Object-wise Individual Appearance Manipulation with Layer Detection

Takahiro Nagata[1,*], Toshiyuki Amano[1]

[1] *Wakayama University, 930, Sakaedani, Wakayama-shi, Wakayama, Japan*

Abstract
Appearance Manipulation enables to change the perceptual color, texture, and shape with illumination projection. However, it is still unclear how to apply the manipulation for each object independently, not unique manipulation for the whole scene. This paper proposes a method to independently apply appearance manipulation to foreground and background with the layer in a scene which can detect from the pixel correspondence among two projector-camera systems. Furthermore, our method removes cast-shadow-like illumination unevenness created by the foreground from its layer detection.

Keywords
Spatial augmented reality, Projector-camera system, Human-centered computing

1. Introduction

The Shader Lamps, which enabled color manipulation of the buildings by texture mapping on white walls [1], presented the potential of spatial augmented reality (SAR) through light projection. Since then, various techniques have been proposed for SAR applications [2].

Unlike conventional projection mapping, Amano et al. proposed an alternative projection technique to manipulate apparent object color with illumination projection in a projector-camera feedback manner [3]. It has another potential to hack an appearance of the real world and our visual perception. Currently, many applications of appearance manipulation are proposed [4, 5, 6].

However, they apply uniform appearance manipulation for the whole area of the scene. The object-wise individual manipulation pushes the boundary of appearance control and potentially other applications ever attempted. This study aims to extend the concept of appearance manipulation to enables object-wise individual appearance manipulations. This paper specifically focused on the detection of each object region of in the scene which consists of multiple objects and exploring techniques to manipulate object appearance for each object individually. For instance, this technique could be applied to illumination in theaters, amusement parks, photography, etc.

Semantic segmentation [7] is a key technology for computer vision, enabling the precise detection of each object with a label. However, it is not guaranteed to work correctly under the illumination projection, which changes the apparent color or texture. Meanwhile, the appearance manipulation system owns a pixel correspondence between the camera and projector which can validate the assumptions of placement of the object in which layer.

When a foreground object exists in a scene, cast-shadows occur on the background object. Solving this problem has been investigated as a longstanding research challenge in SAR. Sukthankar et al. [8] proposed a method that can remove shadows caused by occluding objects by using two projectors in overlap projection. Audet et al. [9] achieved it with tracking for dynamic scenes, and Flagg et al. [10] proposed another adaptive technique with an IR camera. However, these methods aim to display a given video source on the screen and do not involve appearance manipulation.

This paper proposes a method that discriminates between foreground and background objects by independently working two projector–camera systems and achieving separate appearance manipulations for each object. Since Appearance Manipulation comprises projectors and cameras, its system owns pixel correspondings among cameras and projectors. This paper attempt to identify the shadow areas caused by the foreground. Furthermore, we address adjusting the light intensity in the superimposed regions to eliminate the brightness difference without unaffected by changes in appearance.

2. Related Work

Amano et al. achieved appearance manipulation of objects by using a projector-camera feedback. In this method, a refrectance estimation is introduced to generate control reference for Model Predictive Control (MPC) [3]. Figure 1 illustrates the block diagram of appearance manipulation. The main processing step involves firstly creating an estimated appearance image C_{est} under white projection from the captured image C and the previous step's projected image P. Next, the desired image processing is applied to C_{est} to create the target image R. Then, the latest projected image P is adjusted by the difference between C

APMAR'23: The 15th Asia-Pacific Workshop on Mixed and Augmented Reality, Aug. 18-19, 2023, Taipei, Taiwan
*Corresponding author.
s246198@wakayama-u.ac.jp (T. Nagata); amano@wakayama-u.ac.jp (T. Amano)
https://kmitd.github.io/ilaria/ (T. Amano)
0009-0005-0542-6267 (T. Nagata); 0000-0003-4146-1375 (T. Amano)

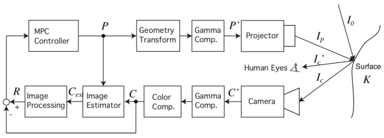


Figure 1: Feedback control of a projector-camera system.

and R in consideration of robustness with the MPC controller. Finally, appearance manipulation is realized with the illumination projection from the projector after the geometrical deformation based on pixel correspondence.

