# A NEW MULTIDIRECTIONAL EXTRAPOLATION HOLE-FILLING METHOD FOR DEPTH-IMAGE-BASED RENDERING

*Lai-Man Po*[1], *Shihang Zhang*[1,2], *Xuyuan Xu*[1], *Yuesheng Zhu*[1,2]

[1] Department of Electronic Engineering, City University of Hong Kong, Kowloong, Hong Kong, China
[2]Communication and Security Lab, Shenzhen Graduate School of Peking University, Shenzhen, China

## ABSTRACT

Depth-Image-Based Rendering (DIBR) is widely used to generate virtual view of a scene from a known view with associated depth map in 3D video applications. However, disocclusion arises in image warping of DIBR. Many hole-filling methods have been proposed such as constant color, horizontal interpolation, horizontal extrapolation, and variational inpainting, but they cause different types of annoying artifact for large holes with complex texture background. In this paper, a novel multidirectional extrapolation hole-filling method is proposed to enhance visual quality for large hole-filling with complex texture background. The proposed method uses neighbor pixels' texture features to estimate hole-filling direction in a pixel-by-pixel manner. Experimental results demonstrated that the proposed method could provide better visual quality compared with conventional methods for virtual views synthesis with high-quality depth map.

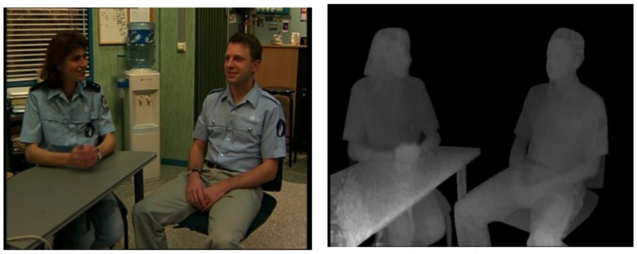
***Index Terms -*** Depth-Image-Based-Rendering, DIBR, Hole-Filling, Disocclusion, Directional Extrapolation.

## 1. INTRODUCTION

The next emerging revolution after the high-definition video will be 3D video applications such as 3DTV [1] and Free-view TV (FTV) [2]. One of the high potential 3D video representations is based on Depth-Image-Based Rendering (DIBR) [3-6] with use of video-plus-depth and multiview plus depth formats. In video-plus-depth format, the 3D video is represented by one view of the video with its depth map as shown in Fig. 1. To obtain a stereo pair, a second view is synthesized by DIBR in the decoder side. Multiple views can also be synthesized for autostereoscopic 3D displays. The video-plus-depth representation provides high compressibility of the 3D video signal and flexibility to render views with variable baseline by DIBR. However, DIBR has intrinsic limitations with artifacts arising in the rendering process. This particularly occurs in areas that become disoccluded in the rendering process. To overcome this drawback, advanced 3D video representations such as multiview-plus-depth and layered depth video (LDV) have been proposed. The video-plus-depth format, however, has shown its practicality in large-scale subjective evaluations especially for mobile devices with small display.

The disocclusion problem of DIBR is due to areas covered by objects in the reference view may appear in the synthesized view. Newly exposed areas (termed 'disocclusions' or 'holes') do not have correspondence to the reference view, their texture and depth attributes are uncertain. In DIBR, these holes have to be filled properly otherwise annoying artifacts will appear on the large discoccluded regions. Many hole-filling methods [7] have been developed with use of simple and sophisticated image processing techniques. Most of these conventional hole-filling methods are based on the neighbor pixels to fill-up the holes by linear interpolation, extrapolation, or diffusion techniques. These methods vary in complexity from a very simple method that filling with a constant color to a complex method with use of inpainting technique. Well-known hole-filling methods are Constant Color, Horizontal Interpolation, Depth-Aided Horizontal Extrapolation (DAHE), and Variational Inpainting [8]. The characteristics of these methods are quite different with special types of artifact especially in filling up large holes. To tackle this problem, a new multidirectional extrapolation hole-filling method for disocclusion restoration is proposed. In which vertical raster scans are used to fill in the disoccluded regions pixel-by-pixel and the extrapolation direction for filling up the hole's pixel is estimated based on background texture. C-shaped templates formed with surrounding background and non-hole pixels are employed to estimate the best extrapolation direction for matching the background textures.

