

Simultaneous Projection and Positioning of Laser Projector Pixels

Yuki Kitajima, Daisuke Iwai, *Member, IEEE*, and Kosuke Sato, *Member, IEEE*

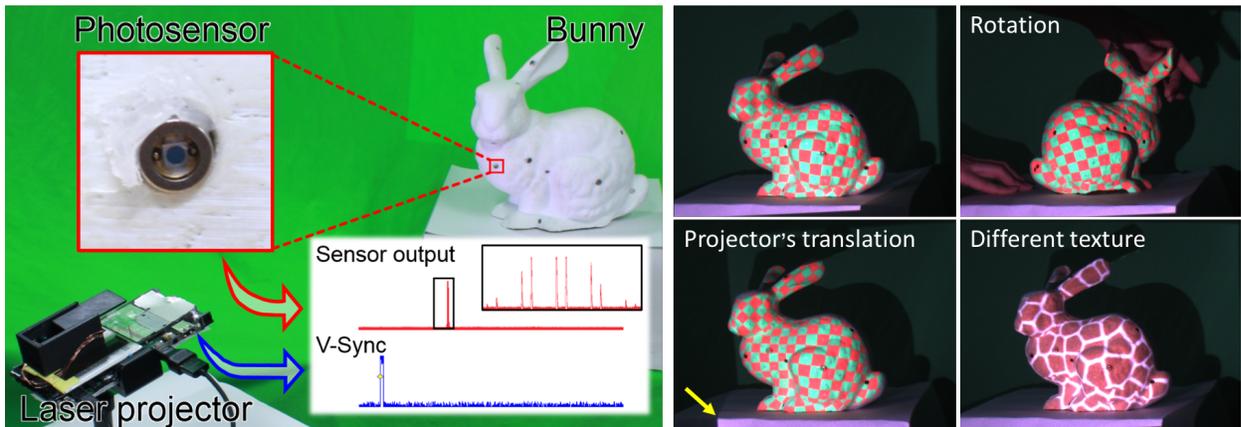


Fig. 1. Dynamic projection mapping of the SPAP (Simultaneous Projection And Positioning) approach: (left) a prototype system of the direct method consisting of a laser projector and photosensors embedded in a projection surface and (right) projection mapping results for different surface poses (rotation and translation) and different textures. The yellow arrow indicates the left edge of the projected image that is horizontally shifted towards right due to the projector's translation.

Abstract—This paper presents a novel projected pixel localization principle for online geometric registration in dynamic projection mapping applications. We propose applying a time measurement of a laser projector raster-scanning beam using a photosensor to estimate its position while the projector displays meaningful visual information to human observers. Based on this principle, we develop two types of position estimation techniques. One estimates the position of a projected beam when it directly illuminates a photosensor. The other localizes a beam by measuring the reflection from a retro-reflective marker with the photosensor placed in the optical path of the projector. We conduct system evaluations using prototypes to validate this method as well as to confirm the applicability of our principle. In addition, we discuss the technical limitations of the prototypes based on the evaluation results. Finally, we build several dynamic projection mapping applications to demonstrate the feasibility of our principle.

Index Terms—Dynamic projection mapping, spatial augmented reality, laser projector, light pen, geometric registration.

1 INTRODUCTION

Dynamic projection mapping (PM), or spatial augmented reality (AR), is a technology that controls the appearance of an arbitrarily shaped surface by projecting computer-generated images onto it while the surface and/or projector is moving [7, 8, 11, 36, 40]. Even though multiple systems have been proposed, the projectors applied in the majority of these prior systems, such as DLP (digital light processing) and LCD (liquid crystal display), are not suitable in some dynamic PM applications. In these projectors, a light ray from a light source is spatially modulated on a SLM (spatial light modulator) and then projected through an objective lens. To increase the brightness of the projected imagery, the aperture of the projector is normally designed to be large, which consequently leads to a narrow depth-of-field (DOF). A projection surface is expected to move over a wide area range in a dynamic PM scenario; however, conventional projectors cannot project sharp images over such wide areas. Previously developed defocus blur compensation or DOF extension techniques can improve this situation but only to a limited extent [12, 22, 35]. Conversely, in theory, a laser

projector has an infinite DOF [17] and therefore is suitable for dynamic PM [19]. In a dynamic PM system with a laser projector, it is inevitable to assume a situation where the distance between the projector and the surface varies significantly. Therefore, we need to geometrically register the projector so that the pixel alignment on the surface is consistently accurate with respect to the distance.

In this paper, we propose the novel geometric registration approach of a laser projector for dynamic PM. Our approach is an extension of the principle of the light pen [3]. The light pen, in which a photosensor is embedded, can measure its position on a CRT (cathode ray tube) screen at each frame while displaying meaningful image content. The pen detects changes in the brightness of nearby screen pixels when the CRT electron beam scans across them and communicates the timing of this event to the computer. Because a CRT scans the entire screen one pixel at a time, the computer can estimate the pen's position from the latest timestamp. The image forming mechanism of the laser projector is also based on raster scanning. A MEMS (Micro-Electro-Mechanical Systems) mirror adjusts the direction of a projected beam, and the color of each projector pixel is controlled by modulating the laser diode intensities of different primary colors. Leveraging this mechanism at each frame, we measure the time when a projected beam scanning over a projection surface hits a photosensor that is, for example, embedded in the surface. The time information is then used to estimate the position of the beam in the projector's screen coordinate system when it illuminates the sensor. Because the time information is invariant to the distance from the projector to the sensor, this method does not depend on the distance and therefore meets the requirement for dynamic PM

- Yuki Kitajima, Daisuke Iwai, and Kosuke Sato are with Graduate School of Engineering Science, Osaka University. E-mail: see <http://www.sens.sys.es.osaka-u.ac.jp/>.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxx

described in the previous paragraph. This principle allows us to measure the position of a photosensor in each frame while projecting meaningful image content and, therefore, is referred to here as **Simultaneous Projection And Positioning (SPAP)**.

Based on the SPAP principle, we develop two types of geometric registration methods. One estimates the position of a projected beam when it directly illuminates the photosensor. The other localizes a beam that is reflected from a retro-reflective marker by placing the photosensor in the optical path of the projector. We implement prototype systems of these two methods (Figure 1(left) shows a system with the former method) and investigate their geometric registration performances by evaluating the estimation accuracies. The estimation errors are measured by varying the distance from the laser projector to the photosensor or a retro-reflective marker, the projected light intensity, and other critical factors. Finally, we implement various dynamic PM applications to show the feasibility of our techniques. They include a projector-based texture mapping for a moving 3D surface (Figure 1(right)), a magic lens, a warehouse management application with a handheld projection system, and a drone projection. In addition, we demonstrate the geometric alignment of multiple overlapping projectors based on our method.

Our methods can be clearly distinguished from the conventional light pen technology for a CRT display. We shift the locus of the synchronizing signal from the CRT display surface to more general real world geometry. We tackle several issues unique to a projection display. These include the recognition of multiple projectors in an overlapping area and the position estimation of a pixel projected onto a retro-reflective marker by a reflected light measurement. In addition, we investigated the effects of sensor distance and the intensities of both environmental light and the projected beam on the position estimation accuracy.

To summarize, this paper provides the following contributions:

- We propose two types of novel laser projector geometric registration techniques for dynamic PM based on the principle of a light pen.
- We investigate the geometric registration performance of the techniques by evaluating the estimation accuracies of prototype systems under various conditions.
- We demonstrate several dynamic PM applications showing the feasibility of the techniques.

2 RELATED WORK

Following the pioneering work [11], geometric registration techniques for dynamic PM have been widely studied. Online geometric registration can be achieved by applying an external sensor to measure the six degrees of freedom (DOF) poses of a surface relative to a projector and geometrically transform the original projection textures to align with the surface based on the pose information. Even though this earliest study uses a magnetic tracking sensor to measure the pose, the majority of previous studies have applied either computer vision- or photosensor-based tracking technologies, which roughly fall into two categories: pose estimation and code projection approaches.

2.1 The pose estimation approach

As other types of AR systems (e.g., video see-through and optical see-through) [27], dynamic PM has also applied marker-based tracking technologies. Several studies have attached printed 2D markers onto planar projection surfaces, and have estimated poses from captured images using nearly the same technique as in typical video see-through AR [26, 40]. Other studies have applied the principle of a motion capturing system [5, 7, 8, 33]. These studies attached multiple markers to a non-planar rigid surface, applied multiple cameras to measure the markers' 3D positions using stereo reconstruction techniques, and estimated the pose of the surface from the measured information. Marker-less techniques have also been investigated. Zheng et al. provides a general model-based approach that iteratively optimizes the pose of the

projected image to minimize the 2D image differences between the captured (real) and expected (simulated) appearances of the projected surface [47]. Instead of using the entire image, other works estimate the pose of the surface by comparing discrete features detected in the captured and expected appearances [25, 41]. Leveraging recent tremendous advancements in depth imaging, Siegl et al. used a RGB-D camera to estimate the pose of a projection surface employing a projective iterated closest point (ICP) algorithm [42].

A projector can be geometrically calibrated using a camera calibration technique [46] due to the optical duality in these devices. However, projectors cannot directly measure fiducial calibration markers, and therefore, a camera is needed to acquire the correspondences between the projected pixels and markers [10, 45]. The increment of the number of system components generally leads to an accumulation of calibration errors. Therefore, such geometric calibration methods are inherently error-prone. Furthermore, as indicated in [45], accurate registration is achieved only when a projection surface is located in an area where the calibration is performed (e.g., where the markers are placed). However, as described in Section 1, a laser projector has an infinite DOF and, therefore, always displays sharp images on a surface, even outside the area where the calibration is conducted, which leads to significant misalignments of the projected imagery.