This study aims to achieve independent appearance manipulation for both the foreground and background objects using layers in the scene that can be detected from the pixel correspondence between the two projector-camera systems. Additionally, we aim to remove brightness variations caused by cast shadows created by foreground objects in the projection.

3. Proposed method

We assume a foreground object is positioned in front of the background within a two pairs of projectors (Prj1, Prj2) and cameras (Cam1, Cam2), as shown in Figure 2. In this situation, we have three projection states: Region I, where the foreground obstructs one projection without affecting the other. For this region, the conventional method can be applied. In Region Ⅱ, overlapping projection causes over-illumination requiring novel compensation techniques, and in Region Ⅲ, where no projection can reach and cast shadows can not be removed. In this paper, we focus on region discrimination and illumination suppression in Region Ⅱ to address cast-shadow-like projection unevenness.

3.1. Layer-based Discrimination

To identify the aforementioned regions (I, Ⅱ, and Ⅲ) as well as foreground or background in the Cam1 and Cam2 images, this study uses pre-acquired pixel mapping with both the foreground and background planes. The pixel mapping is stored with a look-up table that describing the pixel correspondings between camera and projector.

When the captured pixel coordinates are (x_c, y_c) and the projected pixel coordinates are (x_p, y_p), pixel mapping from the camera to the projector is denoted as

$$(x_p, y_p) = f_{C2P}(x_c, y_c), \tag{1}$$

and those obtained in the foreground and background planes are denoted as f_{fC2P} and f_{bC2P} respectively, as shown in Fig. 2.

In this study, foreground and background are discriminated by comparing the geometric calibration results of the actual scene with the intermediate value

$$(x_m, y_m) = \{f_{fC2P}(x_c, y_c) + f_{bC2P}(x_c, y_c)\}/2 \tag{2}$$

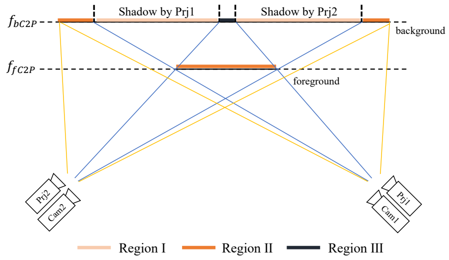


Figure 2: Scene placement and some projection regions.

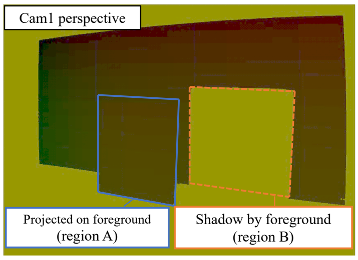


Figure 3: Pixel mapping showing with x and y coordinate values visualized in R and G.

of these pixel mapping as a threshold value.

Figure 3 shows the pixel mapping $f_{C_1 2P_2}$ obtained from Cam1 with Prj2. Region B in Cam1 has a shadow of projection caused by a foreground object. With another pixel map $f_{C_1 2P_1}$ obtained with Prj1, we can distinguish the nature of each region (x_c, y_c) as follows:

- Region I: One of pixel map has corresponding.
- Region Ⅱ: Both pixel maps have corresponding.
- Region Ⅲ: Neither of the pixel maps has corresponding.

It should be noted the pixel mapping is acquired by gray code projection.

3.2. Illumination Suppression

In this section, we briefly present an illumination suppression method proposed by Uesaka et al.[11], specifically designed to address the illumination overlapping region denoted as Region Ⅱ in Section 3.1. Given a captured image $\mathbf{C}_1$ from Cam1, ambient light $\mathbf{C}_0$, and projected light $\mathbf{P}_1$, the reflectance K of the object surface can be estimated to be

$$\hat{K} = diag[\mathbf{C}_1./(M\mathbf{P}_1 + \mathbf{C}_0)], \tag{3}$$

where $M \in \mathbb{R}^{3\times3}$ denotes the color mixing matrix between projector and camera, ./ denotes component-wise division. However, in the overlapping regions, the reflectance K is overestimated due to the projected light from the two units, resulting in overprojection. Therefore, considering the projection $\mathbf{P}_2$ from Prj2, $\mathbf{C}_1$ can be estimated as

$$\mathbf{C}_1 = \cos(\theta_1)KM\mathbf{P}_1 + \cos(\theta_2)KM\mathbf{P}_2 + K\mathbf{C}_0 \tag{4}$$

