

Hole-Filling for DIBR Based on Depth and Gradient Information

Regular Paper

Dan Wang¹, Yan Zhao^{1*}, Zheng Wang² and Hexin Chen¹

¹ Jilin University, Changchun, China

² Changchun Institute of Optics, Fine Mechanics and Physics, Changchun, China

*Corresponding author(s) E-mail: zhao_y@jlu.edu.cn

Received 18 June 2014; Accepted 17 December 2014

DOI: 10.5772/60060

© 2015 The Author(s). Licensee InTech. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Depth image-based rendering (DIBR) is a method for generating new virtual images from known viewpoints. However, holes often appear in the rendered virtual images due to occlusion and inaccurate depth information. In this paper, we present a novel hole-filling algorithm to improve the image quality of DIBR. In the proposed method, depth information is added to the priority calculation function when determining the order of hole-filling. Then, the gradient information is used as auxiliary information when searching for the optimal matching block. Experimental results show that the proposed algorithm achieves better objective quality and also improves the subjective quality of the rendered images.

Keywords hole-filling, depth, gradient, depth image-based rendering

1. Introduction

By using monoscopic colour video and its associated per-pixel depth information, the DIBR technique is able to generate an arbitrary video view through the three-dimensional image transform technique (3D image warping) [1]. However, holes may appear unexpectedly due to

inaccurate depth maps and changes in occlusion resulting from the transformation of viewpoints. Therefore, an effective hole-filling algorithm is the key to obtaining high quality images of virtual views. Various methods have been proposed for hole-filling in DIBR [2]. Layered depth images (LDIs) [3-5] have been proposed as a method to remove holes. In this instance, several layers of depth image are needed to provide sufficient information. However, the LDIs method is computationally expensive. An approach that combines the technology of preprocessing depth maps and the method of targeting image postprocessing [6, 7] is therefore proposed to render the depth value smoothly. However, this approach does not work well when the virtual viewpoint is far away from the source reference point. Filling holes with multiple reference images and their associated depth maps is proposed in [8], which uses two viewpoints nearest to the virtual view as reference images to create target images. Then, holes are reduced by applying image fusion. Obviously, hole-filling using multiple reference images and depth maps is computationally expensive. Hierarchical hole-filling and depth adaptive hierarchical hole-filling approaches are proposed in [9], which can avoid geometric distortion in the rendered image through the use of a pyramid-like approach for estimating the hole pixels from the lower resolution estimation of the 3D warped image. Holes in the scene are filled using the temporal correlation of texture and depth

information from one view, while the regions along the static objects are filled using the inpainting technique in [10]. The algorithm in [10] is more suitable for 3D video systems with single-view-plus-depth format. In [11], a novel multidirectional extrapolation hole-filling method, which uses neighbour pixels' texture features to estimate hole-filling direction in a pixel-by-pixel manner, is proposed. This method has better visual quality for the synthesis of virtual views with a high-quality depth map. A classic exemplar-based hole-filling method, which is referred as the EBI algorithm in this paper, is proposed in [12]. The EBI algorithm focuses on filling both structural and textural information, and it pays more attention to the effects caused by filling order. It avoids producing discontinuous and uncompleted results by using image isophotes for determining the filling order. Therefore, the EBI algorithm is widely used in a large number of applications such as image inpainting, video inpainting and film stunts, due to its better performance. A similar method, referred to as the DAII algorithm in this paper, is provided in [13], where the calculating priority is modified and aided depth is used to enhance discovering the best matching patch.

The rest of this paper is organized as follows. As the basis of the proposed algorithm, works related to the EBI and DAII algorithms are introduced in section 2. Section 3 provides the proposed algorithm. Experimental results are given in section 4 and the paper is concluded in section 5.

2. Related works

2.1 EBI algorithm

There are three primary steps in the EBI algorithm: (i) calculating the priority of each point at the hole boundary; (ii) finding a matching patch with the best filling priority; (iii) updating priority and repeating the step (ii) until every hole point is filled.