A coaxial projector-camera system (ProCams) relaxes this constraint. Because the projector and camera are optically identical in a coaxial ProCams, the pixel correspondences between the devices are consistent even when the distance between the projector and the surface is changed [6, 23]. It has been demonstrated that the coaxial setup can register projected textures with respect to even a non-rigid, deformable surface [13, 36, 37]. However, cameras generally have limited DOFs, which constrains the area where user interactions are available.

2.2 The code projection approach

Another geometric registration approach is spatiotemporal pattern projection. A series of patterns encode the position of each pixel in the projector's screen coordinate system as binary codes (e.g., graycode). Therefore, this approach directly acquires the position by decoding the series of projected patterns, which are measured either by photosensors or cameras. Lee et al. and Raskar et al. embedded photosensors in a projection surface to measure the graycode patterns projected onto them [30, 38, 43]. The decoded information could then be used to transform projection textures to align with the surface. Camera-based techniques capture projected patterns reflected on a projection surface and directly determine the correspondences of the projector pixels and the surface points [16, 49]. This approach achieves more accurate registration than the pose estimation approach because geometric calibration to know the relationship between different coordinate systems, which is generally error-prone as described above, is not necessary. However, it is not suitable for dynamic PM because the surface and/or the projector cannot move while multiple patterns are being projected frame by frame. Furthermore, the patterns must be projected again if the surface or the projector moves. For example, a usual XGA (1024 × 768 pixels) projector requires 20 patterns (= 10 + 10 bits) to be projected. Therefore, these systems cannot continually register the projector with a moving surface at an interactive rate.

To track the surface, Lee et al. proposed reducing the number of required graycode patterns in their photosensor-based system [31]. This method projects graycode patterns in only a small area around each photosensor, and therefore only a small number of patterns are sufficient. Consequently, this system can track a surface moving at a slow speed. Zhou et al. embedded a unique temporal code in the mirror flip pattern of each pixel of a DLP (digital light processing) projector, which is detected by a photosensor [48]. Because each mirror flips at thousands of hertz, the code is imperceptible to human observers. Even though this technique can align projected textures with moving surfaces, it requires expensive, custom hardware setups (DLP Discovery Kit and a high-performance FPGA). In another study, a more affordable programmable high frame-rate projector (DLP LightCrafter) projected graycode patterns to estimate the positions and orientations of small robots in which photosensors were embedded [29]. Raskar

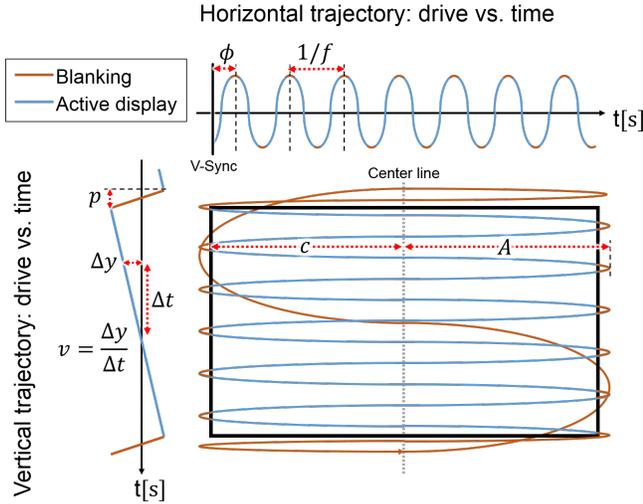


Fig. 2. Bidirectional raster scanning mechanism for a laser projector.

et al. built cost efficient high-speed binary code projectors to measure the position of photosensors [39]. Koderá et al. showed that smaller number of gradient patterns than that of graycode were sufficient for the photosensor positioning [28]. However, these systems do not display meaningful images for human observers.

2.3 Our approach

Our approach is fundamentally different from those previously described. Leveraging the raster scanning mechanism of a laser projector, we extend the touch position estimation principle of a light pen to estimate a projected pixel position by measuring the time when this pixel hits a photosensor. Our approach has the following advantages. First, it does not require error-prone extrinsic calibrations between the projector and the photosensors. Second, the geometric registration is consistently accurate with respect to the distance between the projector and a surface. Third, the registration is performed at each frame and therefore keeps registering the projector with the surface, even when they move. Fourth, it can be implemented with simple and inexpensive hardware setups. Finally, our technique can estimate the position of the projected beam while displaying meaningful visual information to human observers.

3 SIMULTANEOUS PROJECTION AND POSITIONING

To geometrically register a laser projector, we apply a time measurement of a projected beam using a photosensor. This section describes two types of registration techniques. The first one, referred to here as the **direct** method, estimates the position of the projected beam from the time when it directly illuminates on the photosensor. The second one, referred to as **reflection** method, estimates the position of the beam reflected back from a retro-reflective marker from the time measurement by placing the photosensor in the optical path of the projector.

In a laser projector, RGB beams are emitted from laser diodes, which are combined with beam combiners or prisms. The combined laser beam is then reflected on a two-axis scanning MEMS mirror and draws a 2D image on a surface based on the raster scanning principle (Figure 2). The horizontal scanning motion of the MEMS mirror is created by running the horizontal axis at its resonant frequency (e.g., 18 KHz [17]). The scan velocity varies sinusoidally with position. The vertical scan direction is driven at a video rate (=60 Hz) with a sawtooth waveform to provide constant velocity from the top to the bottom of the image and a rapid retrace back to the top on beginning a new frame. Therefore, we can assume a simple mathematical model of the vertical and horizontal positions ($x(t)$ and $y(t)$, respectively) as follows:

$$x(t) = A \cos(2\pi f t - \phi) + c, \quad (1)$$

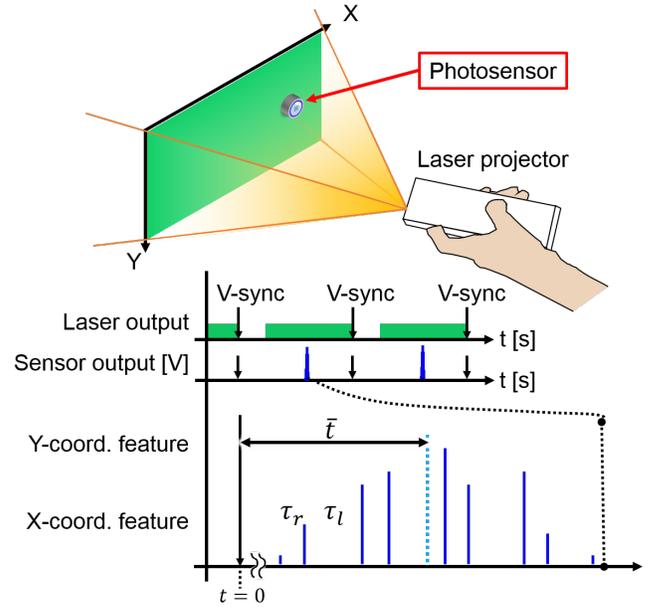


Fig. 3. Position estimation principle for the direct method.

$$y(t) = v(t - p), \quad (2)$$

where A , f , ϕ , and c represent the amplitude, frequency, phase, and the central position of the horizontal sinusoidal wave, respectively, and v and p represent the vertical velocity of the top-to-bottom movement and the time offset, respectively. In each frame, the time t is reset (i.e., $t = 0$) in synchronization with the V-Sync signal of the projector.

3.1 The direct method

In the direct method, photosensors are embedded in the projection surfaces. For the projected beam measurement, we developed a sensor node that consists of a photosensor, a low performance microcontroller, a battery, an amplifying circuit, and other electronic circuit modules. Each node converts the output current from the photosensor to a voltage, amplifies it, and sends the amplified voltage data to a host node.

3.1.1 Single projection system

Suppose a situation where we place a photosensor in the field of view (FOV) of a laser projector that projects a spatially uniform color as shown in Figure 3. Because the light receiving section of the sensor is, in general, not infinitesimally small, the sensor detects the beam several times during each frame. As shown in Figure 3, the output signal of the sensor in a frame consists of multiple pulses indicating that the projected beam passes across different portions of the light receiving section.

For the horizontal position estimation, we do not directly use the time from the V-Sync until a pulse. Instead, focusing on three adjacent pulses, we use the ratio of the time between the first and second pulses to that between the second and third pulses (Figure 3). We find that this strategy provides better estimation results because it is not susceptible or sensitive to a fluctuation of f due to a disturbance such as a deviation in the heat inside the projector and the environmental magnetic field [17]. Note that this method is simple enough to work at an interactive rate (60 Hz) on a sensor node controlled by a low performance microcontroller.