This paper is organized as follows. Section 2 introduces conventional hole-filling methods and the detail of the multidirectional extrapolation hole-filling method is described in section 3. Performance evaluation of the proposed method is provided in section 4. Finally, a conclusion is drawn in section 5.



(a) 2D Video frame (b) Depth map

**Fig. 1.** The first frame of the well-known test video of "Interview" in video-plus-depth format.

## 2. CONVENTIONAL HOLE-FILLING METHODS

In this section, the four well-known conventional hole-filling methods [7] of Constant Color, Horizontal Interpolation, DAHE, and Variational Inpainting [8] are briefly reviewed.

### 2.1. Constant Color Hole-Filling

In constant color hole-filling method, the hole $H$ is filled with a single constant color $C$. The constant color is commonly selected as the average color of the hole boundary's pixels, which can be defined as:

$$I[n] = C, \forall n \in H \tag{1}$$

$$C = \frac{\sum_{\forall m \in \partial H} I[m]}{\sum_{\forall m \in \partial H} 1} \tag{2}$$

where $\partial H$ are the boundary of the mask defining the hole $H$. In these equations, the boundary is defined as pixels that are not part of the hole but are neighbors to pixels in the hole. In this paper, an 8-point neighborhood is used to calculate the constant color.

### 2.2. Horizontal Interpolation Hole-Filling

In the horizontal interpolation hole-filling method, the holes are filled by horizontally interpolating the pixel values in the hole's boundary. Each pixel of the hole $H$ is filled by horizontal interpolation of pixels in the boundary as defined below:

$$I[n] = (1-w)I_l[n] + wI_r[n], \forall n \in H \tag{3}$$

$$I_l[n] = I[m_l],\ m_l \in \partial H \tag{4}$$

$$I_r[n] = I[m_r],\ m_r \in \partial H \tag{5}$$

$$\mathrm{w} = \frac{n-m_l}{m_r-m_l} \tag{6}$$

where $m_l$ and $m_r$ represent the positions of the first boundary pixel to the left and right respectively of the point in the hole. The weight $w$ is the normalized distance between the left border of the hole and the point to be filled. As depth information is not taken into account in this method, foreground and background objects might be fused together within these horizontal interpolated lines.

### 2.3. Depth-Aided Horizontal Extrapolation Hole-Filling

The unnatural linear change of colors in horizontal direction of horizontal interpolation hole-filling can be solved by making use of the depth values of the pixels. In DAHE hole-filling, the depth information is taken into account to avoid using foreground objects, which are visible and not occluded. Thus, DAHE fills the hole by horizontally extrapolating into the hole using the values of pixels on the boundary that with larger depth (or smaller depth values) as expressed below:

$$I[n] = \begin{cases} I[m_l] & if\ d_r[m_l] < d_r[m_r] \\ I[m_r] & if\ d_r[m_r] < d_r[m_l] \end{cases} \tag{7}$$

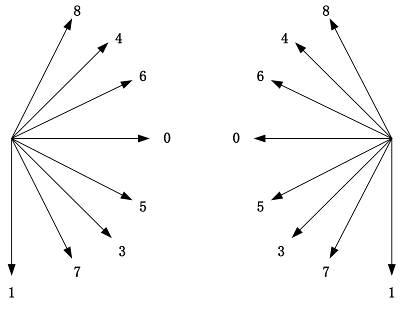
where $m_l$ and $m_r$ represent the positions of the first image pixel to the left and right respectively of the point being filled. This method expands horizontally occluded object by using the known information on the border of the occluded region. However, the disadvantage is that it will certainly fail if the border of the occluded region has non-horizontally structures because the extrapolation is done horizontally and other direction information can be lost.

