A Monocular Projector-Camera System using Modular Architecture

Kenta Yamamoto, Daisuke Iwai, *Member, IEEE,* Ikuho Tani, and Kosuke Sato, *Member, IEEE*

Abstract—This paper presents a monocular projector-camera (procam) system using modular architecture based on relay optics. Conventional coaxial procam systems cannot support (1) online changes to lens settings (zoom and focus) and (2) wide-angle projection mapping. We develop design guidelines for a proposed procam system that would solve these restrictions and address the proposed system's unique technical issue of crosstalk between the camera and projector pixels. We conducted experiments using prototypes to validate the feasibility of the proposed framework. First, we confirmed that the proposed crosstalk reduction technique worked well. Second, we found our technique could achieve correct alignment of a projected image onto a moving surface while changing the zoom and focus of the objective lens. The monocular procam system also achieved radiometric compensation where a surface texture was visually concealed by pixel-wise control of a projection color based on the captured results of offline color pattern projections. Finally, we demonstrated the high expandability of our modular architecture, through the creation of a high dynamic range projection.

Index Terms—Projector-camera system, augmented reality, projection mapping.

1 Introduction

Projection mapping seamlessly merges real and virtual worlds by projecting a computer-generated image onto a physical surface [1], [2]. Compared with other augmented reality (AR) approaches, such as optical and video see-through AR, projection mapping can provide a better user experience because it does not require the user to wear or hold a display device. It has been applied in fields such as medicine [3], industrial design [4], online conferencing [5], office work [6], [7], and entertainment [8], [9]. Led by the recent advances in high-speed and low-latency projector hardware [10], the latest research tends toward dynamic projection mapping, where the geometrical relationship between a projection target and a projector changes during the projection [11], [12], [13], [14], [15]. To control the appearance of a target surface, accurate alignment of the projected image onto the target and pixel-wise color correction are crucial; thus, geometric registration and radiometric compensation have been the primary technical challenges [1]. To tackle these, a projector-camera (procam) system in which sensing capability is virtually provided to each projector pixel is widely applied. A procam system is effective only when the pixel correspondences between the devices are known. However, a normal procam system is not suitable for dynamic projection mapping applications because multiple structured light patterns need to be projected and captured to acquire the correspondences after the target moves. When a depth camera is applied in the system, the pixel correspondences can be estimated from the geometrical relationships among the projector, the camera, and the target [16], [17]. However, the estimation accuracy is not sufficient for the radiometric compensation.

• *K. Yamamoto, D. Iwai, I. Tani, and K. Sato are with the Graduate School of Engineering Science, Osaka University, Japan. E-mail: see https://www.sens.sys.es.osaka-u.ac.jp*

Manuscript received April 19, 2005; revised August 26, 2015.

A promising solution for this technical issue is a coaxial procam system in which the projector and camera share the same optical axis [18], [19], [20]. In this way, the coaxial system is implemented by placing the projector and the camera behind different sides of a half mirror. The pixel correspondences between the coaxial devices do not change during the movement of the target surface; thus, the projection of the structured light patterns is required only once in offline calibration. However, when the zoom or focus setting of the camera or the projector lens is changed, the calibrated pixel correspondences are no longer valid, and additional calibration is required. In addition, the field of view (FOV) of the coaxial setup is limited by the half mirror, as shown in Fig. 1(a). This technical limitation prevents the use of a wide-angle lens, such as a fisheye lens, in a procam system, although researchers have recently explored convincing applications of wide-angle projection mapping, such as displaying 360 VR videos [21], [22], [23], [24], [25]. Therefore, this limitation potentially narrows the range of application fields.

In this paper, we propose a monocular procam system in which the projector and the camera share the same objective lens, as shown in Fig. 1(b). As the primary technical contribution, we propose applying relay optics to optically transfer the image panels of the camera and the projector at the focal point of the objective lens, such that the transferred image panels are overlaid with each other. The overlaid pixels have both sensing and displaying capabilities; thus, we call them *bilateral pixels*. The pixel correspondences between the camera and the projector do not depend on the state of the objective lens; that is, they do not change even when the zoom or focus setting of the lens is changed. In addition, a lens of any focal length, including a fisheye lens, can be used as the objective. We develop design guidelines for the monocular procam system, including solutions for unique technical issues, such as the crosstalk between the

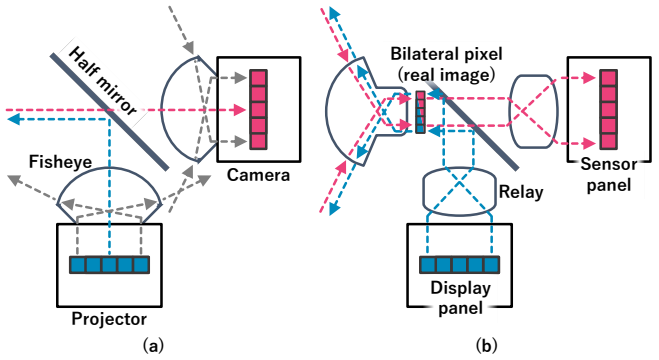


Fig. 1. Schematic diagram of procam systems: (a) a conventional coaxial procam system cannot support a wide-angle lens, such as a fisheye lens, because part of the FOV of such devices cannot be covered by the half mirror, while (b) the wide-angle lens can be used in the proposed monocular procam system applying relay optics that virtually transfer the camera's sensor panel and the projector's display panel to the focal point of the objective lens, which form bilateral pixels.

camera and the projector pixels. Furthermore, the proposed optical framework is designed to support modular architecture, by which we can combine additional imaging and displaying devices with our procam system via relay optics. This modularity enables the extension of the projection performance beyond the technical limitation of projector hardware. Specifically, we realize a high dynamic range (HDR) procam system in this research. This paper describes the detailed design of the monocular procam system, and demonstrates its feasibility in dynamic projection mapping with various lens settings (zooming and focusing) through experiments using prototypes.