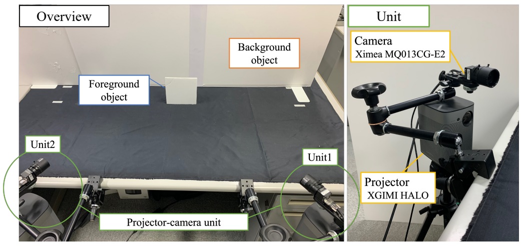
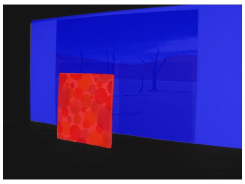


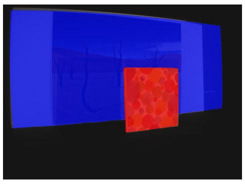
Figure 4: Arrangement of experimental setup.

(a) Captured image C_1

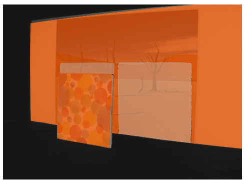
(b) Captured image C_2



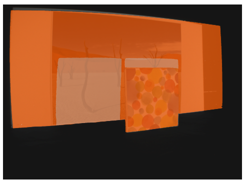
(c) Cam1 perspective



(d) Cam2 perspective



(e) Cam1 perspective



(f) Cam2 perspective

Figure 5: (a)(b)Captured images taken from cameras, (c)(d)Red is the foreground, blue is the background, (e)(f)Region I, II, and III and corresponding color in Fig. 2.

where θ_1 and θ_2 are the angles between the camera and the surface normal respectively, accounting for the attenuation based on Lambert's cosine law. In this case, the relation

$$KM\mathbf{P}_2 = \frac{\cos(\theta_1)}{\cos(\theta_2)} KM\mathbf{P}_1 \tag{5}$$

can be established because we know that $\mathbf{P}_1 = \mathbf{P}_2$ when $\theta_1 = \theta_2$ in the study by Shimana et al. [5]. Based on this, the reflectance K can be estimated as

$$\hat{K} = \frac{\cos(\theta_1)}{\cos(\theta_1) + \cos(\theta_2)} diag[\mathbf{C}_1./(M\mathbf{P}_1 + \mathbf{C}_0)]. \tag{6}$$

Similarly, by utilizing the captured image $\mathbf{C}_2$ from Cam2, our calculation process enables the accurate estimation of reflectance using only own unit information. A key advantage of our approach is that the reflectance estimation in each unit is performed solely using its own captured image and projection image. In this study, for simplicity, we approximate $\theta_1 = \theta_2$ and suppress overprojection by estimating reflectance in each system and correcting for it.

4. Results

4.1. Experimental setup

Figure 4 shows that styrofoam board coated with white paper regarded as object with Lambert reflection was placed to get pixel mapping of foreground and background. The distance between their plane was 150mm. The unit is composed so that the shooting and projection directions of the camera and projector are aligned. Afterward, we performed an experiment on matte photo paper printed illustration to validate the proposed method.

We performed foreground and shadow region identification. Figure 5 demonstrated that foreground or background and shadow regions were correctly discriminated.

4.2. Manipulation results

From the results of the previous section, we controlled the projection in region II and performed separate appearance manipulation for the foreground and background. The results are shown in Figure 6. From the manipulation results of the color chart (matte photo paper) shown in upper row, we confirmed that independent image processing as bright saturation, monolize and color phase applied to the foreground and background. As shown in middle row, the results have potential applications in stage effects. Moreover, we verified that an object like origami with specular reflection shown in bottom row can be correctly manipulated when it doesn't reflect to the camera or viewer's perspective. The brightness difference problem is improved compared to that under white illumination, but it has not been fully eliminated.