The detailed process for this method is as follows. We denote the time spent scanning back and forth on the left (or right) side of the sensor as τ_l (or τ_r). Because a single sensor measures multiple pulses in each frame, as mentioned above, multiple τ_l and τ_r values are obtained. Therefore, we use the median values (denoted as $\bar{\tau}_l$ and $\bar{\tau}_r$, respectively) to compute the ratio. Suppose the ratio is $d = \bar{\tau}_r / \bar{\tau}_l$, the horizontal

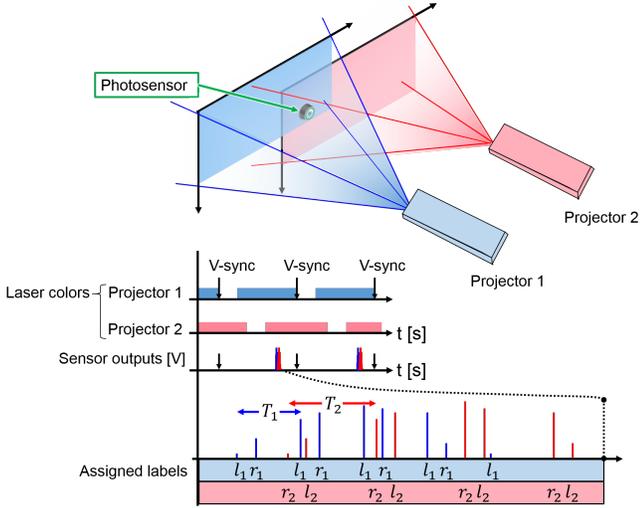


Fig. 4. Projector recognition principle in an overlapping area of a multi-projection system.

position is estimated by rewriting Equation (1) such that

$$\hat{x}(t) = A \cos\left(2\pi \frac{1}{\tilde{\tau}_l + \tilde{\tau}_r} \frac{\tilde{t}_l}{2}\right) + c, \quad (3)$$

$$= A \cos\left(\frac{\pi}{1+d}\right) + c. \quad (4)$$

Note that V-Sync and the scanning direction are obtained directly from the driving circuit of the laser projector. In the following, we refer to the latter signal as H-Drct, which takes two states, left and right. The vertical position is estimated using Equation (2). Here, we directly use the time t because v and p in the equation are less sensitive to environmental disturbances than f . Because multiple t values are measured by the sensor in each frame, we estimate $y(t)$ with the averaged value of t , which is denoted as \bar{t} :

$$\hat{y}(t) = v(\bar{t} - p). \quad (5)$$

3.1.2 Multi-projection system

Multi-projection systems are useful to increase the FOV as well as the maximum luminance of the projected imagery [32, 44]. To realize such systems, we propose a method to separately estimate the positions of projected beams from multiple projectors on a photosensor when they overlap each other. In the following, we assume the simplest case where two non-synchronized projectors illuminate a single sensor without any loss of generality. In such a case, the output signal of the sensor is a mixture of the pulses of two projectors as shown in Figure 4. Once we can correctly classify the pulses into two clusters and recognize the correct relation between each cluster and its corresponding projector, the position of the sensor in each screen coordinate system can be independently computed using the method described in Section 3.1.1.

The pulses are clustered and their projectors are recognized using the period of the wave in Equation (1). We denote the period as T_i , where $i \in \{1, 2\}$ represents the projector's ID. Because the period T_i fluctuates over time as described in Section 3.1.1, we estimate it as the sum of τ_l and τ_r of the previous frame. First, we focus on the first pulse of the current frame. We can determine that the first pulse is caused by projector i if there is another pulse that is detected T_i from it. Then, from the H-Drct information, we can check whether a projected beam hitting the sensor to the left or right caused the first pulse. If it is to the left, we assign the label l_i to the first pulse. Otherwise, r_i is assigned. Then, we repeatedly assign the same label l_i (or r_i) to pulses that are T_i from the previous pulse labeled as l_i (or r_i). Once this process is completed, we repeat this process for the rest of the pulses and assign r_i (or l_i) to the pulses that are caused by the same beam (i.e., pulse

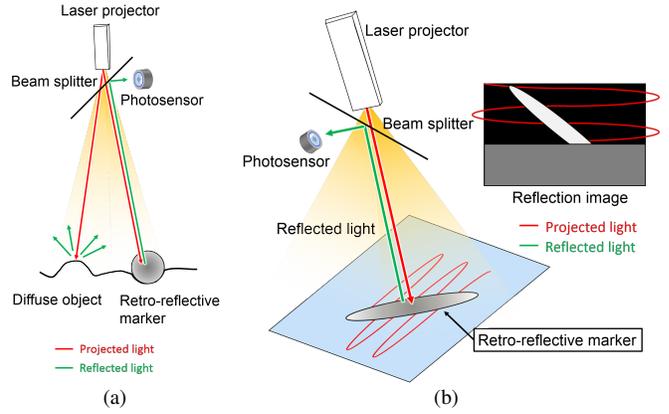


Fig. 5. Principle of the reflection method: (a) an overview of the reflection method and (b) the generation of a reflection image.

interval is T_i) hitting from the other direction. Next, we change our focus from projector i to j ($j \neq i$) and repeat the same processes with T_j for the rest of the pulses assigning l_j or r_j to them. Finally, we can compute τ_l and τ_r for each projector from the clustered pulses.

This method works well, except in a case where $T_1 = T_2$. In this case, it is not guaranteed that pulses belonging to l_1 and r_1 are caused by the beams of projector 1 and vice versa. To solve this problem, we use the previous estimated positions, $[\hat{x}_1(t-1) \hat{y}_1(t-1)]^t$ and $[\hat{x}_2(t-1) \hat{y}_2(t-1)]^t$. In particular, we compute four sensor positions, i.e., $[\hat{x}_k(t) \hat{y}_k(t)]^t$ ($k = 1, 2, 3, 4$), using the pulses belonging to l_1 and r_1 , l_1 and r_2 , l_2 and r_1 , and l_2 and r_2 , respectively. Then, for each projector, the sensor position in the current frame is determined by choosing one of the four positions so that the Euclidean distance between the previous and current positions are the closest. Failures still occur when $T_1 = T_2$ and the sensor positions in the screen coordinate systems of the two projectors are very close. However, even this critical situation, which is very rare, does not last long because the period T_i fluctuates over time as mentioned above.

3.2 Reflection method

The direct method relies on sensor nodes connected to a host node. Conversely, the reflection method is a stand-alone approach that does not require the embedding of any sensors in the surfaces; rather, it requires that only a single sensor be embedded in the laser projector. Instead of photosensor-based markers, the reflection method applies retro-reflective markers. When a projected beam illuminates the retro-reflective markers, most of the reflected light returns to the projector along the same path of the incident beam. Conversely, when the beam is projected onto the other regions whose reflectance properties are not retro-reflective, only a small fraction of the reflected light returns. Therefore, if a photosensor is placed in the projector's optical path as shown in Figure 5(a), its output shows significantly higher values only when the beam illuminates a retro-reflective marker.

The reflection measurement of a raster scanned beam using a photosensor has been applied to a barcode reader [1], a flying-spot scanner [2], and a range finder [4, 34]. Other studies worked on displaying meaningful images with a single color for human observers, while measuring the reflection of a projected beam [15, 24]. Compared to these prior works, our method realizes the reflection measurement while displaying full color images.

Consider a 2D image with the same spatial resolution as the projector. Each pixel value is determined by the photosensor's output signal, which is measured when a projector pixel of the same position is projected. In each frame, this process generates an image, referred to here as a **reflection image**, which is the equivalent of an image captured by a camera with the projector's FOV (Figure 5(b)). The reflection image shows the retro-reflective markers in a manner similar to a motion capture camera consisting of an IR camera and an IR LED

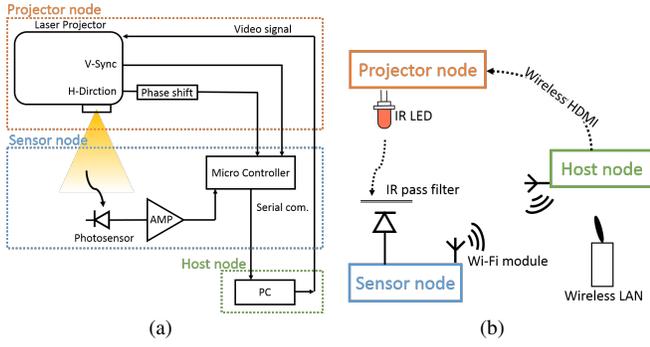


Fig. 6. System diagram of the direct method: (a) the wired system and (b) the wireless system.

array ring attached around its lens. Because the reflection image can be processed as a normal camera image, any computer vision algorithm can be applied without modifications. The scalability of this method is high because it only requires attaching widely used retro-reflective markers on surfaces. A projection system of this method currently needs to be connected to a normal PC to handle heavy processes that cannot be handled by the microcontroller applied in the direct method. But, we believe this does not make the scalability significantly lower.

The position estimation method of a reflected projected beam is different from that of the direct method. If there are multiple markers, it is difficult to know if two timestamps, in which high values are measured, are caused by the same marker. Therefore, the ratio of the time spent scanning on a marker's left side and that spent on its right side cannot be computed. If we naïvely use the time from V-Sync, we fail to accurately estimate the position of a beam due to the fluctuation in the frequency f in Equation (1). Instead, we find that the H-Drcr signal correctly encodes the frequency f , and therefore we can estimate f from the previous H-Drcr signal at each frame. Based on the estimated \hat{f} , we can estimate the horizontal position as follows:

$$\hat{x}(t) = A \cos(2\pi\hat{f}(t - t_H) - \phi) + c, \quad (6)$$

where t_H represents the time when the H-Drcr state changes (i.e., left to right or vice versa). The vertical position is estimated using Equation (2). One might think that Equation (6) is simpler than Equation (4), and that we should use it for the direct method as well. However, we assumed the use of a low performance microcontroller in the direct method, as mentioned in Section 3.1, and the computational cost required to measure t_H at each horizontal line is generally too high for current microprocessors.