2.1.1 Calculating priority

As shown in Figure 1, the image to be inpainted is denoted by I . The target region is denoted by Ω and its contour is denoted by $\delta\Omega$; ϕ is referred to as the known region (source region), p is a point on $\delta\Omega$ and Ψ_p is a target patch centred at point p .

Priority in the EBI algorithm is defined by:

$$P(p) = C(p)D(p) \quad (1)$$

where $C(p)$, which represents the reliability of Ψ_p , is called the confidence term and usually defined by:

$$C(p) = \frac{\sum_{q \in \Psi_p \cap \phi} C(q)}{\text{area}(\Psi_p)} \quad (2)$$

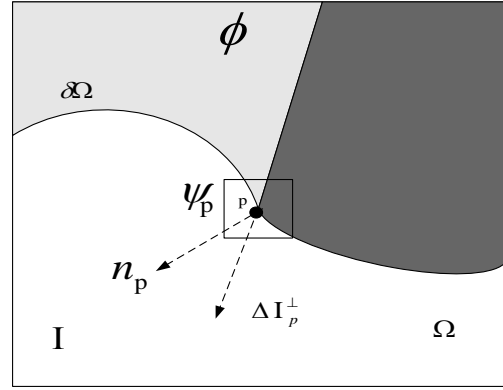


Figure 1. Illustration of the EBI algorithm

where $\text{area}(\Psi_p)$ is the area of Ψ_p and $C(p)$ is often initialized to:

$$C(p) = \begin{cases} 0, & \forall p \in \Omega \\ 1, & \forall p \in I - \Omega \end{cases} \quad (3)$$

$D(p)$, which represents the strength of the isophote, is referred to as the data term and usually defined by:

$$D(p) = \frac{\left| \left(\nabla I_p \right)^\perp \cdot n_p \right|}{\alpha} \quad (4)$$

where n_p is the normal vector at point p , ∇I_p is the gradient term, \perp is the orthogonal operator and α is a normalizing factor.

2.1.2 Finding the best matching patch

The hole pixel \hat{p} , which has the greatest priority and the most optimal target block $\Psi_{\hat{p}}$ will be obtained after calculating the priority of all points at the hole boundary. Then, the best matching patch $\Psi_{\hat{q}}$ for $\Psi_{\hat{p}}$ will be found in the image of the source region. The best matching block $\Psi_{\hat{q}}$ usually satisfies:

$$\Psi_{\hat{q}} = \arg \min_{\Psi_q \in \Phi} d(\Psi_{\hat{p}}, \Psi_q) \quad (5)$$

where $\Psi_{\hat{p}}$ and Ψ_q are referred to as pixel blocks with centred points \hat{p} and q , respectively. The pixel similarity $d(\Psi_{\hat{p}}, \Psi_q)$ between the target block $\Psi_{\hat{p}}$ and the matching block Ψ_q is calculated by the sum of squared difference (SSD), which is defined by:

$$SSD = \sqrt{\sum_{i=1}^m \sum_{j=1}^n (p_{ij} - q_{ij})^2} \quad (6)$$

Additionally, the pixel colours are represented in the colour space of CIE Lab, due to the nature of perceptual uniformity.

2.1.3 Updating priority

After Ψ_p^\wedge has been filled, the priority of p will be updated to:

$$C(p) = C(\hat{p}) \quad \forall p \in \Psi_{\hat{p}} \cap \Omega \quad (7)$$

Additionally, the boundary involved will be updated and the process is repeated until the entire region is completed.

2.2 DAI algorithm

According to the analysis in [12], the confidence term $C(p)$ in (1) will gradually become smaller or even zero, so that no matter how large $D(p)$ is, the product $P(p)$ will likely remain zero. This will lead to negative impacts on the spreading of the isophote, as well as the inpainting effect. Therefore, in the DAI algorithm, (1) is modified to:

$$P(p) = C(p) + D(p) \quad (8)$$

Furthermore, in the DAI algorithm, depth-aided texture estimation is used to search for the best matching patch and the definition of pixel similarity is changed to:

$$d(\Psi_{\hat{p}}, \Psi_q) = SAD(\Psi_{\hat{p}}, \Psi_q) + \lambda \left| \text{ave}(\sum \text{depth}(p_i)) - \text{ave}(\sum \text{depth}(q_0)) \right| \quad (9)$$

