



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

M.S. THESIS

Boundary Decision-Based Hole Filling Method For Depth Image-Based Rendering

경계면 판단을 이용한 깊이 영상 기반 렌더링의 효과적인
홀 채움 방법

BY

CHO JEA HYUNG

FEBRUARY 2016

DEPARTMENT OF ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

M.S. THESIS

Boundary Decision-Based Hole Filling Method For Depth Image-Based Rendering

경계면 판단을 이용한 깊이 영상 기반 렌더링의 효과적인
홀 채움 방법

BY

CHO JEA HYUNG

FEBRUARY 2016

DEPARTMENT OF ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

Boundary Decision-Based Hole Filling Method For Depth Image-Based Rendering

경계면 판단을 이용한 깊이 영상 기반 렌더링의
효과적인 홀 채움 방법

지도교수 김 태 정

이 논문을 공학석사 학위논문으로 제출함

2015년 11월

서울대학교 대학원

전기컴퓨터 공학부

조 재 형

조재형의 공학석사 학위 논문을 인준함

2015년 11월

위 원 장: _____

부위원장: _____

위 원: _____

Abstract

In 3D display systems, depth image-based rendering (DIBR) is the most common technique to generate virtual images through reference views. However, disoccluded hole filling problems remain a challenging issue and the quality of the synthesized views depends on the hole filling process. Image inpainting-based hole filling is a popular approach to fill the hole region but the conventional method generates annoying artifacts along the boundaries of the foreground. This paper presents an efficient hole filling method that fills the boundary with the high visual quality. The proposed method fills the background region prior to the boundary region by using the depth values. Also, based on the boundary decision, boundary regions are filled with background-related data by clustering the pixels on the boundary into background and foreground. Experimental results show that the proposed method has reduced boundary artifacts and achieves higher visual comfort.

keywords: Depth image based rendering, View synthesis, Image in-painting, Hole filling

student number: 2014-21690

Contents

Abstract	i
Contents	ii
List of Figures	iv
List of Tables	v
1 Introduction	1
2 Related Work	4
2.1 Depth Image-Based Rendering(DIBR)	4
2.2 Hole Filling	6
2.2.1 Image In-painting	7
2.2.2 In-painting Based Hole Filling	8
3 Proposed Method	10
3.1 Pre-processes Of Hole Filling	10
3.2 Priority Computation	12
3.2.1 Local Property	12
3.2.2 Neighborhood Property	12
3.3 Patch Matching	15

3.3.1	Boundary Decision	15
3.3.2	Distance Function	16
4	Experimental Result	18
4.1	Experiments for generating stereoscopic view	19
4.2	Experiments for generating free-viewpoint view	22
5	Conclusion	25
	Abstract (In Korean)	29

List of Figures

2.1	Sample set of DIBR input. (a) Reference color image. (b) Corresponding depth map	5
2.2	Warped images. (a) Warped with small parallel shift for stereoscopic view. (b) Warped with virtual viewpoint camera matrix.	6
2.3	Notation diagram.	7
3.1	Flowchart of DIBR	10
3.2	Crack filled images.	11
3.3	Illustration of principle of neighborhood property. (a) A sample warped image. (b) Depth map of (a). (c) Depth patch Z_p on foreground. (d) Depth patch Z_p on background.	14
4.1	A set of input sample for experiment. (a) Reference color image. (b) Pre-processed image for hole filling.	19
4.2	Result images. (a), (c) Result images using existing method. (b), (d) Result images using proposed method.	20
4.3	Enlarged images. (a) Enlarged images of result using existing method. (b) Enlarged images of result using proposed method.	21
4.4	Input set and result. (a) Warped image (camera 4 to camera 3). (b) Reference image. (c) Result image of existing method. (d) Result image of proposed method.	23

List of Tables

4.1	PSNR comparison of Ballet sequence.	24
4.2	PSNR comparison of Breakdancers sequence.	24

Chapter 1

Introduction

As 3D display technique has advanced, 3D contents have received much attentions. Most of the times 3D display systems such as 3D television(3DTV) and free viewpoint television(FTV) need not only realistic views from cameras but also views from virtual viewpoints. For these reasons, the importance of technique for generating virtual views is significant for advanced 3D systems.

Depth image-based rendering(DIBR)[1, 2] is the most recognized technique to synthesize the virtual view. Because of its high flexibility of rendering the viewpoint, DIBR is considered as an essential technique for generating virtual view. It uses reference realistic view and corresponding per-pixel depth map as input. DIBR consists of three main steps.

First, depth map is pre-processed to protect foreground objects from misalignment of depth map and color image. Second step is 3D warping. Each pixel of reference image is warped according to virtual viewpoint and the depth value. Warped image contains disoccluded region also noted as holes that has no information [2]. Final step of DIBR is to fill the holes properly that viewers can have accurate 3D perception. The quality of the synthesized view depends on the hole filling process.