5. Discussions

5.1. Over illumination supression

To evaluate the effectiveness of our illumination suppression in overlapping areas, we compared our method using Equation (6) with the conventional method (Figure 7). Observing the boundary between regions I and II, we can see that the boundary between the regions is clearly visible in Fig. 7c, while the brightness difference is improved in Fig. 7d. However, Fig. 7b still shows a marked color difference. Because this is due to the fact $\theta_1 = \theta_2$ was assumed for simplicity in Eq. (6), the change in radiance due to Lambert's cosine law was not considered. In addition, the individual color difference of projector is not considered, which leads colored shadow. Future research will be addressing these issues and finding solutions.

5.2. Adaptive foreground detection

In the current calibration phase, the discrimination between the foreground and background is determined. However, this static scene assumption becomes insufficient when

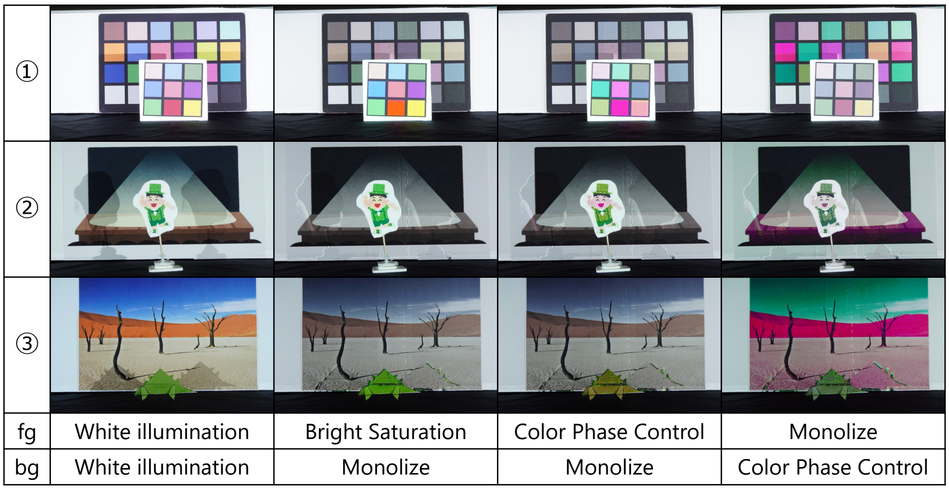
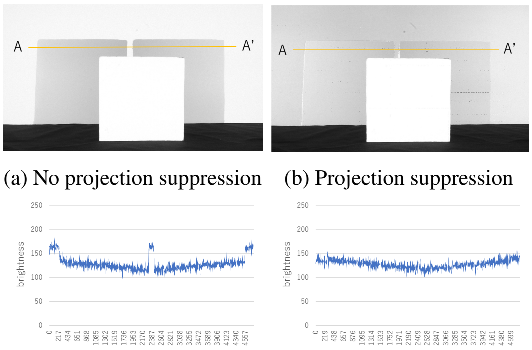


Figure 6: Appearance Manipulation results by the proposed method (fg:foreground, bg:background).Over illumination is not suppressed during white illumination.



(a) No projection suppression (b) Projection suppression

(c) A-A' in (a) (d) A-A' in (b)

Figure 7: Comparing lightness transition with white projection.

(a) Estimated image C_{est1^f} (b) Estimated image C_{est1^b}

Figure 8: Comparison of images obtained using $f_{C_2 2C_1}$ of each plane.

foreground objects are in motion during operation. Therefore, our next step involves the development of an adaptive discrimination. If the foreground and background objects can be assumed to be planar, a possible solution is to compare C_2, which is deformed using $f_{fC_2 2C_1}$ and $f_{bC_2 2C_1}$ obtained by geometric calibration, with the actual C_1. Specifically, when the captured image is as shown in Fig. 5, the estimated images C_{est1^f} and C_{est1^b} are obtained as shown in Figure 8. Then it is possible to determine which of image C_{est1^f} or C_{est1^b} is closer to C_1 and identify foreground and background. Moreover the homography transformation is an alternative solution for discrimination. However, it does not provide accurate pixel mapping due to the lack of lens distortion consideration. Therefore, pixel mapping is still required to achieve precise results.