where p_1 denotes the non-hole pixel point and q_0 in the source block corresponds to the point of the unknown part in the target block; $SAD(\Psi_{\hat{p}}, \Psi_q)$ represents the sum of absolute difference (SAD) between the target block $\Psi_{\hat{p}}$ and the matching block Ψ_q :

$$SAD = \sum_{i=1}^m \sum_{j=1}^n |p_{ij} - q_{ij}| \quad (10)$$

λ is a Lagrange multiplier, which is defined by:

$$\lambda = \begin{cases} \lambda_b & \text{ave}(\sum \text{depth}(p_i)) \leq \text{ave}(\sum \text{depth}(q_0)) \\ \lambda_f & \text{ave}(\sum \text{depth}(p_i)) > \text{ave}(\sum \text{depth}(q_0)) \end{cases} \quad (11)$$

where λ_f and λ_b are two constants and $\lambda_f > \lambda_b$.

3. Proposed algorithm

3.1 Calculating priority

In the EBI algorithm, the amount of structural information in the hole region is measured by the gradient or variation of the grey level. The greater the $D(p)$, the higher the

priority will be. However, the data term $D(p)$ may not fully reflect the complex edge structural information when the variation of the grey level is not large, which will bring about error texture growth. The depth information reflects the distance between objects in the scene and the camera. The depth values of objects at different positions are different. Additionally, depth values will change dramatically at the edge of the two objects with different distances. Therefore, depth factor $Z(p)$ is used as auxiliary information for determining the edge structure when calculating priority in the proposed algorithm. Even if the variation of the grey level is not obvious, which means the data term $D(p)$ is not good enough to reflect the information around hole points, a more reasonable priority can be obtained for each hole point with the help of depth factors. In the proposed algorithm, the priority is redefined by:

$$P(p) = C(p) + D(p) + Z(p) \quad (12)$$

where $Z(p)$ is the depth term, which reflects the change rate of the average depth of the pixels in the target block and $Z(p)$ is defined by:

$$Z(p) = \sum_{q \in z_p \cap \Omega} \left(\frac{z_p(q) - |\bar{z}_p|}{\alpha} \right)^2 \quad (13)$$

where $|\bar{z}_p|$ represents the average depth value of the pixels known in the target block and α is a normalizing factor.

3.2 Searching for the best matching patch

The sum of squared difference (SSD) is the criterion for selecting a best source patch in the EBI algorithm, which puts particular importance on the colour difference between pixels. However, it cannot distinguish the structure and texture information in the image very well, which results in mismatching and makes the hole-filled image look unnatural. In terms of an image, structural information reflects its macroscopic features, while textural information reflects its microcosmic features. Inaccurate texture growth will appear during the process of filling if only the SSD is considered.

The DAI algorithm has a better effect for regions containing natural texture (weak structure) than for those with strong complex structural information and might therefore naturally lead to structural fracture. Furthermore, the DAI algorithm assumes that all hole points appear in the background region, which is not true for all cases. In other words, when using the DAI algorithm, all of the hole pixels will be filled by the background pixels, even if they belong to the foreground or the junction between foreground and background. Therefore, it may inevitably cause unexpected discontinuity and distortion at the object's edge.

To solve this problem, we combined the gradient information into the definition of pixel similarity in the proposed method, which is given by:

$$d(\Psi_p, \Psi_q) = SSD(\Psi_p, \Psi_q) + Grad(\Psi_p, \Psi_q) \quad (14)$$

Generally, the warped image is processed in YUV space and the image is often provided in 4:2:0 format. Thus, $SSD(\Psi_p, \Psi_q)$ and $Grad(\Psi_p, \Psi_q)$ are simplified by considering only the Y component:

$$SSD(\Psi_p, \Psi_q) = \sqrt{\sum_{i=1}^m \sum_{j=1}^n (p_{ij}^Y - q_{ij}^Y)^2} \quad (15)$$

$$Grad(\Psi_p, \Psi_q) = \sum_{i=1}^m \sum_{j=1}^n (|p_{(i+1)j}^Y - p_{ij}^Y| - |q_{(i+1)j}^Y - q_{ij}^Y|) \quad (16)$$

where $SSD(\Psi_p, \Psi_q)$ represents the sum of squared difference between Ψ_p and Ψ_q and $Grad(\Psi_p, \Psi_q)$ represents the sum of absolute difference of gradient between Ψ_p and Ψ_q .

