2016 IEEE International Symposium on Mixed and Augmented Reality Adjunct Proceedings

Diminishing Real Objects and Adding Virtual Objects Using a RGB-D Camera

Hajime Sasanuma*
Chiba University

Yoshitsugu Manabe†
Chiba University

Noriko Yata‡
Chiba University

ABSTRACT

AR technology is often used in applications that simulate an arrangement of furniture. These applications superimpose CG furnitures on any place. These applications cannot replace the CG furniture with a real furniture if there is a real furniture. To solve this problem, this paper proposes a method that can erase real object from the real environment (Diminished Reality) and add virtual object to the real environment (Augmented Reality) including the region of erased object using RGB-D camera.

Keywords: Augmented Reality, Diminished Reality, RGB-D camera.

Index Terms: H.5 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities

1 INTRODUCTION

AR (Augmented Reality)[1] technology is used to various products and fields. For example, there is an application that can simulate an arrangement of furniture by using AR technology[2]. User puts the marker in where user wants to display a CG furniture, then a virtual furniture is seen through the display. However, the application cannot be used in where a real furniture is already placed because the exact position and posture of a virtual furniture cannot be obtained.

This research proposes a method that erase a real object and display the virtual object to solve this problem. To erase a real object, DR (Diminished Reality)[3] technology is used. The algorithm of the technology is to inpaint the region of a replacing object that the user wants to erase based on information of around region of the object. By combining this technology and AR technology, the replacing object is erased and the virtual object with appropriate posture is displayed in selected position.

Siltanen proposed a method that erase a real object and add virtual object using a inpainting method that uses pixels around the target region to inpaint that region [4]. To extract a wall and floor of the room, user has to put a marker on the floor and select a border line of the wall and floor. The target region must be defined in three-dimensions. Therefore, the floor and wall must be solid-colored because the pixels used to inpaint are obtained from around the target region. This paper proposes the method that can apply to background with texture, and a marker and border line selection are unnecessary.

Kawai et al. proposed a diminished reality method applying to background with texture [5]. By an exemplar-based image inpainting method, the target region which includes the background is inpainted. This method is independent of texture of background. However, the calculation cost for image rectification is high. In addition, if the background wall is solid-colored, the proposed method is not available because Visual-SLAM which is method to segment the background surfaces uses some feature points of the background.

Meerits et al. proposed a diminished reality method using RGB-D camera [6]. This method uses RGB-D camera, RGB camera and a marker. The main scene inpainting the target region is captured by RGB camera, and RGB-D camera takes the occluded area from different view. The three-dimensional structure is obtained by RGB-D camera, and that information is converted by aligning between two cameras using the marker and used to inpaint the target region of the main scene. By doing so, the occluded area doesn ' t require to be planar and the background structure can change and move freely. However, if the occluded surface and the replacing object are in close contact, this method is not available because the image of occluded area cannot be taken.

2 PROPOSED METHOD

Figure 1 shows the processing flow of the proposed method. Overview of this research is to erase a replacing object and to display a virtual object with appropriate posture by estimating the background of the replacing object. That background is floor, wall and ceiling because this method is assumed to be used in the room.

RGB and depth images are obtained using RGB-D camera. Depth image is used to estimate surfaces of wall or floor of the room and to keep geometric consistency of displayed virtual object. User selects a rectangular region including a replacing object in the RGB image. This region is a target region (Figure 2). A region around the target region is divided for each surface corresponding to floor and wall, and the orientation and texture information of each surface is obtained in initializing process. Using these information, the replacing object is erased by inpainting process to the target region. User selects a position which user wants to display a virtual object in the RGB image which the replacing object is erased, and the virtual object is superimposed using the position information of the appropriate background surface. These processes are only executed in first frame. The position and texture information of each surface which constructs the background of the replacing object is updated from the second frame with the information of the previous frame.

2.1 Initializing process

Initializing process consists of three processes consisting of the region segmentation, the estimation of background and the obtaining texture image. User selects the target region that includes the replacing object that the user wants to erase in RGB image. To inpaint the target region, orientation and texture information of background of the replacing object is necessary. So, the three-dimensional point cloud obtained from the depth image is segmented with the normal vector of each point, and the point groups on the background surfaces of the replacing object are extracted by excluding some unnecessary point groups. From these point groups on the background surfaces, the orientation and texture information is calculated. At the end of this process, three-dimensional vertex coordinates of an around region of the replacing object are calculated from the information of background of the replacing object. These calculated vertex coordinates are used in the update process.