Exemplar based image in-painting scheme by Criminisi et al. [6] is widely utilized in hole filling process. Oh et al. [3] proposed boundary inversion for in-painting based hole filling. Since patch based in-painting scheme uses local patch data, patches on the foreground region propagate foreground data into holes. [3] replaced boundary pixels on foreground to background pixels on the opposite site of the hole. However their algorithm is not suitable for complex background or misalignments of color image and depth map. This occurs distortions of background textures. Also foreground pixels still dominantly affects the patches on the foreground so that artifacts appear. Luo et al. [4] used reference image to find the best matching patch with depth aided term. They added Lagrange multiplier and the absolute difference of depth between source region depth of the target patch and unknown region depth of a candidate patch. Their method is not able to solve the problem of selection of foreground patches prior to background. Daribo et al. [5] added depth related term for computing priority of filling order of patches. They also considered depth similarities between patches to find the best matching patch. This method improved the quality of synthesized view compared to Criminisi’s image in-painting method. However annoying artifacts still appear on the boundary of foreground and background with the method. Since human visual system is sensitive to the high frequency component, boundary artifacts raise uncomfortableness to 3D perception. There are two reasons causing this problem. First, patches on the foreground regions are selected to fill prior to patches on the background regions. Second, patches on the boundary region contains both foreground and background information so that the patches from unrelated region are matched.

In this paper faithful hole filling method that is effective on the boundary is proposed. To overcome the problems of existing methods, we focus on how to fill disocclusions naturally to feel visual comfort. Two main procedures of hole filling are patch priority computation and patch matching. Patch priority computation step is developed to fill background region prior to foreground region. Local depth of patch

and neighboring depth of patch are used in this procedure. In patch matching step, the patch is determined whether its on the boundary or not. If so, background pixels are regarded as more important pixels and we assign bigger weights to background pixels than foreground pixels. This allows to fill disocclusions well associated with background.

This paper is organized as follows : In chapter 2, overview of related works of DIBR will be introduced. Details of the proposed method will be described in Chapter 3. Experimental results and a conclusion are shown in Chapter 4 and Chapter 5 respectively.

Chapter 2

Related Work

2.1 Depth Image-Based Rendering(DIBR)

In this section, more detail of each step of DIBR is introduced. As already mentioned in previous chapter, DIBR is the technique for generating virtual viewpoint using one or more reference views and corresponding depth map. Figure 2.1 shows the sample of input set for DIBR provided by Microsoft Research[7] that is widely used in 3D research field.

First step of DIBR is depth map pre-processing. Depth map smoothing is commonly used by existing methods. This reduces the number of hole pixels and robust to noise in the depth map. But smoothing the depth map causes the boundary pixels to be warped to wrong location. These mis-located pixels have bad effect on filling big holes and cause the distortions on the synthesized view. Lu et al. [8] proposed foreground object protected depth map pre-processing method to reduce these problems. They predict the location of holes using the depth value and find the end points of foreground using Laplacian operator on the reference color image. Then by assigning depth of boundary ramp pixels to foreground depth the method reduces the misalignments of depth map and color image and leads entire foreground object to be



Figure 2.1: Sample set of DIBR input. (a) Reference color image. (b) Corresponding depth map

warped.

In the 3D warping step pixels in the reference view are projected to the virtual view. For the purpose of generating stereoscopic image using monocular image (e.g. 2D-to-3D conversion), it is assumed that virtual viewpoint has small parallel disparity from reference viewpoint [2]. In this case the baseline distance between two views is considered as the distance between left-eye and right-eye. Each pixel is shifted with the parallel directions based on its depth value (see Figure 2.2 (a)).

For generating free-viewpoint view, camera matrices are used so that position of virtual viewpoint relative to the reference camera is required [1]. Using the camera matrix and depth map, reference view pixels are mapped to 3D space through the camera coordinate system as follows:

$$sm = A_v[R_v|t_v]M \quad (2.1)$$

where m refers to the image coordinate of a pixel and M is the world coordinate of the point. s is an normalization scalar value and matrix A_v specifies the intrinsic parameters of virtual viewpoint. Matrix R_v and t_v indicates the extrinsic parameters which are the rotation parameters and translation parameters respectively. Then syn-

thesized image can be initialized by projecting each pixel in the reference view onto the virtual view(see Figure 2.2 (b)).

Figure 2.2 shows the warped image of both cases. Black regions in the figures are disoccluded areas so called holes. The proposed method to fill the hole region in this paper can be adopted in both cases.



Figure 2.2: Warped images. (a) Warped with small parallel shift for stereoscopic view. (b) Warped with virtual viewpoint camera matrix.

Final and most crucial step is hole filling. Quality of result of synthesized view depends on how to fill the hole region. This procedure is introduced in following section.

2.2 Hole Filling

Essential background knowledge of hole filling step is image in-painting. Image in-painting is an technique to fill in the missing regions in the image or to remove objects from the image. Image in-painting scheme is utilized to fill the holes in virtual view of DIBR.