A simpler implementation of the bilateral pixels has been done just by placing a half mirror between the objective lens and the image panels of the projector and the camera [26]. However, the size of the sensor panel of the camera is generally not the same as that of the projector. When the sensor panel is smaller, a part of the projector's FOV cannot be measured. In addition, when the backfocus of the objective lens is shorter than the optical path of the half mirror, captured and projected images inevitably become blurred. In contrast, our relay optics-based system that optically transfers the image panels can resize the transferred panels and place them exactly at the focal point of the objective lens, regardless of its backfocus. However, our system suffers from a severe vignetting problem because of its limited aperture [27].

2 Related Work

Researchers have developed coaxial procam systems for the accurate alignment of projected images onto moving surfaces. A simple but rather novel application is to visualize pixel-wise invisible information, such as temperature [28] and the normal [12] of the surface using projected imagery, where an unconventional camera (e.g., thermography) is used as the coaxial camera. In more general applications, the pose of the target surface is required to render a projection texture as if the projected texture attaches to the moving surface. The coaxial procam systems estimate the pose by analyzing a captured image and geometrically register the projector to the surface [11], [13], [20]. Near infrared cameras are usually applied to avoid the interference of a projected image with a captured image. Another important role of the coaxial camera is to obtain the reflectance property of a target surface in each frame. Because the pixel correspondences between the camera and the projector do not change during the movement of a projection surface, the coaxial setup allows pixel-wise color correction through which a target's appearance can be reproduced on a textured moving surface [18], [19], [29]. Recent advances in calibration techniques have made the coaxial procam system more accessible for non-expert researchers and developers [30], [31].

As discussed in Sect. 1, the coaxial setup cannot support (1) a situation whereby the lens settings (zoom and focus) are changed online and (2) wide-angle projection mapping. These restrictions can be solved using a procam system in which the projector shares the same objective lens with the camera. In this paper, we propose optically transferring the image panels using relay optics to realize a monocular procam system. Researchers have previously applied relay optics in a coaxial procam system to quickly redirect projector and camera FOVs using a tiny galvo mirror system [14]; however, the projector and the camera did not share an objective lens, and thus, this system could not address the first restriction. To the best of our knowledge, our research is the first to introduce relay optics to realize a monocular procam system that addresses both the first and the second restrictions.

Relay optics have also been utilized to overcome technical limitations of projector hardware. For instance, previous works have shown that the double modulation of light can extend dynamic range [32] and spatiotemporal resolutions [33]. Double modulation was realized by cascading two spatial light modulators (SLMs), such as a liquid crystal on silicon (LCoS) panel using relay optics in the optical path of a projector. Such a high performance projector based on the double modulation principle would be highly compatible with our monocular procam system because both apply relay optics. We realize a modular architecture using relay optics. To demonstrate the high expandability of the modular architecture, we integrate an HDR projector into the monocular procam system.

3 Method

3.1 Monocular procam system using relay optics

Relay optics consisting of two convex lenses optically transfer an object to another location. Suppose that the object is placed in front of the first lens such that the distance between the object and the first lens is its focal length f_1, rays emitted from the object are collimated by the lens. The collimated beam passes through the second lens and converges on a spot behind the lens at its focal point. Consequently, a *real image* of the object is formed at the focal point. The size of the real image relative to the original object is $\frac{f_2}{f_1}$, where f_2 is the focal length of the second lens. Therefore, we can control the size of a relayed object by adjusting the focal lengths of the lenses. The relay optics

can transfer not only a physical object but also a real image at the focal point of the first lens.

In our system, an objective lens is placed in front of a half mirror, while two sets of relay optics are installed behind different sides of the half mirror (Fig. 1(b)). On the projector side, rays emitted from the projector's display panel pass through the first relay optics, reflect at the half mirror, and converge at the focal point of the objective lens, where a real image of the display panel is consequently formed. The objective lens then projects this to a projection surface in a scene. On the camera side, the objective lens forms the real image of the surface by concentrating the rays reflected from the surface at the focal point. The real image then passes through the half mirror and the second relay optics and finally converges at the camera's sensor panel. In principle, the sensor panel and the display panel can be swapped.

An important property of the monocular procam system is that the pixel correspondences between the coordinate systems of the camera and the projector are consistent even when the focusing and zooming settings of the objective lens are changed. We use a homography transformation that has been widely used in conventional procam systems to map a camera pixel to a projector pixel [20], [28], [30]. For the calibration of the homography transformation, we obtain the pixel correspondences by projecting gray code patterns onto a physical surface and capturing them using the proposed procam system. The correspondences are then used to estimate the parameters of the homography matrix. This calibration process needs to be performed only once in advance.

3.2 Crosstalk reduction

The capturing component of our monocular procam suffers from crosstalk. Part of the rays emitted from the projector's display panel specularly reflect at the back surface of the objective lens, travel back toward the half mirror, and reach the camera's sensor panel. The stray light significantly decreases the contrast of a captured image. When the procam system is used to visually cancel the texture of a surface (i.e., radiometric compensation), the surface appearance must be captured in full color, and thus the crosstalk needs to be reduced. We apply the following two optical solutions to alleviate the crosstalk. First, we use polarization to cut off the stray light. We place a linear polarizer in front of the display panel and apply a polarizing beam splitter (PBS) as our half mirror. This optical setup prevents the stray light from passing through the PBS while allowing incoming rays from the scene to reach the sensor panel. Second, we slightly shift and tilt the objective lens to redirect the stray light toward a blank space rather than the PBS. The amount of shift and tilt required for reducing the crosstalk is dependent on the shape of the back surface of the objective lens. Because the lens shape information is not always available, especially for a compound lens, we manually searched for the optimal amount of shift and tilt by which the crosstalk would be the most significantly reduced.

When it is not required to capture the surface appearance in full color, we can adopt another strategy. LEDs and lasers have recently been used as the light sources of projectors, instead of conventional lamps. These light sources have narrow spectral band properties and thus do not emit infrared (IR) light. Thus, we install an optical filter that passes IR light while cutting visible light in front of the sensor panel to prevent the stray light from being incident on the sensor panel. Although this setup does not allow us to perform a radiometric compensation, it can geometrically correct a projected image onto a moving target. Specifically, once we attach fiducial markers on the target, we can estimate the geometrical relationship between the procam system and the target by analyzing the captured markers.

