerpots, provided by Middlebury database [8, 22, 23] are used in our experiments. All the depth images provided by Middlebury database have the same resolution with the corresponding color images. The original images are shown in Fig. 6. T_0 is set as 5, and n is set as 3. The depth map of Venus after being refined by the above algorithms is shown in Fig. 7a. Since the disparity of the target region decreases linearly, after rendering by the preprocessed depth map, the large

Table 1 Comparison of PSNR (dB) for Middlebury database

Sequences	Without preprocessing	Lee's method [14]	Xu's method [30]	Proposed method
Venus	28.35	28.39	28.41	28.58
Middl1	32.64	31.26	33.89	35.00
Rocks1	31.48	29.70	33.76	35.40
Lampshade1	31.11	30.74	35.19	35.80
Flowerpots	24.88	24.20	26.82	29.37



Fig. 12 The original images of real world test sequences. (a) Balloons (1024*768), (b) Door Flowers (1024*768), (c) Lovebird1 (1024*768), and (d) Exit (640*480)

disocclusion on the virtual image is divided into several small uniform holes as shown in Fig. 7b. The small holes are easily filled by various inpainting methods [1, 20, 26]. In this paper, we show the result with Telea image inpainting [26]. To further improve the quality of the synthesized images of the proposed method, the pixels which are located in the regions of small holes are processed by a horizontal median filter [9] after image inpainting. The $1*5$ smoothing window for the horizontal median filter was adopted in our experiment. To further show the advantages of the divide and conquer method with linear interpolation, the intermediate results of the proposed method are shown in Fig. 8. It can be seen from the figure, after rendering with the preprocessed depth map, the large disocclusion on the virtual image is divided into small holes, which are easily filled by image inpainting.

For comparison subjectively, results of different methods are also shown in Figs. 9, 10 and 11. The virtual view images are synthesized with the original depth maps (Figs. 9–11a), the preprocessed depth maps by Lee’s method [14] (Figs. 9–11b), Xu’s method [30] (Figs. 9–11c), and the proposed divide-and-conquer hole-filling method (Figs. 9–11d), and the groundtruth (Figs. 9–11e), respectively. The top row shows the results of dealing with the disocclusion by the above methods, and the bottom row shows some parts of the virtual images which are labeled with red rectangles. It can be seen that the proposed method can fill the disocclusion regions more effectively than others. Meanwhile, the visual effect of the virtual image is improved remarkably. Compared to Xu’s method [30], the annoying artifact and diffusion effect are reduced and even removed by our proposed method.

To compare with the ground-truth provided by Middlebury database objectively, PSNR values are computed between the virtual images and the corresponding ground-truth. The results of PSNR for Middlebury database are shown in Table 1. From the table, the virtual images rendered with the proposed method achieve higher PSNR values compared to the images rendered without preprocessing, Lee’s method [14], and Xu’s method [30]. For

Table 2 Comparison of PSNR (dB) for Real World Test Sequences

Sequences	Without preprocessing	Lee’s method [14]	Xu’s method [30]	Proposed method
Balloons	23.38	23.92	24.44	24.79
Door Flowers	26.28	27.13	27.72	27.96
Lovebird1	28.90	28.95	28.97	29.09
Exit	20.68	20.25	20.84	21.02

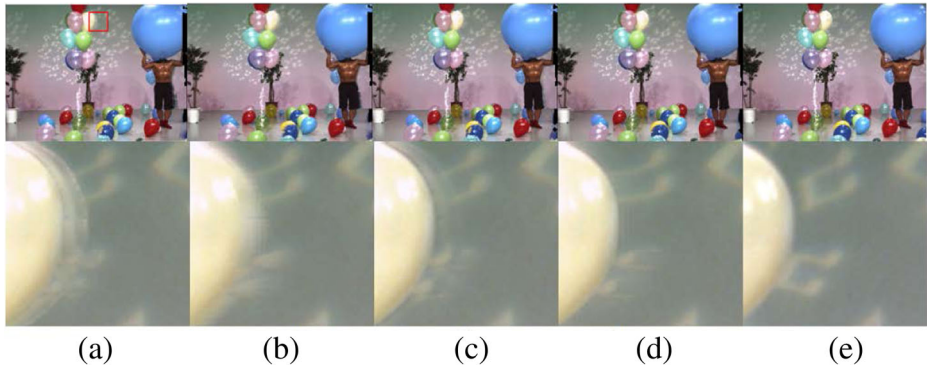


Fig. 13 Synthesized virtual images of Balloons based on the produced depths. (a) Without preprocessing, (b) Lee's method, (c) Xu's method, (d) proposed divide-and-conquer hole-filling method, and (e) the groundtruth