2.2.1 Image In-painting

Criminisi et al.[6] proposed exemplar-based image in-painting that fills the missing region at the patch level. They propagate textures and structures by data sampled from the remainder of the image.

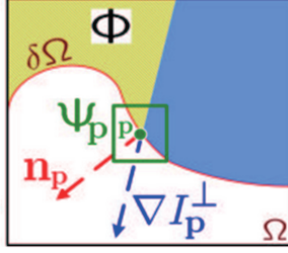


Figure 2.3: Notation diagram.

To fill the missing region with composite textures and structures, patch priority is defined to decide the filling order. They demonstrate that the filling order highly affect the quality of the filled region. The missing region would be filled better when patches on the structure region higher priority. This will lead to propagate structural information properly. Figure 2.3, captured from [6], illustrates the notation of the algorithm. Given an image I , the source region and the missing region are noted as Φ and Ω respectively. The boundary of the missing region is defined as $\delta\Omega$. For each patch Ψ_p centered at point p where $p \in \delta\Omega$ the priority $P(p)$ is computed as follows:

$$P(p) = C(p) \cdot D(p) \quad (2.2)$$

where $C(p)$ is called the *confidence* term and $D(p)$ the *data* term. They are defined as follows:

$$C(p) = \frac{\sum_{q \in \Psi_p \cap (I - \Omega)} C(q)}{|\Psi_p|} \quad \text{and} \quad D(p) = \frac{|\nabla I_p^\perp \cdot \mathbf{n}_p|}{\alpha} \quad (2.3)$$

where $|\Psi_p|$ is the area of Ψ_p , α is a normalization factor, \mathbf{n}_p is a unit vector orthogonal to the front $\delta\Omega$ in the point p , and ∇I_p^\perp is the isophote at point p . The *confidence* term

indicates the reliability of the patch by a measure of the amount of reliable information in the patch. During initialization, the function $C(p) = 0, \quad \forall_p \in \Omega$ and $C(p) = 1, \quad \forall_p \in \Phi$. The *data* term gives higher priority to the patch with strong isophote to encourage filling in from structure region.

Once all the priorities of patches on $\delta\Omega$ is computed, the patch $|\Psi_{\hat{p}}|$ with highest priority is found. Then to fill the missing region of the patch $|\Psi_{\hat{p}}|$, data from source region is sampled. By searching candidate patches in the source region the most similar patch $|\Psi_{\hat{q}}|$ is found as follows:

$$\Psi_{\hat{q}} = \arg \min_{\Psi_q \in \Phi} d(\Psi_{\hat{p}}, \Psi_q) \quad (2.4)$$

where the distance $d(.,.)$ is defined as the sum of squared differences(SSD) between two patches. Having found the most similar patch $\Psi_{\hat{q}}$, missing region in the patch $\Psi_{\hat{p}}$ is copied from its corresponding position inside $\Psi_{\hat{q}}$. After updating *confidence* values, these process is repeated until whole region is filled.

Xu et al.[9] introduced patch structure sparsity to measure the confidence of a patch located at the image structure. They are inspired from nonlocal means approach for image in-painting by Wong et al.[10] and the representation of an image with sparse combination. The structure sparsity indicates a similarity of a certain patch with its neighboring patches. Textures has large similarities with its neighborhood while patches on the structures(e.g. edges and corners) are less similar with neighboring patches. Also most similar patch $|\Psi_{\hat{q}}|$ is represented by the sparse linear combination of candidate patches under local consistency constraints.

2.2.2 In-painting Based Hole Filling

Daribo et al. [5] developed exemplar based in-painting algorithm using depth map to fill disocclusions. They set the focus that holes are the result of the displacement of foreground objects. Therefore the exposed region should be filled with background

information. Their method is called depth-aided patch priority computation and patch matching using virtual depth map.

In priority computation 'Level regularity' term $L(p)$ is added on the conventional priority computation(2.2). This can be written as follows:

$$P(p) = C(p) \cdot D(p) \cdot L(p) \quad (2.5)$$

where

$$L(p) = \frac{|Z_p|}{|Z_p| + \sum_{q \in Z_p \cap \Phi} (Z_p(q) - \bar{Z}_p)^2}. \quad (2.6)$$

Z_p is depth patch centered at p and q is the pixel location. $|Z_p|$ and \bar{Z}_p are denoted as area of the depth patch and the mean value of depth inside the patch. This gives higher priority to the patch laying on the same depth level. Variance of depth value inside the patch is used and the term has bigger value on background region.

However in Daribo's method, if a patch lays on the foreground region with small number of holes, variance of depth patch cannot represent the location of the patch well. This often causes the foreground patches to have higher priority than the background patches.

Patch matching with minimum distance computation (2.4) is updated using depth value as follow:

$$\Psi_{\hat{q}} = \arg \min_{\Psi_q \in \Phi} \{d(\Psi_{\hat{p}}, \Psi_q) + \beta \cdot d(Z_{\hat{p}}, Z_q)\} \quad (2.7)$$

β is a parameter to control the importance of depth distance. This favors patches with same depth level with target patch in matching process.