*e-mail:aafa2765@chiba-u.jp
†e-mail:manabe@faculty.chiba-u.jp
‡e-mail:yata@chiba-u.jp

978-1-5090-3740-7/16 $31.00 © 2016 IEEE
DOI 10.1109/ISMAR-Adjunct.2016.48

IEEE computer society

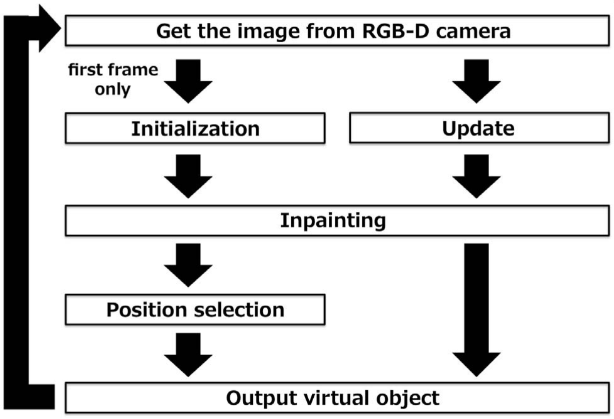


Figure 1: Processing flow

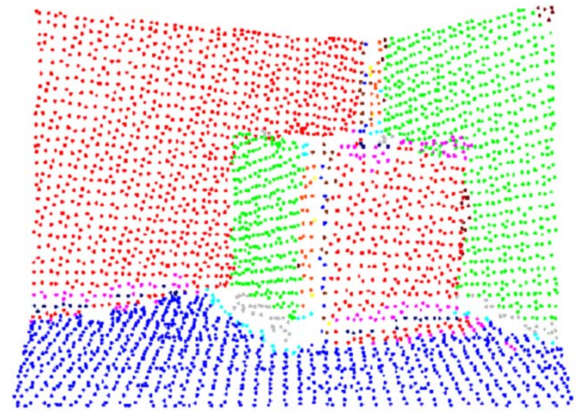


Figure 3: Classified point cloud

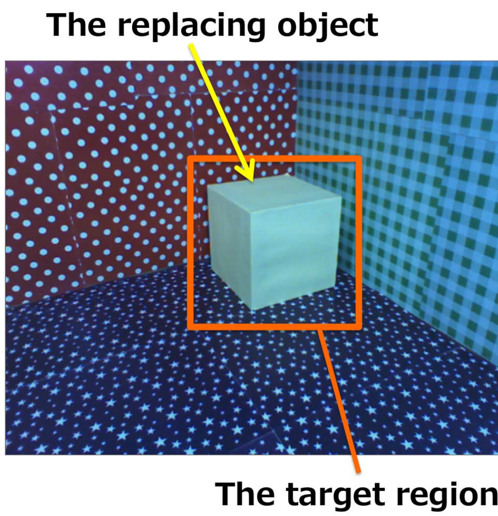


Figure 2: Description of the definition

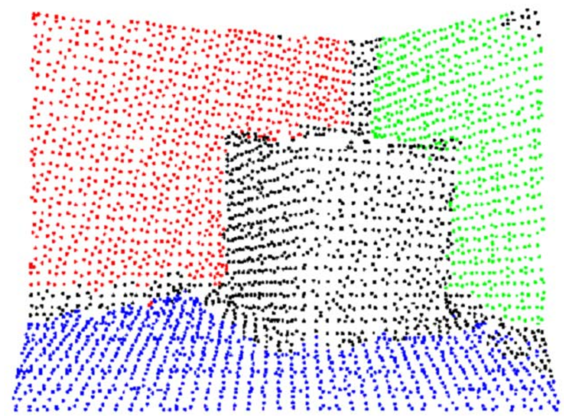


Figure 4: Point cloud of background surface

Region segmentation The three-dimensional point cloud data is obtained from the depth image. The normal vector of each point is calculated by using around points of the target point. Then, all points are segmented by classifying the point group in accordance with the value of the normal vector. Figure 3 shows a classified point cloud, and the colors express the classified each points group [7].