### 2.4. Variational Inpainting Hole-Filling

The first three methods are very simple while a more sophisticated method based on variational inpainting has also been employed for DIBR's hole-filling. Basically, inpainting is a family of hole-filling techniques which is used to restore the specks and scratches in damaged paintings. As there is a natural similarity between damaged holes in paintings and disocclusions in DIBR warped images, inpainting has the potential to exhibit its power in disocclusion restoration. Here the Bertalmio's method [8] is implemented for comparison. In general, this sophisticated method performs better than the three previously reviewed simple methods, but annoying artifact looking like rubber-sheet effect still appears in the large holes.

## 3. MULTIDIRECTIONAL EXTRAPOLATION

To overcome the disadvantage of conventional methods, annoying artifacts on large disoccluded, a novel Multi-Directional Extrapolation Hole-Filling (MDEHF) method is proposed in this section. The main idea is to making use of directional correlation of background pixels surrounding the holes to fill a target hole's pixel, which is similar to the intra_4x4 prediction [9] of H.264/AVC video coding. The proposed MDEHF uses two sets of 9 extrapolation directions (including DC) to fill the holes for left-view and right-view warped images as shown in Fig. 2(a) and Fig.1(b). These directions are determined by the characteristic of DIBR that background of left-view and right-view warped image are always on the left-side and right-side of the holes, respectively. Thus, the hole's pixels extrapolated from these 8 directional extrapolation and one DC (constant color) have much higher chance to match the texture or structure of the background.

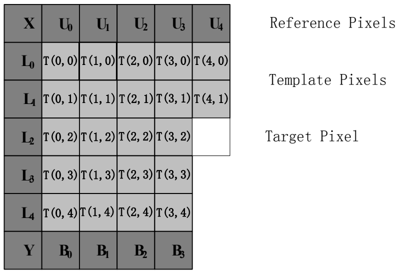


(a) (b)

**Fig. 2**: Extrapolation directions: (a) left-view, and (b) right-view.

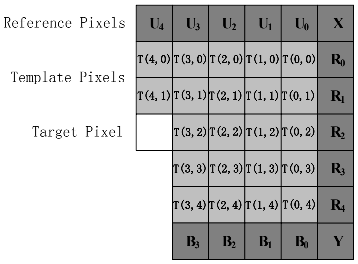
| Left-View/Right-View Extrapolation Direction (*i*) | Hole-Filling Extrapolation Equation |
|---|---|
| Horizontal_Right /Horizontal_Left(0) | H=T(3, 2) |
| Vertical_Down (1) | H=T(4, 1) |
| DC (2) | H= [T(3, 1)+T(3, 2)+T(3, 3)+T(4, 1)]/4 |
| Diagonal_Down_Right /Diagonal_Down_left (3) | H= T(3, 1) |
| Diagonal_Up_Right /Diagonal_Up_Left (4) | H= T(3, 3) |
| Horizontal_Down_Right /Horizontal_Down_Left (5) | H=[T(3, 1)+T(3, 2)]/2 |
| Horizontal_Up_Right /Horiaontal_Up_Left (6) | H= [T(3, 2)+T(3, 3)]/2 |
| Vertical_Down_Right /Vertical_Down_Left (7) | H= [T(3, 1)+T(4, 1)]/2 |
| Vertical_Up_Right /Vertical_Up_Left (8) | H=T(3, 4) |

**Table 1**: Extrapolation table for left-view and right-view.



| X | $U_0$ | $U_1$ | $U_2$ | $U_3$ | $U_4$ |
|---|---|---|---|---|---|
| $L_0$ | T(0, 0) | T(1, 0) | T(2, 0) | T(3, 0) | T(4, 0) |
| $L_1$ | T(0, 1) | T(1, 1) | T(2, 1) | T(3, 1) | T(4, 1) |
| $L_2$ | T(0, 2) | T(1, 2) | T(2, 2) | T(3, 2) | |
| $L_3$ | T(0, 3) | T(1, 3) | T(2, 3) | T(3, 3) | |
| $L_4$ | T(0, 4) | T(1, 4) | T(2, 4) | T(3, 4) | |
| Y | $B_0$ | $B_1$ | $B_2$ | $B_3$ | |