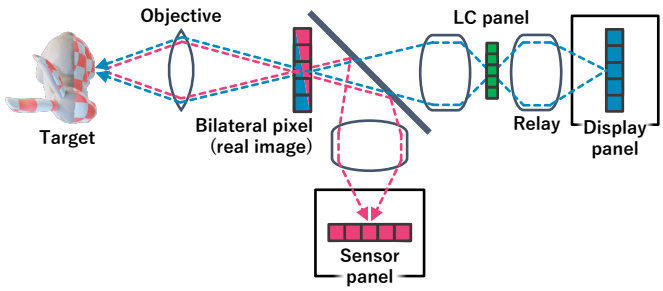


Fig. 2. Schematic diagram of a monocular procam system with HDR projection capability.

3.3 Extension of projector performance

As discussed in Sect. 2, our modular architecture using the relay optics enables cascading multiple SLMs to enhance the projector's performance. Specifically, we realize HDR projection by spatially modulating the light twice with two different SLMs (Fig. 2). The double modulation boosts the contrast of a projection system in a multiplicative manner [32]. Although the first SLM cannot completely block the light even when a pixel value of zero (i.e., black) is inputted, the second SLM can further block the leaked light, which realizes the contrast boosting. Suppose that the contrast ratio of the first SLM is $c_1 : 1$ and that of the second SLM is $c_2 : 1$, then that of the double modulation system becomes $c_1 c_2 : 1$.

Between the half mirror and the projector's display panel, we insert a grayscale liquid crystal (LC) panel that spatially modulates the transparency. A projection image on the display panel is transferred to the LC panel using relay optics where the contrast is boosted. Rays emitted from the LC panel converge at the focal point of the objective lens using different relay optics. An original HDR image is split into the display panel and the LC panel as follows. First, we decompose the original color O into its luminance L and chrominance C components. Without the loss of generality, we suppose that both SLMs spatially modulate light with 8-bit tonal resolution. Then we assign the transparency value on the LC panel as $\sqrt{L}$. The displayed RGB color on the display panel is computed from the luminance $\sqrt{L}$ and the chrominance C.

4 Experiment

4.1 Crosstalk reduction

We validated the effectiveness of the applied crosstalk reduction strategy using a prototype (Fig. 3(a)). The prototype was implemented using the sensor panel of an industrial

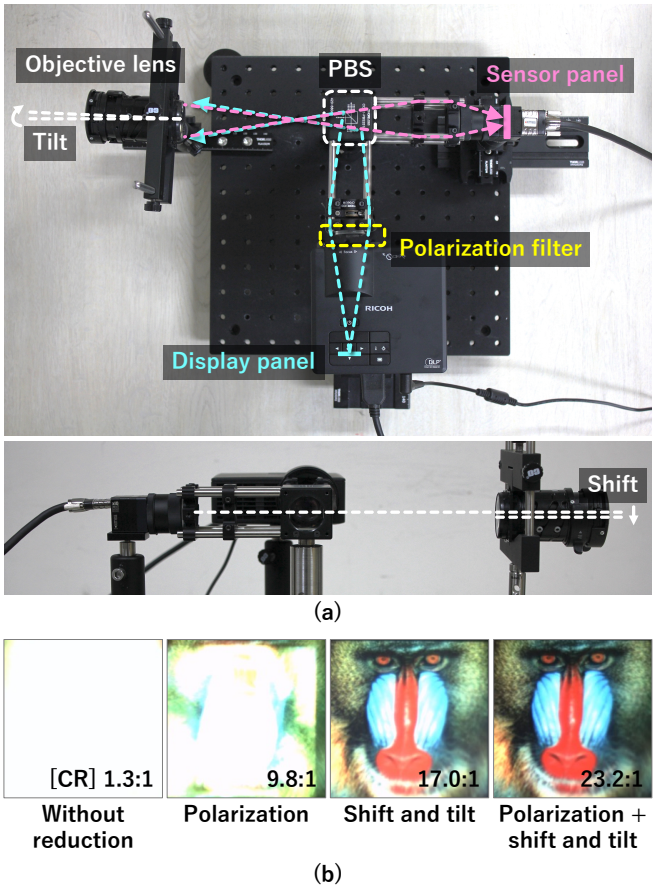


Fig. 3. Evaluation of the crosstalk reduction technique: (a) the prototype system, and (b) captured results of a screen on which a mandrill image was projected in different crosstalk reduction conditions. The metric indicated in each sub-figure represents the contrast ratio (CR).

camera (Ximea MQ013CG-ON, resolution: 1264×1016 pixels, panel size: 0.31 inch) and a digital light processing (DLP) projector (Ricoh PJ WXC1110, resolution 1280×800 pixels, panel size: 0.45 inch). We disassembled the projector and removed the built-in objective lens. These two devices were placed behind the different sides of a PBS (Thorlabs CCM1-PBS251/M). In front of the PBS, we placed a zoom lens (focal length: 24–29 mm) that was removed from a projector (Epson EMP-1710) as the objective lens of the monocular procam system. Because the size of the camera's sensor panel was smaller than that of the projector's display panel, we reduced the size of the real image formed by the objective lens. Specifically, the relay optics for the camera consisted of lenses of focal length 50 mm (Thorlabs AC254-050-A) and 100 mm (Thorlabs AC254-100-A) to reduce the real image to half its original size. Besides, the focal lengths of the lenses in the relay optics for the projector were both 100 mm (Thorlabs AC254-100-A) to transfer the display panel at equal magnification. A polarizer (Thorlabs LPVISE100-A-ϕ1) was inserted between the projector's display panel and its relay optics. We shifted the objective lens by about 4 mm and tilted it by about 3 degrees in the prototype.

Figure 3(b) shows captured images of a projected result of a mandrill texture on a flat, uniformly white screen with and without the crosstalk reduction methods, which are adding a polarizer and shifting and tilting the objective lens. We computed the contrast ratio of each result by dividing the maximum luminance value by the minimum value. As shown in the figure, crosstalk was significantly reduced using both methods. Comparing the image quality improvements using the two methods, we found that shifting and tilting the objective lens was more effective, and the improvement using the other method was also unignorable. Therefore, we confirmed that both methods are essential for reducing crosstalk.