example, with the proposed method, we can achieve 2.55 dB gain for Flowerpots as well as 0.17 dB gain for Venus, compared with Xu's method [30]. The reasons why the proposed method is effective than other methods are as follows: 1) The important foreground information has less geometrical distortion compared with the smoothing based methods. 2) It is reasonable to fill holes using adjacent background texture, since the missing information of the disocclusion region in the synthesized view, which is occluded by the foreground object, is background information. 3) The disocclusion is divided into several small holes by the proposed divided and conquer strategy. Compared with big holes, small holes are easily to be filled, since there are background pixels which are adjacent with the small holes in the horizontal direction and the missing pixels on the small holes are similar with their horizontal neighbors.

It should be noted that the PSNR gain of smoothing based methods sometimes are lower than the results without preprocessing, as smoothing based methods cause geometrical

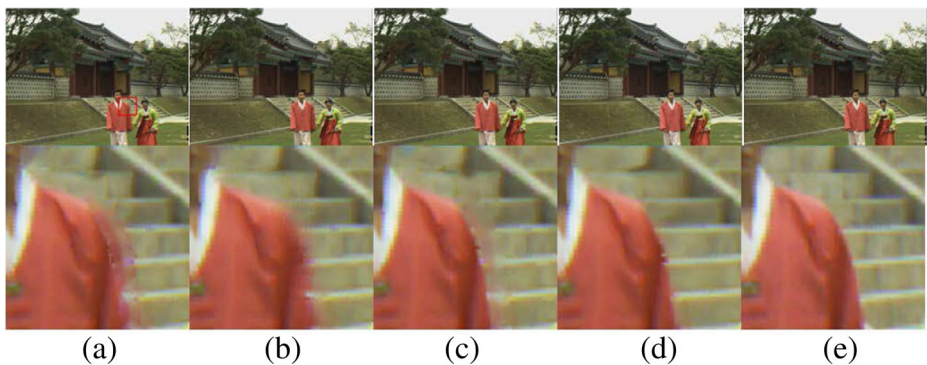


Fig. 14 Synthesized virtual images of Lovebird1 based on the produced depths. (a) Without preprocessing, (b) Lee's method, (c) Xu's method, (d) proposed divide-and-conquer hole-filling method, and (e) the groundtruth

Table 3 Comparison of the Running Time (in Seconds)

Sequences	Lee's method [14]	Xu's method [30]	Proposed method
Venus	0.350	0.002	0.039
Midd11	0.432	0.002	0.040
Rocks1	0.487	0.002	0.039
Lampshade1	0.451	0.002	0.039
Flowerpots	0.415	0.002	0.039
Balloons	0.928	0.004	0.063
Door Flowers	0.988	0.006	0.079
Lovebird1	1.172	0.007	0.091
Exit	1.276	0.007	0.089

distortion in texture. These methods are generally used for improving the visual effect of the synthesized views or boosting the computational speed. The dilation based method with a sophisticated in-painting method or texture synthesis method can reduce more errors and obtain a higher PSNR gain. It is mainly because the important foreground information has less geometrical distortion compared with the smoothing based methods, and the missing textures located in the disocclusion region are similar with the adjacent background textures. In addition to inheriting the advantages of the dilation based method, the proposed method divides the disocclusion into several small holes and thus makes holes to be filled easily and accurately.

To further evaluate the performance of the proposed method for real word cases, four test sequences, Balloons, Door Flowers, Lovebird1, and Exit [4, 7, 25, 27], are used in our experiments. The original images are shown in Fig. 12. The resolution for Balloons, Door Flowers, and Lovebird1 is 1024*768, and the resolution for Exit is 640*480. The 1st and 3rd viewpoints of Balloons, the 8th and 9th viewpoints of Door Flowers, the 6th and 7th viewpoints of Lovebird1, and the 1st and 2nd viewpoints of Exit are adopted. The PSNR values of the four test sequences are shown in Table 2, and the subjective comparison of different methods are shown in Figs. 13 and 14. From these results, it can be observed that the proposed method can improve the quality of the synthesized view objectively and subjectively.

In order to investigate the additional computation complexity of the proposed method, we performed experiments under the simulation environments of Intel(R) Core(TM) i5-4590 CPU @3.30GHz with 8.0GB memory and 64 bit Windows 7 operating system. The comparisons of running time are shown in the Table 3. It can be observed that the running time of the proposed algorithm is less than Lee's method, and a little more than Xu's method, which is acceptable.