(a) C-shaped template for Left-view Hole-Filling.



| $U_4$ | $U_3$ | $U_2$ | $U_1$ | $U_0$ | X |
|---|---|---|---|---|---|
| T(4, 0) | T(3, 0) | T(2, 0) | T(1, 0) | T(0, 0) | $R_0$ |
| T(4, 1) | T(3, 1) | T(2, 1) | T(1, 1) | T(0, 1) | $R_1$ |
| | T(3, 2) | T(2, 2) | T(1, 2) | T(0, 2) | $R_2$ |
| | T(3, 3) | T(2, 3) | T(1, 3) | T(0, 3) | $R_3$ |
| | T(3, 4) | T(2, 4) | T(1, 4) | T(0, 4) | $R_4$ |
| | $B_3$ | $B_2$ | $B_1$ | $B_0$ | Y |

(b) C-shaped template for Right-view Hole-Filling.

**Fig. 3.** The C-shaped templates used in the extrapolation direction estimation of the proposed MDEHF.

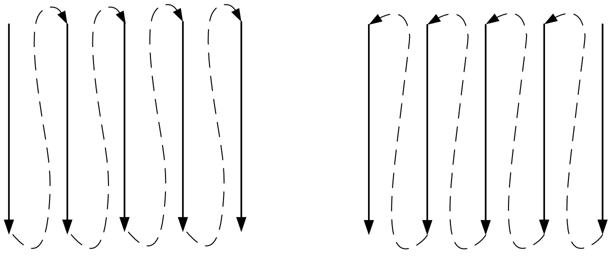

(a) (b)

**Fig. 4.** The proposed (a) vertical-left-raster scan, (b) vertical-right-raster scan for left-view and right-view hole-fillings, respectively.

### 3.1. Pixel-based Extrapolated Hole-Filling

The proposed MDEHF fills up the hole pixel-by-pixel because hole's pixels are empty and we cannot estimate the extrapolated direction from the hole. Thus, the proposed method estimates the extrapolation direction from C-shaped templates around the target hole's pixel as shown in Fig. 3(a) and Fig. 3(b) for left-view and right-view rendering. Considering the accuracy of direction estimation and the computational complexity, the template size is chosen as 22 pixels. With this design, it can make sure that the pixels within the templates are available for direction estimation.

### 3.2. Extrapolation Direction Estimation

The cost function for selecting the best extrapolation direction is based on sum of absolute differences (SAD) as defined below:

$$Distortion(i) = \frac{\sum_{x,y\in template}|T(x,y)-P_i(x,y)|}{templateSize} \quad (8)$$

$$BestDirection = arg\,min\{Distortion(i)\} \quad (9)$$

where $i$ is the extrapolation directions from 0 to 8 as shown in Fig. 2 and Table 1, ($x$, $y$) is the position of the pixels, the $T(x,y)$ is pixel of the template, the *templateSize* is the number of pixels in template and $P_i(x,y)$ are predicted pixels from $i$ extrapolation direction. After obtaining the best direction based on equation (9), the target pixel is filled up based on its neighbor's template pixels of $T(3,4)$, $T(3,3)$, $T(3,2)$, $T(3,1)$, and $T(4,1)$ using the hole-filling equations as listed in Table 1.