4.2 Spatial accuracy of projection mapping in different focus and zoom settings

In the proposed system, the pixel correspondences between the camera's sensor panel and the projector's display panel are invariant to the objective lens condition (i.e., focus and zoom). Thus, the correspondences need to be obtained only once in advance via homography, as described in Sect. 3.1. Using the same prototype built in Sect. 4.1, we validated the spatial accuracy of projection mapping across different lens conditions with fixed pixel correspondences.

We prepared three focus conditions by placing a target surface at different locations (400, 600, and 800 mm away from the objective lens) and adjusting the focus ring of the objective for each location such that the target was in focus. We labelled the three focus conditions *near* (400 mm), *middle* (600 mm), and *far* (800 mm). The zoom setting of the objective was fixed at a focal length of 24 mm. In addition, we prepared three zoom conditions by manually adjusting the focal length of the objective lens at 24 mm, 26.5 mm, and 29 mm while fixing the target surface at 600 mm away from the objective. We labelled the three zoom conditions *wide* (24 mm), *neutral* (26.5 mm), and *tele* (29 mm). In all the zoom conditions, the focus ring was adjusted such that the target surface placed 800 mm away from the lens was in focus. In summary, there were five experimental conditions of the focus/zoom combinations in total: near/wide, middle/wide, far/wide, far/neutral, and far/tele.

The target surface was an A4-format flat screen on which a grid pattern and AR marker were printed. Once the procam system captured the target surface, we estimated its pose relative to the system by analyzing the captured AR marker. The pose information was then used to estimate the 2D positions of the grid intersections in the coordinate system of the camera's sensor panel. Hereinafter, the estimated points are referred to as camera points. The intrinsic parameters required for the AR marker's pose estimation were calibrated in advance using Zhang's method [34]. The estimated camera points were finally transformed to the coordinate system of the projector's display panel. Hereinafter, the estimated points are referred to as projector points. We projected green circles on the projector points.

Figure 4 shows images captured by the procam system in the five conditions, where the red dots have been digitally overlaid at the camera points for reference. We manually fitted a circle to each of the projected blobs and computed their centers to obtain the positions of the projector points in the sensor plane coordinate system. We then computed the distances between the grid intersections and the corresponding projector points. Figure 4 shows the averages

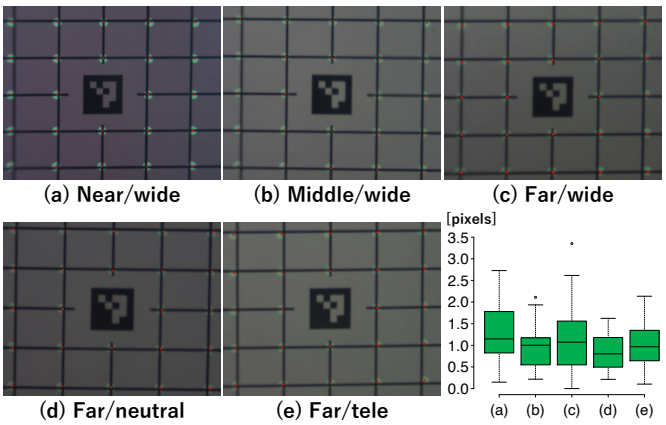


Fig. 4. Spatial accuracy evaluation: (a)–(e) captured images of the green circles projected onto the estimated grid intersection positions. The red dots are overlaid onto the estimated intersections in the camera image for reference. The graph on the bottom-right shows the box plot of the errors. Each small caption (e.g., near/wide) indicates the focus condition (near) and the zoom condition (wide), respectively.

and standard deviations of the distances across the five conditions. We confirmed that the spatial accuracy of projection mapping was not significantly affected by the lens condition. Therefore, this result indicated that our system is suitable in a dynamic projection mapping application for which the focus and zoom of the objective lens are adjusted online. We demonstrate this in Sect. 4.3.

In the current system, we calibrated the intrinsic parameters for each zoom condition. Although this approach is not practical, we expect to use an online updating method for the intrinsic parameters [35] in actual applications.

4.3 Dynamic projection mapping with a varifocal wide-angle lens

Coaxial procam systems cannot support (1) changes in the zoom and focus settings of the projector's lens and (2) wide-angle projection mapping, as discussed in Sect. 2. In this section, we demonstrate that our monocular procam system can support these situations.

In this experiment, we built another prototype (Fig. 5(a)), in which a wide-angle lens (Canon EF8-15mm F4L Fisheye USM) was applied as the objective of the system. We used the same devices of the camera sensor panel, the projector display panel, and the PBS as the ones used in Sect. 4.1. Considering that the objective lens was originally designed for a full-frame digital monocular reflex (DSLR) camera, we determined the focal lengths of the lenses in the relay optics. Specifically, the relay optics for the camera consisted of lenses of focal length 30 mm (Thorlabs AC254-030-A) and 150 mm (Thorlabs AC254-150-A) to reduce the real image formed by the objective lens to one-fifth its original size at the sensor panel. The focal lengths of the lenses in the relay optics for the projector were 100 mm (Thorlabs AC254-100-A) and 200 mm (Thorlabs AC254-200-A) to transfer the display panel 2.0 times larger at the focal point of the objective lens. The measured horizontal and vertical FOVs of the camera were 120 and 100 degrees, respectively. Those of the projector were 94 and 66 degrees, respectively.