Chapter 3

Proposed Method

3.1 Pre-processes Of Hole Filling

Figure 3.1 shows the flow chart of the DIBR proposed in this paper. The main contribution is the hole filling process in the virtual view. In this section, overviews of rest of steps before hole filling process are introduced.

As a first step of DIBR, we adopted Lu's depth map pre-processing method [8]. Entire object including the boundary pixels should be warped. Otherwise if boundary pixels remain on the background side, hole would be filled with improper data by propagating boundary color.

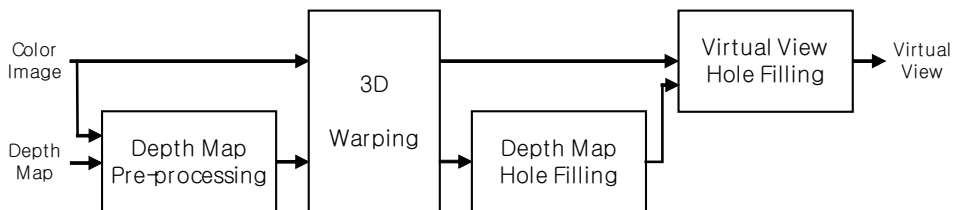


Figure 3.1: Flowchart of DIBR

Once all the pixels are warped, the virtual view contains various size of holes as

shown in Figure 2.2. How to fill large holes is highly influential to visual comfort system. Also they are much challenging problems to fill properly. On the other hand the single pixel-size holes also called as cracks can be easily filled with a simple interpolation. For large size holes, we dilate hole region to remove boundary pixels remaining on the background side.

Depth map is also warped in same way as reference color map in the step of 3D warping. Since depth map consists of textureless gray scale value and we know that holes should be filled with background information, holes in the warped depth map can be filled easily. Depth of holes are filled by minimum value of nearest source region depth on the parallel axis of the hole. It can be written as follows:

$$z_h = \min(z_{l_h}, z_{r_h}) \quad \text{where} \quad h \in \Omega \quad (3.1)$$

where z_p indicates depth value for pixel p and Ω denotes set of hole pixels keeping the similar notation with image in-painting. l_h and r_h are the nearest pixels in the source region to the left and right direction of pixel h respectively. Then the virtual view with large holes and its depth map is given as shown in Figure 3.2.

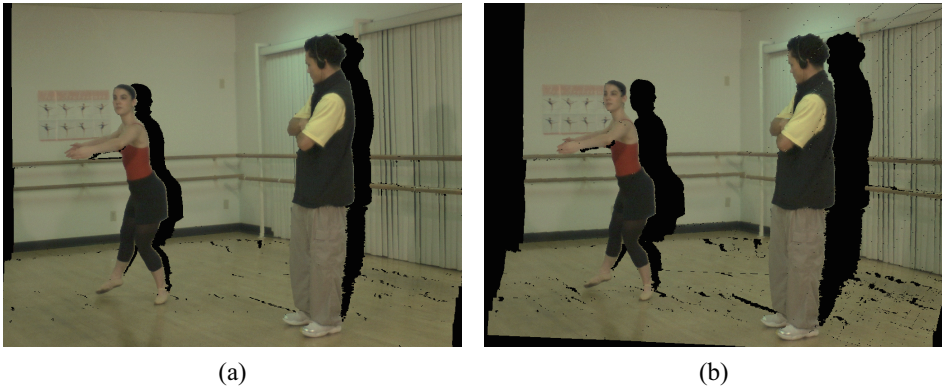


Figure 3.2: Crack filled images.

3.2 Priority Computation

In priority computation, main goal of the procedure is to give higher priority to background region. We develop the (2.5) to ensure the background regions to be filled prior to foreground or boundary region. We added the *background* term $B(p)$ so the equation for priority computation is as follows:

$$P(p) = C(p) \cdot D(p) \cdot L(p) \cdot B(p) \quad \text{where} \quad B(p) = B_l(p) \cdot B_n(p). \quad (3.2)$$

$B(p)$ consists of two terms. $B_l(p)$ and $B_n(p)$ denote *local* and *neighborhood* properties of *background* term respectively. Details of these terms is introduced in following subsections.

3.2.1 Local Property

First, depth patch is utilized. For each patch, depth value of the patch is given by the virtual depth map. Since background regions are further from the camera than the foreground region, depth value of background regions is smaller than foreground regions. Average value of depth patch that uses only local values inside the patch is considered. Local feature of *background* term $B_l(p)$ is modeled with following equation:

$$B_l(p) = \frac{\max(z) - \bar{Z}_p}{\alpha} \quad (3.3)$$

where $\max(z)$ is the maximum value of depth value in the depth map and α is the normalization factor(e.g. In 8-bit depth map, $\alpha = 255$).