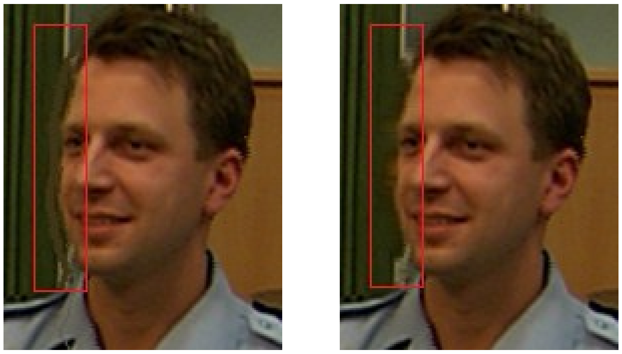
### 3.3 Vertical Raster Scans

As shown in Fig. 3(a), the C-shaped template for the left-view image rendering consists of diagonal lower left pixels of the target pixel. The situation is different from the conventional intra_4x4 prediction, which consists of the diagonal upper right pixels for the extrapolation, in which conventional horizontal raster scan (left-to-right from top-to-bottom) cannot make sure that the pixels are available for prediction and encoding processes. However, in the proposed MDEHF with the C-shaped templates as shown Fig. 3, vertical raster scans is used to make sure that pixels within the templates are available for direction estimation and hole-filling. Thus, vertical-left-raster (VLR) scan (top-to-bottom from left-to-right) and vertical-right-raster (VRR) scan (top-to-bottom from right-to-left) as shown in Fig. 4(a) and Fig. 4(b) are proposed for left-view and right-view hole-fillings, respectively.

### 3.4 Multidirectional Extrapolation Hole-Filling Procedure

The overall procedure of the proposed MDEHF can be summarized as:

**Step 1:** For the left-view rendering, VLR scan is used to detect hole's pixel. Once a hole's pixel is detected, a left-view rendering C-shaped template is selected based on the location of the hole. For the right-view rendering, VRR scan is used to detect hole's pixels.


(a) Constant Color (b) Horizontal Interpolation

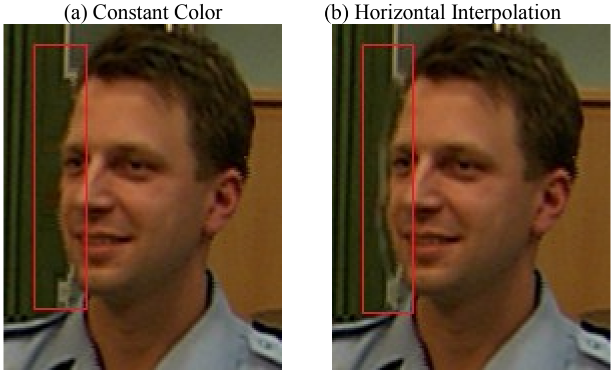

(c) DAHE (d) Variational Inpainting

**Fig. 5.** The enlarged rendered left-view images of the first frame of 'Interview' sequence using the four conventional hole-filling methods.

**Step 2:** The SAD of template using prediction directions (0-8) are calculated and select the minimum cost direction as the hole-filling direction for the target pixel.

**Step 3:** With use of the selected direction, the color components and depth value of the target hole's pixel are filled up based on the equations as shown in Table 1.

**Step 4:** Repeat steps 1- 3 until all the holes are filled in.

## 4. EXPERIMENTAL RESULTS

The performance of the proposed MDEHF method is evaluated by the first frame of the well-known test video "Interview" in video-plus-depth format. Its monoscopic 2D color image and associated depth map are shown in Fig. 1(a) and 1(b), respectively. The edge in depth map is preprocessed to match the color edge to reduce the artifact in image warping and prediction error in hole filling. To compare the performance of different hole-filling methods on DIBR especially on filling large holes, the enlarged left-view of the rendered "Interview" image using the four conventional hole-filling methods of constant color, horizontal interpolation, DAHE, and variational inpainting are shown in Fig.5(a) to 5(d), respectively. These figures demonstrated that various types of annoying artifacts are appeared on large holes tapped in the Fig with red rectangle.

For constant color hole-filling method as shown in Fig. 5(a), it is easy for a viewer to notice the artificial regions, in which a flat-color artifact appears on the left side of the human face of the "Interview" test image, which is a very large hole. For horizontal interpolation hole-filling method as shown in Fig. 5(b), the horizontal-stripe artifact appears on the left side of the face. It is because the horizontal interpolation hole-filling method leads to unnatural linear change in horizontal direction and viewers are quite sensitive to these gradual color changes. For DAHE hole-filling method as shown in Fi.g 5(c), the horizontal-stripe artifact is still noticeable and not very matches the vertical texture of the background. For the variational inpainting hole-filling method as shown in Fig. 5(d), in general, it performs better than three previously methods, but annoying artifact looking like rubber-sheet effect still appears in the large holes. These artifacts become very annoying when large holes occur on complex background especially with non-horizontal textures. It is because these methods do not have mechanism to match the background textures or structures.

