Real-Time Diminished Reality using Multiple Smartphones

Toshihiro Honda*
Keio University

Takuya Inoue†
Keio University

Hideo Saito‡
Keio University

Abstract

In this paper, we present a system for real-time Diminished Reality with multiple smartphones. In this system, we assume multiple smartphones capture the same scene that is occluded by obstacles. Areas of the obstacles are extracted from each camera image replaced with image of the hidden areas which are captured using different viewpoint camera. In the proposed method, we suppose that the target scene can be approximated as a plane. Therefore, we compute homography matrices between each camera image by using natural features. Then, obstacle area which is not approximated as a plane can be removed by synthesizing the image warped with the homography matrix and the user viewpoint image. We can perform real-time processing because we send each camera image to PC which returns obstacle-removed images at every frame. We experimentally demonstrate the effectiveness of the proposed method using three viewpoint images.

Keywords: Diminished Reality, Multiple Smartphones, Real-Time, Homography

1 Introduction

Diminished Reality(DR) is a technique for removing obstacles and replacing the area with a proper target scene image. Enomoto et al. proposed DR using multiple web cameras[1]. In their system, they allow each web camera connected each PC to share data through a network, and compute homography matrices between each camera using ARTag, then remove obstacles by blending process.

In the proposed method, no ARTag is required because homography matrices are computed in real-time by matching natural feature points. We also apply median process in order to determine the color of the replaced pixels, which is more simple algorithm than blending process. Users can move freely by using smartphones as a device for capturing camera image.

2 Proposed System

Figure 1 shows our proposed system. We use multiple smartphones in our system. In this paper, we introduce our method using three smartphones.

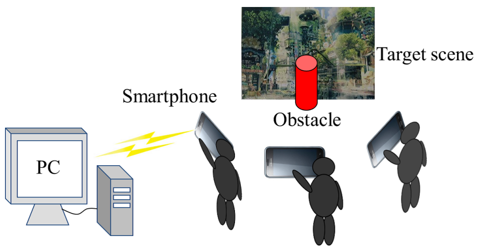


Figure 1: Experimental system.

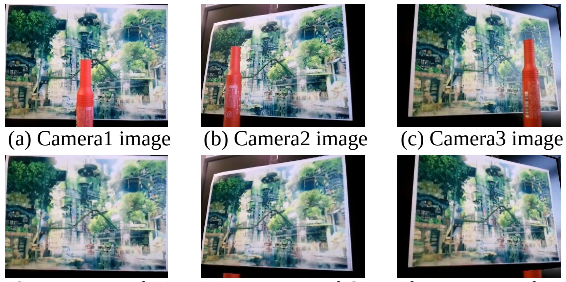
*e-mail:t-honda@hvrl.ics.keio.ac.jp
†e-mail:inoue@hvrl.ics.keio.ac.jp
‡e-mail:saito@hvrl.ics.keio.ac.jp

C_1, C_2, C_3 are camera images captured by the three smartphones. We detect natural feature points of C_1, C_2, C_3 using CenSurE[2], then describe feature values of the points using BRIEF[3]. The detected feature points in each camera are matched with the feature points detected in the other cameras based on the descriptors. According to the corresponded feature points, we can compute homography matrices $\mathbf{H}_{21}$ (betweeen C_1 and C_2) and $\mathbf{H}_{31}$ (between C_1 and C_3). In this case, C_1 is a basis image that relates C_2 and C_3, so that the obstacles in C_1 can be replaced with the object scene images captured in C_2 and C_3.

C_2 is warped to an image which is seen from the smartphone capturing C_1 by $\mathbf{H}_{21}$. Similarly, C_3 is warped by $\mathbf{H}_{31}$. At this time, locations of obstacles are different among the three images because obstacles are not approximated as a plane. Therefore, pixel values of the other two images in the location of obstacles which are projected on one of three images are pixel values of the target scene. Because of this, we can get the image removing obstacles if we use pixel values in the middle of ones of three images for every pixel.

3 Experimental Result

Figure 2 shows the result of removing an obstacle by using this system. The experimental platform was implemented on GALAXY S and 3.4GHz Intel Core i7-2600 desktop with 3.49GB RAM. The image size is 3x240. We can see that the obstacle is removed and the entire background poster can be seen. In addition, proposed system runs in approximately 5-6 fps, which can be approximately considered as real-time.



(a) Camera1 image (b) Camera2 image (c) Camera3 image
(d) DR image of (a) (e) DR image of (b) (f) DR image of (c)

Figure 2: Removing an obstacle.

4 Conclusion

Through experiments aimed at the real scene, we confirmed that we could remove obstacles in real-time.

References

[1] Akihito Enomoto and Hideo Saito. Diminished Reality using Multiple Handheld Cameras. ACCV'07 Workshop on Multi-dimensional and Multi-view Image Processing, pp. 130-135, 2007.

[2] Motilal Agrawal, Kurt Konolige and Morten Rufus Blas. CenSurE: Center Surround Extremas for Realtime Feature Detection and Matching. ECCV, volume 5305, pages 102-115, 2008.

[3] Michael Calonder, Vincent Lepetit, Christoph Strecha and Pascal Fua. BRIEF: Binary Robust Independent Elementary Features. ECCV, volume 6314, pages 778-792, 2010.