3.2.2 Neighborhood Property

In addition, *background* term also considers the location of the patch by comparing depth value with neighboring patch across the hole. Axis of the hole is determined by the axis of warping. For simplicity in this paper, warping axis is set to horizontal direction that is commonly used.

For each patch Ψ_p and its depth patch Z_p centered at pixel $p(i, j)$ with i^{th} row and j^{th} column, the direction to the hole is decided either left or right. The patch is divided into left and right half. $NH_l(\Psi_p)$ and $NH_r(\Psi_p)$ indicates the number of hole pixels for left and right side of the patch Ψ_p respectively. This determines the direction of holes where a side with bigger number is regarded as hole direction. Then the depth patch across the hole in the filled region Z_q is found. By comparing the Z_p and Z_q , we can distinguish the patch Ψ_p is laying on either foreground region or background region. The pixel q is found as follows:

$$q(i, j) = \begin{cases} p(i, j + \gamma) & \text{if } NH_r(\Psi_p) > NH_l(\Psi_p) \\ p(i, j - \gamma) & \text{otherwise} \end{cases} \quad (3.4)$$

where γ is a minimum integer value that satisfies $\Psi_q \in \Phi$. Then to utilize this to priority computation the neighborhood feature of *background* term is defined as follows:

$$B_n(p) = \frac{1}{1 + \exp\left(-\frac{\bar{Z}_q - \bar{Z}_p}{\lambda}\right)} \quad (3.5)$$

where λ is a normalization factor. Using difference of depth patch this term favors the background region to be filled prior to foreground region. Figure 3.3 shows the illustration of the principle of neighborhood property. Let us assume that green ellipse is a foreground on yellow background. Given a warped view(Figure 3.3 (a)) and the depth map(Figure 3.3 (b)), red squares in Figure 3.3 (c) and Figure 3.3 (d) indicate the depth patch Z_p to compute priority. In Figure 3.3 (c), $NH_r(\Psi_p) > NH_l(\Psi_p)$ satisfies so that Z_q is marked as blue square while Figure 3.3 (d) is the opposite direction. Depth patch on Figure 3.3 (c) is on the foreground and Z_q is on background while Figure 3.3 (d) is opposite. Therefore $B_n(p)$ in Figure 3.3 (c) has lower value than the one in Figure 3.3 (d). Especially this term is powerful to fill boundary region. This will leads the patches on the boundary region to contain enough background pixels.

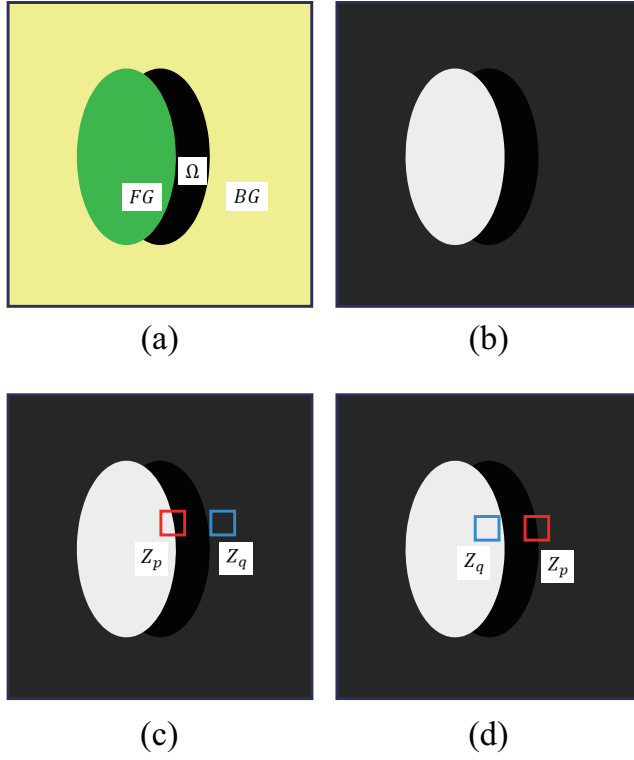


Figure 3.3: Illustration of principle of neighborhood property. (a) A sample warped image. (b) Depth map of (a). (c) Depth patch Z_p on foreground. (d) Depth patch Z_p on background.

3.3 Patch Matching

As we focused on reducing artifacts on boundary patch matching procedure is designed to fill holes well associated with background. To measure the distance of two patches for finding most similar patch, (2.7) was used. But the patches on boundary regions contains both foreground and background pixels. So if Ψ_p is on the boundary, best matched patch Ψ_q with minimum distance is often from irrelevant location and this causes artifacts. To reduce the artifact, we determines whether the patch Ψ_p is on the boundary or not. Then we apply different distance function for two cases. After that, Xu's method [9] is applied to find Ψ_q as a linear combination of candidates patches. How to determine the boundary region and modified distance function is introduced in following subsections.