To demonstrate the outstanding performance of the proposed MDEHF method, the rendered left-view image of the first frame of 'Interview' sequence using proposed method is shown in Fig. 6(a) and Fig. 6(b). The enlarged rendered image Fig. 6(b) demonstrated the vertical line on top of human head is recovered by the matching the vertical texture of the background behind the human face and multidirectional prediction of disocclusion region. Experimental results demonstrate that the proposed method achieves better background texture matching in various parts of the rendered "Interview" image especially on the large hole area with non-horizontal textures. We have also performed experiments on other well-known video sequences and similar improvements are also obtained as compared with conventional methods for virtual views synthesis with high-quality depth map.

## 5. CONCLUSION

In the DIBR based on conventional hole-filling methods due to no mechanism to match the background textures or structures, various of annoying artifacts generated on large holes such as flat-color artifact of constant color hole-filling, horizontal-stripe artifact of horizontal interpolation or DAHE Hole-Fillings, and rubber-sheet artifact of variational inpainting Hole-Filling. In this paper, a novel multidirectional extrapolation hole-filling method is proposed to achieve better background texture matching in DIBR. The extrapolation direction for the hole-filling is estimated based on C-sharp template on the available background pixels. To make sure the availability of these background pixels for direction estimation and hole-filling in various directions, vertical raster scans and pixel-by-pixel hole-filling are used in the proposed MDEHF method. Experimental results show improvements in large hole with complex background, especially multiple direction texture background, are observed and perceptually better with comparison of annoying artifacts in conventional hole filling methods.

## ACKNOWLEDEMENT

The work described in this paper was substantially supported by a GRF grant with project number of 9041501 (CityU 119909) from City University of Hong Kong, Hong Kong SAR, China.

(a) Rendered left-view frame using proposed MDEHF.

(b) Enlarge image of the rendered frame using proposed MDEHF.

**Fig.6.** The rendered left image of the first frame of 'Interview' sequence using proposed MDEHF method**.**

## REFERENCES

[1] A. Kubota, A. Smolic, M. Magnor, M. Tanimoto, T. Chen, and C. Zhang, "Multiview imaging and 3DTV", IEEE Signal Processing Magazine, 24(6):10-21,2007.

[2] T. Fujii and M. Tanimoto, "Free-viewpoint TV system based on ray-space representation," in Proc. SPIE ITCom, vol. 4864-22, 2002, pp. 175-189.

[3] C. Fehn, E. Cooke, O. Schreer, P. Kauff, "3D analysis and image-based rendering for immersive TV applications," Signal Process. Image Commun., 17(9):705-715, 2002.

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

[5] 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.

[6] P. Kauff, N. Atzpadin, C. Fehn,, M. Muller, O. Schreer, A. Smolic, R. Tanger, "Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability," Signal Process. Image Commun., voi. 22(2), pp. 217-234, 2007.

[7] C. Vazquez, W. J. Tam and F. Speranza, "Sterescopic Imaging: Filling Disoccluded Areas in Depth Image-Based Rendering," Proc. Of SPIE, Vol. 6392, 2006.

[8] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in Proc. ACM SIGGRAPH, Intern. Conf. on Computer Graphics and Interactive Techniques, 174, pp. 417–424, (New Orleans, LA), July 2000.

[9] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 560-576, July 2003.

[10] Yu Han, Aidong Men, and Ziyi Quan, "Combined spatial correlation of coded neighboring blocks for intra-frame coding mode prediction," in Proc. of Multimedia and Ubiquitous Engineering 2009, MUE 2009, pp. 534-538. Jun. 2009.