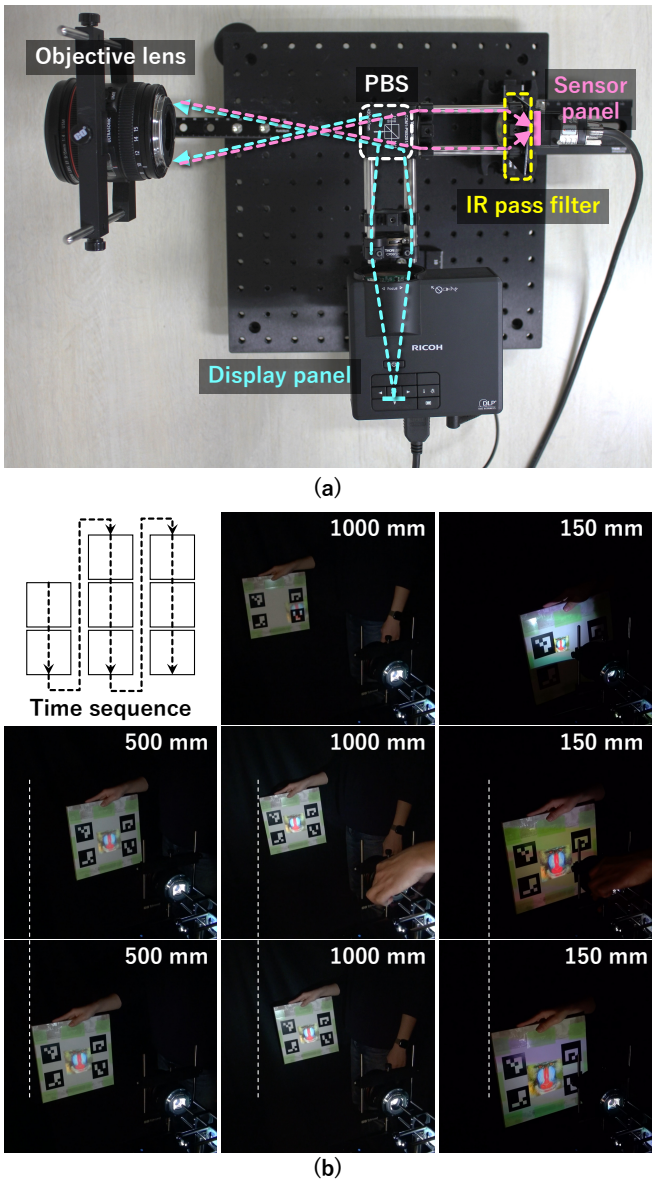


Fig. 5. Experimental results of the dynamic projection mapping: (a) the prototype system applying a wide-angle lens as its objective lens, and (b) the sequence of projected results of the dynamic projection mapping, as recorded by an external camcorder. In (b), the number on the top-right of each sub-image indicates the distance between the objective lens and the projection surface, and the white dashed lines are overlaid to indicate the movement of the surface.

This system was used only for geometric registration and did not need to capture a full color image; thus, we applied the second crosstalk reduction strategy and inserted an IR pass and visible light cut filter in front of the sensor panel. The projection target was an A4-format flat surface, and four AR markers were printed on its corners. Analyzing captured AR markers, the system projected a mandrill texture onto the center area of the surface.

Figure 5(b) shows the captured sequence of the projected results on the target being manually moved back and forth with respect to the procam system. We recorded the sequence using an external camcorder. First, we placed the target 500 mm away from the procam system and adjusted its objective lens such that all the markers were within the

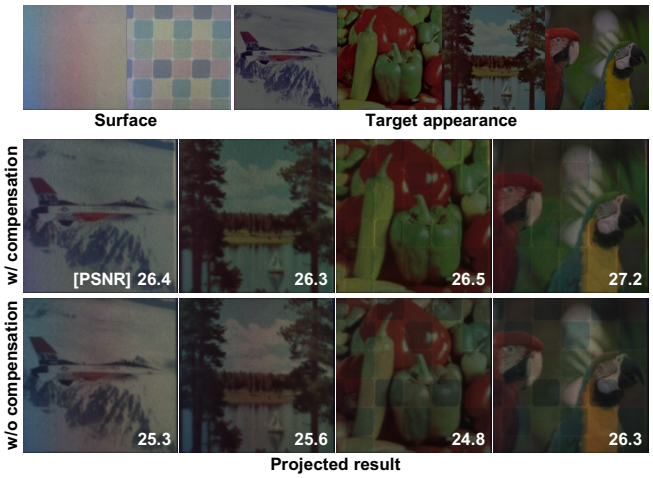


Fig. 6. Experimental results of radiometric compensation. The metric indicated in each sub-figure represents PSNR. A higher PSNR indicates that the projected result is closer to the target appearance.

FOV and appeared focused. When we moved the target in the frontal plane with respect to the procam system, the projected texture successfully followed the surface movement. Next, we moved the target farther to a place that was 1000 mm away from the system. Then, the projected texture failed to follow the target because the markers became too small and blurred. After we adjusted the lens settings, the texture started to follow the target again. Last, we moved the target closer to the system (150 mm). Then, the projected texture failed again to follow the target, because the target became too large and blurred. After we adjusted the lens settings, the texture started to follow the target again. Throughout the experiment, we used the same pixel correspondences between the sensor and display panels, which had been obtained in advance. From this experiment, we confirmed that our monocular procam system can overcome the limitations of conventional procam systems; that is, it supports a dynamic projection mapping application in which a wide-angle lens is applied and/or zooming and focusing settings are dynamically adjusted.

4.4 Radiometric compensation

Radiometric compensation, which visually conceals the texture of a target surface, is an essential technology to control its appearance in projection mapping [1]. Using the experimental setup introduced in Sect. 4.1, we experimentally evaluated if a conventional radiometric compensation technique [36] would work in the proposed procam system. Specifically, our system projected twenty-seven color patterns onto a target surface and captured the projected results. Analyzing the captured appearances, we obtained the reflectance property of the surface per pixel. Then, we computed a compensation image by dividing a target color by the reflectance at each pixel.