3.3 Estimated position refinement

Via a preliminary test, we found that our mathematical model (Equations (1) and (2)) caused non-negligible errors in the estimated pixel positions and that an accurate alignment of a projected image onto a surface was not achieved. However, we also found that the errors were (1) spatially varying, (2) consistent regardless of the distance from the projector to the photosensor or retro-reflective marker, and (3) did not change over time. Therefore, we refined the estimated positions as follows. In an offline process, we measure the errors at all projector pixels (or interpolate them from sparse measurements) and store them as a look-up-table (LUT). In an online process, we refine an estimated pixel position by subtracting the corresponding error in the LUT from the estimated position.

4 PROTOTYPE

We built prototype systems of the direct and reflection methods. This section describes the system configurations followed by the parameter identifications for the position estimation model.

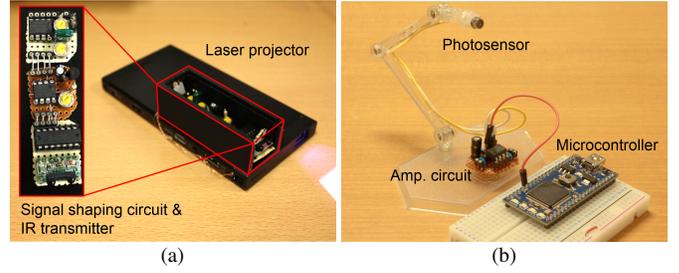


Fig. 7. Prototype of the direct method: (a) the projector node and (b) the sensor node.

4.1 Direct method

In the direct method, the system consists of three components: a projector node, a sensor node, and a host node. The system diagram is shown in Figure 6.

Projector node: Figure 7(a) shows the system overview for our projector node in which we apply a Sony MP-CL1 (1280 × 720 pixels) as the laser projector. The projector receives a video signal from the host node. While displaying the received image, we acquire V-Sync and H-Drcr signals from the driving circuit of the projector. The H-Drcr signal passes through a phase-shift circuit where the phase is shifted by 90° to synchronize the laser output. The processed V-Sync and H-Drcr signals are then sent to the sensor nodes.

Sensor node: Figure 7(b) shows the system overview for our sensor node. The received V-Sync and H-Drcr signals from the projector node are sent to GPIOs of a microcontroller (mbed LPC1768 96 MHz). The output signal from a high-speed photodiode (Hamamatsu Photonics S5971) is amplified and then sent to the microcontroller. The microcontroller detects a pulse when the voltage of the signal is above a predetermined threshold. Then, it computes τ_r and τ_l using the internal timer and transmits them to the host node.

Host node: The host node is a PC (Intel Core i7 3.4GHz, 16GB memory, Geforce GTX 650), which receives the time information from the sensor nodes. It then computes the position of each sensor node using Equation (4). The projection images are geometrically transformed using the position information to align the projection surfaces in which the sensor nodes are embedded. For example, when the positions of four sensor nodes embedded in a planar surface are acquired, a projection image is generated via the homography transformation. The generated image is then transmitted to the projector node.

Wireless system: Considering the applicability and scalability of the system, a wireless system is preferable. Even though most of the experiments in this paper are conducted using a system with wired connections, we also built a wireless system. A projector node receives a video signal from a host node via a wireless HDMI module (Logitech LDE-WHDI202TR). It then broadcasts the V-Sync and H-Drcr signals to sensor nodes via a wireless IR communication module (ROHM RPM851A). This does not cause any significant delays in the communication and does not suffer interference from the projected imagery. τ_r and τ_l are transmitted from each sensor node to the host node via a WiFi communication module (ESPRESSIF ESP-WROOM-02) using the UDP protocol. The latency caused in the WiFi module is currently 12 ms.

4.2 Reflection method

Figure 8 shows the prototype system of the reflection method. We use the same projector and host node as for the prototype of the direct method described in Section 4.1 with another photosensor (Hamamatsu Photonics S5973-01), which has a faster response speed but a smaller light receiving section. The specification sheet of the projector states that the time spent illuminating a single pixel is only approximately 10 ns. Therefore, we require very fast measurements of the output signal from the photosensor to precisely generate a reflection image.

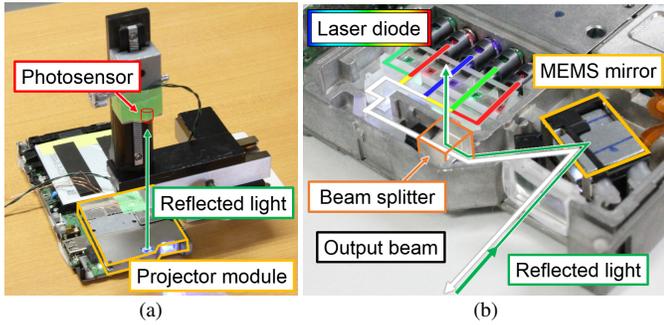


Fig. 8. Prototype of the reflection method: (a) overview and (b) close up.

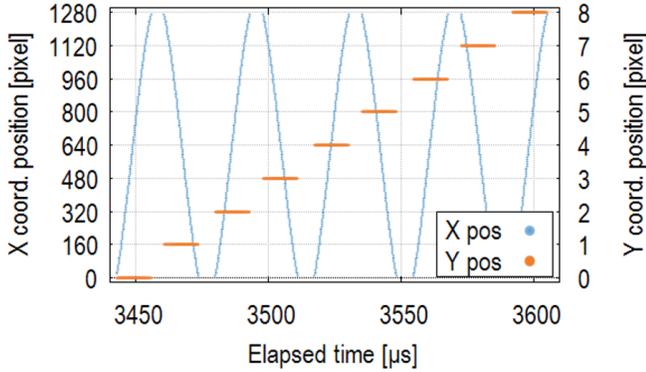


Fig. 9. Relationship between the measured time and the horizontal/vertical coordinate value.

To this end, we connected the photodiode to a USB Oscilloscope (Pico Technology Picoscope 6402C), which performs an A/D conversion of the input signal and directly sends the converted data to the host node. The host node computes a reflection image from the measured data and then applies various computer vision techniques to the image to estimate the poses of the retro-reflective markers. We used a retro-reflective material (3M Scotchlite 983-10) for our markers.

The essential assumption of the reflection method is that the photosensor is placed in the optical path of the projector. We achieve this by applying a beam splitter. One of the simplest methods is to place the beam splitter in front of the aperture of the projector as shown in Figure 5. However, we found that the signal-to-noise (S/N) ratio of such a naïve setup becomes low not only due to the reflected beam from the retro-reflective marker but also due to the environmental light incident on the photosensor. Furthermore, an image of the MEMS mirror always emerges in the reflection image due to second-order reflection in the beam splitter. These artifacts negatively affect the computer vision-based marker detection process. Instead, we propose inserting a beam splitter (Edmund, Non-Polarizing cube beam splitter 47-007) in the projector between the MEMS mirror and the laser diodes (Figure 8). This solves the problems of the naïve setup. First, because light reaching the photosensor is limited to the direction of the projected beam by the MEMS mirror and, consequently, the environment light outside the path of the beam does not reach the sensor, the S/N ratio becomes high. Second, in theory, the second-order reflection of the mirror should not emerge.

4.3 Parameter estimation

We estimated the parameters of our mathematical model in Equations (1) and (2). The same projector was used for the prototype systems of both the direct and reflection methods. Therefore, the estimated parameters are used in both systems.

We placed a sheet of the retro-reflective material in front of the projector entirely covering its FOV. Then, using the prototype system of the reflection model, we measured the time from V-Sync to a timestamp

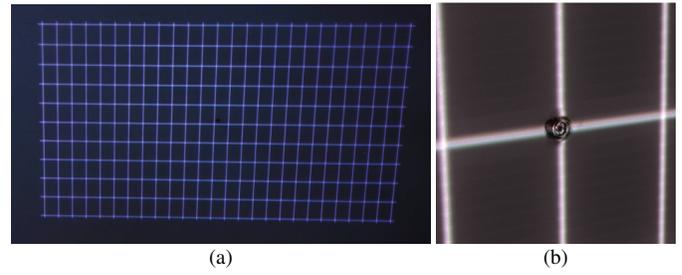


Fig. 10. LUT measurement: (a) grid pattern projection and (b) close up of photosensor.

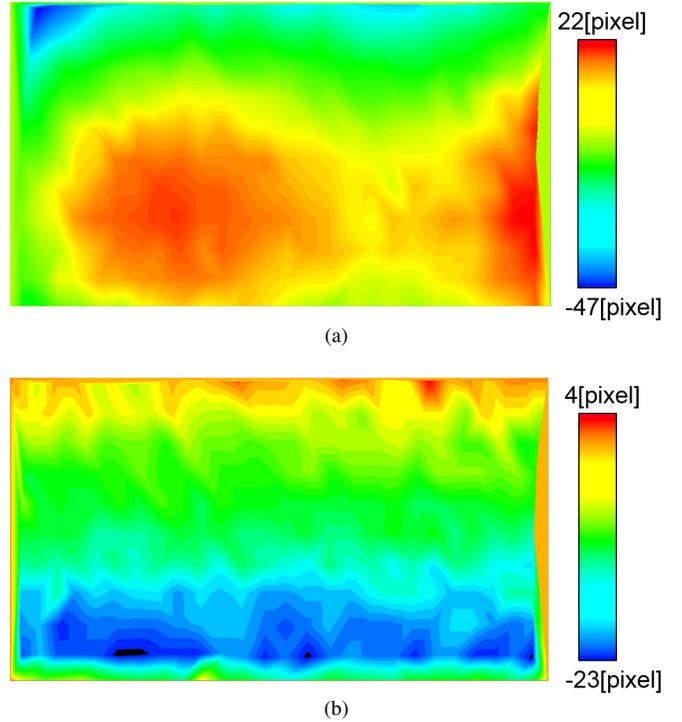


Fig. 11. Estimation errors: (a) horizontal direction and (b) vertical direction.

when a pulse was measured while the projector displayed a single dot pattern with a white (=255) pixel on a black (=0) background. This measurement was repeated every ten pixels in the horizontal direction and every pixel in the vertical direction, i.e., 128×720 pixels in total. For each pixel, we made 30 measurements of the time and averaged them. Figure 9 shows the averaged time data for the first 1000 pixels. We then found the parameters by fitting the data using the Levenberg–Marquardt Method. The estimated parameters are $A = -691.3$ pixels, $f = 26.76 \times 10^3$ Hz, $\phi = 9.3 \times 10^{-2}$ rad, $c = 651.3$ pixel, $v = 53.4 \times 10^3$ pixel/s, and $p = -3427.375$ μ s. The coefficient of determination R^2 of the fitting is 0.992 for Equation (1) and 0.999 for Equation (2).