3.3.1 Boundary Decision

We assumed that the patch is on the boundary if it is the member of set BD . Once a patch with highest priority $\Psi_{\hat{p}}$ is selected, next step is to check if the patch is a member of BD . Depth patch Z_p is used to determine whether the patch is on boundary region. Boundary depth patches contain discontinuity of depth value. With this property we set the two criteria. First, variances of the depth patches is used. Variances of the depth patches on boundary region are larger than those of on background region. But considering only the local variance often determines background patches as boundary patches. This is because background depth also can have depth transitions such as gradual depth change. To solve this problem similarity between Z_p and neighboring depth patches of the horizontal direction are also considered. Set of center pixels of horizontal neighborhood of p with larger area than the width of the patch are defined as $N(p)$ and similarity μ is measured as follows:

$$\mu_{p,q} = \exp\left(-\frac{MSE(Z_p, Z_q)}{\sigma^2}\right) \quad \text{where } q \in N(p) \quad (3.6)$$

where $MSE(., .)$ indicates Mean Squared Error of two patches. The range of $N(p)$ is set to 3 times of the width of $\Psi_{\hat{p}}$ in this paper. Then the set BD is defined as follows:

$$BD = \{\Psi_p : var(Z_p) > T_1, \bar{\mu}_p < T_2\} \quad (3.7)$$

where $var(.)$ and $\bar{\mu}$ indicate variance and mean value of μ respectively with each threshold T_1 and T_2 . In the experiments in this paper, T_1 and T_2 are assigned as 4 and 0.4 respectively.

3.3.2 Distance Function

We address that the distance function to find the most similar patch should be different for boundary region. Distance function (2.7) computes the SSD of both color and depth data between two patches where it often generates artifacts on the boundary region. Suppressing artifacts occurred by boundary patch containing both foreground and background pixels is necessary. Since we want to fill most of the holes on the boundary with background relevant data, different weights are assigned for background and foreground pixels to the patches on boundary.

(2.7) is used to measure the distance between two patches and to find the most similar patch to the target patch for background region. Otherwise if $\Psi_{\hat{p}}$ is on the boundary, pixels inside the $Z_{\hat{p}}$ are classified into background and foreground by K-means algorithm. The patch with only each clustered pixels of background and foreground is defined as $\Psi_{\hat{p},bg}$ and $\Psi_{\hat{p},fg}$ respectively. Different weights are assigned for background and foreground pixels for measuring the distance as follows:

$$dist(\Psi_{\hat{p}}, \Psi_q) = \tilde{d}(\Psi_{\hat{p}}, \Psi_q) + \beta \cdot \tilde{d}(Z_{\hat{p}}, Z_q) \quad \text{if } \Psi_{\hat{p}} \in BD \quad (3.8)$$

where

$$\tilde{d}(\Psi_{\hat{p}}, \Psi_q) = \omega_{bg} \cdot d(\Psi_{\hat{p},bg}, \Psi_q) + \omega_{fg} \cdot d(\Psi_{\hat{p},fg}, \Psi_q), \quad \omega_{bg} > \omega_{fg}. \quad (3.9)$$

ω_{bg} and ω_{fg} are weights for background pixels and foreground pixels of the patch respectively. By giving higher weights to background pixels, best matching patch

is found from background associated region and this enables the holes to be filled properly.

Chapter 4

Experimental Result

In this chapter, we apply the proposed hole filling method and compare with existing method. We test our algorithm to number of images with both stereoscopic warping and camera matrix utilized warping processes. To verify the hole filling performance of the algorithm, rest of the processes including pre-process of hole filling are equally applied to other existing method. In the experiments, we set the patch size to 9x9. In (3.9), weights for background and foreground pixels of patches on the boundary are assigned as $\omega_{bg} = 9$ and $\omega_{fg} = 1$. As we introduced in Chapter 2, the proposed method is suitable for generating both free-viewpoint view and stereoscopic view. Usually in application, baseline distance for generating free-viewpoint view using camera matrix is larger than baseline distance for generating stereoscopic view. Our experiments is conducted on both small baseline distance and large baseline distance to verify that the proposed method is suitable for both stereoscopic view and free-viewpoint view synthesis. In both stereoscopic view and free-viewpoint view generation, single reference image and depth map is used. Also the proposed method is compared to Daribo's depth aided in-painting method.

4.1 Experiments for generating stereoscopic view

We used Microsoft Research data set [7] to generate stereoscopic views. Size of the image is 1024×768 pixels and the baseline distance is set to 60 which is 6% of the width of image. Figure 2.1 is an input of one of the samples we used where Figure 3.2 (a) shows the pre-processed image for hole filling. Another set of input is shown in Figure 4.1.



Figure 4.1: A set of input sample for experiment. (a) Reference color image. (b) Pre-processed image for hole filling.

As shown in Figure 4.2, the proposed hole filling method generates the synthesized view with better quality. Especially on the boundary region, the proposed method shown in Figure 4.2 (b) and (d) filled the holes accurately with background associated data while the existing method shown in Figure 4.2 (a) and (c) shows artifacts along the boundary. This is achieved by propagating background texture to holes and fills background associated data on the boundary region.