4. Experimental results

In order to evaluate the proposed method, we compared the performance of the proposed algorithm with that of the EBI algorithm, DAI algorithm and the algorithm in [11], which is referred to as the MDE algorithm in the comparison results. In the experiments, the multiview video sequence *Breakdancers*, with camera parameters and depth maps provided by Microsoft Research, were used for the performance evaluation. There were eight views in total and each view had 100 frames of image at a resolution of 1024×768. In our proposed algorithm, the block size was 9×9 and the same Lagrange multipliers were used as those in the DAI algorithm. The normalizing factor in (13) was 255 in our experiments.

In the first experiment, we took the first frame of view 3 or camera 3 as reference and generated view 4. Experimental results are shown in Figure 2. Figure 2 (a) shows the original image of view 4; (b) is the DIBR generated image; (c) to (f) are the hole-filled images created by using different hole-filling algorithms. Comparing the results of the left side of the middle individual (region A) in Figure 2 (c), 2 (d) and 2 (f), we can see clearly that at the left shoulder position, the EBI and DAI algorithms chose to fill the hole with a weak structure and texture information according to the edge of the white scarf, rather than the foreground of the left shoulder. The reason for this is that the variance of the grey level between the white scarf and the black sleeveless cloth was greater than that between the black sleeveless cloth and the background wall. Only the variance of the grey level can misguide the priority algorithm based on isophotes, especially for texture areas with large contrasts around the

to-be-filled target. Therefore, the EBI and DAI algorithms could not identify the real important structural edge, which finally resulted in the wrong extension of image structure. The priority calculating in our proposed algorithm changed the filling order and gave filling preference to the target with the maximum mixed priority by adding depth information. Consequently, the proposed algorithm restricted the wrong texture growth to some extent, as shown in Figure 2 (f). Regarding the results of the MDE algorithm in Figure 2 (e), the visual effect of region A looked better than those provided by the EBI and DAI algorithms; however, the pixels in the regions of the white scarf and the black sleeveless cloth around the left shoulder position were wrongly extended to the background. In addition, comparing the results of region B in Figure 2(c), 2(d) and 2(f), the image quality of hole-filling by using the proposed method was significantly better than those produced using the EBI and DAI algorithms, because the structure and texture information could be considered in full by adding the gradient factor.

The corresponding PSNR comparisons are shown in Table 1. The results show that the proposed algorithm was superior to the EBI, DAI and MDE algorithms.