We measured the estimation errors using the direct method in a dark room to prepare the LUT described in Section 3.3. We placed the photodiode of a sensor node 2 m from the projector and displayed a grid pattern consisting of 11×26 white grid lines and a gray background. We manually adjusted the direction of the projector so that each grid intersection overlaid the photodiode while maintaining the distance between the projector and the photodiode (Figure 10(a, b)). Then, we estimated the position of a projected pixel on the photodiode for each intersection point using Equations (4) and (5) with the estimated parameters. The estimated positions were then compared to the ground truth. Because the manual adjustment of the projector direction does

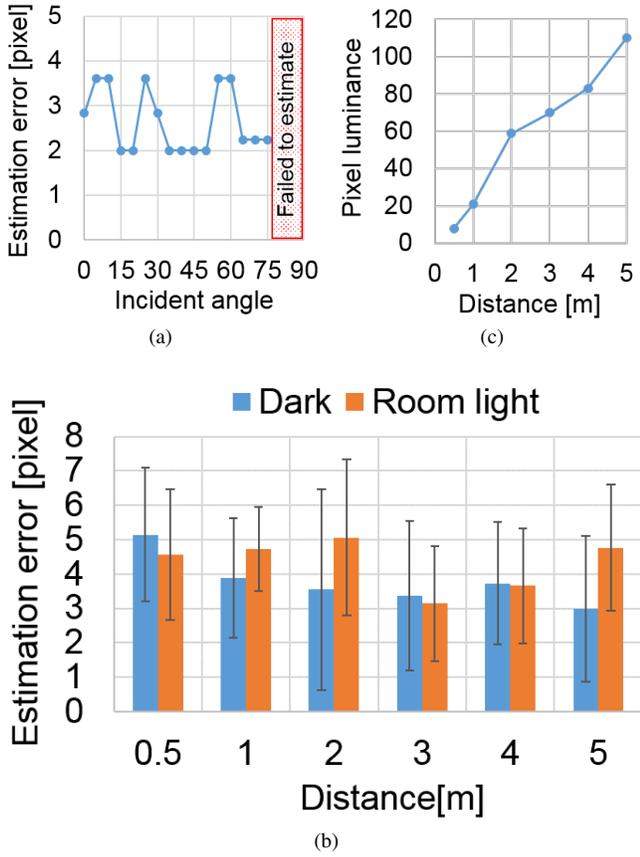


Fig. 12. Evaluation results of the direct method: (a) position estimation errors according to the incident angle, (b) position estimation errors according to the sensor distance and environment light, and (c) relationship between the minimum necessary luminance of a projected light and the sensor distance.

not always guarantee that the intersection point accurately hits the photodiode, we measured the ground truth using a graycode pattern projection technique [30]. Figure 11 visualizes the estimation errors in both the horizontal and vertical directions. Note that the errors for pixels that were not measured in this process are linearly interpolated. These error maps were stored as the LUT.

5 SYSTEM EVALUATION

We evaluated how accurately our methods could estimate the position of a projected beam on a photosensor or a retro-reflective marker using the prototype systems.

5.1 Direct method

There are several factors that could affect the accuracy of our pixel position estimation. We conducted the following four experiments to investigate these effects.

5.1.1 Effect of incident angle

In general, the sensitivity of a photodiode varies directionally. In particular, a projected beam hitting a sensor from a very shallow angle does not generate a current and, therefore, cannot be measured by the system. We measured the estimation errors by changing the incident angle of a projected beam onto the photodiode of a sensor node in a dark room. We placed a projector node 1 m from the photodiode and directed it so that the center of the projection image hit the sensor. Then, we projected a uniformly white image to estimate the pixel position using the direct method. We compared the estimated position with the ground truth measured using the graycode projection technique [30]. This measurement was repeated for different incident angles. In

particular, we changed the incident angle (defined as the angle from the normal) from 0° to 90° at intervals of 5°. Figure 12(a) shows the result. The estimation errors are less than 4 pixels from 0° to 75°, and the direct method fails to estimate the pixel positions at 80°, 85°, and 90°. Therefore, the pixel position estimation works when the incident angle is not too shallow ($\leq 75^\circ$).

5.1.2 Effect of sensor distance

The illuminance of the projected light decreases with the increase in the distance between the projector and the sensor. This may lead to a worse S/N ratio. We measured the estimation errors when changing the distance between the projector and the sensor under a dark room condition. Specifically, the measured distances were 0.5 m, 1 m, 2 m, 3 m, 4 m, and 5 m. We chose 5 m as the maximum distance because the projected result was too dark to understand the image content when the projection surface was placed more than 5 m from the projector. At each distance, we placed the sensor at nine positions so that they were evenly distributed in the projector's screen coordinate. Specifically, the positions were $(x, y) = (100, 60), (640, 60), (1180, 60), (100, 360), (640, 360), (1180, 360), (100, 660), (640, 660),$ and $(1180, 660)$. Then, we projected a uniformly white image and estimated the pixel positions ten times for each distance and each position. The estimation was compared to the ground truth measured using the graycode projection technique. Figure 12(b) shows the average and standard deviation of the errors at each distance. A one-way analysis of the variance (ANOVA) with repeated measurements showed that there is no statistically significant difference in the sensor distance ($p < 0.05$). Therefore, the sensor distance does not significantly affect the estimation accuracy when it is less than 5 m.

5.1.3 Effect of environmental light

The S/N ratio of the system may become worse with increases in the illuminance of the environmental light. We conducted the same experiment described in Section 5.1.2 under normal room light conditions (220 lx). Figure 12(b) shows the average and standard deviation of the measured errors at each distance. We conducted a two-tailed paired t -test for each pair of mean errors measured at the same distance in the dark room and in the normally lit room. No significant difference was found in the pairs ($p \geq 0.05$). Therefore, environmental light does not significantly affect the estimation accuracy when it is less than 220 lx.

5.1.4 Effect of the projected light intensity

The S/N ratio may become worse with decreases in the luminance of the projected beam. We investigated the minimum necessary luminance of a projected light with which we can estimate the position. We placed the sensor 0.5 m from the projector in the center of the projected area. We then projected a uniformly gray image and estimated the pixel position. We changed the gray level and found the minimum (darkest) value with which the pixel position could be estimated. This process was repeated for different sensor distances. Figure 12(c) shows the minimum necessary luminance of a projected light almost linearly increases as the sensor distance increases.

5.1.5 Projector recognition in a multi-projection system

We investigated if our projector recognition technique described in Section 3.1.2 works when projected images overlap each other in a multi-projection system. In this experimental setup, projected images from two projectors overlapped each other on a surface that was placed at a distance of 2 m. As shown in Figure 13(a), one of the projectors was rotated by 90° so that its projected image was in portrait orientation while the other was in landscape orientation. Then, we manually moved a sensor node on the surface as it traced the letter "S". Figure 13(b) shows the measured trajectory of the sensor in each projector's coordinate system. Because each trajectory draws the letter "S" in the correct direction, we confirmed that the projector recognition works well.

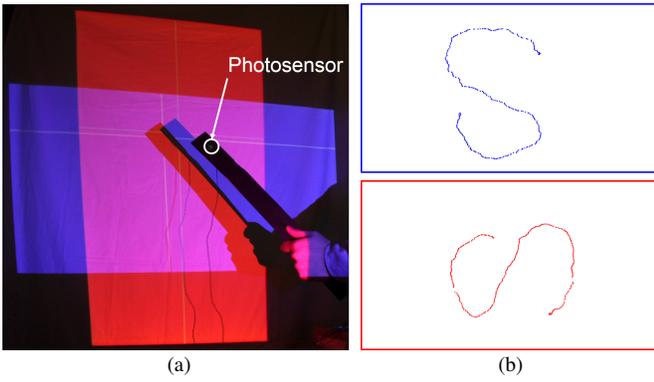


Fig. 13. Projector recognition result in a multi-projection system: (a) a captured scene of the evaluation and (b) the trajectory of the estimated pixel position in a projector in landscape orientation (top) and in portrait orientation (bottom).

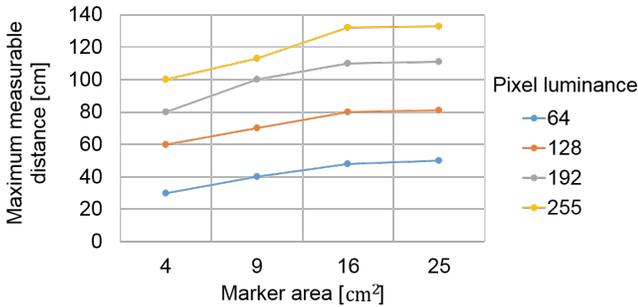


Fig. 14. Evaluation results for the reflection method.