As shown in Fig. 6, we used flat papers as the projection surfaces, on which different color patterns were printed. Two target appearances were prepared for each of the surfaces; thus, there were four different combinations of the surface and the target appearance. The figure shows

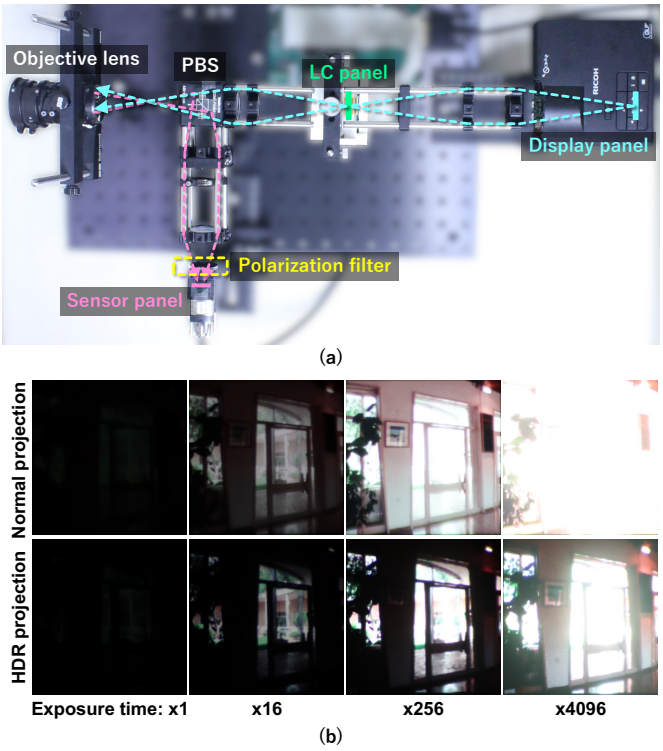


Fig. 7. Experimental results of the HDR projection: (a) the prototype system, and (b) the comparison of the projected results between the proposed HDR projection system and a normal projection system.

the projected results with and without the compensation technique across all four combinations. We computed the peak signal-to-noise ratio (PSNR) value by comparing each projected result with the corresponding target appearance. We confirmed that the surface textures were significantly concealed by the compensation. The compensation would have failed if the crosstalk had remained in the system. Therefore, this result indicated that our crosstalk reduction methods were sufficiently effective to perform radiometric compensation. Besides, we could also see artifacts around the edge regions of the surface texture. We discuss a solution for this technical limitation in Sect. 5.

4.5 HDR projection

To show the high expandability of our modular architecture based on relay optics, we realized HDR projection by extending the prototype built in Sect. 4.1. As described in Sect. 3.3, we inserted a grayscale LC panel and different relay optics between the PBS and the relay optics of the display panel by which the double modulation of light was achieved (Fig. 7(a)). The inserted relay optics were of equal magnification using two convex lenses of focal length 80 mm (Thorlabs AC254-80-A). Because the LC panel passes polarized light, we removed the polarizer that was inserted in front of the display panel. In a preliminary experiment, we found that this led to a small amount of stray light on the sensor panel; thus, we attached the polarizer in front of the sensor panel, which prevented the stray light and improved the capturing quality.

We compared the peak luminance and the contrast of the HDR projection system with those of a normal projection

system, which was the original prototype in Sect. 4.1. First, we measured the peak luminance of the HDR projection system. We sent the uniform white images to both the display panel and the LC panel, and projected the light onto a white screen. The luminance of the projected result was measured using a spectroradiometer (Topcon SR-LEDW). Then we sent the uniform black images to the devices and measured the luimnance of the projected result. This measurement was the lowest luminance of the HDR projection system. The contrast was computed by dividing the peak luminance by the lowest one. As a result, the peak luminance and the contrast of the HDR projection system were 900 cd/m^2 and 58000:1, respectively. We measured those values of the original prototype, which were 5200 cd/m^2 and 1100:1, respectively. Therefore, we confirmed that our HDR projection system can boost the contrast to be more than 50 times greater than the normal projection system, while it reduced the peak luminance by one-fifth.

Figure 7(b) shows the projected results with and without the spatial light modulation by the LC panel. In the former condition, the spatial pattern of the transparency on the LC panel was computed using the algorithm described in Sect. 3.3. In the latter condition, we inputted a uniform white image to the LC panel. The results were captured by the camera sensor of the procam system with different exposure times. From the captured results, we confirmed the contrast of the projected image to be significantly boosted by the second spatial light modulation at the LC panel.

5 Discussion

This paper presents a monocular procam framework by applying relay optics and experimentally validated its performance. As the primary achievement, the experiments in Sect. 4.2 and 4.3 showed that the geometric transformation between the sensor panel and the display panel were not significantly affected by the lens conditions (i.e., zoom and focus). To the best of our knowledge, this is the first time projection mapping has successfully aligned a projected image with a moving physical surface even when the zoom and focus settings of the objective lens were dynamically adjusted. In actual applications, dynamic zooming and refocusing are achieved by electrically controlling the objective lens, such as a professional camera lens (e.g., Canon EF lens) [37], [38] or a multipurpose focus tunable lens [39].

Our system shares the same limitations with conventional coaxial procam systems. The first limitation is its larger form factor than a normal projector device. As the second limitation, the beam splitter blocks half of the projected light, as does the polarization filter. However, the polarization filter did not affect the brightness of the final projected result in our setup. Since our beam splitter was a PBS, the incident light was totally reflected or passed because the polar directions of the PBS and the incident light were set to be identical.

Our crosstalk alleviation technique successfully reduced the stray light and, consequently, significantly increased the contrast of the captured image (Sect. 4.1). Although the shift and tilt of the objective lens had more effect than the polarization method, they generally caused a tilt of the plane of focus, based on the Scheimpflug principle. An interesting future extension would be to integrate an active lens actuation mechanism into the objective to realize an optimal shift and tilt for each application while avoiding critical contrast degradation by the stray light.

Radiometric compensation successfully worked in the proposed system (Sect. 4.4), although it was not perfect. The surface textures were visually concealed in most of the surface areas, but we found visible artifacts around their edges. We consider two reasons for the artifacts. First, the geometric transformation model (i.e., homography) that we used to align the sensor and display panels was not sufficiently accurate. Specifically, there might have been nonlinear misalignment due to lens distortion in the relay optics, which could not be corrected by the homography transformation. We consider that this problem can be solved by applying a non-linear distortion correction technique, such as one used in a popular camera calibration method [34]. Second, the size of a projected pixel was not small enough to compensate for the edge of the texture. A simple solution is to increase the spatial resolution of the display panel, which would, however, incur additional cost. Another solution that would not involve additional cost is applying an edge-aware radiometric compensation technique [40], by which we can alleviate the artifacts without increasing the spatial resolution of the display panel.