Enlarged images in Figure 4.3 clearly shows the difference between the result of existing method and proposed method. Annoying artifacts on boundary produced by existing method is suppressed in proposed method. This helps the viewer to feel the visual comfort.



(a)



(b)

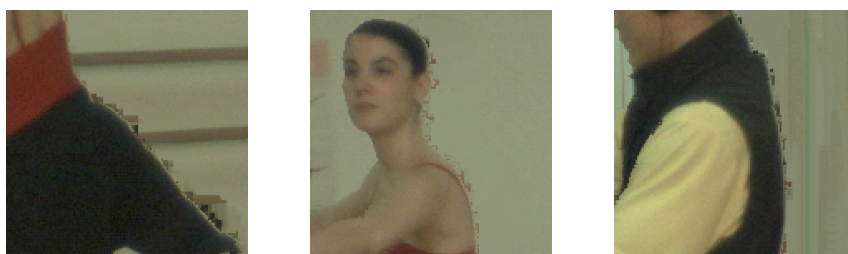


(c)

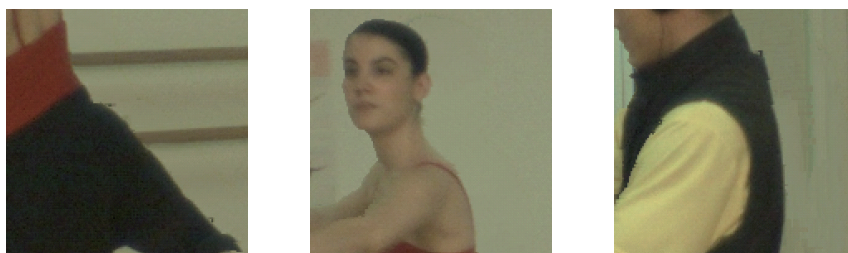


(d)

Figure 4.2: Result images. (a), (c) Result images using existing method. (b), (d) Result images using proposed method.



(a)



(b)

Figure 4.3: Enlarged images. (a) Enlarged images of result using existing method. (b) Enlarged images of result using proposed method.

4.2 Experiments for generating free-viewpoint view

For generating free-viewpoint views, Microsoft Research data set is also used. Also in this section an objective evaluation is measured to verify the quality increment of proposed method. The data set provides images from 8 different viewpoints with depth maps and camera parameters. Using an image and corresponding depth map from camera 4, virtual view from camera 3 viewpoint is synthesized for the test. The ground truth view of camera 3 is compared to evaluate the synthesized view. 5 randomly selected sequences of both Ballet and Breakdancers are tested.

Image captured from camera 4 is warped to camera 3. Input set of hole filling and evaluation combines the warped view and image captured from camera 3. Figure 4.4 shows one of the input set and result images of both existing method and proposed method. For free-viewpoint synthesized image also shows similar result to stereoscopic view. Result of existing method contains artifacts along boundary while result of proposed method preserves boundary of foreground and background by filling background associated data. This enhances the visual comfort.

For an objective evaluation, PSNR is measured. We considered both PSNR for entire image and PSNR for only hole region to compare two methods. Table 4.1 and Table 4.2 shows the comparison result of Ballet and Breakdancers set respectively. Diff indicates the difference of PSNR of proposed method and existing method. Ballet sequence contains average 15% of hole area in the entire image, while 9% of hole region appears in Breakdancers set. Proposed method achieves an increment of PSNR. Average increment of PSNR for hole region is 0.6 and 0.5 for each set.

We observe that proposed method enhances the quality of synthesized view. Also the proposed method is applicable for both stereoscopic and free-viewpoint view generation.



(a)



(b)



(c)



(d)

Figure 4.4: Input set and result. (a) Warped image (camera 4 to camera 3). (b) Reference image. (c) Result image of existing method. (d) Result image of proposed method.

Table 4.1: PSNR comparison of Ballet sequence.

Sequence Number	Entire image PSNR			Hole region PSNR		
	Proposed	Existing	Diff	Proposed	Existing	Diff
Ballet 1	29.102	28.773	0.3288	23.972	23.331	0.6408
Ballet 2	29.760	29.664	0.0956	25.432	25.205	0.2278
Ballet 3	29.348	28.544	0.8042	24.341	22.795	1.5460
Ballet 4	29.597	29.377	0.2198	25.035	24.546	0.4884
Ballet 5	29.401	29.343	0.0587	25.044	24.902	0.1414

Table 4.2: PSNR comparison of Breakdancers sequence.