5.2 Reflection method

Via a preliminary test using the prototype system of the reflection method, we found several factors that affected the maximum distance when a retro-reflective marker is visible in reflection images. Because some were already discussed in Section 5.1, we evaluate the effects of the other unique factors here. In particular, we investigate how the area of a retro-reflective marker and the intensity of a projected pixel affect the maximum distance.

We prepared four rectangular markers with different areas ($20 \times 20 \text{ mm}^2$, $30 \times 30 \text{ mm}^2$, $40 \times 40 \text{ mm}^2$, and $50 \times 50 \text{ mm}^2$) and four uniformly gray projection images of different gray levels (64, 128, 192, and 255). Therefore, there were 16 ($=4 \times 4$) experimental conditions. In each condition, we measured the maximum distance where we could extract the marker in a reflection image. In the extraction, we applied a simple threshold and connected component analysis. If the center of the largest connected component corresponded to the center of the marker up to the Euclidean distance of 5 pixels, we considered that the marker to be correctly extracted. By changing the distance between the projector and the marker, we found the maximum distance.

Figure 14 shows the experimental result. From this result, we confirmed that both the marker area and the pixel intensity affect the maximum distance. Concerning the marker area, the result shows that the maximum distance increases as the area increases. However, the maximum distance does not change between the marker areas of $40 \times 40 \text{ mm}^2$ and $50 \times 50 \text{ mm}^2$. Therefore, it is confirmed that a large marker area of up to $40 \times 40 \text{ mm}^2$ is preferable to provide a wide working space for a dynamic PM interaction system. Concerning the pixel intensity, the result shows that the maximum distance increases as the intensity increases. Therefore, it is always recommended to use a bright image when using the reflection method.

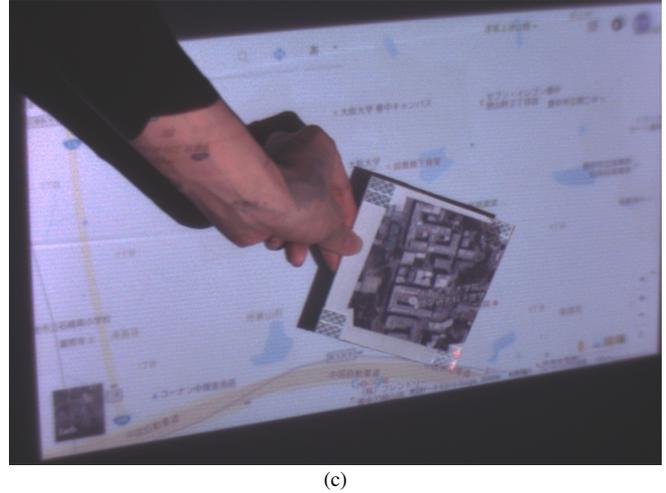
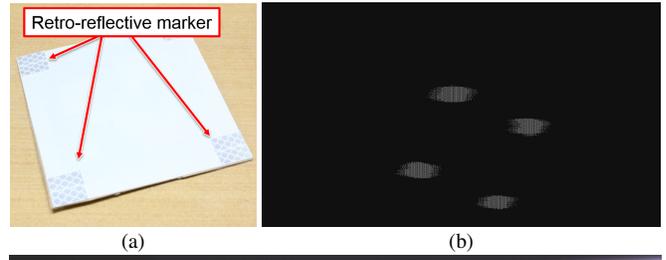


Fig. 15. Magic lens application: (a) a rectangular surface on which four retro-reflective markers are attached, (b) the reflection image, and (c) the projected result of the magic lens application in map viewing.

6 APPLICATIONS

We implemented several application systems to demonstrate the feasibility and usefulness of our approach.

6.1 Magic lens

We implemented a magic lens [14], through which a user can see another layer of displayed graphical information. We attached four retro-reflective markers on a rectangular projection surface (Figure 15(a)). The positions of the pixels projected on the markers were estimated using the reflection method, and these estimations were used to compute the homography matrix at each frame. The original graphical information was projected onto a large screen. A user can see another layer of information by inserting the surface above the projection screen. We replaced the original texture with another with a different layer of information using the estimated homography matrix. Figure 15(b) shows the reflection image, and (c) shows an implemented example where a user can see an aerial photograph of a map through the magic lens.

6.2 Light pen

We implemented a pen-shaped pointing interface similar to the light pen [3]. With a conventional light pen, a photosensor is embedded in the pen tip. Conversely, we embedded a photosensor so that the sensor is visible from the projector because the pen tip faces the projection screen and does not receive a projected beam. In addition, we integrated three push buttons for various clicking manipulations (Figure 16(a)). Figure 16 shows the prototype of our light pen based on the direct method. We demonstrate several interaction techniques such as drawing, 3D modeling, and map browsing as shown in Figures 16(b-d).

6.3 Dynamic PM

Because the main motivation of this work is to realize dynamic PM with a laser projector, we implemented an application where a texture image is geometrically aligned onto the surface of a Stanford bunny, as

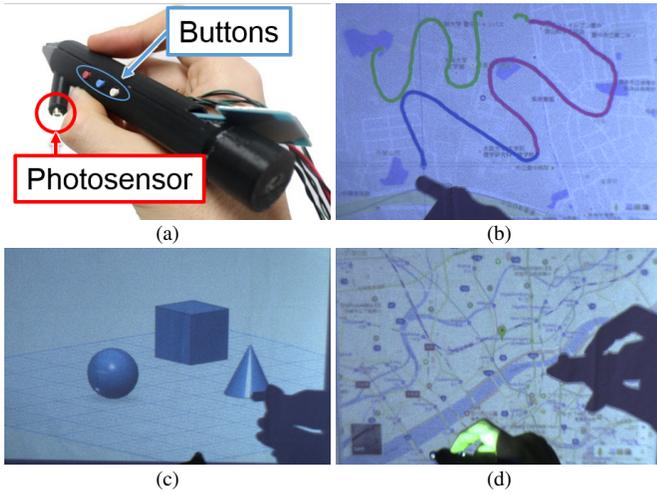


Fig. 16. Light pen application: (a) the light pen prototype, (b) drawing application, (c) 3D modeling application, and (d) map browsing application.

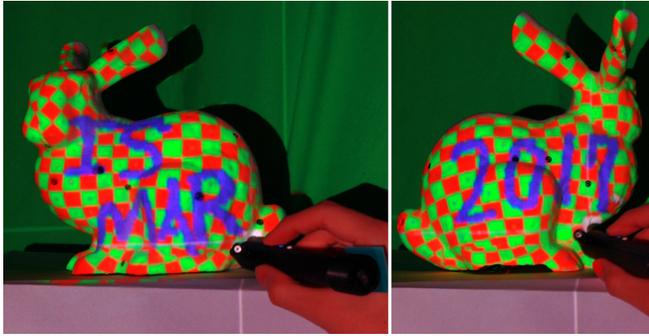


Fig. 17. Drawing on a Stanford bunny surface in the dynamic PM application.

shown in Figure 1. We fabricated the surface from a 3D printer with small holes that are used to embed the photosensors. The positions of the holes were determined in advance so that at least six sensors were visible from the projector regardless of the pose of the surface. We determined the positions of the 24 sensors using a technique similar to that proposed in one of our previous studies [8]. Using more than six sensor positions, we can estimate the 3×4 projection matrix \mathbf{P} by which we can compute which projector pixel (x, y) hits each surface point (X, Y, Z) as $h[x \ y \ 1]^t = \mathbf{P}[X \ Y \ Z \ 1]^t$, where t represents the transpose of the matrix. The pixel positions on the sensors were estimated using the direct method, and the projection matrix estimation was performed in each frame. Figure 17 shows other projected results. It was confirmed that the projected images are consistently aligned on the surface while it is translating and rotating. The figure also shows the combination of our light pen with this application, by which a user can draw on the surface.

6.4 Warehouse scenario

We implemented an application system assuming a warehouse scenario inspired by [38]. When a user directs a handheld projector node at inventory in which a wireless sensor node is embedded, s/he can see a projection of retrieved information collocated with the physical object. Figure 18 shows an implementation example. The system does not obtain the 3D pose of the inventory, but projects each texture image so that its center corresponds to each photosensor. The previous system [38] required several seconds to project spatial code patterns to acquire sensor positions. Conversely, our system achieves the same function while performing the sensor localization in each frame.

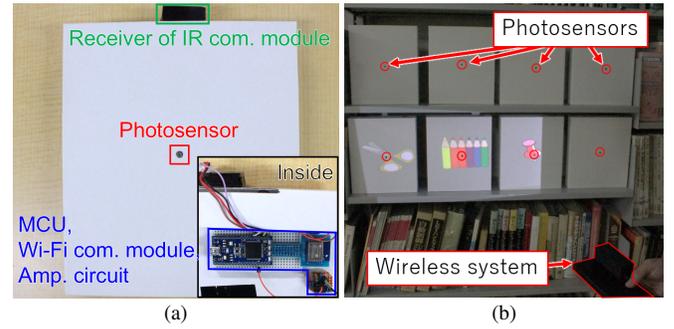


Fig. 18. Warehouse application: (a) an inventory object with a sensor node and (b) the projected result.