We showed that our modular architecture based on relay optics could support the extension of projector performance beyond the capability of projector hardware and, in particular, the HDR projection. The modular architecture would allow us to build other types of high performance projection systems, such as multi-band projection, by combining more than two display panels with slightly different spectral properties [41]. However, adding optical components also leads to additional aberrations and reduction of peak luminance. Based on our calculations, even in our simplest prototype (Sect. 4.1, Fig. 3), the peak luminance was reduced to 7.9% of that of the original projector. In addition, the relay optics introduced vignetting in which the luminance of the corner pixel became 90.3% of that of the center pixel. In our HDR projection setup, significant luminance reduction occurred at the LC panel. A solution for this particular issue is to apply a light steering technique where the first modulator reallocates the light energy of the light source rather than blocking it [32]. In general, careful system design is required to balance performance and image quality.

6 Conclusion

This paper presents a novel procam framework, namely a monocular procam system in which the projector and camera share the same objective lens and, consequently, the pixel correspondences between the camera and projector do not change even when the zoom and focus of the lens are changed. We realized this by applying relay optics through which the camera's sensor panel and the projector's display panel were optically transferred to the focal point of the objective lens where these transferred panels overlapped. Through an experiment using a prototype, we demonstrated that the proposed system maintained alignment of a projected image onto a moving surface while the lens settings

were dynamically changed. We also showed that radiometric compensation worked on our prototype, and that the proposed framework could support extending the dynamic range of the projector.

As mentioned in Sect. 5, an interesting future work would be to apply functional optics, such as an electrically focus tunable lens (ETL), to our system to solve the current technical limitation in the crosstalk reduction methods. The ETL also has potential to extend the depth of field of the procam system by applying the focal sweep technique [42].

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI grant number JP20H05958 and JST, PRESTO Grant Number JPMJPR19J2, Japan.

REFERENCES

[1] O. Bimber and R. Raskar, *Spatial Augmented Reality: Merging Real and Virtual Worlds*. Natick, MA, USA: A. K. Peters, Ltd., 2005.

[2] A. Grundhöfer and D. Iwai, "Recent advances in projection mapping algorithms, hardware and applications," *Computer Graphics Forum*, vol. 37, no. 2, pp. 653–675, 2018.

[3] H. Nishino, E. Hatano, S. Seo, T. Nitta, T. Saito, M. Nakamura, K. Hattori, M. Takatani, H. Fuji, K. Taura, and S. Uemoto, "Real-time navigation for liver surgery using projection mapping with indocyanine green fluorescence: Development of the novel medical imaging projection system," *Annals of Surgery*, vol. 267, no. 6, pp. 1134–1140, 2018.

[4] T. Takezawa, D. Iwai, K. Sato, T. Hara, Y. Takeda, and K. Murase, "Material surface reproduction and perceptual deformation with projection mapping for car interior design," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2019, pp. 251–258.

[5] D. Iwai, R. Matsukage, S. Aoyama, T. Kikukawa, and K. Sato, "Geometrically consistent projection-based tabletop sharing for remote collaboration," *IEEE Access*, vol. 6, pp. 6293–6302, 2018.

[6] D. Iwai and K. Sato, "Limpid desk: See-through access to disorderly desktop in projection-based mixed reality," in *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '06, 2006, p. 112–115.

[7] K. Matsushita, D. Iwai, and K. Sato, "Interactive bookshelf surface for in situ book searching and storing support," in *Proceedings of the 2nd Augmented Human International Conference*, 2011.

[8] M. R. Mine, J. van Baar, A. Grundhofer, D. Rose, and B. Yang, "Projection-based augmented reality in disney theme parks," *Computer*, vol. 45, no. 7, pp. 32–40, 2012.

[9] B. R. Jones, H. Benko, E. Ofek, and A. D. Wilson, "Illumiroom: peripheral projected illusions for interactive experiences," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 869–878.

[10] Y. Watanabe, G. Narita, S. Tatsuno, T. Yuasa, K. Sumino, and M. Ishikawa, "High-speed 8-bit image projector at 1,000 fps with 3 ms delay," in *The International Display Workshops*, 2015, pp. 1064–1065.

[11] G. Narita, Y. Watanabe, and M. Ishikawa, "Dynamic projection mapping onto deforming non-rigid surface using deformable dot cluster marker," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 3, pp. 1235–1248, 2017.

[12] L. Miyashita, Y. Watanabe, and M. Ishikawa, "Midas projection: Markerless and modelless dynamic projection mapping for material representation," *ACM Trans. Graph.*, vol. 37, no. 6, Dec. 2018.

[13] A. H. Bermano, M. Billeter, D. Iwai, and A. Grundhöfer, "Makeup lamps: Live augmentation of human faces via projection," *Computer Graphics Forum*, vol. 36, no. 2, pp. 311–323, 2017.

[14] T. Sueishi, H. Oku, and M. Ishikawa, "Lumipen 2: Dynamic projection mapping with mirror-based robust high-speed tracking against illumination changes," *Presence: Teleoperators and Virtual Environments*, vol. 25, no. 4, pp. 299–321, 2016.

[15] D. Tone, D. Iwai, S. Hiura, and K. Sato, "Fibar: Embedding optical fibers in 3d printed objects for active markers in dynamic projection mapping," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 5, pp. 2030–2040, 2020.

[16] C. Siegl, M. Colaianni, L. Thies, J. Thies, M. Zollhöfer, S. Izadi, M. Stamminger, and F. Bauer, "Real-time pixel luminance optimization for dynamic multi-projection mapping," *ACM Trans. Graph.*, vol. 34, no. 6, oct 2015.

[17] M. T. Ibrahim, A. Majumder, and M. Gopi, "Dynamic projection mapping on deformable stretchable materials using boundary tracking," *Computers & Graphics*, vol. 103, pp. 61–74, 2022.

[18] T. Amano, K. Komura, T. Sasabuchi, S. Nakano, and S. Yamashita, "Appearance control for human material perception manipulation," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, 2012, pp. 13–16.

[19] K. Fujii, M. Grossberg, and S. Nayar, "A projector-camera system with real-time photometric adaptation for dynamic environments," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, 2005, pp. 814–821 vol. 1.

[20] P. Punpongsanon, D. Iwai, and K. Sato, "Projection-based visualization of tangential deformation of nonrigid surface by deformation estimation using infrared texture," *Virtual Reality*, vol. 19, no. 1, pp. 45–56, Mar 2015.