Sequence Number	Entire image PSNR			Hole region PSNR		
	Proposed	Existing	Diff	Proposed	Existing	Diff
Breakdancers 1	31.536	31.398	0.1380	28.820	27.968	0.8516
Breakdancers 2	31.474	31.412	0.0618	29.108	28.655	0.4531
Breakdancers 3	30.976	30.339	0.6367	26.740	24.439	2.3016
Breakdancers 4	30.959	30.906	0.0527	25.700	25.503	0.1968
Breakdancers 5	31.375	31.260	0.1148	26.855	26.361	0.4938

Chapter 5

Conclusion

An efficient hole filling method of DIBR has been presented in this paper. Since foreground objects are warped so that the holes appear between the foreground and background, holes are likely to be filled with background data to achieve visual comfort. Exemplar based image in-painting scheme is modified to fill the holes properly. The conventional method causes problem of filling holes with background data especially on the boundary region.

We designed a priority computation that fills background region prior to boundary region using depth map. Local and neighboring properties of depth value are considered to assign higher priority to background. Then, to fill boundary region with background associated data, we checked whether the patch is on the boundary or not by our boundary decision. Decision is based on local depth variance and similarity of neighboring depth information. By weighting background region for boundary patches, we solve the problem of occurrence of boundary artifact.

Experimental results show that the proposed method reduces boundary artifacts and generate natural synthesized image compared to existing method. The proposed method is expected to gain the visual comfort by filling boundaries naturally. This

method can be utilized in various 3D display system to generate novel virtual views such as 2D-to-3D conversion and free-viewpoint television.

Bibliography

- [1] C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV," *International Society for Optics and Photonics.*, In Electronic Imaging 2004., pp. 93-104.
- [2] L. Zhang, and W. J. Tam, "Stereoscopic image generation based on depth images for 3D TV," *IEEE Transactions on Broadcasting*, vol. 51, no. 2, pp. 191-199, 2005.
- [3] K. J. Oh, S. Yea, and Y. S. H, "Hole filling method using depth based in-painting for view synthesis in free viewpoint television and 3-d video," in *Picture Coding Symposium (PCS) 2009* , 2009. IEEE, pp. 1-4.
- [4] K. Luo, D. Li, Y. Feng, and M. Zhang, "Depth-aided inpainting for disocclusion restoration of multi-view images using depth-image-based rendering," *Journal of Zhejiang University SCIENCE A.*, vol. 10, no.12, pp. 1738-1749, June 2009.
- [5] I. Daribo, "Depth-aided image inpainting for novel view synthesis," in *IEEE International Workshop on Multimedia Signal Processing (MMSP) 2010* , 2010. pp. 167-170.
- [6] A. Criminisi, P. Perez, and K. Toyama "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200-1212, 2004.

- [7] Microsoft Research, <http://research.microsoft.com/en-us/downloads/5e4675af-03f4-4b16-b3bc-a85c5bafb21d/>
- [8] X. Lu, F. Wei, and F. Chen, "Foreground-object-protected depth map smoothing for DIBR," in *IEEE International Conference on Multimedia and Expo (ICME) 2012*, 2012. pp. 339-343.
- [9] Z. Xu, and J. Sun, "Image inpainting by patch propagation using patch sparsity," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1153-1165, 2010.
- [10] A. Wong, and J. Orchard, "A nonlocal-means approach to exemplar-based inpainting," in *IEEE International Conference on Image Processing (ICIP) 2008*, 2008. pp. 2600-2603.

국문 초록

3차원 디스플레이 시스템에서 가상 시점의 영상을 생성하는 방법 중, 깊이 영상 기반 렌더링(Depth Image-Based Rendering, DIBR)이 주목을 받고 있다. 하지만 이 과정에서 나타나는 새롭게 드러나는 부분인 홀을 채우는 방법은 어려운 과제로 남아있다. 홀을 채우는 과정은 가상 시점에서 생성되는 최종 결과물의 품질을 좌우하는데 가장 중요한 역할을 한다. 영상 내 정보가 없는 부분을 채우는 방법인 이미지 인페인팅 기법이 홀을 채우는데 사용된다. 이 때 일반적인 인페인팅 방법을 적용하면 전경과 배경의 경계면에서 눈에 거슬리는 아티팩트가 생성된다. 본 논문에서는 경계면을 효과적으로 채우는 DIBR의 홀 필링 방법을 제안한다. 제안된 방법은 깊이 정보를 이용하여 배경 영역을 경계 영역보다 우선되게 채운다. 또한 패치를 채울 때 해당 패치가 경계 영역 위에 놓인 여부를 판단하고 경계 영역에 놓인 패치의 화소는 전경과 배경화소로 군집화한다. 배경 화소로 판별된 화소는 전경 화소보다 더 큰 가중치를 부여하여 가장 비슷한 패치를 찾는다. 이러한 과정은 홀이 배경과 어울리는 정보로 위지도록 유도하면서 경계면의 아티팩트 생성을 억제한다. 실험을 통해 제안된 방법이 효과적으로 아티팩트 생성을 줄이는 것을 확인하였다.

주요어: 깊이 영상 기반 렌더링, 가상 시점 영상 생성, 영상 인페인팅, 홀 필링

학번: 2014-21690