4 Conclusion

In this paper, we have presented a novel and effective divide-and-conquer method for handling disocclusion of the synthesized image in single-view rendering. Firstly, a binary mask

is used to mark the disocclusion region. Then, the depth pixels located in the disocclusion region are modified by a linear interpolation process. Finally, a median filtering is adopted to remove the isolated depth pixels. With the proposed method, disocclusion regions in the synthesized virtual view are divided into several small holes after DIBR, and are easily filled by image inpainting. Experimental results demonstrate that the proposed method can effectively improve the visual effect of the synthesized view. Furthermore, the proposed method gives higher PSNR values of rendered virtual images compared to previous methods.

Acknowledgments This research was partially supported by the Natural Science Foundation of China (Nos. 61271324, 61520106002, 61471262, 91320201, and 61471260).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Bertalmio M, Sapiro G, Caselles V, Ballester C (2000) Image inpainting. In: Proceedings ACM International Conference on Computer Graphics and Interactive Techniques, New Orleans, America, pp 417–424
2. Chaurasia G, Duchene S, Sorkine-hornung O, Drettakis G (2013) Depth synthesis and local warps for plausible image-based navigation. *ACM Trans Graph* 30(3):1–12
3. Chen K-H, Chen C-H, Chang C-H, Liu J-Y, Su C-L (2015) A shape-adaptive low-complexity technique for 3D free-viewpoint visual applications. *Circ Syst Signal Process* 34(2):579–604
4. Common test conditions of 3DV core experiments, Document E1100, JCT3V, Vienna, AT, 2013
5. Fehn C (2004) Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV. *SPIE* 5291:93–104
6. Feng Y-M., Li D-X, Luo K, Zhang M (2009) Asymmetric bidirectional view synthesis for free viewpoint and three-dimensional video. *IEEE Trans Consum Electron* 55(4):2349–2355
7. Feldmann I, Muller M, Zilly F, Tanger R, Mueller K, Smolic A, Kauff P, Wiegand T (2008) HHI Test Material for 3D Video, ISO, Archamps, France, ISO/IEC JTC1/SC29/WG11, Document M15413
8. Hirschmiller H, Scharstein D (2007) Evaluation of cost functions for stereo matching. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Minneapolis, America, pp 1–8
9. Huang T, Yang G, Tang G (1979) A fast two-dimensional median filtering algorithm. *IEEE Trans Acoust Speech Signal Process* 27(1):13C18
10. Jung S-W (2013) modified model of the just noticeable depth difference and its application to depth sensation enhancement. *IEEE Trans Image Process* 22(10):3892–3903
11. Kellnhofer P, Ritschel T, Myszkowski K, Seidel H-P (2013) Optimizing disparity for motion in depth. *Comput Graph Forum* 32(4):143–152
12. Kim HG, Ro YM (2016) Multi-view stereoscopic video hole filling considering spatio-temporal consistency and binocular symmetry for synthesized 3D video. In: *IEEE Transactions on Circuits and Systems for Video Technology*. doi:10.1109/TCSVT.2016.2515360
13. Koppel M, Makhlof MB, Miller M, Ndjiki-Nya P (2013) Temporally consistent adaptive depth map preprocessing for view synthesis. In: Proceedings of the IEEE International Conference on Visual Communications and Image Processing, Kuching, Malaysia, pp 1–6
14. Lee P-J (2011) Effendi, Nongeometric distortion smoothing approach for depth map preprocessing. *IEEE Trans Multimed* 13(2):246–254
15. Lei J, zhang C, Fang Y, Gu Z, Ling N, Hou C (2015) Depth Sensation Enhancement for Multiple Virtual View Rendering. *IEEE Trans Multimed* 17(4):457–469
16. Loghman M, Kim J (2015) Segmentation-based view synthesis for multi-view video plus depth. *Multimed Tools Appl* 74(5):1611–1625
17. Lu S, Hanca J, Munteanu A, Schelkens P (2013) Depth-based view synthesis using pixel-level image inpainting. In: Proceedings of the IEEE International Conference on Digital Signal Processing, Fira, Greece, pp 1–6