6. Conclusion

In this study, we proposed a method to achieve the independent appearance manipulation of objects by distinguishing foreground and background objects, as a preliminary step toward moving away from uniform processing of appearance. We also proposed a method to identify shadow caused by the foreground and to suppress luminance differences between overlapping regions.

The experimental results confirmed that foreground and background could be independently manipulated. Moreover, we were able to improve the brightness difference issue in the background caused by projections by suppressing the projected light intensity based on the overlapped projection determination.

However, the distinction between foreground and background relies on pre-acquired pixel maps of the actual scene, which limits our ability to handle dynamic movements of foreground objects. Additionally, due to the simplification of not considering the radiance change caused by the cosine of the incident angle, visible brightness differences remained in the overlapped regions.

Future research will work on implementing dynamic foreground object distinction and more accurate methods for improving brightness differences.

References

[1] R. Raskar, G. Welch, K.-L. Low, D. Bandyopadhyay, Shader lamps: Animating real objects with image-based illumination, in: Rendering Techniques 2001: Proceedings of the Eurographics Workshop in London, United Kingdom, June 25–27, 2001 12, Springer, 2001, pp. 89–102.

[2] A. Grundhöfer, D. Iwai, Recent advances in projection mapping algorithms, hardware and applications, Computer Graphics Forum 37 (2018) 653–675. doi:10.1111/cgf.13387.

[3] T. Amano, H. Kato, Appearance control using projection with model predictive control (2010) 2832–2835. URL: https://cir.nii.ac.jp/crid/1050577309352921344.

[4] T. Amano, Projection based real-time material appearance manipulation, in: 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE Computer Society, Los Alamitos, CA, USA, 2013, pp. 918–923. URL: https://doi.ieeecomputersociety.org/10.1109/CVPRW.2013.135. doi:10.1109/CVPRW.2013.135.

[5] T. Amano, I. Shimana, S. Ushida, K. Kono, Successive Wide Viewing Angle Appearance Manipulation with Dual Projector Camera Systems, in: T. Nojima, D. Reiners, O. Staadt (Eds.), ICAT-EGVE 2014 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments, The Eurographics Association, 2014. doi:10.2312/ve.20141364.

[6] K. Murakami, T. Amano, Materiality Manipulation by Light-Field Projection from Reflectance Analysis, in: G. Bruder, S. Yoshimoto, S. Cobb (Eds.), ICAT-EGVE 2018 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments, The Eurographics Association, 2018. doi:10.2312/egve.20181321.

[7] F. Cao, Q. Bao, A survey on image semantic segmentation methods with convolutional neural network, in: 2020 International Conference on Communications, Information System and Computer Engineering (CISCE), 2020, pp. 458–462. doi:10.1109/CISCE50729.2020.00103.

[8] R. Sukthankar, T. J. Cham, G. Sukthankar, Dynamic shadow elimination for multi-projector displays, Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR (2001). URL: https://cir.nii.ac.jp/crid/1362544420996161536. doi:10.1109/cvpr.2001.990943.

[9] S. Audet, J. Cooperstock, Shadow removal in front projection environments using object tracking, 2007 IEEE Conference on Computer Vision and Pattern Recognition (2007). URL: https://cir.nii.ac.jp/crid/1363388844278556416. doi:10.1109/cvpr.2007.383470.

[10] M. Flagg, J. Summet, J. Rehg, Improving the speed of virtual rear projection: A gpu-centric architecture, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005 - Workshops, IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, IEEE Computer Society, United States, 2005. doi:10.1109/CVPR.2005.476.

[11] S. Uesaka, T. Amano, Cast-Shadow Removal for Cooperative Adaptive Appearance Manipulation, in: H. Uchiyama, J.-M. Normand (Eds.), ICAT-EGVE 2022 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments, The Eurographics Association, 2022. doi:10.2312/egve.20221271.