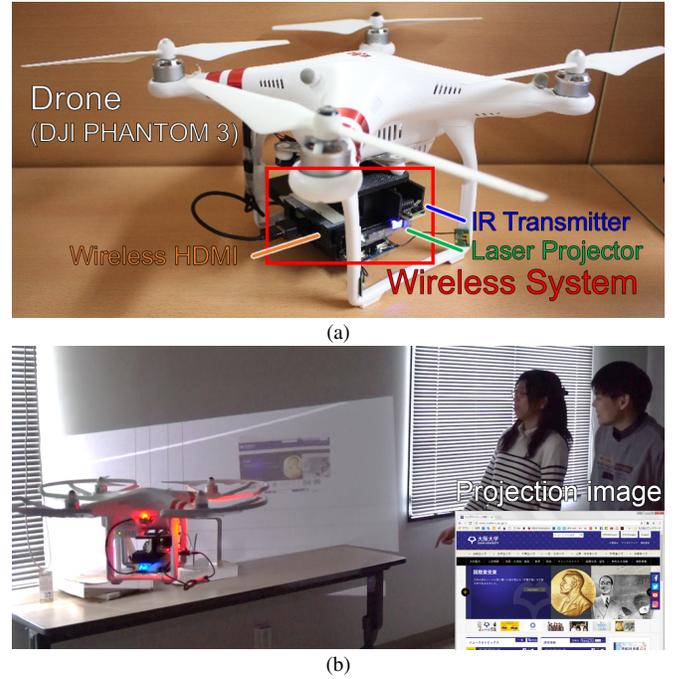


Fig. 19. A prototype of the drone projection (a) and its projected result (b).

6.5 Drone projection

Drone projection is an emerging research topic in which researchers attempt to realize an autonomous projection system that provides SAR anywhere [20, 21]. Because a drone on which a projector is mounted generally flies with fluctuations, the geometric registration parameters must be estimated in each frame. In our application, we mounted our projector node on a drone (DJI PHANTOM 3) and embedded four sensor nodes in a projection surface (Figure 19). We applied the wireless system to this application. Estimated pixel positions on the sensors were used to compute the homography matrix used to transform the projection image to align the surface. Figure 19 shows the result. We confirmed that the wireless system works and that a projection mapping can be achieved even when a projector is flying.

6.6 Stitching

It is possible to realize a larger projection display in a multi-projection system than in a single projection system [32]. To accurately stitch multiple projected images together, we need to know the pixel correspondences between multiple projectors in an overlapping area. We applied the direct method to a stitching application where a user sweeps a sensor node in the overlapping area of two projectors to acquire their pixel correspondences (Figures 20(a) and (b)). The correspondences are

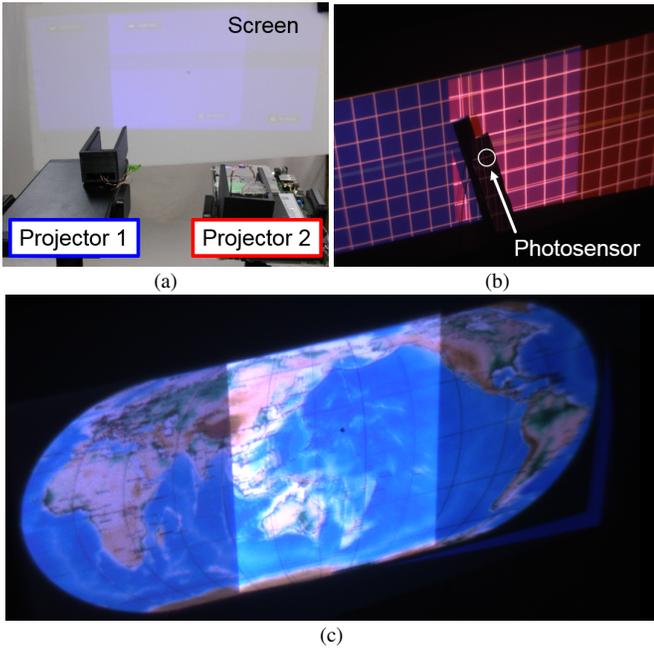


Fig. 20. Stitching of multiple overlapping projections: (a) system overview, (b) a retro-reflective marker sweeping in the overlapping area, and (c) the stitching result.

then used to stitch the projected images together. Figure 20(c) shows the experimental result, by which it was confirmed that our technique could be used in this application scenario.

7 DISCUSSION

From the system evaluation of the direct method (Section 5.1), we confirmed that it could estimate pixel positions on photosensors with an average estimation error of 4 pixels (equivalent to 0.32 % of the width of the projection image). The estimation error does not increase when either the sensor distance or the illuminance of the environmental light differs. The evaluation also shows that the projector recognition technique works correctly when two projectors overlap each other. Via the system evaluation of the reflection method (Section 5.2), we confirmed that the area of a retro-reflective marker and the projected pixel intensity affect the maximum distance where the marker can be accurately detected in a reflection image. We believe that our method can estimate pixel positions with a sufficient accuracy for various dynamic PM applications, as shown in Section 6.

As shown in the supplementary movie, the registration fails from time to time in the direct method, which clearly decreases the user experiences. This happens due to the limited performance of the microcontroller containing only one CPU that sometimes cannot handle signals arrived from multiple photosensors at the same time. This could be solved by applying a higher performance controller containing multiple CPUs. The movie also shows a delay of the registration even in the wired system. This is mainly caused by the latency in the video signal transfer via HDMI from the PC to the projector, which normally takes more than 60 ms. In addition, a relative motion of a surface to the projector might cause noise for the geometric registration in the movie. A motion perpendicular to the optical axis of the projector potentially causes wrong t measurements. The noise becomes larger when the surface and projector are closer to each other, because it makes the motion in the projector’s screen coordinate system larger. However, in practice, only extremely fast motion will cause significant noise, as it takes only 10 μ s to scan each line. A motion along the optical axis does not cause any noise because it does not affect t .

There are a couple of limitations in the current implementation. First, as shown in Section 5.1.2, 5 m is the maximum working distance of a wired prototype system for the direct method. When the projector of the

system is more than 5 m from the photosensor and the projection screen on which the sensor is attached, the projected beam is not sufficiently bright to form an image, and consequently, the sensor cannot detect the beam. The maximum working distance is also limited to approximately 1.3 m for the reflection method. This limitation is governed by the S/N ratio, i.e., the ratio of the reflected light intensity from a retro-reflective marker to the stray light. The S/N ratio is currently very low due to imperfections in the optical design. This could be improved by a more sophisticated system than the current “proof-of-concept” prototype.

Second, when the distance from the projector to the sensor or retro-reflective marker increases, the minimum necessary luminance of the projected light increases (see Section 5.1.4). A naïve solution to this limitation is to convert the luminance range of an original image from between 0 and 1 to between L_{min} and 1, where L_{min} ($0 \leq L_{min} \leq 1$) is the minimum necessary luminance at the maximum sensor distance (i.e., 5 m). Note that the original luminance range here is normalized. However, this method significantly decreases the contrast of a projected result. Another solution is to adaptively change the L_{min} value according to the sensor distance reducing the contrast degradation. A more sophisticated solution would be to increase the luminance only around the areas where the photosensors are placed. Even though this solution preserves the contrast in most areas, unnatural spatially varying luminance modifications could possibly be perceived by observers. Context aware, imperceptible, spatially varying luminance modification techniques might solve this issue [9, 18]. Integrating such adaptive luminance modification techniques with our method is an interesting direction for our future investigations.

8 CONCLUSIONS

This paper presents a novel online geometric registration technique for dynamic PM applications based on a novel measurement principle, SPAP, the time measurement of a raster scanning beam of a laser projector while projecting meaningful visual information to human observers. Via system evaluations, we validated the applicability of our approach to dynamic PM as well as its limitations. Based on the evaluations, we built several application examples to demonstrate the feasibility of the proposal. As discussed in Section 7, there are two directions for our future works. The first is building a system with a more sophisticated optical design to reduce the stray light in the reflection method. The second is to develop an adaptive luminance modification technique to improve the contrast of a projected image.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number JP16H02859. The authors would like to thank Yuta Itoh for providing us the information on disassembling the laser projector, and Takumi Kawahara for helping us in the drone projection experiment.

REFERENCES

- [1] Barcode Reader. https://en.wikipedia.org/wiki/Barcode_reader. Accessed: 2017-06-20.
- [2] Flying-Spot Scanner. https://en.wikipedia.org/wiki/Flying-spot_scanner. Accessed: 2017-06-20.
- [3] Light Pen. https://en.wikipedia.org/wiki/Light_pen. Accessed: 2017-03-13.
- [4] MicroVision PSE-0403Li-101 3D Sensing Mid-Range LiDAR. <http://www.microvision.com/microvision-ces-2017-recap/>. Accessed: 2017-06-20.
- [5] E. Akaoka, T. Ginn, and R. Vertegaal. DisplayObjects: Prototyping Functional Physical Interfaces on 3D Styrofoam, Paper or Cardboard Models. In *Proceedings of International Conference on Tangible, Embedded, and Embodied Interaction*, TEI ’10, pp. 49–56. ACM, New York, NY, USA, 2010.
- [6] T. Amano. Projection Based Real-Time Material Appearance Manipulation. In *Proceedings of IEEE International Workshop on Computational Cameras and Displays*, pp. 918–923, June 2013.
- [7] H. Asayama, D. Iwai, and K. Sato. Diminishable Visual Markers on Fabricated Projection Object for Dynamic Spatial Augmented Reality. In *Proceedings of ACM SIGGRAPH Asia Emerging Technologies*, SA ’15, pp. 7:1–7:2. ACM, New York, NY, USA, 2015.