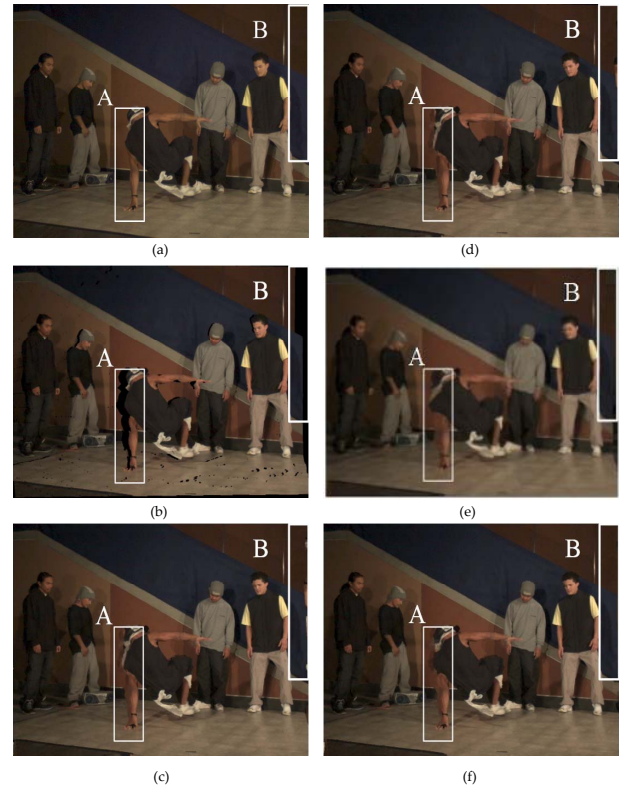


Figure 2. Subjective comparisons of the first frame of view 4: (a) the original image; (b) the DIBR generated image; (c) the result of the EBI algorithm; (d) the result of the DAI algorithm; (e) the result of MDE algorithm; (f) the result of the proposed algorithm

Algorithm	EBI	DAII	MDE	Proposed
PSNR(dB)	31.3651	31.7728	31.2711	32.3714

Table 1. PSNR comparisons of the first frame of view 4

In the second experiment, we took the first frame of view 5 as reference for generating view 6. The comparisons of partially enlarged images are shown in Figure 3. We can see from Figure 3(c) and 3(d) that the grey-white strip edges of the corner show an obviously fractured phenomenon, while the fracture phenomenon is avoided in the image shown in Figure 3(f), further illustrating the advantages of the proposed algorithm.

In the final experiment, we included all of the different views of frame 2 in order to evaluate the performance of the proposed algorithm in more depth. The frame of view 2 was rendered from the corresponding frame of view 1; the frame of view 3 was rendered from the corresponding frame of view 2; the frame of view 4 was rendered from the corresponding frame of view 3; the frame of view 5 was rendered from the corresponding frame of view 6; the frame of view 6 was rendered from the corresponding frame of view 7.

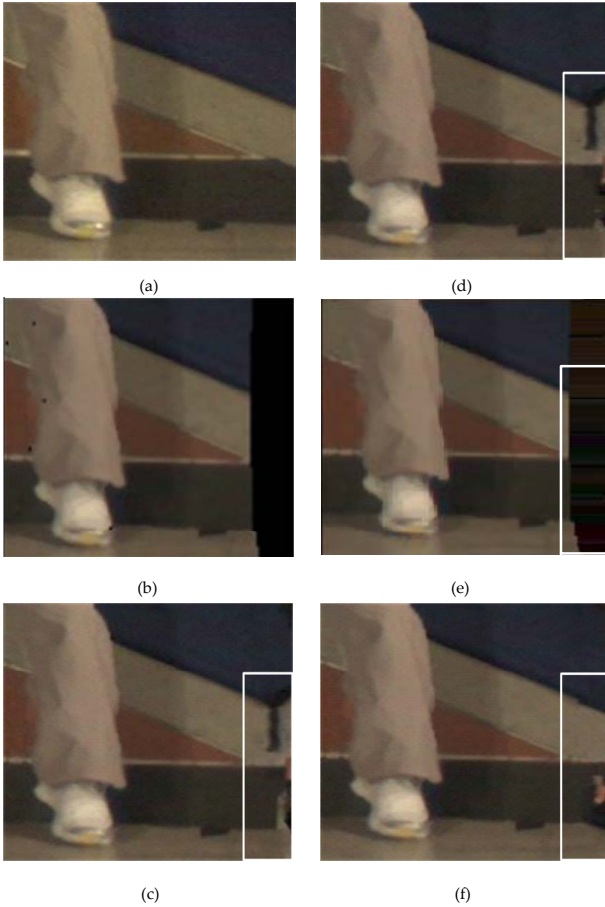


Figure 3. The comparisons of partially enlarged images of view 6: (a) the original image; (b) the DIBR generated image; (c) the result of the EBI algorithm; (d) the result of the DAII algorithm; (e) the result of the MDE algorithm; (f) the result of the proposed algorithm

Table 2 provides detailed PSNR (dB) results for the different views of frame 2. From these results, we can clearly see that the objective performance of the proposed hole-filling algorithm outperformed the EBI, DAII and MDE algorithms in all instances.