Estimation of background region After the region segmentation process, there are some point groups which are same class but exist on different planes. To extract the point groups on the background surfaces of the replacing object, some unnecessary point groups are excluded. The extracted point groups exist outside the target region. The exclusion conditions are that the number of points is larger than threshold and largest in the class with same normal vector. In Figure 4, the colors express each background surface of the replacing object. The black means the excluded points.

The equation of each background surface is calculated by using the coordinate information of the point group.

Obtaining texture image To inpaint the target region, texture images are needed for compensation of the region. In each of the segmented background region, rectangular area of a certain size (30 $\times$ 30px) is selected from each segmented region which doesn't include the target region. Then, the texture image is obtained from selected region (Figure 5). And the coordinate of the center point in each selected region is used in the update process.

2.2 Update process

Update process updates the target region and information of the background of the replacing object.

Update of target region To update target region, feature points are extracted from this frame and previous frame. ORB (Oriented Fast and Rotated Brief)[8] is used as feature detector. Using these points, a three-dimensional transformation matrix to the current frame from the previous frame is calculated. Then, the transformation matrix is applied to the vertex coordinates of target region, and the target region is updated.

Update of background information The coordinates information that is calculated in section 2.1 are defined as coordinates of representative points in surfaces which are estimated to be the background of the replacing object in previous frame. These coordinates are converted to the coordinates of the current frame by using the transformation matrix. Rectangular regions of a certain size (30 $\times$ 30px) around the converted point are obtained. The information of background surfaces of the replacing object are updated by using that rectangular regions. That information includes equations and texture images of each background surfaces. The reason for applying only to the rectangular regions of a certain size is to improve the processing speed.

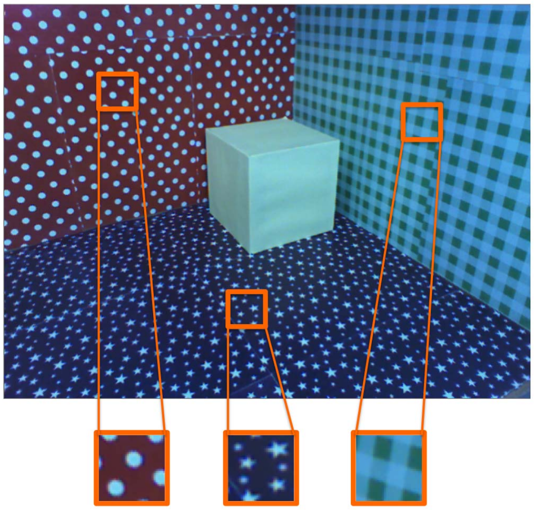


Figure 5: Texture images for inpainting

2.3 Inpainting process

Inpainting is the process of reconstructing lost or deteriorated parts of images and videos [9]. There are two types of approaches to reconstruct. One is to get the actual background by using a plurality of cameras. The other is to estimate around the deteriorated parts. This inpainting process is very important for DR method and this study uses the latter approach.

Each of the background plane is reproduced in the virtual space by using the equations of background surface. The texture images are mapped to the background surfaces in the virtual space. Thus, it is possible to add texture information to the background plane. Then, the virtual space image is superimposed on the input image, and the inpainting processing is realized.

2.4 Selecting virtual object display position

User selects an arbitrary position in the image which the replacing object is erased to appoint a position displaying a virtual object. To display a virtual object with the correct posture, three-dimensional coordinate and normal vector of the position are necessary. The coordinate is obtained from the depth image and the normal vector is calculated by using around points of the position. If the position is inside of the target region, the three-dimensional coordinate and the normal vector are obtained from the information of the background region.