- [8] H. Asayama, D. Iwai, and K. Sato. Fabricating Diminishable Visual Markers for Geometric Registration in Projection Mapping. *IEEE Transactions on Visualization and Computer Graphics*, PP(99):1–1, 2017.
- [9] M. Ashdown, T. Okabe, I. Sato, and Y. Sato. Robust Content-Dependent Photometric Projector Compensation. In *Proceedings of IEEE International Workshop on Projector Camera Systems*, pp. 6–6, June 2006.
- [10] S. Audet and M. Okutomi. A User-Friendly Method to Geometrically Calibrate Projector-Camera Systems. In *Proceedings of IEEE International Workshop on Projector Camera Systems*, pp. 47–54, June 2009.
- [11] D. Bandyopadhyay, R. Raskar, and H. Fuchs. Dynamic Shader Lamps: Painting on Movable Objects. In *Proceedings of IEEE and ACM International Symposium on Augmented Reality*, pp. 207–216, 2001.
- [12] A. Bermanno, P. Brüscheweiler, A. Grundhöfer, D. Iwai, B. Bickel, and M. Gross. Augmenting Physical Avatars Using Projector-based Illumination. *ACM Transactions on Graphics*, 32(6):189:1–189:10, Nov. 2013.
- [13] A. H. Bermanno, M. Billeter, D. Iwai, and A. Grundhöfer. Makeup Lamps: Live Augmentation of Human Faces via Projection. *Computer Graphics Forum*, 36(2):311–323, 2017.
- [14] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Tool-glass and Magic Lenses: The See-through Interface. In *Proceedings of Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, pp. 73–80. ACM, New York, NY, USA, 1993.
- [15] A. Cassinelli, A. Zerroug, Y. Watanabe, M. Ishikawa, and J. Anglesleva. Camera-Less Smart Laser Projector. In *Proceedings of ACM SIGGRAPH Emerging Technologies*, SIGGRAPH '10, pp. 9:1–9:1. ACM, New York, NY, USA, 2010.
- [16] D. Cotting, M. Naef, M. Gross, and H. Fuchs. Embedding Imperceptible Patterns into Projected Images for Simultaneous Acquisition and Display. In *Proceedings of IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 100–109, Nov 2004.
- [17] M. Freeman, M. Champion, and S. Madhavan. Scanned Laser Pico-Projectors: Seeing the Big Picture (with a Small Device). *Optics & Photonics News*, 20(5):28–34, May 2009.
- [18] A. Grundhöfer and D. Iwai. Robust, Error-Tolerant Photometric Projector Compensation. *IEEE Transactions on Image Processing*, 24(12):5086–5099, Dec 2015.
- [19] C. Harrison, H. Benko, and A. D. Wilson. OmniTouch: Wearable Multitouch Interaction Everywhere. In *Proceedings of ACM Symposium on User Interface Software and Technology*, UIST '11, pp. 441–450. ACM, New York, NY, USA, 2011.
- [20] Y. Hosomizo, D. Iwai, and K. Sato. A Flying Projector Stabilizing Image Fluctuation. In *Proceedings of IEEE Global Conference on Consumer Electronics*, pp. 31–32, Oct 2014.
- [21] W. A. Isop, J. Pestana, G. Ermacorra, F. Fraundorfer, and D. Schmalstieg. Micro Aerial Projector - stabilizing projected images of an airborne robotics projection platform. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5618–5625, Oct 2016.
- [22] D. Iwai, S. Mihara, and K. Sato. Extended Depth-of-Field Projector by Fast Focal Sweep Projection. *IEEE Transactions on Visualization and Computer Graphics*, 21(4):462–470, April 2015.
- [23] D. Iwai and K. Sato. Optical Superimposition of Infrared Thermography through Video Projection. *Infrared Physics & Technology*, 53(3):162–172, 2010.
- [24] S. Jeon, H. Fujita, and H. Toshiyoshi. A MEMS Interactive Laser Projection Display with a Built-in Laser Range Finder. In *Proceedings of International Conference on Optical MEMS and Nanophotonics*, pp. 15–16, Aug 2013.
- [25] T. Johnson and H. Fuchs. Real-Time Projector Tracking on Complex Geometry Using Ordinary Imagery. In *Proceedings of IEEE International Workshop on Projector Camera Systems*, pp. 1–8, June 2007.
- [26] T. Karitsuka and K. Sato. A Wearable Mixed Reality with an On-Board Projector. In *Proceedings of IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '03, pp. 321–. IEEE Computer Society, Washington, DC, USA, 2003.
- [27] H. Kato and M. Billinghurst. Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System. In *Proceedings of IEEE and ACM International Workshop on Augmented Reality*, pp. 85–94, 1999.
- [28] T. Kodera, M. Sugimoto, R. T. Smith, and B. H. Thomas. 3D Position Measurement of Planar Photo Detector using Gradient Patterns. In *Proceedings of IEEE Virtual Reality*, pp. 211–212, March 2015.
- [29] M. Le Goc, L. H. Kim, A. Parsaei, J.-D. Fekete, P. Dragicevic, and S. Follmer. Zooids: Building Blocks for Swarm User Interfaces. In *Proceedings of ACM Symposium on User Interface Software and Technology*, UIST '16, pp. 97–109. ACM, New York, NY, USA, 2016.
- [30] J. C. Lee, P. H. Dietz, D. Maynes-Aminzade, R. Raskar, and S. E. Hudson. Automatic Projector Calibration with Embedded Light Sensors. In *Proceedings of ACM Symposium on User Interface Software and Technology*, UIST '04, pp. 123–126. ACM, New York, NY, USA, 2004.
- [31] J. C. Lee, S. E. Hudson, J. W. Summet, and P. H. Dietz. Moveable Interactive Projected Displays Using Projector Based Tracking. In *Proceedings of ACM Symposium on User Interface Software and Technology*, UIST '05, pp. 63–72. ACM, New York, NY, USA, 2005.
- [32] A. Majumder and M. S. Brown. *Practical Multi-projector Display Design*. A. K. Peters, Ltd./CRC Press, Natick, MA, USA, 2007.
- [33] M. R. Marner, R. T. Smith, J. A. Walsh, and B. H. Thomas. Spatial User Interfaces for Large-Scale Projector-Based Augmented Reality. *IEEE Computer Graphics and Applications*, 34(6):74–82, Nov 2014.
- [34] L. Miyashita, R. Yonezawa, Y. Watanabe, and M. Ishikawa. 3D Motion Sensing of Any Object Without Prior Knowledge. *ACM Transactions on Graphics*, 34(6):218:1–218:11, Oct. 2015.
- [35] M. Nagase, D. Iwai, and K. Sato. Dynamic Defocus and Occlusion Compensation of Projected Imagery by Model-Based Optimal Projector Selection in Multi-Projection Environment. *Virtual Reality*, 15(2):119–132, 2011.
- [36] 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*, PP(99):1–1, 2016.
- [37] P. Punpongsonon, D. Iwai, and K. Sato. Projection-Based Visualization of Tangential Deformation of Nonrigid Surface by Deformation Estimation Using Infrared Texture. *Virtual Reality*, 19(1):45–56, 2015.
- [38] R. Raskar, P. Beardsley, J. van Baar, Y. Wang, P. Dietz, J. Lee, D. Leigh, and T. Willwacher. RFIG Lamps: Interacting with a Self-describing World via Photosensing Wireless Tags and Projectors. *ACM Transactions on Graphics*, 23(3):406–415, Aug. 2004.
- [39] R. Raskar, H. Nii, B. deDecker, Y. Hashimoto, J. Summet, D. Moore, Y. Zhao, J. Westhues, P. Dietz, J. Barnwell, S. Nayar, M. Inami, P. Bekaert, M. Noland, V. Branzoi, and E. Bruns. Prakash: Lighting Aware Motion Capture Using Photosensing Markers and Multiplexed Illuminators. *ACM Transactions on Graphics*, 26(3), July 2007.
- [40] R. Raskar, J. van Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines. iLamps: Geometrically Aware and Self-Configuring Projectors. *ACM Transactions on Graphics*, 22(3):809–818, July 2003.
- [41] C. Resch, P. Keitler, and G. Klinker. Sticky Projections—A New Approach to Interactive Shader Lamp Tracking. In *Proceedings of IEEE International Symposium on Mixed and Augmented Reality*, pp. 151–156, Sept 2014.
- [42] 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 Transactions on Graphics*, 34(6):237:1–237:11, Oct. 2015.
- [43] M. Sugimoto, K. Kodama, A. Nakamura, M. Kojima, and M. Inami. A Display-Based Tracking System: Display-Based Computing for Measurement Systems. In *Proceedings of International Conference on Artificial Reality and Telexistence*, pp. 31–38, Nov 2007.
- [44] J. Tsukamoto, D. Iwai, and K. Kashima. Radiometric Compensation for Cooperative Distributed Multi-Projection System Through 2-DOF Distributed Control. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1221–1229, Nov 2015.
- [45] L. Yang, J. M. Normand, and G. Moreau. Practical and Precise Projector-Camera Calibration. In *Proceedings of IEEE International Symposium on Mixed and Augmented Reality*, pp. 63–70, Sept 2016.
- [46] Z. Zhang. A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, Nov 2000.
- [47] F. Zheng, R. Schubert, and G. Weich. A General Approach for Closed-Loop Registration in AR. In *Proceedings of IEEE Virtual Reality*, pp. 47–50, March 2013.
- [48] L. Zhou, S. Fukushima, and T. Naemura. Dynamically Reconfigurable Framework for Pixel-Level Visible Light Communication Projector. In *Proceedings of SPIE*, p. 89790J (14 pages), 2014.
- [49] S. Zollmann and O. Bimber. Imperceptible Calibration for Radiometric Compensation. In *Proceedings of Eurographics Short Papers*, pp. 61–64. The Eurographics Association, 2007.