5. Conclusion

This paper proposed a novel exemplar-based hole-filling algorithm for improving virtual images in a DIBR system. Based on the widely used EBI algorithm, depth information was added to the priority computation to obtain a more accurate order of hole-filling and gradient information was introduced to the searching process for finding best matching blocks. The comparative experimental results showed that our proposed algorithm exhibited advantages in terms improving both the objective quality and subjective effect over the traditional exemplar-based hole-filling algorithms. The fracture areas in the rendered image were effectively reduced and the texture around the holes was more reasonably extended. Furthermore, the experimental results also showed that changing the priority and taking the depth and gradient information into consideration were efficient for improving the performance of image hole-filling. In future work, the proposed algorithm can be further optimized in order to reduce mismatching and its efficiency can be improved by using adaptive methods.

Algorithm	EBI	DAII	MDE	Proposed
View 2	29.5082	28.9963	29.3939	30.7612
View 3	30.3584	30.9452	30.0759	31.9389
View 4	31.2703	31.1518	31.4564	31.8760
View 5	29.3829	29.2462	29.2323	30.8589
View 6	28.4823	28.3633	25.8721	28.8059

Table 2. PSNR (dB) results of frames 2 of different views

6. Acknowledgements

This work was supported by the National Natural Science Foundation of China project (No.61271315 and No. 61171078) and in part by the National High Technology Research and Development Program of China (No. SS2012AA010805).

7. References

- [1] Fehn, C. (2004) Depth-Image-Based Rendering (DIBR), Compression, and Transmission for a New Approach on 3D-TV, Proceedings of SPIE 5291:93-104.
- [2] Vazquez, C., Tam, W. J., Speranza, F. (2006) Stereoscopic image filling disoccluded areas in depth image-based rendering, Proceedings of SPIE 6392:63920D-12.

- [3] Shade, J., Gortler, S., He, L. W., Szeliski, R. (1998) Layered depth images. Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Technique, NY, USA, pp. 231-242.
- [4] Huang, T. S., Zhang, C (2008) A layered method of visibility resolving in depth image-based rendering, Proceedings of the 19th International Conference on Pattern Recognition. pp.1-4.
- [5] Jantet, V., Guillemot, C., Morin, L. (2011) Object-based Layered Depth Images for improved virtual view synthesis in rate-constrained context, Proceedings of the 18th IEEE International Conference on Image Processing. pp.125-128.
- [6] Chen, W. Y., Chang, Y. L., Lin, S. F., Ding, L. F., Chen, L. G. (2005) Efficient Depth Image Based Rendering with Edge Dependent Depth Filter and Interpolation, Proceedings of the IEEE International Conference on Multimedia and Expo. pp.1314-1317.
- [7] Zhang, L., Tam, W. J. (2005) Stereoscopic image generation based on depth image for 3DTV, IEEE Transactions on Broadcast, 51(2): 191-199.
- [8] Narayanan, P. J., Sashi, Kumar P., Sireesh Reddy K. (2004) Depth+texture representation for image based rendering, Proceedings of the 4th Indian Conference on Computer Vision, Graphics and Image Processing. pp.113-118.
- [9] Solh, M., AlRegib, G. (2012) Hierarchical Hole-Filling For Depth-Based View Synthesis in FTV and 3D Video, IEEE Journal of Selected Topics in Signal Processing, 6(5): 495-504.
- [10] Yao, C., Tillo, T., Zhao, Y., Xiao, J., Bai, H., Lin, C. (2014) Depth Map Driven Hole Filling Algorithm Exploiting Temporal Correlation Information, IEEE Transactions on Broadcast, 60(2): 394-404.
- [11] Po, L., Zhang, S., Xu, X., Zhu, Y. (2011) A new multidirectional extrapolation hole-filling method for depth-image-based rendering, Proceedings of the IEEE International Conference on Image Processing. pp.2637-2640.
- [12] Criminisi, A., Pérez, P., Toyama, K. (2004) Region filling and object removal by exemplar-based image inpainting, IEEE Transactions on Image Processing, 13(9):1200-1212.
- [13] Luo, K., Li, D., Feng, Y., Zhang, M. (2009) Depth-aided inpainting for disocclusion restoration of multi-view images using depth-image-based rendering, Journal of Zhejiang University SCIENCE A, 10(12):1738-1749.