[21] Z. Li, L. Chan, T. Teo, and H. Koike, "Omniglobevr: A collaborative 360° communication system for vr," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '20, 2020, p. 1–8.

[22] K. Maeda, M. Piekenbrock, T. Sato, and H. Koike, "A tabletop system using an omnidirectional projector-camera," in *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, ser. ISS '18, 2018, p. 311–314.

[23] S. Toyohara, T. Sato, and H. Koike, "Balloonygen: Extended tabletop display embedded with balloon-like deformable spherical screen," in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '18, 2018.

[24] Z. Li, S. Miyafuji, T. Sato, H. Koike, N. Yamashita, and H. Kuzuoka, "How display shapes affect 360-degree panoramic video communication," in *Proceedings of the 2018 Designing Interactive Systems Conference*, ser. DIS '18, 2018, p. 845–856.

[25] Z. Li, S. Miyafuji, E. Wu, H. Kuzuoka, N. Yamashita, and H. Koike, "Omniglobe: An interactive i/o system for symmetric 360-degree video communication," in *Proceedings of the 2019 on Designing Interactive Systems Conference*, ser. DIS '19, 2019, p. 1427–1438.

[26] S. Pyo, J. Shim, and G. Lee, "Iprocam: A lens-sharing projector-camera system for augmented reality applications," in *CHI '07 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '07, 2007, p. 2615–2620.

[27] L. Mignard-Debise and I. Ihrke, "A vignetting model for light field cameras with an application to light field microscopy," *IEEE Transactions on Computational Imaging*, vol. 5, no. 4, pp. 585–595, 2019.

[28] D. Iwai and K. Sato, "Optical superimposition of infrared thermography through video projection," *Infrared Physics & Technology*, vol. 53, no. 3, pp. 162–172, 2010.

[29] H. Asayama, D. Iwai, and K. Sato, "Fabricating diminishable visual markers for geometric registration in projection mapping," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 2, pp. 1091–1102, 2018.

[30] T. Amano, "Projection center calibration for a co-located projector camera system," in *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 449–454.

[31] S. Huang, L. Xie, Z. Wang, Z. Zhang, F. Gao, and X. Jiang, "Accurate projector calibration method by using an optical coaxial camera," *Appl. Opt.*, vol. 54, no. 4, pp. 789–795, Feb 2015.

[32] G. Damberg, J. Gregson, and W. Heidrich, "High brightness hdr projection using dynamic freeform lensing," *ACM Trans. Graph.*, vol. 35, no. 3, may 2016.

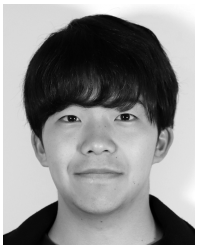
[33] F. Heide, D. Lanman, D. Reddy, J. Kautz, K. Pulli, and D. Luebke, "Cascaded displays: Spatiotemporal superresolution using offset pixel layers," *ACM Trans. Graph.*, vol. 33, no. 4, jul 2014.

[34] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[35] T. Taketomi, K. Okada, G. Yamamoto, J. Miyazaki, and H. Kato, "Geometric registration for zoomable camera using epipolar constraint and pre-calibrated intrinsic camera parameter change," in *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2013, pp. 295–296.

[36] T. Yoshida, C. Horii, and K. Sato, "A virtual color reconstruction system for real heritage with light projection," in *Proceedings of*

the International Conference on Virtual Systems and Multimedia 2003, 2003, pp. 161–168.
[37] S. Choi, M. Gopakumar, Y. Peng, J. Kim, and G. Wetzstein, "Neural 3d holography: Learning accurate wave propagation models for 3d holographic virtual and augmented reality displays," *ACM Trans. Graph.*, vol. 40, no. 6, dec 2021.
[38] Y. Bando, H. Holtzman, and R. Raskar, "Near-invariant blur for depth and 2d motion via time-varying light field analysis," *ACM Trans. Graph.*, vol. 32, no. 2, apr 2013.
[39] D. Iwai, H. Izawa, K. Kashima, T. Ueda, and K. Sato, "Speeded-up focus control of electrically tunable lens by sparse optimization," *Scientific Reports*, vol. 9, no. 1, p. 12365, Aug 2019.
[40] S. Mihara, D. Iwai, and K. Sato, "Artifact reduction in radiometric compensation of projector-camera systems for steep reflectance variations," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 9, pp. 1631–1638, 2014.
[41] Y. Li, A. Majumder, D. Lu, and M. Gopi, "Content-Independent Multi-Spectral Display Using Superimposed Projections," *Computer Graphics Forum*, 2015.
[42] D. Iwai, S. Mihara, and K. Sato, "Extended depth-of-field projector by fast focal sweep projection," *IEEE transactions on visualization and computer graphics*, vol. 21, no. 4, pp. 462–470, 2015.



Kenta Yamamoto received his B.S. and M.S. degrees from Osaka University, Japan, in 2020 and 2022, respectively. His research interests include projection mapping and augmented reality.

Daisuke Iwai received his B.S., M.S., and Ph.D. degrees from Osaka University, Japan, in 2003, 2005, and 2007, respectively. He was a visiting scientist at Bauhaus-University Weimar, Germany, from 2007 to 2008, and a visiting Associate Professor at ETH, Switzerland, in 2011. He is currently an Associate Professor at the Graduate School of Engineering Science, Osaka University. His research interests include spatial augmented reality and projector-camera systems. He is a member of the IEEE.

Ikuho Tani received his B.S. degree from Osaka University, Japan, in 2022. He is currently a master course student in Graduate School of Engineering Science, Osaka University, Japan. His research interests include projection mapping and augmented reality.

Kosuke Sato received his B.S., M.S., and Ph.D. degrees from Osaka University, Japan, in 1983, 1985, and 1988, respectively. He was a visiting scientist at the Robotics Institute, Carnegie Mellon University, from 1988 to 1990. He is currently a Professor at the Graduate School of Engineering Science, Osaka University. His research interests include image sensing, 3D image processing, digital archiving, and virtual reality. He is a member of ACM and IEEE.