2.5 Superimposing virtual object

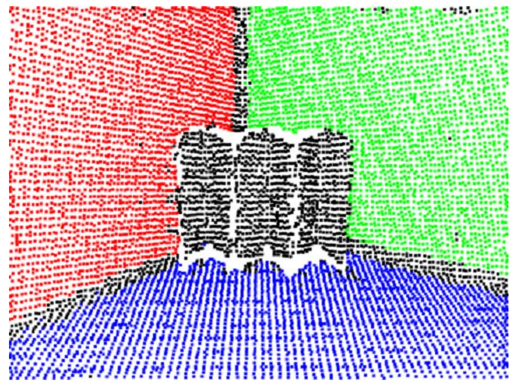
Using the three-dimensional coordinate and the normal vector information which is obtained in section 2.4, the virtual object is superimposed with correct position and posture. However, the posture of the virtual object may not fit the intention of the user. Therefore, the process which the user adjusts the posture by manual operation is added.

3 Experiments

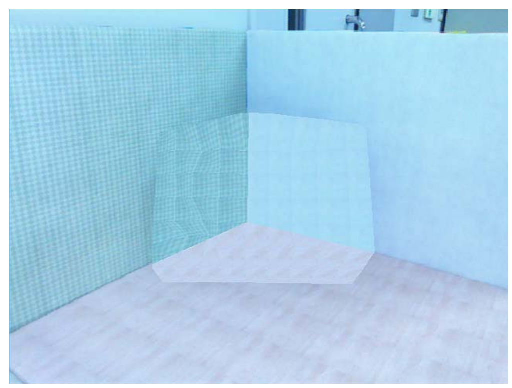
In experiments, we use a PC with Core i7-6700 3.4 GHz CPU, 8 GB of memory for input images with resolution 640 $\times$ 480 captured by a RGB-D camera (Microsoft Kinect).

Figure 6 shows results of some processes. The proposed method successfully segmented the background surfaces of the replacing object (Figure 6(b)). Figure 6(c) shows the inpainting result for the target region using the textures obtained from the around target region. Figure 6(d), 7 shows that the virtual object was superimposed

(a) input image



(b) Segmented point cloud data



(c) Inpainted image

(d) Virtual object superimposed image

Figure 6: Results of proposed method

Figure 7: Different viewpoint result image

with correct posture and position which was selected from the RGB image including the inpainted region. The inpainted region in Figure 6(c) provide a sense of incongruity. It is because of the texture images are obtained with disregard the pattern of the background surface design.

4 Conclusion

This paper proposed a method that can erase a real object from real environment and add a virtual object to real environment including the region which the replacing object is erased by using a RGB-D camera. To erase the replacing object, a region includes the replacing object is inpainted by using texture and orientation information which is obtained from around the region. In future work, we will develop a method for obtaining the texture information which inpaints the target region by recognizing the pattern of texture of the background surface, and solve the problem that this method cannot use in the scene where there are other furnitures around the target furniture.

References

[1] Asahi Suzuki, Yoshitsugu Manabe, Noriko Yata, "Design of an AR Marker for Cylindrical Surface," Proc. of International Symposium on Mixed and Augmented Reality, pp.293-294, 2013

[2] IKEA Home Page : http://www.ikea.com/jp/ja/

[3] J. Herling, W. Broll, "PixMix: A real-time approach to high-quality diminished reality," Proc. of International Symposium on Mixed and Augmented Reality, pp.141-150, 2012

[4] S. Siltanen, "Diminished reality for augmented reality interior design," The Visual Computer, pp.1-16, 2015

[5] N. Kawai, T. Sato, N. Yokoya, "Diminished reality considering background structures," Proc. of International Symposium on Mixed and Augmented Reality, pp.259-260, 2013

[6] S. Meerits, H. Saito, "Real-Time Diminished Reality for Dynamic Scenes," Proc. of International Symposium on Mixed and Augmented Reality, pp.53-59, 2015

[7] K. Matsumoto, F. De Sorbier, H. Saito, "Real-time enhancement of RGB-D point clouds using piecewise plane fitting," Proc. of European Workshop on Visual Information Processing, pp.1-6, 2014

[8] Q. Yanyan, X. Hongke, C. Huiru, "Image feature points matching via improved ORB," Proc. of International Conference on Progress in Informatics and Computing, pp.204-208, 2014

[9] J. Herling, W. Broll, "Advanced Self-contained Object Removal for Realizing Real-time Diminished Reality in Unconstrained Environments," Proc. of International Symposium on Mixed and Augmented Reality, pp.207-212, 2010