18. McMillan L, Bishop G (1995) Plenoptic modeling: an image-based rendering system. In: Proceedings ACM Conference On Special Interest Group on Computer Graphics and Interactive Techniques, New York, USA, pp 39–46
19. Oh K-J, Yea S, Vetro A, Ho Y-S (2010) Virtual view synthesis method and self-evaluation metrics for free viewpoint television and 3D video. *Int J Imaging Syst Technol* 20(4):378–390
20. Po L-M, Zhang S, Xu X, Zhu Y (2011) A new multidirectional extrapolation hole-filling method for depth-image-based rendering. In: Proceedings of the IEEE International Conference on Image Processing, Brussels, Belgium, pp 2637–2640
21. Reel S, Cheung G, Wong P, Dooley LS (2013) Joint texture-depth pixel inpainting of disocclusion holes in virtual view synthesis. In: Proceedings of the IEEE Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, Kaohsiung, Taiwan, pp 1–7
22. Scharstein D, Pal C (2007) Learning conditional random fields for stereo. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Minneapolis, America, pp 1–8
23. Scharstein D, Szeliski R (2002) A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int J Comput Vis* 47(1-3):7–42
24. Solh M, AlRegib G (2012) Hierarchical hole-filling for depth-based view synthesis in FTV and 3D video. *IEEE J Sel Topics Signal Process* 6(5):495–504
25. Tanimoto M, Fujii T, Fukushima N (2008) 1D Parallel Test Sequences for MPEG-FTV, ISO, Archamps, France, ISO/IEC JTC1/SC29/WG11, Document M15378
26. Telea A (2004) An image inpainting technique based on the fast marching method. *J Graph Tools* 9(1):25–36
27. Vetro A, McGuire M, Matusik W, Behrens A, Lee J, Pfister H (2005) Mitsubishi Electric Research Labs(USA), Multiview video test sequences from MERL, ISO/IEC JTC1/SC29/WG11, m12077, Busan, Korea
28. Wang L, Hou C, Lei J, Yan W (2015) View generation with DIBR for 3D display system. *Multimed Tools Appl* 74(21):9529–9545
29. Xu X, Po L-M, Cheung K-W, Ng K-H, Wong K-M, Ting C-W (2012) A foreground biased depth map refinement method for DIBR view synthesis. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Kyoto, Japan, pp 805–808
30. Xu X, Po L-M, Cheung K-W, Ng K-H, Wong K-M, Ting C-W (2013) Depth map misalignment correction and dilation for DIBR view synthesis. *Signal Process Image Commun* 28(9):1023–1045
31. Yao C, Tillo T, Zhao Y, Xiao J, Bai H, Lin C (2014) Depth map driven hole filling algorithm exploiting temporal correlation information. *IEEE Trans Broadcast* 60(2):394–404
32. Zhao Y, Zhu C, Chen Z, Tian D, Yu L (2011) Boundary artifact reduction in view synthesis of 3D video: from perspective of texture-depth alignment. *IEEE Trans Broadcast* 57(2):510–522



Jianjun Lei received the Ph.D. degree in signal and information processing from Beijing University of Posts and Telecommunications, Beijing, China, in 2007. He is currently a Full Professor with the School of Electronic Information Engineering, Tianjin University, Tianjin, China. He was a visiting researcher at the Department of Electrical Engineering, University of Washington, Seattle, WA, from August 2012 to August 2013. His research interests include 3D video processing, 3D display, and computer vision.



Cuicui Zhang received the B.S. degree in communication engineering from Tianjin University, Tianjin, China, in 2013. She is currently pursuing the M.S. degree at the School of Electronic Information Engineering, Tianjin University, Tianjin, China. Her research interests include 3D image processing and 3D display.



Min Wu received the B.S. degree in communication engineering from Tianjin University, Tianjin, China, in 2014. She is currently pursuing the M.S. degree at the School of Electronic Information Engineering, Tianjin University, Tianjin, China. Her research interests include 3D image processing and 3D display.



Lei You received the Ph.D. degree in electronic circuits and systems from Beijing University of Posts and Telecommunications, Beijing, China, in 2010. He is currently a Lecturer with the School of Electronic Information Engineering, Tianjin University, Tianjin, China. His research interests include video/image sensing, processing and adaptive transmission.



Fan Kefeng received the received Ph.D. degree in test signal processing from Xidian University, Xi'an, China in 2007. He is currently a senior engineer with the China Electronics Standardization Institute, Beijing 100007, China. His research interests include multimedia content protection and 3D video processing.



Chunping Hou received the M.Eng. and Ph.D. degrees, both in electronic engineering, from Tianjin University, Tianjin, China, in 1986 and 1998, respectively. Since 1986, she has been the faculty of the School of Electronic and Information Engineering, Tianjin University, where she is currently a Full Professor and the Director of the Broadband Wireless Communications and 3D Imaging Institute. Her current research interests include 3D image processing, 3D display, wireless communication, and the design and applications of communication systems.